Politecnico di Torino

Politecnico di Torino

Master's Degree in Mathematical Engineering
Academic Year 2025/2026

Beyond the Staircase Method: Model-Based Estimation of Fatigue Parameters under Limited Data.

The Case of High-Strength Bolts

Supervisors:
Prof. Mauro Gasparini
Prof. Carlo Rosso

Candidate:
Vera Sorrentino

Contents

List of Figures

Abstract

This thesis arises from an mechanical engineering problem addressing the critical evaluation of fatigue strength estimation procedures in very high cycle regimes. It investigates the possibility that ISO 12107 based procedures may underestimate the B_{50} parameter, which represents the median of the fatigue strength distribution, thereby producing conservative assessments of component performance. To this end, the study examines different methodological approaches to the analysis of B_{50}, clarifying the relationship between normative references, statistical assumptions, and the information content of the available experimental dataset.

After an overview of the ISO 12107 framework, the estimation of B_{50} is carried out through parametric S–N modeling based on the Stromeyer formulation, expressed within a bivariate normal framework that captures stress and fatigue life variability. The analysis is further extended to the Random Fatigue Limit (RFL) model, which accounts for intrinsic variability of the fatigue limit among specimens and provides a more flexible representation in high cycle regimes.

The work provides a critical assessment of the estimation process, highlighting how the evaluation of B_{50} is the outcome of specific modeling and inferential choices. The limited size of the dataset, consisting of fifteen specimens, represents the primary cause for high uncertainty in the estimation process, suggesting that a larger experimental basis would allow for a more stable characterization of the fatigue strength distribution.

Acknowledgements

I would like to express my sincere gratitude to Prof. Carlo Rosso for providing the experimental staircase fatigue data used in this thesis. By sharing these data and presenting the practical problem from which they originated, together with its industrial context, he made it possible to investigate the statistical aspects of fatigue testing and to analyze the phenomenon from a consistent statistical perspective.

I am deeply indebted to my supervisor, Prof. Mauro Gasparini, for his guidance and support throughout the development of this thesis. I am particularly grateful for his availability, patience, and continuous encouragement. His constructive comments and suggestions were extremely valuable and helped to clarify and strengthen several aspects of this work.

Chapter 1

Introduction: Statistical Evaluation of Fatigue Testing Methods

The design of reliable and durable mechanical components is a fundamental aspect of modern engineering, with direct economic and environmental implications. Among these components, high-strength fastening elements, such as fasteners, play a critical role in nearly every mechanical system and are frequently subjected to cyclic loads that determine their fatigue life.

The pioneering work of August Wöhler [1] introduced the fundamental tools for describing fatigue: the *S–N curve* and the concept of the *fatigue limit*, a stress level below which the material seemingly withstands loading indefinitely. However, experimental results consistently exhibit significant scatter, due to inherently stochastic phenomena related to microstructure, surface finish, defects, or residual stresses introduced during manufacturing (for example, differences between rolled and cut threads). For this reason, design cannot rely on deterministic values but must instead be grounded in a statistical description of fatigue strength.

This variability can be interpreted from two complementary perspectives:

1. the distribution of *fatigue life* (N) at a given stress level (S);
2. the distribution of *fatigue strength* (S) for a prescribed life (N).

For components such as screws, the second analysis is generally the most relevant: the engineering objective is not to ask "how long" a component lasts at a random load, but "what is the maximum safe stress level" (S) it can withstand for a target design life N_{ref} (e.g., 10^7 cycles).

To this end, industrial practice commonly employs the *staircase method* (or *up and down method*), which efficiently estimates the statistical distribution of fatigue strength, and in particular the median fatigue limit defined as the stress level for a 50% probability of failure (industrially termed B_{50}), and the standard deviation (σ), which quantifies the statistical scatter.

The need to compare results obtained from different laboratories, materials, or suppliers has made the development of shared procedures for planning and analyzing fatigue tests essential. In this context, the standard *ISO (International Organization for Standardization) 12107:2012, "Metallic materials - Fatigue testing - Statistical planning and analysis of data"* [2] was introduced to provide an official methodological framework for determining fatigue properties with adequate confidence and a practical number of specimens.

A detailed analysis of the standard, however, reveals a fragmentation of the statistical approach.

- **Staircase method** (Section 7): This procedure is designed to handle runouts and estimate the strength distribution at a fixed life (N_{ref}), but at the cost of discarding other information. In particular this method is analyzed using the Dixon and Mood analysis [3], important historically but outdated.

- **S-N curve estimation** (Section 8): It uses Least Squares regression to model the S-N relationship. Its critical drawback is that standard regression cannot handle "runout" data (unfailed specimens).

The contribution of this thesis is to critically analyze the causes of the inconsistencies inherent in this approach and to propose a unified statistical model capable of overcoming these limitations, providing a more robust and integrated characterization of the fatigue behavior of fastening elements.

1.1 Problem Statement

In the industrial context of fasteners manufacturing, the determination of the fatigue limit is carried out according to the ISO 12107:2012 standard, based on the *staircase* methodology. This approach provides the estimation of the mean (associated with the B_{50} value, corresponding to a 50% probability of failure) and of the standard deviation using the formulas of Dixon and Mood [3]. From these quantities, a lower tolerance limit for a given survival probability is then calculated according to the following relation:

$$B_p = B_{50} - k_p\,\hat{\sigma},$$

where B_{50} is the estimated mean, $\hat{\sigma}$ the estimated standard deviation, and k_p a statistical factor associated with the target survival probability p, to be discussed at the end of chapter 2. The lower tolerance limit employed by the company corresponds to the B_{90} value, representing the stress level at which a 90% survival probability is expected. However, this is not the final value used in practice, as it is further scaled by a safety factor, resulting in a final value that is lower than the initially estimated B_{90}.
However, industrial experience shows that the results obtained with this method are often overly conservative: several components tested in fatigue exhibit failures at stress levels higher than those predicted by the standard estimation.

At the same time, the S-N representation commonly used in the industrial field assumes the existence of a lower horizontal asymptote, interpreted as the fatigue strength corresponding to theoretically infinite life, known as the *fatigue limit*. Since this value cannot be directly measured, it is typically estimated in an engineering manner and, in the case of our client, approximated by the B_{90} value obtained from the *staircase* method. However, if this estimate proves excessively conservative, the resulting value may not realistically represent the actual stress level below which the component can operate without failure.

This framework therefore motivates the need to investigate alternative methodologies capable of estimating the fatigue limit and the B_{50} fatigue strength with greater realism and reduced conservatism.

1.1.1 Additional Contributing Factors

The literature indicates that the main statistical weakness of the method is the systematic *underestimation of the standard deviation*. This issue becomes particularly evident when only small sample sizes are available (typically $N < 30$), where the staircase procedure tends to yield a significantly biased estimate of the standard deviation, here denoted as $\hat{\sigma}$ [4, 5, 6].

While the underestimation of the variance represents the primary statistical limitation, several practical and methodological factors also contribute to producing a biased estimate of the mean. The literature identifies the following secondary mechanisms:

1. **Information Loss:** The staircase method is a pass/fail procedure. It uses censored data: if a specimen breaking at 10^5 cycles is treated identically to one breaking at 9×10^6 cycles (for $N_{\text{ref}} = 10^7$) and if it survive we don't know its actual fatigue life. This constitutes a substantial loss of information. Furthermore, the analysis often considers only the specimens of the less frequent event.

2. **Small Sample Size:** the sample size depends on the purpose of the test and the availability of test material. Fatigue tests are expensive, leading industry to use the minimum number of specimens. ISO 12107 suggests sample sizes of at least 28 specimens for reliability purpose, while a smaller number for exploratory test. Statistical literature consistently classifies these sample sizes "insufficient" for stable parameter estimation.

3. **Parameter Sensitivity:** With small N, the estimated distribution becomes highly sensitive to the initial stress level S_{init} and to the step size d.

4. **Limit of Applicability (Non-Homogeneity):** ISO 12107 explicitly states that its statistical framework is valid only for materials exhibiting *homogeneous behavior governed by a single failure mechanism.* If the tested screws experience mixed failure modes (e.g., both internal and surface initiation), applying a single-distribution staircase model may lead to biased estimates.

1.2 Methodology

This thesis investigates the statistical properties of traditional fatigue strength estimation procedures and evaluates alternative model-based approaches for the analysis of fatigue data in the presence of limited samples and runouts.

The methodology is structured into the following phases:

1. **Description of Traditional Fatigue Strength Estimation Methods.** The S-–N curve, the staircase testing procedure, and the Dixon–Mood estimator are introduced, with emphasis on their underlying assumptions and practical implementation. Their statistical formulation is outlined as a basis for the subsequent numerical investigation.

2. **Simulation-Based Evaluation of the Staircase Method.** Monte Carlo simulations are performed to study the statistical behavior of the staircase method under controlled conditions. The effects of key design parameters (step size, initial stress level, and sample size) on the bias and variability of the estimated mean and standard deviation are quantified.

3. **Assessment of Alternative Statistical Models for S–N Data.** Existing probabilistic S–N modeling approaches are analyzed and applied to the same context. In particular, the bivariate formulation and the Random Fatigue Limit (RFL) model are considered. These approaches allow the joint treatment of failures and runouts through likelihood-based inference and provide a coherent framework for estimating fatigue-strength distributions.

4. **Application to Experimental Fatigue Data.** The different methods are applied to real-world experimental datasets from fatigue tests on high-strength screws provided by an industrial partner. The results are compared in terms of stability, uncertainty quantification, and lower-tail fatigue strength estimation.

1.3 Structure of the Thesis

This thesis is organized as follows:

Chapter 1 (Introduction): Introduces the technical context of fatigue testing for screws, the governing ISO 12107 standard, and the central industrial problem of fatigue limit underestimation.

Chapter 2 (Literature Review: ISO's analysis): Reviews the statistical procedures recommended by the ISO 12107 standard and discusses the main fatigue concepts with particular attention to the assumptions underlying staircase-based estimation.

Chapter 3 (Simulation Methodology): Describes the construction of the Monte Carlo simulation model, including assumptions used for data generation and the implementation of the staircase protocol to analyze the estimation bias as a function of test parameters.

Chapter 4 (Bivariate Normal Model): Develops an alternative approach based on Maximum Likelihood Estimation (MLE) and a Stromeyer-type relationship, treating stress and fatigue life within a gaussian normal framework to use data more efficiently.

Chapter 5 (Random fatigue limit): Introduces the Random Fatigue Limit (RFL) model, in which the fatigue limit is treated as a specimen-specific random variable. Both frequentist and Bayesian formulations are presented.

Chapter 6 (Results and discussion): Applies the developed models to real-world experimental data of high-strength screws, comparing the estimates of B_{50} and the fatigue limit γ to evaluate their stability and reliability.

Chapter 7 (Conclusions and Future Work): Summarizes the main findings of the study, discusses the limitations imposed by the available data, and outlines possible directions for future research and experimental design improvements.

Chapter 2

Literature Review: Statistical Framework of ISO 12107

The International Organization for Standardization (ISO) is a worldwide federation of national standards bodies whose main task is to define International Standards. Its role is fundamental because these standards ensure consistency, reliability, and comparability of technical procedures and measurements across different countries and industries.

In this work, we analyze the second edition of *ISO 12107, Metallic materials - Fatigue testing -Statistical planning and analysis of data.*[2] This International Standard presents methods for the experimental design of fatigue tests and for the statistical analysis of the resulting data. Its objective is to determine the fatigue properties of metallic materials with both a high degree of confidence and a practical number of specimens.
One aspect that is not fully clarified in the ISO document is the use of logarithmic transformations for stress (S) or number of cycles (N); this point will therefore be examined and discussed in detail. Moreover, all statistical procedures prescribed in the standard implicitly rely on the assumption that the tested specimens form a representative random sample of the material population of interest, so that the estimated fatigue properties can be meaningfully interpreted at the population level.

Within this framework, a widely used operational approach for defining the fatigue limit is the *Constant Amplitude Fatigue Limit* (CAFL), where components are subjected to cyclic loading with fixed stress amplitude up to a predetermined life target. [7] In the specific industrial practice considered in this work, focusing on bolts, components are tested up to 5×10^6 cycles at constant amplitude. If no failure occurs within this life, the test is classified as a *runout*, and the component is considered to have survived the fatigue limit threshold for that specific load level. This procedure aligns with current best-practice guidelines for fatigue data analysis and provides the reference framework against which the ISO 12107 recommendations and alternative test methods (such as the Staircase procedure) will be critically assessed in the following sections. [7]

In this chapter, we define the relevant fatigue properties and outline the statistical methods recommended by the standard.

2.1 Fatigue properties

In the field of structural integrity design and analysis for metallic components, one of the most critical and complex failure mechanisms is that which derives from repeated loads over time. This mechanism, known as **fatigue**, is the process of progressive damage that manifests in materials subjected to cyclic loads, even when such loads are significantly lower than the material's yield strength or ultimate strength. Each stress cycle induces an accumulation of microscopic damage, primarily related to the mechanisms of microcrack nucleation and growth. As the cycles proceed, this damage evolves into the formation of a macroscopic crack and, finally, the fracture of the component. To analyse and predict this behaviour, experimental characterization is performed on specimens, i.e. carefully machined and surface-controlled samples. These specimens are tested under well-defined cyclic loading conditions so that the mechanisms of crack initiation and growth can be isolated, allowing the acquisition of reproducible and statistically meaningful fatigue data. The fatigue response of a material is represented by *S–N curves*, which correlates the stress amplitude (S) to the number of cycles to failure, N. From this description are derived the core concepts used in fatigue analysis:

- **Fatigue life**: number of stress cycles[1] applied to a specimen, at an indicated stress level, before it attains a failure[2] criterion defined for the test.

[1] Smallest segment of the stress time function which is repeated periodically.

[2] The failure can be defined in different ways, depending on the application. Examples includes

- **Fatigue strength**: value of stress level S at which a specimen would fail at a given number of cycles.

- **Fatigue limit**: fatigue strength at long life.

Because fatigue life and fatigue strength are influenced by numerous intrinsic and extrinsic factors, such as microstructure, defects, surface conditions, and variability in specimen preparation, fatigue data inevitably exhibit significant scatter, which typically increases as fatigue life increases. For this reason, statistical analysis plays an essential role in characterizing the fatigue response. Fatigue life and fatigue strength may, in fact, be treated as random variables, whose probability distribution allows for the definition of S–N curves at different failure probability levels, or for the estimation of the parameters of the strength distribution at a fixed number of cycles.

2.1.1 Distributions

There is no universal agreement in the literature on which probability family best models fatigue variability: fatigue life has often been modeled as log-normal (with $log(N)$ approximately normal), Weibull, or other distributions, while fatigue strength has been described using Weibull, normal, log-normal and additional alternatives. This diversity arises from differences in microstructure, surface conditions, defects, and testing protocols across datasets. The choice of a statistical model therefore often depends on the underlying mechanical assumptions and on the conventions associated with the adopted testing methodology. For mechanically complex components such as screws and threaded fasteners, the selection typically converges toward two main distribution families. Among the available alternatives, the log-normal distribution is widely used in fatigue analysis because it provides a convenient representation of the variability observed in fatigue data and is consistent with many statistical formulations of S–N models.
In this work we will consider:

- Fatigue life $\sim LogN(\mu_N, \sigma_N^2)$;

- Fatigue strength $\sim LogN(\mu_S, \sigma_S^2)$;

time to fracture of a specimen, crack initiation or a specimen experiencing irreversible deformation.

2.2 S-N curve

The S-N curve (or *Wöhler curve*) is a curve fitted to the fatigue life for p per cent survival values at each of several stress levels. It is an estimate of the relationship between applied stress and the number of cycles-to-failure that p per cent of the population would survive; p may be any number, such as 95, 90, etc. In [2] p=50% probability of failure. It describes the average or characteristic fatigue behaviour of a material under repeated loading and is one of the fundamental tools in fatigue analysis. The S-N curve is a deterministic functional model that captures how the mean fatigue life varies as a function of the applied stress level.

The experimental data points used to construct the curve typically exhibit significant specimen-to-specimen scatter. Consequently, the S-N curve represents only the mean trend, while the variability around it is handled separately through appropriate statistical models as said in the previous section.[8] Two main approaches can be used to model fatigue-life data:

- *Fatigue-life model specification:* define a model $F_N(t; S, \theta)$ for the fatigue life N conditional on a given stress amplitude S. This choice implicitly induces a corresponding model for fatigue strength.
- *Fatigue-strength model specification:* define a model $F_S(t; N, \theta)$ for the fatigue strength S conditional on a given number of cycles N. This choice, in turn, induces a corresponding model for fatigue life.

The first approach is the traditional one and remains the most widely adopted in practice, mainly because N is directly observable. The second approach is more recent and becomes especially appealing when the S-N relationship is nonlinear. [8] In this thesis, the first approach is adopted.
After selecting the modeling approach, two additional components must be specified:

- a functional S-N regression relationship, typically continuous and decreasing;
- a probabilistic model describing the scatter around the S-N relationship; in this work, a lognormal distribution is assumed.

2.2.1 Models Used in ISO 12107

ISO 12107 adopts a deliberately simple and standardised approach. The purpose of the standard is not to introduce the most accurate or mechanically sophisticated fatigue model, but to ensure that laboratories and industries produce results that are repeatable, comparable, and statistically consistent. For this reason, the standard mainly advocates for the use of *linear* (log-log) or *quadratic* regression models to describe the relationship between stress amplitude and fatigue life. The parameters for these models are typically estimated using the *least-squares method*, as the standard provides a straightforward procedural framework for their calculation. However, while the *ISO 12107* provides a robust practical guide, it acknowledges its own simplified nature by referencing more sophisticated models in its bibliography, such as the one proposed by *Bastenaire* and *Stromeyer*. Unlike the basic polynomial fits provided by the standard, the Bastenaire model (and similar curvilinear approaches) offers a more realistic representation of the curvature typically observed in S–N relationships, particularly in the transition region close to the fatigue limit. While the Stromeyer relationship, in particular, is useful when modelling S–N curves that exhibit the presence of a fatigue limit. Despite these mentions, the standard does not provide an analytical procedure for these advanced models, favoring mathematical simplicity over statistical optimality. This gap suggests that traditional ISO-compliant estimates may not be optimal, especially when compared to modern approaches like the *Random Fatigue-Limit Model (RFLM)* which utilizes *Maximum Likelihood Estimation (MLE)* to better account for the inherent run-outs and the stochastic nature of the fatigue limit. We will see them in 2.2.2.

2.2.1.1 Basquin model: Log-linear S–N model

The relationship between the applied stress amplitude S and the number of cycles to failure N is traditionally assumed to be linear in a logarithmic scale. This relationship, known as the *Basquin model* (or Wöhler-Basquin), is expressed as:

$$\log_{10} N = A + B \log_{10} S, \tag{2.1}$$

or, equivalently, in the power-law form

$$N = CS^B, \qquad \text{where } C = 10^A.$$

This formulation is the oldest and most widely adopted in fatigue literature due to its simplicity and the ease with which its parameters can be estimated via linear regression. As typically presented in engineering practice, the model represents the relationship between a specific quantile of the failure-time distribution (usually the median, or 50% probability of failure) and the stress level S.

The *ISO 12107* standard provides a formalized procedure for estimating the parameters of this linear relationship using the *least-squares method*. Defining $X = \log(S)$ as the independent variable and $Y = \log(N)$ as the dependent variable, the slope B and the intercept A are determined as follows:

$$B = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n}}$$

$$A = \frac{\sum_{i=1}^{n} Y_i - B \sum_{i=1}^{n} X_i}{n}$$

To evaluate the precision of the fit, the standard defines the estimated standard deviation $\hat{\sigma}$ for the regression:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n - l}}$$

where n is the sample size and l represents the number of parameters estimated in the model (in this linear case, $l = 2$). The denominator $n - l$ represents the degrees of freedom of the error.

Finally, the quality of the correlation is assessed through the coefficient of determination R^2, which represents the proportion of variation explained by the model:

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

Generally, a value of $R^2 \geq 0.9$ is considered indicative of an adequate fit for fatigue data.

While these equations provide a statistically consistent way to fit experimental data, they are inherently limited by the assumption of a constant variance and a purely linear trend in the log-log domain. In literature, this approach is often criticized because the *Basquin model* fails to capture the physical reality of the fatigue limit. By comparing these results with more advanced methodologies, such as the *Random Fatigue-Limit Model*, it can be demonstrated that the ISO-compliant regression may lead to non-optimal life predictions, particularly in the high-cycle fatigue regime where the curve should asymptotically approach a fatigue limit distribution.

2.2.1.2 The quadratic S–N model.

Beyond the linear Basquin relationship, the *ISO 12107* introduces a *quadratic S-N model* to account for potential curvatures in the experimental data. This model is expressed as a second-order polynomial in the log-log domain:

$$\log_{10} N = b_0 + b_1 \log_{10} S + b_2(\log_{10} S)^2 \tag{2.2}$$

where b_0, b_1, b_2 are the regression parameters. The solution for these parameters is obtained through linear algebra using the general linear problem matrix form:

$$b = (X'X)^{-1}X'Y \tag{2.3}$$

where b is the matrix of the calculated regression parameters, X is the matrix of independent variables, and Y is the matrix of dependent variables.

To quantify the variation and the quality of the fit, the standard introduces the following sums of squares:

- **Sum of Squares Total (R_{SST}):** measures the total variation of the data.

$$R_{SST} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

- **Sum of Squares Regression (R_{SSR}):** represents the variation explained by the quadratic model.

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

- **Sum of Squares Error (R_{SSE}):** accounts for the residual variation (scatter) not captured by the model.

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

The estimated standard deviation for the quadratic regression is then defined as:

$$\hat{\sigma} = \sqrt{\frac{R_{SSE}}{n - l}} \tag{2.4}$$

where $l = 3$ for the quadratic model, representing the three estimated parameters (b_0, b_1, b_2). Finally, the generalized coefficient of determination R^2 is calculated as:

$$R^2 = \frac{R_{SSR}}{R_{SST}} \tag{2.5}$$

While this formulation provides a better statistical fit (R^2) for non-linear data compared to the standard Basquin model, it is often criticized for its lack of physical basis. Unlike the *Stromeyer* or *Bastenaire* models, the quadratic fit does not represent a physical asymptotic fatigue limit; rather, it is a purely empirical approach.

Furthermore, as noted in the critique of ISO standards by authors like *Pascual*, these polynomial regressions remain limited by their inability to correctly handle censored data (run-outs) and their assumption of constant variance, which is rarely observed in the high-cycle fatigue regime.

2.2.2 Modern Models in the Literature

In contrast to the normative framework, recent scientific literature [9], [8] proposes a variety of more complex S-N models designed to better capture the behaviour in the low-cycle, high-cycle, and ultra-high-cycle regimes. Examples include the models of Kim–Zhang, Sendeckyj (in its various formulations), Kohout–Vechet, Richard, as well as energetic or damage-based approaches and modified exponential or power-law models.

These models introduce more realistic curvature in transition regions, provide improved descriptions of the potential fatigue limit, and achieve greater accuracy in the gigacycle regime, often with physically interpretable parameters. However, their application requires high-quality datasets, introduces increased computational and methodological complexity, and is difficult to standardise. For these reasons, despite their relevance in research, such models are not adopted in ISO 12107, which remains based on the traditional log-linear formulation to ensure consistency, simplicity, and reproducibility in industrial and normative contexts.

This work focuses on models with a horizontal asymptote, as this is characteristic of bolts.

2.2.2.1 The Stromayer relationship

In 1914 Stromayer introduced a fatigue-limit S-N model, which is a generalization of 2.1 and has the following relationship:

$$log(N) = \beta_0 + \beta_1 log(S - \gamma) + \sigma_N \epsilon, \qquad S_{min} > \gamma, \tag{2.6}$$

where ${\sigma_N}$[3] is a random-error term and ϵ has a location-scale distribution with $\mu = 0$ and $\sigma = 1$, and γ is known as *endurance-limit*. It describes for $\gamma > 0$ the concave up curvature commonly seen in S-N data when plotted on log log scales. When the stress S is low enough and it is close to γ, the lifetime is infinite, and it doesn't cause permanent damage.

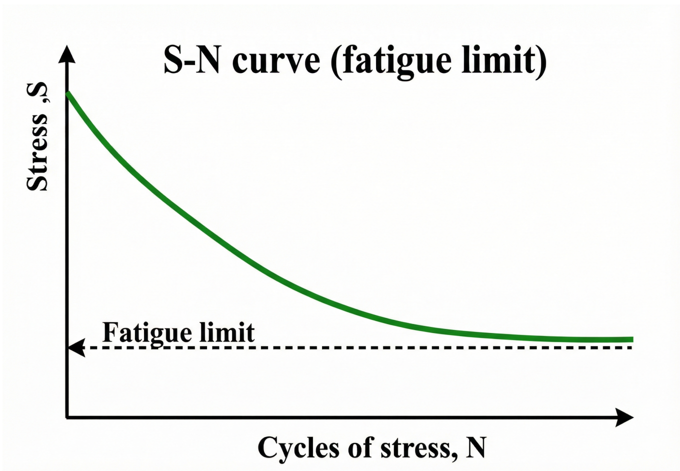


Figure 2.1: S–N with a fatigue limit γ

2.2.2.2 The Random fatigue limit

The Random Fatigue-Limit (RFL) model, proposed by Pascual and Meeker (1999) [10][4], represents a stochastic extension of the classical Stromeyer relationship (Eq. 2.6). This approach stems from the observation that assuming a constant fatigue limit γ for a whole population is often physically unreasonable, given the inevitable micro-structural inhomogeneities that affect threshold stress from specimen

[3]This model account for a constant σ_N, but was observed that often is necessary to be nonconstant.

[4]A revised version of this paper was later published in 2026 [11]

to specimen. Consequently, the RFL model treats γ as a random variable rather than a fixed material constant. Conditionally on a specific fatigue limit γ, the fatigue life N is modeled as:

$$\log(N) = \beta_0 + \beta_1 \log(S - \gamma) + \epsilon, \quad \text{for } S > \gamma \tag{2.7}$$

where S is the applied stress, ϵ is a random error term with constant scale σ_ϵ, and γ follows a distribution usually Lognormal with parameters μ_γ and σ_γ.

Since the individual limit γ is unobservable, the unconditional cumulative distribution function of fatigue life, $F_N(t; S)$, is obtained by integrating over the distribution of γ:

$$F_N(t; S) = \int_{-\infty}^{\log(S)} \frac{1}{\sigma_V} \Phi\left(\frac{\log(t) - [\beta_0 + \beta_1 \log(S - e^V)]}{\sigma_\epsilon}\right) \phi_\gamma\left(\frac{V - \mu_V}{\sigma_V}\right) dV \tag{2.8}$$

where $V = \log(\gamma)$. The RFL model naturally captures two critical physical phenomena without requiring complex variance functions:

- **Curvature:** It accurately describes the non-linear behavior of the S-N curve on log-log scales near the fatigue limit.
- **Non-constant scatter:** It predicts the increase in fatigue life variability at lower stress levels. As S approaches the random distribution of γ, the term $\log(S - \gamma)$ becomes highly sensitive, naturally modeling the larger scatter observed in the high-cycle fatigue regime.

2.2.2.3 The Nishijima S–N Hyperbola relationship

The Nishijima S-N relationship describes the fatigue behavior using a hyperbolic function on log-log scales defined as:

$$[\log(S) - E][\log(S) + A \log(N) - B] = C \tag{2.9}$$

where the regression parameters are $\beta = (A, B, C, E)$. To specify a fatigue-strength model suitable for inducing a fatigue-life model, this relationship is expressed explicitly for stress S as a function of cycles N:

$$S = h(N;\beta) = \exp\left(\frac{-A\log(N) + B + E + \sqrt{[A\log(N) - (B-E)]^2 + 4C}}{2}\right) \tag{2.10}$$

The physical interpretation of the parameters is as follows:

- A and B: Define the oblique asymptote for the high-stress (finite life) region. Specifically, A is the negative of the slope, and B is the intercept at $\log(N) = 0$.
- E: Represents the horizontal asymptote, corresponding to the fatigue limit (infinite life) on the logarithmic scale.
- C: Is a curvature parameter; $\sqrt{C}$ represents the vertical distance between the S-N curve and the intersection point of the two asymptotes (oblique and horizontal), determining the sharpness of the transition (the "knee") of the curve.

2.3 The Staircase Method

In sensitivity experiments, a fundamental constraint is the inability to perform more than one observation on a given specimen: once tested, the material's state is irreversibly altered. To address this limitation and optimize the estimation of the fatigue limit—thereby overcoming the experimental burden required by full S-N curves the *Staircase Method* [3] (also known as the *Up-and-Down Method*) was developed.

This approach is a sequential testing procedure specifically designed to efficiently estimate the population mean (μ_S), which corresponds to the 50% failure probability stress level (B_{50}) in the *long life* N_0 regime. The primary advantage of the method lies in its adaptive design: the test protocol converges and oscillates around the mean, thus maximizing the statistical information obtained from each specimen and minimizing tests in extreme failure-probability regions. Furthermore, under specific circumstances, the statistical analysis required is relatively simple compared to more traditional methods.

Regarding the sample size, the ISO 12107 standard specifies that while a *minimum* *of 15 specimens* is sufficient for exploratory research, obtaining reliable reliability data necessitates *at least 30 specimens*.

The experimental protocol is defined as follows:

1. An initial stress level (S_0), a *runout* cycle limit (N_0), and a fixed stress step (d) are selected.

2. For each i-th specimen tested at level S_i:

   - If the specimen **fails** (Failure, $N < N_0$), the subsequent test $(i + 1)$ is performed at level $S_{i+1} = S_i - d$. (rapresented by the $\times$ in the figure 2.2)
   - If the specimen **survives** (Runout, $N \geq N_0$), the subsequent test $(i + 1)$ is performed at level $S_{i+1} = S_i + d$. (rapresented by the $\circ$ in the figure 2.2)

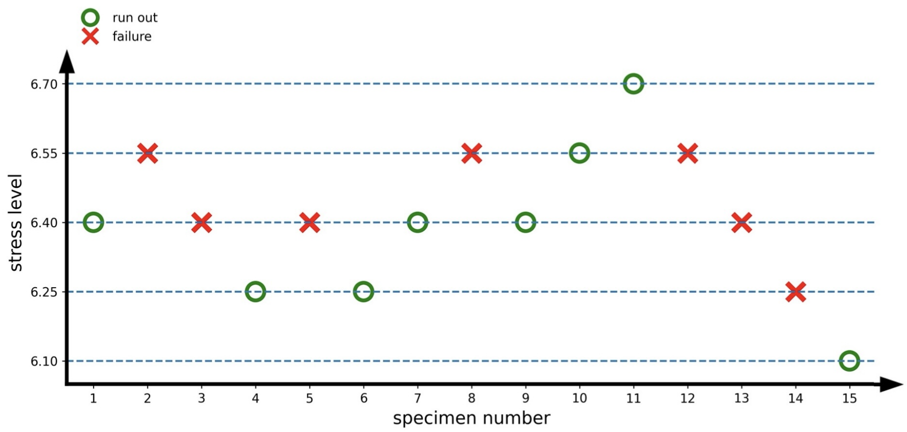


Figure 2.2: Up and down method

The classic statistical analysis of this method, developed by Dixon and Mood (1948) and adopted by the ISO 12107 standard, relies on two key assumptions:

1. The fatigue strength distribution is *Normal*.

2. The parameter analysis uses only the frequencies (f_i) of the *less frequent event* (either failures or runouts).

The estimates of the population mean ($\hat{\mu}_y$) and standard deviation ($\hat{\sigma}_y$) are calculated using the following formulas:

Mean Estimation:

$$\hat{\mu}_y = S_0 + d\left(\frac{A}{C} \pm \frac{1}{2}\right) \tag{2.11}$$

Standard Deviation Estimation:

$$\hat{\sigma}_y = 1,62d(D + 0,029) \quad \text{provided } D > 0.3 \tag{2.12}$$

Where:

- S_0 is the lowest stress level considered in the analysis.
- d is the stress step.
- f_i is the frequency of the less frequent event at level i (where $i = 0$ for S_0).
- $A = \sum i f_i$
- $C = \sum f_i$ (the total count of the less frequent event)
- $B = \sum i^2 f_i$
- $D = \frac{BC - A^2}{C^2}$

In the mean calculation, the $\pm 1/2$ term is *positive* $(+1/2)$ if the analyzed event is the *runout* (non-failure), and *negative* $(-1/2)$ if it is the *failure*.
The condition $D > 0.3$ is necessary to ensure the validity of the standard deviation approximation. When $D \leq 0.3$, the statistical method becomes unstable, often indicating that the chosen stress step d was too large relative to the actual standard deviation of the material.

In such cases, to avoid an unrealistic underestimation of the scatter, the literature [12] suggests fixing the standard deviation using the following lower bound:

$$\hat{\sigma}_y = 0.53d \quad \text{provided } D \leq 0.3 \tag{2.13}$$

The coefficient 0.53 (more precisely 0.533) is derived by substituting the threshold value $D = 0.3$ directly into Equation 2.12:

$$1.62 \cdot (0.3 + 0.029) \approx 0.533$$

The computations described in this section (including the one-sided tolerance factor k and the resulting lower tolerance limit) were implemented in R; the full notebook-style code is reported in Appendix A.

2.3.1 Proof of mean and standard deviation estimation

To compute the mean μ and standard deviation σ, assume that the fatigue strength S at a fixed number of cycles N_0 follows a normal distribution with parameters (μ, σ^2) and that tests are performed at discrete levels

$$S_i = S_0 \pm id, \qquad i = 0, 1, 2, \ldots,$$

with constant step size d and initial stress level S_0. Let n_i denote the number of runouts and m_i the number of failures observed at level S_i. The joint distribution of the counts (n_i, m_i) can be written as

$$\mathbb{P}(n, m \mid S_0, N_0) = K \prod_{i=-\infty}^{\infty} p_i^{n_i} q_i^{m_i}, \tag{2.14}$$

where K is a normalizing constant and

$$p_i = \int_{-\infty}^{S_i} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt = \Phi\left(\frac{S_i - \mu}{\sigma}\right), \qquad q_i = 1 - p_i.$$

In a staircase (up-and-down) test it holds that $|n_i - m_{i-1}| = 0$ or 1 for all i, so that either the sequence (n_i) or (m_i) contains almost all the information in the sample. Defining $N = \sum_i n_i$ and $M = \sum_i m_i$ with $N \leq M$, each observation contributing to n_i can be viewed as arising from a typical two-step sequence: a result at level S_{i-1} with probability q_{i-1}, followed by a result at level S_i with probability p_i. The contribution of level i can thus be grouped as

$$(p_i q_{i-1})^{n_i},$$

and, neglecting the few initial observations used only to move from S_0 to the vicinity of the mean, expression (2.14) can be rewritten in the compact form

$$\mathbb{P}(n, m \mid S_0, N_0, M - N) = K' \prod_i (p_i q_{i-1})^{n_i}, \tag{2.15}$$

where K' is another normalizing constant.[5]

[5] Even if $M - N$ is not small, only a small amount of information is neglected, because these observations are spent in moving from the initial level S_0 to the region of the mean and therefore contribute little to the precise location of μ [3].

Starting from (2.15), the likelihood can be written as

$$L(\mu, \sigma) = K' \prod_i (p_i q_{i-1})^{n_i},$$

with $p_i = \Phi\big((S_i - \mu)/\sigma\big)$ and $q_i = 1 - p_i$, and the corresponding log-likelihood is

$$\ell(\mu, \sigma) = \log L(\mu, \sigma) = \sum_i n_i \Big(\log p_i + \log q_{i-1} \Big).$$

Introducing the standard notation

$$x_i = \frac{S_i - \mu}{\sigma}, \qquad p_i = \Phi(x_i), \qquad q_i = 1 - \Phi(x_i), \qquad z_i = \varphi(x_i),$$

where φ and Φ denote, respectively, the pdf and cdf of the standard normal distribution, the following identities will be used:

$$\frac{\partial}{\partial x} \log \Phi(x) = \frac{\varphi(x)}{\Phi(x)} = \frac{z(x)}{p(x)}, \qquad \frac{\partial}{\partial x} \log \Big(1 - \Phi(x)\Big) = -\frac{\varphi(x)}{1 - \Phi(x)} = -\frac{z(x)}{q(x)}.$$

Since $x_i = (S_i - \mu)/\sigma$, it follows that

$$\frac{\partial x_i}{\partial \mu} = -\frac{1}{\sigma}, \qquad \frac{\partial x_i}{\partial \sigma} = -\frac{S_i - \mu}{\sigma^2} = -\frac{x_i}{\sigma}.$$

Differentiating the log-likelihood with respect to μ yields

$$\frac{\partial \ell}{\partial \mu} = \sum_i n_i \left(\frac{\partial \log p_i}{\partial x_i} \frac{\partial x_i}{\partial \mu} + \frac{\partial \log q_{i-1}}{\partial x_{i-1}} \frac{\partial x_{i-1}}{\partial \mu} \right) = \frac{1}{\sigma} \sum_i n_i \left(\frac{z_{i-1}}{q_{i-1}} - \frac{z_i}{p_i} \right),$$

so that the first likelihood equation (score equation for μ) is

$$\sum_i n_i \left(\frac{z_{i-1}}{q_{i-1}} - \frac{z_i}{p_i} \right) = 0. \tag{2.16}$$

Similarly, differentiation with respect to σ gives

$$\frac{\partial}{\partial \sigma} \log p_i = \frac{z_i}{p_i} \left(-\frac{x_i}{\sigma} \right) = -\frac{x_i z_i}{\sigma\, p_i}, \qquad \frac{\partial}{\partial \sigma} \log q_{i-1} = \left(-\frac{z_{i-1}}{q_{i-1}} \right) \left(-\frac{x_{i-1}}{\sigma} \right) = \frac{x_{i-1} z_{i-1}}{\sigma\, q_{i-1}},$$

and therefore

$$\frac{\partial \ell}{\partial \sigma} = \frac{1}{\sigma} \sum_i n_i \left(\frac{x_{i-1} z_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i} \right). \tag{2.17}$$

The exact score equation for σ is obtained by setting $\partial \ell / \partial \sigma = 0$ in (2.17). Equations (2.16) and (2.17) define the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$.

To study the properties of these estimators, the expected values $E(n_i)$ are considered in place of the observed counts n_i. The up-and-down scheme implies the recursion

$$\frac{E(n_{i+1})}{q_i} = \frac{E(n_i)}{p_i}, \tag{2.18}$$

which expresses a balance of the expected flow of observations between adjacent stress levels. In fact, at level i there are on average $E(n_i)$ tests, each of which leads to a "success" (runout) with probability p_i or to a "failure" with probability q_i. Thus, the expected number of transitions from S_i to S_{i+1} is $E(n_i)p_i$, whereas the expected number of transitions from S_{i+1} back to S_i is $E(n_{i+1})q_i$. In the (quasi-)stationary regime of the staircase, these two flows must balance, so that $E(n_i)p_i = E(n_{i+1})q_i$, which is equivalent to (2.18).

This recursion can be solved by introducing the weights

$$w_i = \mathbf{1}_{\{i=0\}} + \mathbf{1}_{\{i>0\}} \prod_{j=0}^{i-1} \frac{q_j}{p_j} + \mathbf{1}_{\{i<0\}} \prod_{j=i}^{-1} \frac{p_j}{q_j},$$

so that

$$E(n_i) = N \frac{w_i}{\sum_{k=-\infty}^{\infty} w_k}, \tag{2.19}$$

where $N = \sum_i n_i$ is the total number of informative observations in the staircase sequence.

Because the stress levels are equally spaced ($S_i = S_{i-1} + d$), the standardized levels satisfy

$$x_i = \frac{S_i - \mu}{\sigma} = x_{i-1} + \frac{d}{\sigma}.$$

The ratio

$$r = \frac{d}{\sigma}$$

represents the step size expressed in units of the standard deviation and plays a central role in the behaviour of the estimators.

Following Dixon and Mood, for a generic $x \in \mathbb{R}$ and $r = d/\sigma$ the functions

$$\alpha(u) = \frac{z(x)}{q(x)} - \frac{z(x+r)}{p(x+r)}, \qquad u = x + \frac{r}{2}, \tag{2.20}$$

$$\beta(u) = \frac{x\, z(x)}{q(x)} - \frac{(x+r)\, z(x+r)}{p(x+r)}, \tag{2.21}$$

are introduced, where

$$z(x) = \varphi(x), \qquad p(x) = \Phi(x), \qquad q(x) = 1 - \Phi(x).$$

The terms appearing in the score equations (2.16) and (2.17) can be expressed as particular evaluations of $\alpha(u)$ and $\beta(u)$ at consecutive values of x and $x + r$ [3]. It can be shown numerically that, for moderate step size $d < 2\sigma$ (i.e. $r = d/\sigma < 2$), the function $\alpha(u)$ is nearly linear in u, whereas $\beta(u)$ is nearly quadratic in u, as indicated by the behaviour of its first derivative $\beta'(u)$ (see Figures 2.3 and 2.4, obtained by numerical evaluation of (2.20) and (2.21)).

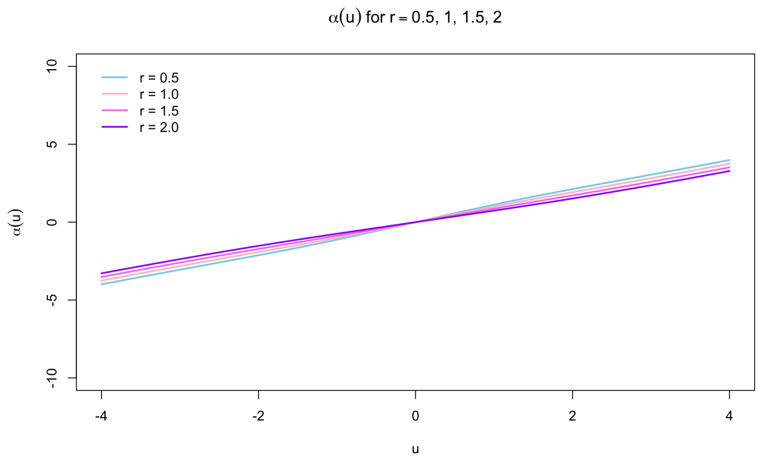


Figure 2.3: Example of the function $\alpha(u)$ for different values of $r = d/\sigma$ (here $r = 0.5, 1.0, 1.5, 2.0$). The curves are almost straight lines, confirming that $\alpha(u)$ is approximately linear in u when $d < 2\sigma$.

Figure 2.3 and 2.4 confirm the key approximation used by Dixon and Mood: for practical step sizes $d < 2\sigma$ the functions $\alpha(u)$ and $\beta(u)$ can be accurately represented by low-order polynomials in u, with coefficients that depend only on the ratio $r = d/\sigma$.

Under this approximation, the contribution at level i in (2.17) can be rewritten in a form involving the difference $x_i - x_{i-1} = d/\sigma$:

$$\frac{x_{i-1}z_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i} \approx \frac{x_i - x_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i},$$

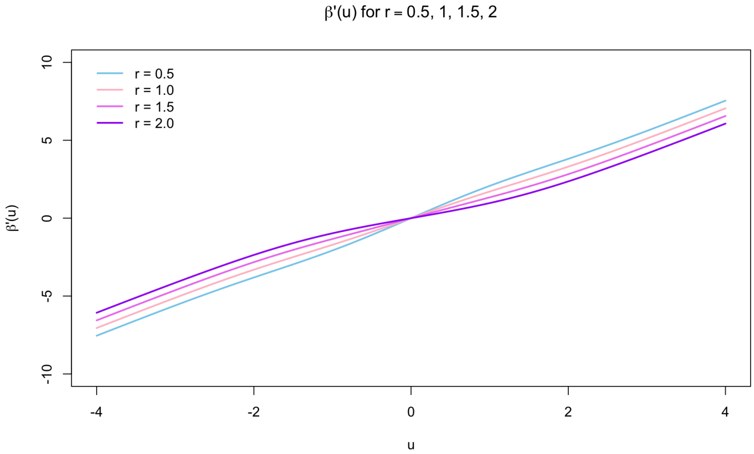


Figure 2.4: Example of the derivative $\beta'(u)$ for different values of $r = d/\sigma$. The almost linear behaviour of $\beta'(u)$ in u implies that $\beta(u)$ is nearly quadratic, consistently with the Dixon-Mood analysis for $d < 2\sigma$.

so that the score equation for σ can be expressed as

$$\sum_i n_i \left(\frac{x_i - x_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i} \right) \approx 0. \tag{2.22}$$

The likelihood equations (2.16) and (2.22) do not admit a simple closed-form solution,but their structure, together with the approximations based on $\alpha(u)$ and $\beta(u)$, shows that the estimators $\hat{\mu}$ and $\hat{\sigma}$ depend essentially on the first two moments of the tested stress levels. To make this dependence explicit, consider the weighted moments

$$\bar{y}_1 = \frac{1}{N} \sum_i n_i y_i, \qquad \bar{y}_2 = \frac{1}{N} \sum_i n_i y_i^2. \tag{2.23}$$

Using the properties of the up-and-down sequence and the approximations based on $\alpha(u)$ and $\beta(u)$, Dixon and Mood show that

$$E(\bar{y}_1) = \mu - \frac{d}{2}, \tag{2.24}$$

and

$$E(\bar{y}_2) - E^2(\bar{y}_1) + \frac{d^2}{4} = \sigma^2. \tag{2.25}$$

To obtain (2.24)–(2.25) is been used (2.19)

$$E(\bar{y}_1) = \sum_i \frac{w_i}{\sum_k w_k} y_i, \qquad E(\bar{y}_2) = \sum_i \frac{w_i}{\sum_k w_k} y_i^2. \tag{2.26}$$

In other words, $E(\bar{y}_1)$ and $E(\bar{y}_2)$ are the first and second moments of the discrete distribution on the levels y_i with stationary weights proportional to w_i. Since the underlying fatigue strength is normal with mean μ and variance σ^2, the standardized levels can be written as

$$y_i = \mu + x_i\sigma, \qquad x_i = x_0 + i\,\frac{d}{\sigma},$$

and the weights w_i define a symmetric discrete distribution in the x-space, concentrated around $x = 0$ but supported on the grid with spacing $r = d/\sigma$. Because the grid does not include the exact mean μ, the stationary distribution is centred halfway between two adjacent levels. This implies that the mean of the discrete distribution is shifted by half a step, leading to

$$E(\bar{y}_1) = \mu - \frac{d}{2},$$

which is (2.24) and it shows that the sample mean of the levels is biased by exactly half a step due to the discrete nature of the testing grid. Correcting for this bias yields the classic mean formula (2.11):

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

For the second moment, we write

$$E(\bar{y}_2) - E^2(\bar{y}_1) = \sum_i \frac{w_i}{\sum_k w_k}\, y_i^2 - \left(\sum_i \frac{w_i}{\sum_k w_k}\, y_i\right)^2,$$

that is, the variance of the discrete distribution on the y_i with weights w_i. Re-expressing y_i in terms of x_i and using the approximations based on $\alpha(u)$ and $\beta(u)$ for $d < 2\sigma$, Dixon and Mood show that this discrete variance underestimates the continuous variance σ^2 by approximately $d^2/4$, so that

$$E(\bar{y}_2) - E^2(\bar{y}_1) + \frac{d^2}{4} = \sigma^2,$$

which is (2.25). The correction term $d^2/4$ compensates for the half-step bias in the mean and for the discretization of the normal distribution on the staircase grid. These relations ((2.24)–(2.25)) link the weighted sample moments of the staircase levels to the underlying normal parameters μ and σ^2, and form the basis of the practical Dixon–Mood estimators used in fatigue and sensitivity testing.[3] Equation 2.24 explains the origin of the term $\pm 1/2$ in the final mean formula (2.11): the stationary distribution of tests is shifted by half a step relative to the grid.

For the standard deviation, the relationship is more complex. The term $E(\bar{y}_2) - E^2(\bar{y}_1)$ corresponds to the variance of the observed levels. However, simply using the sample variance is inaccurate due to the discrete nature of the grid. Dixon and Mood plotted the theoretical relationship between the pseudo-variance $\frac{NB-A^2}{N^2}$ and the true σ for various step sizes d. They observed that for $d < 2\sigma$, the relationship is approximately linear [3].

By fitting a line to this theoretical curve, they obtained the empirical coefficients:

$$\text{Slope} \approx 1.62, \qquad \text{Intercept} \approx 0.029.$$

This linearization allows the implicit Maximum Likelihood solution to be replaced by the explicit formula (2.12) utilized in the standard:

$$\hat{\sigma} \approx 1.62 \cdot d \cdot \left(\frac{NB - A^2}{N^2} + 0.029 \right).$$

This explains why the formula contains these specific numerical constants: they are the result of a linear approximation of the likelihood solution in the region of interest ($d/\sigma \approx 1$).

Asymptotic accuracy of the Dixon-Mood estimators

The previous analysis has shown that the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ can be expressed in terms of the weighted moments (2.23) and that their expectations satisfy (2.24)–(2.25). To quantify the accuracy of these estimators, Dixon and Mood investigate the expected second derivatives of the log-likelihood

$$L = \log P,$$

where P is the likelihood in the compact form (2.15). The relevant quantities are

$$E\left[-\frac{\partial^2 L}{\partial \mu^2}\right], \qquad E\left[-\frac{\partial^2 L}{\partial \sigma^2}\right], \qquad E\left[-\frac{\partial^2 L}{\partial \mu\, \partial \sigma}\right].$$

Using the score equations (2.16), (2.17), the expressions for $E(n_i)$ in (2.19) and the properties of the weights w_i, these expectations can be written in the form

$$E\left[-\frac{\partial^2 L}{\partial \mu^2}\right] = -\frac{N}{\sigma^2} \sum_i \omega_i \left(\frac{z_{i-1}^2}{q_{i-1}^2} + \frac{x_i^2}{p_i^2} \right) / \sum \omega_i = \frac{N}{\sigma^2} G(r), \tag{2.27}$$

$$E\left[-\frac{\partial^2 L}{\partial \sigma^2}\right] = -\frac{N}{\sigma^2}\sum_i \omega_i \left(\frac{x_{i-1}^2 z_{i-1}^2}{q_{i-1}^2} + \frac{x_i^2 z_i^2}{p_i^2}\right) / \sum \omega_i = \frac{N}{\sigma^2} H(r), \tag{2.28}$$

$$E\left[-\frac{\partial^2 L}{\partial \sigma^2}\right] = -\frac{N}{\sigma^2}\sum_i \omega_i \left(\frac{x_{i-1} z_{i-1}^2}{q_{i-1}^2} + \frac{x_i z_i^2}{p_i^2}\right) / \sum \omega_i, \tag{2.29}$$

where $r = d/\sigma$ is the dimensionless step size and $G(r)$, $H(r)$ are functions tabulated and plotted in 2.5. [3]. The mixed second derivative does not vanish exactly, but it

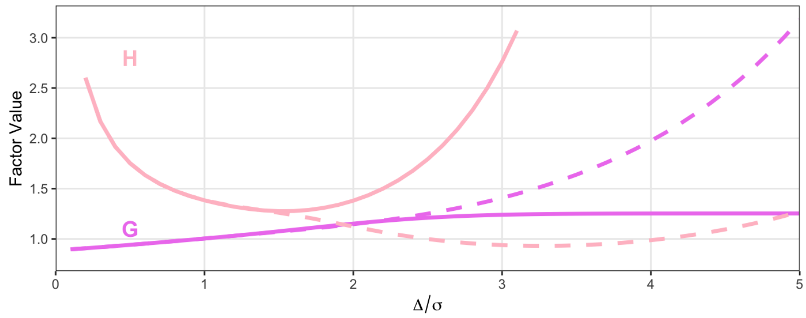


Figure 2.5: The corrective factors G and H as functions of the normalized step size d/σ. The curves illustrate the dependency of the estimates' asymptotic standard error on the position of the true mean relative to the testing levels: the **solid line** represents the scenario where the true mean falls exactly on a testing level, while the **dashed line** corresponds to the mean falling midway between two adjacent levels.

has been shown that the corresponding correlation between $\hat{\mu}$ and $\hat{\sigma}$ is extremely small: of order 10^{-4} when $d = \sigma$ and about 10^{-2} when $d = 2\sigma$.[3] This implies that the off-diagonal element of the information matrix is negligible compared with the diagonal terms, so that, for practical purposes, the covariance between $\hat{\mu}$ and $\hat{\sigma}$ can be ignored and the information matrix may be treated as diagonal with entries given by (2.27) and (2.28). So

$$\sigma_u^2 = G^2 \frac{\sigma^2}{N} \qquad \sigma_d^2 = H^2 \frac{\sigma^2}{N}. \tag{2.30}$$

Under this approximation, the asymptotic variance-covariance matrix of $(\hat{\mu}, \hat{\sigma})$ is approximately the inverse of the information matrix, so that

$$\mathrm{Var}(\hat{\mu}) \approx \frac{\sigma^2}{N\,G(r)}, \qquad \mathrm{Var}(\hat{\sigma}) \approx \frac{\sigma^2}{N\,H(r)}. \tag{2.31}$$

These expressions show that the precision of the Dixon–Mood estimators increases with the number of informative observations N and depends sensitively on the ratio $r = d/\sigma$ through the functions $G(r)$ and $H(r)$ plotted in figure 2.5.

2.3.2 Lower tolerance limit

Through the staircase method, it is possible to determine the lower tolerance limit in accordance with ISO 12107. This limit is utilized by engineers to define the fatigue strength threshold.

The lower tolerance limit relates the estimated mean B_{50} and the sample standard deviation $\hat{\sigma}$ to a stress level B_p that is expected to be exceeded by a specified fraction P of the population with a given confidence level $1 - \alpha$:

$$B_p = \bar{X} - k_p \, \hat{\sigma}.$$

The factor k_p, often called the *tolerance factor*. For small sample sizes ($n < 30$), k_p cannot be approximated simply by a standard normal quantile; instead, it is derived by combining the distributions of the sample mean (t-Student) and the sample variance (chi-squared). A general expression for k_p is:

$$k_p = \frac{t_{\gamma,\nu}\sqrt{\nu\left(1 + \frac{1}{n}\right)}}{\sqrt{\chi^2_{1-P,\nu}}},$$

where:

- n is the sample size which belong to the rare event,
- $\nu = n - 1$ is the number of degrees of freedom,
- $t_{\gamma,\nu}$ is the γ-quantile of the t-Student distribution with ν degrees of freedom (associated with the confidence level $1 - \alpha$),
- $\chi^2_{1-P,\nu}$ is the $(1 - P)$-quantile of the chi-squared distribution with ν degrees of freedom.

As n increases, the uncertainty in $\hat{\sigma}$ decreases and k_p approaches the standard normal quantile corresponding to the desired coverage P.

In practice, ISO 12107 provides tables of k_p values for common confidence and coverage levels.

Chapter 3

Critical analysis of Staircase method: Underestimation of the fatigue strength

The staircase (or "up-and-down") method, originally introduced by Dixon and Mood [3] in 1948 and later standardized (for example in ISO 12107:2012), represents one of the most widespread procedures in industrial applications for determining the fatigue limit of a material. Its popularity stems from its efficiency: by concentrating the tests near the median stress level (B_{50}), the method allows for an estimation of the mean fatigue strength value using a relatively small number of specimens.

However, although the scientific literature[1] widely agrees that the staircase method provides an accurate and un-biased estimate of the mean (B_{50}), its reliability significantly degrades when the objective is to estimate the dispersion (or scatter) of the data, namely the standard deviation (σ). To analyze this phenomenon, a Monte Carlo simulation was used in this work, following the approach discussed in [6].

3.1 Monte Carlo Simulation

A Monte Carlo simulation was implemented to quantify the accuracy and bias of the Dixon-Mood method, with a specific focus on the underestimation of the

[1] [5], [4], [6]

standard deviation.

As a foundational assumption, consistent with the Dixon-Mood analysis and the procedure described in the ISO 12107 standard, the "true" population of the specimens was assumed to follow a Normal distribution of fatigue strength. The fatigue limit region was modeled as a series of equally spaced stress levels, separated by a distance *d* and the first specimen can be placed at an arbitrary level.
To run the simulation, a known "truth" must be defined: a true mean (μ_{true}) and a true standard deviation (σ_{true}) representing the population. Knowledge of these parameters allows the exact calculation of the failure probability ($P_{failure}$) for any given stress level via the Normal cumulative distribution function.

The simulation of a single staircase test faithfully replicates the experimental procedure (Fig 3.1) :

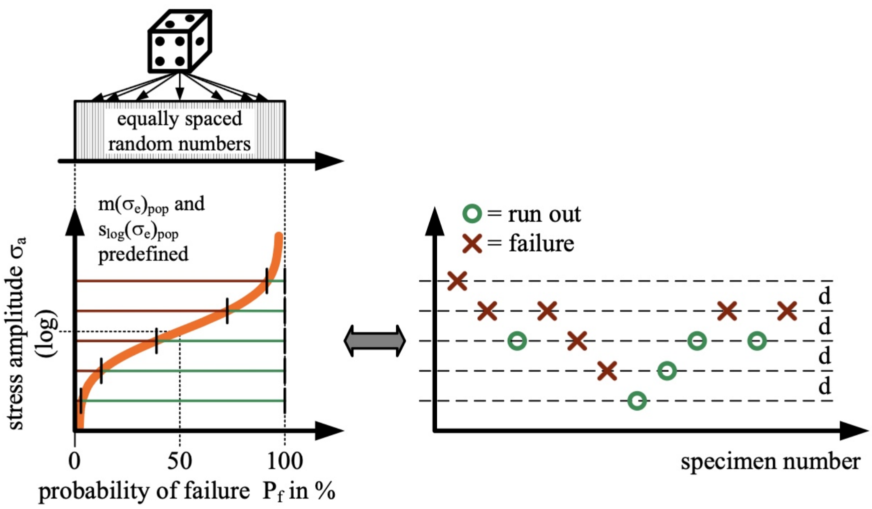


Figure 3.1: Simulation of staircase method using Monte Carlo simulation. [6]

1. **Start:** The test begins at a starting stress level, start_ level.

2. **Virtual Test:** For each specimen, the outcome is determined by generating a uniformly distributed random number between 0 and 1.

3. **Outcome:** This random number is compared to the failure probability ($P_{failure}$) calculated at the current stress level. If the random number is less than $P_{failure}$, the specimen "fails" (outcome 1); otherwise, it is a "run-out" (outcome 0).

4. **Staircase Rule:** The stress level for the next specimen is determined by the standard rule: it is decreased by a step d in case of failure (outcome 1) or increased by d in case of a run-out (outcome 0).

This process is repeated for the established number of specimens n, generating a single complete virtual test, which is then analyzed.

To quantify not only the bias but also the dispersion of the estimates produced by the Dixon-Mood staircase method, this work makes systematic use of the interquartile range (IQR), defined as the interval between the 25th and 75th percentiles of the sampling distribution. The IQR is a robust, distribution-free measure of variability and is particularly well suited for Monte Carlo analyses, as it is insensitive to extreme outliers that may occasionally appear in simulated tests that fail the $D > 0.3$[2] validity criterion. In the context of this study, the IQR is computed separately for both the estimated mean fatigue strength ($\widehat{\mu}$) and the estimated standard deviation ($\widehat{\sigma}$). For each simulation scenario (fixed sample size n or fixed step-size ratio d/σ_{true}), the Monte Carlo replications produce an empirical distribution of $\widehat{\mu}$ and $\widehat{\sigma}$. The corresponding 25th percentile (Q_1), median (Q_2), and 75th percentile (Q_3) are then extracted:

$$Q_1 = \text{quantile}(\widehat{\theta}, 0.25), \qquad Q_2 = \text{median}(\widehat{\theta}), \qquad Q_3 = \text{quantile}(\widehat{\theta}, 0.75),$$

where $\widehat{\theta}$ denotes either $\widehat{\mu}$ or $\widehat{\sigma}$.

These three statistics are used to construct the *ribbon plots* shown in the results section 3.1.2 the median provides a measure of central tendency, while the IQR band ($[Q_1, Q_3]$) visually represents the core variability of the estimator. Boxplots generated for the same data complement these ribbon plots and highlight the shape of the sampling distribution more explicitly.

Using the IQR instead of the full standard deviation of the Monte Carlo samples offers two key advantages:

- avoids giving disproportionate weight to the occasional highly biased estimates produced by invalid tests;
- it provides a stable, reproducible summary of estimator variability even when the distribution of $\widehat{\sigma}$ is skewed due to systematic underestimation.

[2]is necessary for the reliability of Dixon-Mood method as explained in the 2.3

This makes the IQR a particularly appropriate tool for evaluating the performance of the staircase method.

3.1.1 Simulation Parameter Setup

To make the Monte Carlo simulation realistic, it is crucial to model how the test parameters are chosen. In practice, the initial stress level is a "qualified estimate" based on the test engineer's experience.

To account for this uncertainty, the simulation model does not use a fixed starting level. Instead, the initial level (*start_level*) is treated as a random variable. For each simulated test, the starting level is sampled from a Normal distribution centered on the "true" population mean (*mu_true*). The standard deviation for this sampling is set as 7% of the true mean (*mu_true*), following the value suggested by Müller [6] to represent this typical estimation error. The inclusion of a randomized starting level makes the simulation highly realistic.
Furthermore, predicting the true standard deviation (σ_{true}) is difficult, as it can depend on process, material, or component variations. For this reason, randomizing this parameter as well could enhance realism, as discussed in [6]. However, for the scope of this analysis, this parameter is held at a fixed value of $\sigma_{true} = 20$ MPa. The true mean for the population was set to $\mu_{true} = 500$ MPa.

The σ_{true} parameter is critical because, as the simulation results will demonstrate, the accuracy of the estimated standard deviation strongly depends on the ratio between the chosen step size d and this true standard deviation (d/σ_{true}) although this would require an additional layer of randomness in the simulation design. [2] As noted in ISO 12107, this ratio should ideally be within the range of 0.5 to 2 for the analysis to be valid.

Our simulation investigates the accuracy of the Dixon-Mood estimates by analyzing two primary scenarios:

- The dependency on the **sample size (n)**. This analysis was run for $n = \{10, 12, 15, 20, 25, 30, 40, 50\}$, while holding the ratio d/σ_{true} at the optimal value of 1.0.

- The dependency on the **step size ratio (d/σ_{true})**. This analysis was run for ratios of $\{0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0\}$, while holding the sample size fixed at $n = 15$.

3.1.2 Results

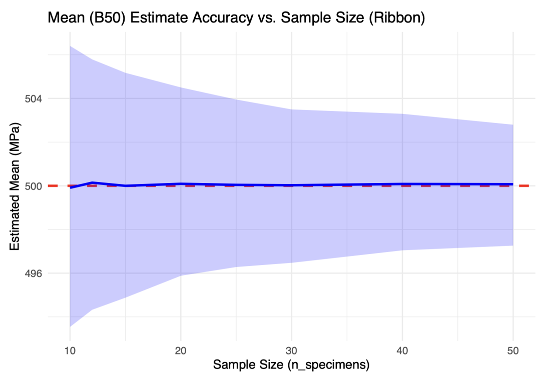
The Monte Carlo simulation was used to analyse the accuracy and robustness of the Dixon-Mood staircase method in two complementary scenarios:
(i) varying the sample size n while keeping the step-size ratio fixed at its optimal value ($d/\sigma_{\text{true}} = 1$)
(ii) varying the step-size ratio while keeping the number of specimens fixed ($n = 15$).
For each scenario, the simulation generated the sampling distributions of the estimated mean fatigue strength ($\widehat{\mu}$) and standard deviation ($\widehat{\sigma}$), together with the fraction of tests that met the Dixon-Mood validity criterion ($D > 0.3$). The statistical behaviour of the estimators is summarized using the median and the interquartile range (IQR), which provide a robust characterization of the central tendency and variability of the estimates.

3.1.2.1 First scenario: varying the sample size

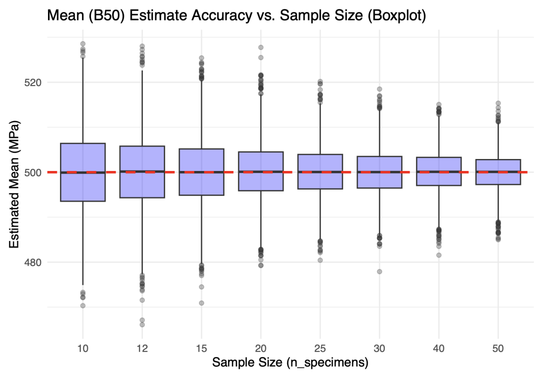
The first scenario is based on the evolution of the staircase estimators as the number of specimens n increases from $n = 10$ to $n = 50$, with the step-size ratio fixed at $d/\sigma_{\text{true}} = 1$.

- **Estimated mean B_{50}**: The results confirm the robustness of the staircase method for estimating the mean. As shown in Figures 3.2a and 3.2b, the median of the estimates (the blue line in the ribbon plot and the center line in the boxplots) remains consistently centered on the "true" value μ_{true} of 500 MPa. The results table 3.1 confirms that the median bias is negligible, regardless of the number of specimens used. The only visible effect of n is, as expected, a reduction in the dispersion of the estimates as the sample size increases.

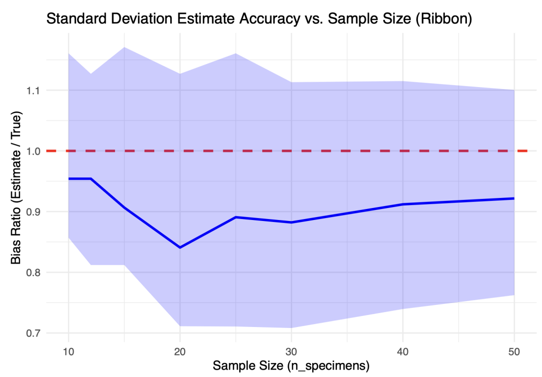
- **Estimated standard deviation $\hat{\sigma}$**: The analysis of the standard deviation reveals, on the contrary, significant critical issues. The "Bias Ratio" shown in Figures 3.2c and 3.2d highlights a clear tendency for underestimation. For



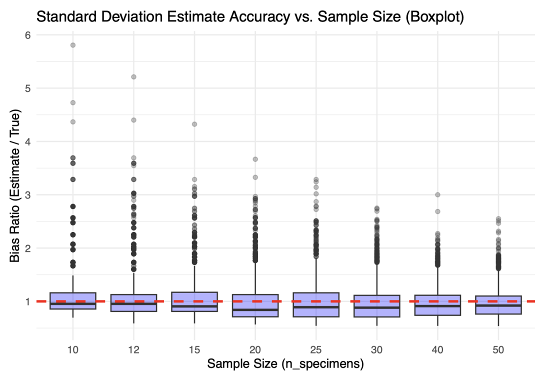
(a) Ribbon of estimated mean vs n



(b) Boxplot of estimated mean vs n



(c) Ribbon of estimated standard deviation vs n



(d) Boxplot of estimated standard deviation vs n

Figure 3.2: Results of first scenario.

$n = 15$, the median estimate is already about 5% lower than the true value (ratio 0.954, as can be seen in Table 3.1), and this underestimation worsens slightly as the sample size increases, settling at a negative bias of about -8% to -10% for $n \geq 40$. However, the most obvious problem is the enormous dispersion of the estimates for low n.

- **Test validity** ($D > 0.3$): A fundamental result of this simulation is the percentage of "valid" tests, i.e., those that respect the $D > 0.3$ condition imposed by the ISO 12107 standard for the validity of the $\hat{\sigma}$ estimate. As shown in the graph in Figure 3.3, with small samples, most tests fail this criterion. For example, with $n = 10$, only 6.16% of the tests are valid, while with $n = 15$, only 15.54% (see Table 3.1). To exceed 50% validity, 30 specimens are required. This implies that, with small samples, the estimate of $\hat{\sigma}$ is not only imprecise but, in most cases, would not even be considered valid according to the standard.

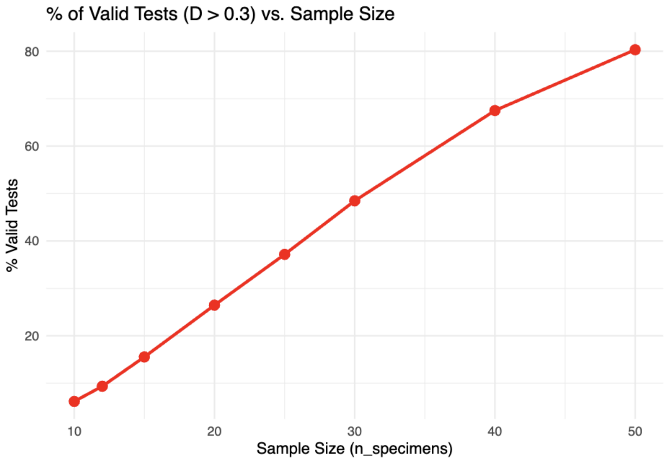


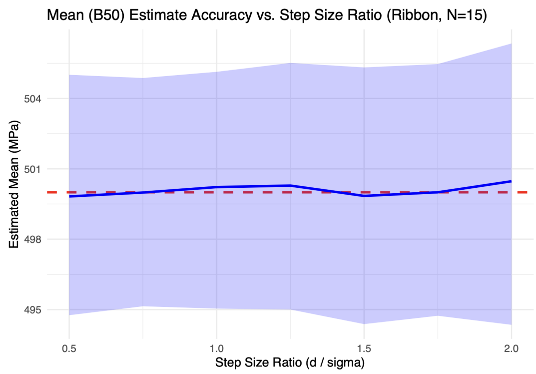
Figure 3.3: Valid test varying the sample size n

n (specimens)	% Valid Tests	Median μ (MPa)	Median σ (MPa)	σ Bias Ratio (Est./True)
10	6.16	500.67	22.54	1.127
12	9.36	500.84	19.84	0.992
15	15.54	500.36	19.08	0.954
20	26.46	500.28	18.94	0.947
25	37.16	499.56	18.94	0.947
30	48.46	500.01	18.58	0.929
40	67.50	500.05	18.17	0.909
50	80.32	499.84	18.45	0.922

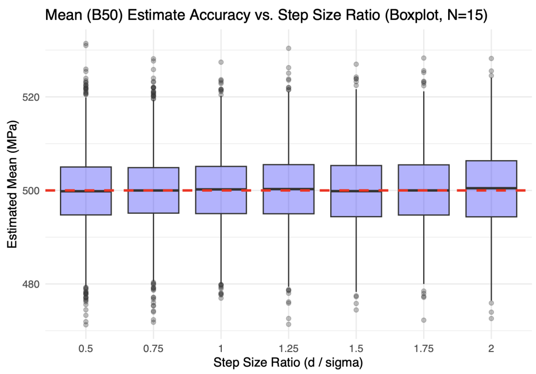
Table 3.1: Results of the mean varying the Sample Size n). Median values and σ bias calculated over 5000 simulations for each n, with $d/\sigma_{true} = 1.0$.

3.1.2.2 Second scenario: varying the ratio d/σ_{true}

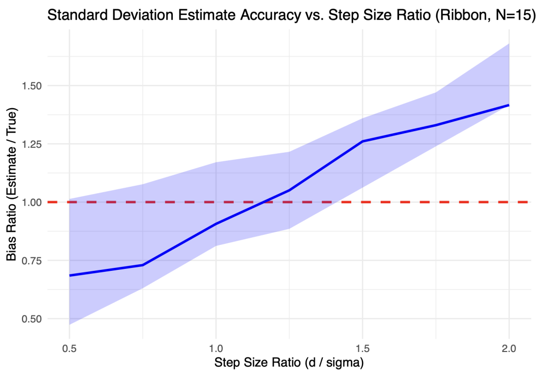
In this second scenario, the sample size was fixed at a realistic value of $n = 15$, while the step size d was varied. This allowed the impact of a non-optimal choice of the step size to be investigated by simulating d/σ_{true} ratios ranging from 0.5 to 2.0.



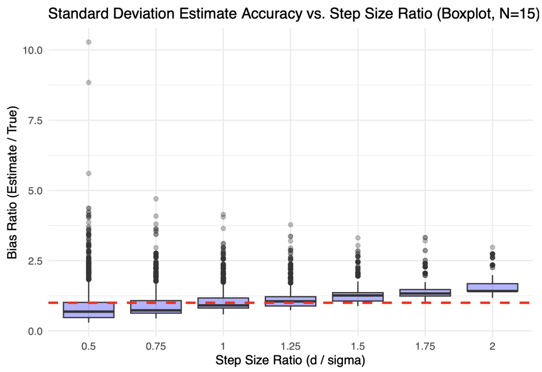
(a) Ribbon of estimated mean vs d/σ_{true}



(b) Boxplot of estimated mean vs d/σ_{true}



(c) Ribbon of estimated standard deviation vs d/σ_{true}



(d) Boxplot of estimated standard deviation vs d/σ_{true}

Figure 3.4: Results of second scenario.

- **Estimated mean** B_{50}: Once again, the estimation of the mean proves to be exceptionally stable. As shown in Figures 3.4a and 3.4b, the median of the μ estimate remains consistently centered on the "true" value of 500 MPa. Table 3.2 confirms that the choice of step size does not introduce any significant bias in the mean estimation.

- **Estimated standard deviation** $\hat{\sigma}$: This is the most significant finding of the entire simulation. The estimate of σ is extremely sensitive to the initial choice of the step size d. Figures 3.4c, 3.4d, and Table 3.2 show a clear linear trend:

  - If $d < \sigma_{\text{true}}$ (ratio < 1.0), the method severely underestimates σ (by up to -31.5% for $d/\sigma = 0.5$).
  - If $d > \sigma_{\text{true}}$ (ratio > 1.0), the method severely overestimates σ (by up to +68.1% for $d/\sigma = 2.0$).
  - The only accurate estimate is obtained when $d \approx \sigma_{\text{true}}$.

- **Test validity** $D > 0.3$: The validity of the tests was also found to be dependent on the d/σ_{true} ratio. As shown in the graph in Figure 3.5, the percentage of valid tests is maximized (at approx. 15.6%, as per Table 3.2) when $d/\sigma_{\text{true}} = 1.0$, and drops if the chosen step is too small (12.6%) or too large (only 8.6%).

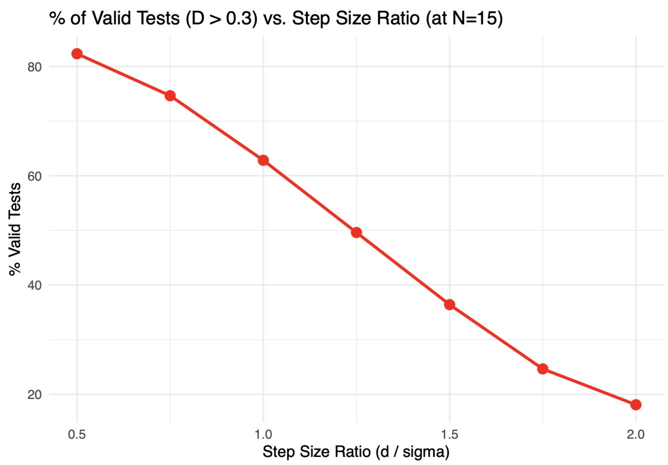


Figure 3.5: Valid test varying d/σ_{true}

3.1.3 Analysis of the results

The behavior observed in the graphs reflects the findings established in the literature. Specifically, the classical staircase analysis (Dixon-Mood) is known to suffer from significant limitations when applied to small sample sizes ($n < 30$). The primary issues are twofold: a systematic error (*bias*) that typically leads to an underestimation of the true standard deviation σ, and a high variability (*scatter*) of the estimate itself, making single point estimates unreliable.

$d/\sigma_{\mathbf{true}}$ **Ratio**	**% Valid Tests**	**Median μ (MPa)**	**Median σ (MPa)**	**σ Bias Ratio (Est./True)**
0.50	12.60	500.53	13.69	0.685
0.75	15.14	500.41	16.90	0.845
1.00	15.60	500.36	19.45	0.973
1.25	15.02	500.02	23.67	1.184
1.50	13.68	499.58	25.71	1.285
1.75	10.88	499.75	29.42	1.471
2.00	8.58	499.93	33.62	1.681

Table 3.2: Results of Analysis 2 (vs. Step Size Ratio d/σ_{true}). Median values and σ bias calculated over 5000 simulations for $n = 15$ specimens.

3.1.3.1 Scatter problem

To address the issue of high variability, the literature proposes distinct strategies. While analytical correction methods are primarily used to compensate for bias, *Bootstrap* techniques are frequently suggested to mitigate scatter. By resampling the data, bootstrapping helps to stabilize the variance of the estimator, providing more robust results.

Furthermore, Wallin in 2011 [4] provided an analytical approximation for the coefficient of variation of the standard deviation estimator to explicitly quantify this statistical scatter. The relative uncertainty is given by:

$$\frac{SD(\hat{\sigma})}{\hat{\sigma}} \approx \left(1.5 + 0.5 \cdot \frac{\hat{\sigma}}{d}\right) \cdot \frac{\beta}{\sqrt{n-5}} \tag{3.1}$$

where d is the step size, n is the number of specimens, and $\beta \approx (n - 3.5)/n$ represents the bias factor of the estimate. This formula demonstrates that the uncertainty increases significantly when the step size d is small relative to $\hat{\sigma}$ or when the sample size n is limited.

3.1.3.2 Bias problem

To compensate for the systematic underestimation of the standard deviation, several analytical correction methods have been proposed in the literature.

Svensson and Lorén *Svensson and Lorén* in 2002 [13] demonstrated that the standard deviation estimated by Dixon and Mood ($\hat{\sigma}_{DM}$) is systematically lower than the true population dispersion, especially for small sample sizes. To correct this bias, they proposed applying a multiplicative correction factor which depends on the number of specimens n:

$$\sigma_{SL} = \sigma_{DM} \cdot \frac{n}{n-3} \tag{3.2}$$

This correction has the effect of increasing the standard deviation estimate found from Dixon Mood analysis, with the increase greater for smaller sample sizes. Since the standard deviation estimate is merely multiplied by a constant greater than 1, this correction will yield slightly more scatter in results compared to Dixon-Mood.

Braam and vann der Zwaag *Braam and vann der Zwaag* in 1998 [5] proposed a correction which accounts both the number of specimens N as well as the step size d:

$$\frac{\sigma_{BZ} - \sigma_{DM}}{d} = (\frac{\sigma_{BZ}}{d} - 0.95) \cdot exp(-\frac{n}{4.93 \cdot \frac{\sigma_{BZ}}{d} + 24.48}) \tag{3.3}$$

This equation can be solved iteratively, graphically, or with the help of a solver.

Pollak et al. To mitigate the systematic underestimation of the standard deviation inherent in the Dixon–Mood method, particularly when dealing with small sample sizes, Pollak et al. [12] proposed a comprehensive correction factor. The corrected standard deviation σ_{PC} is obtained by scaling the Dixon-Mood estimate σ_{DM} using the following empirical relationship:

$$\sigma_{PC} = \sigma_{DM} \cdot A \cdot \left(\frac{n}{n-3}\right) \cdot \left(B\frac{\sigma_{DM}}{d}\right)^m \tag{3.4}$$

This correction is designed to address two distinct sources of bias. The term $\frac{n}{n-3}$ accounts for the limited degrees of freedom when n is small, effectively counteracting the tendency of sequential tests to produce optimistically low scatter estimates. The second component is necessary to mitigate the error sensitivity related to the initialization of the step size d, specifically its ratio to the estimated dispersion. The coefficients A, B, and m are constants dependent on the sample size n, derived from Monte Carlo simulations.

Chapter 4

Alternative model using the maximum likelihood: Bivariate Gaussian model

This chapter presents an exploratory analysis based on a bivariate statistical model for fatigue data. Since the analysis is mainly methodological and exploratory in nature, this section may not be included in the final presentation.

4.1 Maximum Likelihood Estimation for Fatigue Data

Maximum Likelihood Estimation (MLE) is a major parametric inference methodology used to estimate the unknown parameters of a probabilistic model from an observed sample dating back to the origins of modern mathematical statistics. The core idea is to treat the observed data as fixed and to evaluate, for each admissible parameter value, how plausible the observed sample is under the assumed model; this plausibility is quantified by the *likelihood function*.

4.1.1 General Framework

Let $\mathbf{X} = (X_1, \dots, X_n)$ denote the random vector of observations and let $\mathcal{F} = \{F_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ be a statistical model, i.e. a family of probability laws for $\mathbf{X}$ indexed by a (possibly vector-valued) parameter $\boldsymbol{\theta}$ taking values in the parameter space Θ, often specified through a joint probability density function $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$. Once a realization $\mathbf{x} = (x_1, \dots, x_n)$ is observed, the joint density can be regarded as a function of the parameter, thereby defining the likelihood function

$$L(\boldsymbol{\theta}; \mathbf{x}) := f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta.$$

If the observations are assumed independent (e.g., i.i.d.), the joint density factorizes and the likelihood takes the product form

$$L(\boldsymbol{\theta}; \mathbf{x}) = f(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}).$$

The maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ is defined as any value maximizing the likelihood over the parameter space:

$$\widehat{\boldsymbol{\theta}} \in \arg\max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}; \mathbf{x}).$$

For analytical and computational convenience, it is common to maximize the log-likelihood

$$\ell(\boldsymbol{\theta}; \mathbf{x}) = \log L(\boldsymbol{\theta}; \mathbf{x}) = \sum_{i=1}^{n} \log f(x_i; \boldsymbol{\theta}),$$

which leads to the same maximizer because the logarithm is a strictly increasing transformation.

4.1.2 Application to Staircase Fatigue Testing

In fatigue reliability studies, the estimation problem is complicated by the nature of the experimental design and the resulting data structure. The staircase method [3, 2] employs an up-and-down procedure in which the applied stress level is sequentially adjusted based on pass/fail responses evaluated at a fixed threshold lifetime N_0.

Although the experimental protocol is driven by a binary decision rule, the collected data may contain two different types of information. For some specimens, failure occurs before N_0 and the fatigue lifetime N_i is fully observed. For others, the

specimen survives beyond N_0, resulting in a runout and hence in a *right-censored* observation.
Formally, let $\delta_i \in \{0, 1\}$ denote the censoring indicator:

$$\delta_i = \begin{cases} 1, & \text{if failure occurs before } N_0, \\ 0, & \text{if the } i\text{-th observation is a runout } (N_i > N_0). \end{cases} \tag{4.1}$$

Under the assumption that $(\log N, \log S)$ follows a joint bivariate normal distribution, where S denotes the applied stress and N the fatigue lifetime, inference is naturally based on the conditional distribution of N given S. Let $f_{N|S}(\cdot \mid s)$ and $F_{N|S}(\cdot \mid s)$ denote the corresponding conditional probability density function and cumulative distribution function, respectively.

Since the exact fatigue lifetime is observed only for failures, while runouts provide only the information that $N_i > N_0$, the likelihood function must combine density contributions for failures and survival probabilities for censored observations. The resulting likelihood is

$$L(\boldsymbol{\theta}; \mathbf{N}, \mathbf{x}) = \prod_{i=1}^{n} \left[f_{N|S_i}(N_i \mid S_i) \right]^{\delta_i} \left[1 - F_{N|S_i}(N_0 \mid S_i) \right]^{1-\delta_i}, \tag{4.2}$$

where $\boldsymbol{\theta}$ denotes the vector of model parameters.
In this formulation, failure observations contribute through the conditional density evaluated at the observed lifetime, while runouts contribute through the conditional survival function evaluated at the censoring threshold N_0. This structure correctly accounts for right censoring induced by the staircase testing procedure.

Inference is performed by maximizing the corresponding log-likelihood

$$\ell(\boldsymbol{\theta}; \mathbf{N}, \mathbf{x}) = \log L(\boldsymbol{\theta}; \mathbf{N}, \mathbf{x})$$

over the parameter space Θ. The parameters of interest include the mean vector and covariance matrix of the underlying bivariate normal distribution, which characterize the dependence between stress and fatigue life as well as the dispersion of the fatigue process.

4.2 The Bivariate Statistical Model

To extend the analysis beyond the purely binary formulation of the classical *staircase* method, we consider likelihood-based approaches that explicitly model the relationship between the applied stress level S and the fatigue life N, consistently with a probabilistic formulation of the S–N curve.

The objective is to characterize the statistical relationship between stress and fatigue life while estimating the median fatigue strength (B_{50}), their variability, and the correlation between the two variables. To this end, N and S are treated as correlated random variables, and the model is formulated on a logarithmic scale using a Bivariate Normal distribution.

4.2.1 Stromeyer-type model

A Stromeyer-type S–N relationship, introduced in Section 2.2.2, is assumed, in which fatigue life increases asymptotically as the stress level decreases:

$$\log N = \beta_0 + \beta_1 \log(S - \gamma), \qquad S > \gamma, \tag{4.3}$$

where γ represents the horizontal asymptote of the S–N curve and can be interpreted as the fatigue limit.
The parameter γ represents a common fatigue-limit threshold for all specimens. For stress levels $S \leq \gamma$, the model implies that fatigue failure does not occur within any finite number of cycles, corresponding to an idealized infinite fatigue life.

Introducing the transformed variables

$$Y = \ln(N), \tag{4.4}$$

$$Z = \ln(S - \gamma), \tag{4.5}$$

the model is formulated in the (Y, Z) space. The constraint $\gamma < S_{\min}$ ensures that the logarithmic transformation is well defined for all observed stress levels.

The random vector (Y, Z) is assumed to follow a Bivariate Normal distribution:

$$\begin{pmatrix} Y \\ Z \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \sigma_Y^2 & \rho\,\sigma_Y \sigma_Z \\ \rho\,\sigma_Y \sigma_Z & \sigma_Z^2 \end{pmatrix} \right).$$

The model is therefore fully characterized by the parameter vector

$$\boldsymbol{\theta} = (\mu_Y, \mu_Z, \sigma_Y, \sigma_Z, \rho, \gamma),$$

where μ_Y and μ_Z are the marginal means, σ_Y and σ_Z the marginal standard deviations, ρ the linear correlation coefficient ($-1 < \rho < 1$), and γ the threshold parameter associated with the fatigue limit.

4.2.1.1 Likelihood formulation

Parameter estimation is performed via Maximum Likelihood Estimation (MLE). To this end, the conditional distribution of fatigue life given stress is considered. From standard properties of the Bivariate Normal distribution, the conditional distribution of Y given Z is univariate normal:

$$Y \mid Z = z \sim \mathcal{N}\Big(\mu_{Y|Z},\, \sigma^2_{Y|Z}\Big)\,,$$

with

$$\mu_{Y|Z} = \mu_Y + \rho\frac{\sigma_Y}{\sigma_Z}(z - \mu_Z), \tag{4.6}$$

$$\sigma_{Y|Z} = \sigma_Y\sqrt{1-\rho^2}. \tag{4.7}$$

Equation (4.6) defines the median S–N relationship on the log–log transformed scale, since the conditional median of a normal distribution coincides with its mean.

The experimental dataset includes both failures, for which the fatigue life N_i is observed, and run-outs, for which only the information $N_i > N_0$ is available, where N_0 denotes the fixed cycle limit of the test.Introducing the indicator variable defined in Equation 4.1, the likelihood function is constructed by combining density contributions for failures and survival probabilities for right-censored observations:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\Big[f_{Y|Z}(y_i \mid z_i;\boldsymbol{\theta})\Big]^{\delta_i}\left[\mathbb{P}(Y > \log N_0 \mid Z = z_i;\boldsymbol{\theta})\right]^{1-\delta_i}. \tag{4.8}$$

The survival probability can be written explicitly as

$$\mathbb{P}(Y > \log N_0 \mid Z = z_i) = 1 - F\left(\frac{\log N_0 - \mu_{Y|z_i}}{\sigma_{Y|Z}}\right),$$

where $F(\cdot)$ denotes the standard normal cumulative distribution function.

4.2.1.2 Initialization of the model parameters

The numerical maximization of the likelihood requires the specification of suitable starting values for the model parameters. Because the proposed bivariate model is nonlinear and includes the fatigue-limit parameter γ, a data-driven initialization strategy is adopted in order to improve the stability and convergence of the optimization algorithm.

Let $S_{\min}$ and $S_{\max}$ denote the minimum and maximum observed stress levels in the dataset. Since the Stromeyer-type relationship requires $S > \gamma$ for all observations, the fatigue-limit parameter must satisfy the constraint $\gamma < S_{\min}$. To guarantee this condition, the initial value of γ is chosen as a fixed fraction of the minimum observed stress,

$$\gamma^{(0)} = 0.9\, S_{\min},$$

which ensures that the transformation $Z = \log(S - \gamma)$ is well defined for every observation.

Given $\gamma^{(0)}$, the transformed stress variable is computed as

$$z_i^{(0)} = \log(S_i - \gamma^{(0)}),$$

while the fatigue-life variable is defined as $y_i = \log N_i$.
Initial values for the location parameters of the bivariate normal model are obtained as the empirical means of the transformed variables,

$$\mu_Y^{(0)} = \frac{1}{n}\sum_{i=1}^{n} y_i, \qquad \mu_Z^{(0)} = \frac{1}{n}\sum_{i=1}^{n} z_i^{(0)}.$$

Similarly, the scale parameters are initialized using the corresponding sample standard deviations,

$$\sigma_Y^{(0)} = \mathrm{sd}(y), \qquad \sigma_Z^{(0)} = \mathrm{sd}(z^{(0)}).$$

The correlation coefficient between the two variables is initialized at

$$\rho^{(0)} = 0,$$

which represents a neutral starting point that does not impose any prior assumption about the dependence structure between fatigue life and transformed stress.

During numerical optimization it is convenient to work in an unconstrained parameter space. For this reason, the parameters subject to constraints are reparameterized before being passed to the optimizer.

First, the standard deviations σ_Y and σ_Z must be strictly positive. This constraint is enforced by optimizing their logarithms, so that the optimizer works on $\log \sigma_N$ and $\log \sigma_Z$. Because the exponential function maps the real line onto positive values, this guarantees $\sigma_N > 0$ and $\sigma_Z > 0$ for any parameter value explored during the optimization.

Second, the fatigue-limit parameter must satisfy the constraint $\gamma < S_{\min}$. To enforce this condition automatically, γ is expressed through a logistic transformation

$$\gamma = S_{\min} \operatorname{logit}^{-1}(\eta_\gamma),$$

where $\operatorname{logit}^{-1}(x) = \frac{1}{1+e^{-x}}$ maps the real line onto the interval $(0, 1)$. Consequently, the parameter η_γ is unconstrained, while the corresponding value of γ is always restricted to the admissible interval $(0, S_{\min})$.

The initial value of the transformed parameter is therefore

$$\eta_\gamma^{(0)} = \operatorname{logit}\left(\frac{\gamma^{(0)}}{S_{\min}}\right).$$

The resulting starting vector supplied to the numerical optimizer is

$$\mathbf{p}^{(0)} = \left(\mu_N^{(0)}, \mu_Z^{(0)}, \log \sigma_N^{(0)}, \log \sigma_Z^{(0)}, \rho^{(0)}, \eta_\gamma^{(0)}\right).$$

4.2.1.3 Computational implementation and parameter mapping

Numerical estimation of the model parameters is performed by minimizing the negative log-likelihood function using the `optim` routine in the `R` environment. This approach is equivalent to maximizing the log-likelihood but is numerically more convenient for the optimization procedure.

To ensure that all physical and statistical constraints are satisfied during optimization, the theoretical parameter vector $\boldsymbol{\theta}$ is mapped to an unconstrained parameter vector $\mathbf{p} \in \mathbb{R}^6$. The numerical implementation directly follows the likelihood formulation given in Eq. (4.8). For reproducibility, the full R implementation of the likelihood function and the optimization routine is reported in Appendix B.

In particular, logarithmic transformations are used to enforce positivity of the standard deviations, a hyperbolic tangent transformation constrains the correlation coefficient to the interval $(-1, 1)$, and a scaled logistic transformation is employed for the fatigue-limit parameter γ, ensuring the constraint $\gamma < S_{\min}$ required for the definition of $\log(S - \gamma)$.

These transformations are introduced to allow optimization in an unconstrained parameter space and to improve numerical stability. To the best of our knowledge, the fatigue literature does not provide specific recommendations regarding the parameter reparameterization required for numerical estimation in this type of bivariate formulation; the mapping adopted here is therefore motivated primarily by computational considerations.

The correspondence between theoretical parameters and optimization variables is summarized in Table 4.1.

Parameter	**optim variable**	**Code name**	**Transformation**
μ_Y	`par[1]`	`muY`	Identity
μ_Z	`par[2]`	`muZ`	Identity
σ_Y	`par[3]`	`log_sigY`	$\sigma_Y = \exp(\texttt{par[3]})$
σ_Z	`par[4]`	`log_sigZ`	$\sigma_Z = \exp(\texttt{par[4]})$
ρ	`par[5]`	`atanh_rho`	$\rho = \tanh(\texttt{par[5]})$
γ	`par[6]`	`eta_gamma`	$\gamma = S_{\min}\,\text{logit}^{-1}(\texttt{par[6]})$

Table 4.1: Mapping between theoretical parameters and variables used in the R code

4.2.1.4 Definition of the median fatigue strength S_{50} at N_0

Within the proposed bivariate framework, it is natural to define a fatigue–strength quantity directly comparable with the classical staircase estimate B_{50}. To this end, we introduce the median fatigue strength at the fixed cycle level N_0, denoted by S_{50}.

For a given stress level $S > \gamma$, the probability of failure at N_0 is given by the conditional distribution of Y given Z:

$$\mathbb{P}(N \leq N_0 \mid S) = \mathbb{P}(Y \leq \log N_0 \mid Z) = F\left(\frac{\log N_0 - \mu_{Y|Z}}{\sigma_{Y|Z}}\right),$$

where $\mu_{Y|Z}$ and $\sigma_{Y|Z}$ are defined in Eqs. (4.6)–(4.7).

The median fatigue strength S_{50} is defined as the unique stress level satisfying

$$\mathbb{P}(N \leq N_0 \mid S_{50}) = 0.5. \tag{4.9}$$

Equivalently, S_{50} is the solution of the equation

$$F\left(\frac{\log N_0 - \mu_{Y|Z(S_{50})}}{\sigma_{Y|Z}}\right) = 0.5,$$

where $Z(S_{50}) = \log(S_{50} - \gamma)$.

Since the standard normal distribution is symmetric, Eq. (4.9) corresponds to the condition

$$\log N_0 = \mu_{Y|Z(S_{50})},$$

which implicitly defines S_{50} as a function of the model parameters. In the present work, this equation is solved numerically via one–dimensional root finding.

By construction, S_{50} represents the stress level at which the failure probability at N_0 equals 50%. This definition provides a mathematically consistent analogue of the staircase B_{50} estimate, while explicitly accounting for the fatigue–limit parameter γ and for the joint probabilistic structure of fatigue life and stress.

4.2.2 Model formulation with fixed fatigue-limit parameter

The full bivariate formulation introduced above includes the fatigue-limit parameter γ as an unknown quantity to be estimated jointly with the remaining parameters. This specification allows the horizontal asymptote of the S–N curve to be determined directly from the data and represents the most flexible formulation containing a deterministic threshold.

However, in staircase experiments conducted at a fixed life N_0, the available information is often limited to binary failure/runout outcomes concentrated in a narrow stress range. In such settings, the likelihood function may exhibit weak curvature with respect to γ, leading to practical identifiability issues. In particular, strong dependence may arise between the threshold parameter and the location parameters of the bivariate normal distribution, resulting in a relatively flat log-likelihood surface along the γ direction.

For this reason, alongside the full six-parameter specification, a reduced five-parameter formulation can also be considered, in which γ is treated as fixed and maximum likelihood estimation is performed on the remaining parameter vector

$$\boldsymbol{\theta}_\gamma = (\mu_N, \mu_Z, \sigma_N, \sigma_Z, \rho).$$

In this configuration, the probabilistic structure of the model remains unchanged, and only the asymptote position of the S–N curve is imposed a priori rather than estimated from the sample.

The fixed-threshold specification is not introduced as a preferred model, but as a complementary formulation aimed at assessing the sensitivity of the inferred fatigue-strength quantities to the threshold assumption. Its role is therefore primarily diagnostic, especially in small-sample staircase settings where the identifiability of γ may be limited.

Chapter 5

Random Fatigue Limit Model

Classical S–N regression models with a fatigue limit, such as the *Stromeyer relationship*, provide an effective phenomenological description of high-cycle fatigue behavior by introducing a horizontal asymptote in stress. In these formulations, the fatigue limit is treated as a fixed, material-dependent parameter.

However, experimental fatigue data often exhibit substantial variability across specimens, even under nominally identical materials and testing conditions. The presence of runouts and the dispersion observed in the high-cycle regime suggest that fatigue resistance may not be adequately represented by a single deterministic threshold.

The *Random Fatigue Limit* (RFL) model addresses this limitation by extending the classical Stromeyer formulation through the introduction of specimen-specific fatigue limits. In this framework, the fatigue limit is modeled as a latent random variable at the population level, allowing variability in endurance behavior to be explicitly incorporated into the model structure.

The purpose of this chapter is to introduce and formalize the Random Fatigue Limit model from a statistical perspective. Starting from the deterministic Stromeyer relationship discussed previously in section 2.2.2, the model is progressively generalized to account for population-level variability in the fatigue limit. Issues related to parameter estimation, inference, and comparison with alternative formulations are addressed in the subsequent sections.

The numerical implementations of the Random Fatigue Limit model used in this work are reported in the appendices. The frequentist estimation procedure is described in

Appendix C.0.1, while the Bayesian implementation based on MCMC sampling is presented in Appendix C.0.2.

5.1 Statistical formulation of the RFL model

A natural starting point for the Random Fatigue Limit model is the deterministic Stromeyer-type stress–life relationship. In its classical form, the model can be written as

$$\log N_i = \beta_0 + \beta_1 \log(S_i - \gamma) + \varepsilon_i, \tag{5.1}$$

where S_i denotes the applied stress level for specimen i, γ is a fixed fatigue limit common to all specimens, and $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$.

While this formulation captures the observed flattening of the S–N curve, it implicitly assumes that all specimens share the same fatigue limit. Experimental evidence suggests, however, that fatigue resistance varies across specimens due to microstructural heterogeneity, manufacturing variability, and local defects. This motivates treating the fatigue limit as a specimen-specific quantity rather than as a deterministic material constant.

The Random Fatigue Limit (RFL) model extends the Stromeyer relationship by replacing the fixed parameter γ with a specimen-specific fatigue limit γ_i, $i = 1, \ldots, n$.

Let $\Gamma = (\gamma_1, \ldots, \gamma_n)$ denote the vector of specimen-specific fatigue limits. The components γ_i represent latent heterogeneity in fatigue resistance across the population. At the population level, the fatigue limits are assumed to be exchangeable. Working on the logarithmic scale, let

$$V_i = \log \gamma_i, \qquad V_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2). \tag{5.2}$$

with probability density function $f_V(v) = \frac{1}{\sigma_\Gamma}\phi\Big(\frac{v-\mu_\Gamma}{\sigma_\Gamma}\Big)$, where $\phi(\cdot)$ denotes the standard normal density. Conditional on the applied stress S_i and on the specimen-specific fatigue limit γ_i, the logarithm of the fatigue life follows the same Stromeyer functional form,

$$W_i = \log N_i \mid (S_i, \gamma_i) \sim \mathcal{N}\Big(\beta_0 + \beta_1 \log(S_i - \gamma_i),\, \sigma^2\Big), \qquad S_i > \gamma_i. \tag{5.3}$$

The quantity $(S_i - \gamma_i)$ represents the effective stress acting on specimen i, i.e. the excess stress above its individual fatigue limit that drives fatigue damage accumulation.

If the applied stress does not exceed the specimen-specific fatigue limit, that is $S_i \leq \gamma_i$, the specimen is not susceptible to fatigue failure at that stress level. Within the RFL framework, such a specimen can be regarded as *immune* to fatigue damage under continued cycling at S_i. Formally,

$$\mathbb{P}(N_i < \infty \mid S_i \leq \gamma_i) = 0.$$

This immunity is a structural consequence of the latent fatigue limit. Since the individual limits γ_i are unobserved, the immune status of a specimen cannot be directly inferred from the data. Consequently, an observed runout does not necessarily imply immunity: it may correspond either to an immune specimen ($S_i \leq \gamma_i$) or to a susceptible specimen whose realized fatigue life exceeds the experimental censoring threshold. The marginal fatigue-life distribution therefore exhibits a mixture structure, combining an immune fraction with a susceptible population governed by the Stromeyer regression mechanism.

So for a given stress level S_i the probability that specimen i is *active* (i.e. susceptible to fatigue failure) at stress S_i is therefore

$$F_a(S_i) = \mathbb{P}(\gamma_i < S_i) = \mathbb{P}(V_i < \log S_i) = \Phi\left(\frac{\log S_i - \mu_\Gamma}{\sigma_\Gamma}\right), \tag{5.4}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

The complementary probability

$$1 - F_a(S_i) = \mathbb{P}(\gamma_i \geq S_i)$$

represents the immune fraction at stress level S_i.

Since fatigue life is observable only for active specimens, the cumulative distribution function of N_i conditional on S_i follows from the law of total probability:

$$F_N(n \mid S_i) = \mathbb{P}(N_i \leq n \mid S_i) = \mathbb{P}(\gamma_i < S_i)\, \mathbb{P}(N_i \leq n \mid S_i, \gamma_i < S_i). \tag{5.5}$$

Integrating over the distribution of V_i on the active region yields the marginal distribution

$$\mathbb{P}(N_i \leq n \mid S_i) = \int_{-\infty}^{\log S_i} \Phi\left(\frac{\log n - \beta_0 - \beta_1 \log(S_i - e^v)}{\sigma}\right) f_V(v)\, dv. \tag{5.6}$$

This expression corresponds to the fatigue-life distribution induced by the Random Fatigue Limit model. It combines an immune fraction $1 - F_a(S_i)$ with a susceptible population governed by the conditional Stromeyer mechanism.
The integral does not admit a closed-form solution and must therefore be evaluated numerically.

5.1.1 Likelihood contribution with immunity and numerical approximation

Let δ_i denote the censoring indicator, with $\delta_i = 1$ for observed failure and $\delta_i = 0$ for run-out at the censoring level N_0. Moreover, define $x_i := \log S_i$. The likelihood contribution of observation i is given by

$$L_i = \begin{cases} F_a(S_i)\, \mathbb{E}[f_W(w_i \mid S_i, V_i) \mid V_i < x_i]\,, & \delta_i = 1, \\ \\ (1 - F_a(S_i)) + F_a(S_i)\, \mathbb{E}[\Pr(W_i > \log N_0 \mid S_i, V_i) \mid V_i < x_i]\,, & \delta_i = 0, \end{cases} \tag{5.7}$$

where $f_W(\cdot)$ denotes the normal density of W_i conditional on (S_i, V_i).

The likelihood thus has a mixture structure: a discrete immune component corresponding to $\gamma_i \geq S_i$ and a continuous component associated with specimens that are not immune.

In this setting, the marginal likelihood contributions involve integration over an unbounded domain. In order to evaluate these integrals, numerical approximation is required.

In the implementation adopted in this work, the conditional expectations with respect to $V_i \mid V_i < x_i$ are approximated by numerical quadrature. Specifically, the integral is evaluated as an average over K equally spaced quantiles of the truncated normal distribution

$$V_i \mid V_i < x_i \sim \mathcal{N}(\mu_\gamma, \sigma_\gamma^2) \text{ truncated above at } x_i,$$

in full agreement with the likelihood formulation used in the computational implementation. For numerical stability and computational convenience, parameter estimation is performed by maximizing the log-likelihood function obtained by summing the logarithms of the individual contributions L_i.

To reduce sensitivity to arbitrary starting values and to anchor the optimization procedure to the empirical information contained in the data, initial values for the regression parameters are obtained from a preliminary least-squares fit based on the failure observations only. These data-driven starting values improve the robustness of the numerical maximization.
Parameter estimation is carried out by numerical maximization of the resulting log-likelihood with respect to

$$(\beta_0, \beta_1, \sigma, \mu_\gamma, \sigma_\gamma),$$

using the same maximum likelihood principles introduced in Chapter 4.

The RFL model extends the likelihood framework discussed therein by introducing latent variables and numerical integration, while preserving the standard structure for censored observations.

5.1.2 Determination of the median fatigue strength S_{50}

The marginal fatigue-life distribution derived above provides, for each stress level S, the probability of failure by a fixed target number of cycles N_0, namely $F_N(N_0 \mid S)$, defined by the equation (5.6).

Stress quantiles are defined implicitly as the solutions of

$$F_N(N_0 \mid S_p) = p,$$

for a prescribed probability level $p \in (0, 1)$. In particular, the median fatigue strength S_{50} is defined by

$$F_N(N_0 \mid S_{50}) = 0.5.$$

Since the function $F_N(N_0 \mid S)$ does not admit a closed-form expression, its evaluation for a given stress level S requires numerical integration.

The integral is evaluated using the same quantile-based quadrature scheme employed in the likelihood computation. In particular, the expectation with respect to

$$V \mid V < \log S \sim \mathcal{N}(\mu_\gamma, \sigma_\gamma^2) \text{ truncated above at } \log S$$

is approximated as an average over K equally spaced quantiles of the truncated normal distribution. This yields the approximation

$$F_N(N_0 \mid S) \approx \Phi\left(\frac{\log S - \hat{\mu}_\Gamma}{\hat{\sigma}_\Gamma}\right) \cdot \frac{1}{K}\sum_{k=1}^{K} \Phi\left(\frac{\log N_0 - (\hat{\beta}_0 + \hat{\beta}_1 \log(S - e^{v_k}))}{\hat{\sigma}}\right). \quad (5.8)$$

The value S_{50} is then obtained by solving

$$F_N(N_0 \mid S) - 0.5 = 0$$

using a one-dimensional numerical root-finding algorithm. The monotonicity of $F_N(N_0 \mid S)$ with respect to S ensures the existence and uniqueness of the solution.

5.2 Bayesian Formulation of the RFL model

The Random Fatigue Limit model presented above has an inherently hierarchical structure, involving regression parameters, population-level fatigue-limit parameters, and specimen-specific latent fatigue limits, called *random effects* in statistics. In the frequentist analysis, inference relied on the marginal likelihood obtained by integrating out the latent variables. In the Bayesian framework, the structural specification of the model is unchanged, but prior distributions are assigned to all unknown quantities and inference is conducted through the joint posterior distribution.

The stress–life relationship and the structural immunity mechanism are therefore preserved; what changes is the inferential strategy. Rather than optimizing a marginal likelihood, inference is based on posterior sampling, which propagates uncertainty simultaneously across all model components.

The RFL specification is naturally expressed as a hierarchical Bayesian model. Conditional on the latent fatigue limit γ_i and stress level S_i, the observed fatigue life, or its censoring mechanism in the case of runouts, follows a normal distribution on the log-life scale. The specimen-specific fatigue limits γ_i are treated as unobserved random variables drawn from a population distribution capturing material heterogeneity, while the regression coefficients (β_0, β_1), the life dispersion parameter σ, and the fatigue-limit distribution parameters $(\mu_\Gamma, \sigma_\Gamma)$ are assigned prior distributions.

This hierarchical formulation is particularly suitable for fatigue data, which are typically characterized by limited sample sizes, substantial specimen-to-specimen variability, right-censored observations, and latent endurance thresholds. The Bayesian approach avoids direct optimization of a marginal likelihood and instead provides a coherent framework for uncertainty propagation, yielding direct posterior uncertainty quantification for derived quantities such as S_{50}.

Since the resulting posterior distribution is analytically intractable due to the presence of latent variables, censoring, and nonlinear transformations, inference is carried out using Markov Chain Monte Carlo methods.

In the present analysis, *JAGS* (Just Another Gibbs Sampler) is adopted as a computational framework that facilitates the specification and estimation of hierarchical Bayesian models with latent variables and censored observations.

5.2.1 Bayesian implementation in JAGS

The Random Fatigue Limit model with structural immunity is implemented in JAGS by formulating both the population distribution of the latent fatigue limits and the conditional Stromeyer-type stress–life relationship within a fully Bayesian framework. All unknown quantities, including specimen-specific fatigue limits and censored fatigue lives, are treated as stochastic nodes and jointly sampled from their posterior distribution via Markov Chain Monte Carlo.

For each specimen $i = 1, \dots, n$, let S_i denote the applied stress and N_i the observed number of cycles. We define

$$x_i = \log S_i, \qquad W_i = \log N_i.$$

Runouts occur at a fixed censoring level N_0; on the logarithmic scale this corresponds to the threshold $\log N_0$. To formally encode right-censoring in the probabilistic model, we introduce an indicator variable

$$is\_cens_i = \begin{cases} 1, & \text{if specimen } i \text{ is a runout,} \\ 0, & \text{if specimen } i \text{ fails.} \end{cases}$$

This indicator encodes whether the log-life W_i is fully observed or only partially observed due to right-censoring.

The specimen-specific fatigue limits are modeled on the log scale as

$$V_i = \log \gamma_i, \qquad V_i \sim \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2).$$

A fundamental physical constraint must hold: failure can occur only if $S_i > \gamma_i$, or equivalently $V_i < x_i$ as explained in Section 5.1. This condition is enforced directly in the sampling scheme by truncating the normal distribution of V_i for failure observations. Specifically, for failures the support of V_i is restricted to values strictly below x_i, while for runouts no upper restriction is imposed. This guarantees that posterior draws remain physically admissible.

Structural immunity is introduced deterministically through the indicator

$$active_i = \mathbb{I}(x_i > V_i),$$

When $active_i = 1$, the specimen is susceptible to fatigue failure at stress S_i; when $active_i = 0$, the specimen is immune at that stress level.

For susceptible specimens, the Stromeyer-type stress–life relationship is written as

$$\mu_i^{act} = \beta_{0,c} + \beta_1 \Big(\log(S_i - \gamma_i) - \bar{x} \Big),$$

where $\bar{x}$ denotes the empirical mean of $\log S_i$, introduced for centering. To avoid numerical instability when S_i is close to Γ_i, the difference $S_i - \Gamma_i$ is bounded below by a small positive constant in the computation. The regression is expressed in centered form to improve numerical stability and reduce posterior correlation between intercept and slope during MCMC sampling. Specifically, the predictor is written as $\log(S_i - \gamma_i) - \bar{x}$, where $\bar{x}$ is the empirical mean of $\log S_i$. The original intercept parameter is recovered through the deterministic transformation

$$\beta_0 = \beta_{0,c} - \beta_1 \bar{x}.$$

For immune specimens, failure is structurally excluded by assigning a mean log-life far beyond the censoring threshold,

$$\mu_i^{imm} = \log N_0 + k_{Imm} \sigma,$$

with k_{Imm} a fixed large multiplier. This construction provides a computational representation of the theoretically infinite fatigue life associated with immunity: by shifting the mean sufficiently above the censoring level, the probability of observing a finite failure within the experimental range becomes negligible. The effective mean used in the likelihood is therefore

$$\mu_i = active_i \, \mu_i^{act} + (1 - active_i) \, \mu_i^{imm}.$$

This construction induces a deterministic mixture between the Stromeyer-type stress–life regression component, governing susceptible specimens, and a high-life component representing structural immunity.

The log-life is modeled as $W_i \sim \mathcal{N}(\mu_i, \sigma^2)$.

Right censoring is incorporated through an explicit probabilistic representation of the censoring mechanism,, which enforces the consistency between the latent log-life W_i and the censoring threshold. Specifically, when $is\_cens[i] = 1$, the model constrains W_i to lie above $\log N_0$, thereby inducing a likelihood contribution equal to the corresponding upper-tail probability of the normal distribution. For failure observations, the exact value of W_i is observed and directly enters the likelihood. For runouts, W_i remains unobserved and is constrained to satisfy $W_i > \log N_0$, so that its likelihood contribution reduces to $\mathbb{P}(W_i > \log N_0 \mid \mu_i, \sigma)$.

The prior specification is defined consistently with the hierarchical structure of the model. The population mean of the fatigue limit is assigned

$$\mu_\Gamma \sim \mathcal{N}(\mu_{\Gamma_0},\, 10)\; T(-\infty, \min x_i),$$

to avoid degenerate configurations in which the population fatigue limit exceeds all tested stress levels, which would be incompatible with the occurrence of failures. The population variability of the fatigue limit is

$$\sigma_\Gamma \sim \text{Uniform}(0.001, 0.5).$$

The centered intercept is assigned

$$\beta_{0,c} \sim \mathcal{N}(\log N_0,\, 100),$$

reflecting that the relevant log-life scale is naturally anchored around the experimental target life N_0, while the large variance keeps the prior diffuse relative to the observed variability. The slope parameter is specified as

$$\beta_1 \sim \mathcal{N}(-2,\, 10)\; T(-\infty, 0),$$

encoding the physical requirement that fatigue life decreases as stress increases. The prior mean -2 reflects a typical decreasing trend observed in high-cycle fatigue applications, whereas the large variance preserves flexibility. Finally, the residual standard deviation

$$\sigma \sim \text{Uniform}(0, 2).$$

Overall, the prior specification ensures posterior propriety in the presence of censoring and latent variables, contributes to stable MCMC sampling under limited data, and constrains parameters to physically meaningful regions without introducing strong prior information. The resulting structure provides mild regularization while allowing inference to remain primarily data-driven.

Initial values for the regression parameters are derived from a preliminary linear regression fit based on the failure observations. This strategy reduces arbitrariness in the initialization step by anchoring starting values to the empirical information contained in the data, rather than to externally imposed guesses[1]. The resulting estimates are translated into the centered parametrization and slightly perturbed across three parallel chains to obtain dispersed, data-driven initial values.

The model is then fitted using three independent MCMC chains, each including an adaptation phase followed by burn-in and sampling iterations. Convergence is evaluated through trace plots and Gelman–Rubin diagnostics.

Posterior inference relies on samples drawn from the joint distribution

$$p(\beta_{0,c}, \beta_1, \sigma, \mu_\Gamma, \sigma_\Gamma, \{V_i\} \mid \text{data}),$$

5.2.2 Posterior computation of the S_{50}

In the Bayesian analysis, the median fatigue strength S_{50} is defined exactly as in the frequentist formulation, namely as the stress level satisfying

$$F_N(N_0 \mid S_{50}) = 0.5,$$

where $F_N(N_0 \mid S)$ denotes the marginal probability of failure by the target life N_0 under the RFL model.

The difference concerns the inferential treatment of the model parameters. In the frequentist approach, S_{50} is obtained by substituting maximum likelihood estimates into $F_N(N_0 \mid S)$. In the Bayesian framework, the parameters are random and described by their posterior distribution. Accordingly, for each posterior draw

$$(\beta_0, \beta_1, \sigma, \mu_\Gamma, \sigma_\Gamma),$$

[1] 14.

the equation

$$F_N(N_0 \mid S) - 0.5 = 0$$

is solved numerically.

For a fixed parameter draw, evaluating $F_N(N_0 \mid S)$ requires integration over the truncated normal distribution

$$V \mid V < \log S \sim \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2),$$

which is approximated using the same quantile-based quadrature scheme adopted in the likelihood computation. The root of the equation is then determined using a one-dimensional numerical solver.

Repeating this procedure across posterior draws produces a posterior sample of S_{50} values. Posterior summaries, such as the posterior median and 95% credible interval, are subsequently computed from this sample. In this way, uncertainty in S_{50} directly reflects posterior uncertainty in all regression parameters and fatigue-limit hyperparameters.

Chapter 6

Results and discussion

In this chapter, the statistical models developed in the previous chapters are applied to experimental fatigue test data on high-strength bolts, obtained using a staircase procedure at a fixed life of $N_0 = 5 \times 10^6$ cycles.

The analysis is carried out on bolts data provided by Prof. Rosso. The dataset consists of 15 specimens, of which 7 failed before reaching N_0, while 8 reached runout. Therefore, the available information follows the typical structure of a staircase experiment, where each observation is classified as failure or runout at the target cycle level N_0, and the exact number of cycles to failure is also available for the specimens that failed.
The applied stress levels range is from approximately 6.1 to 6.7 kN. All the analyzed

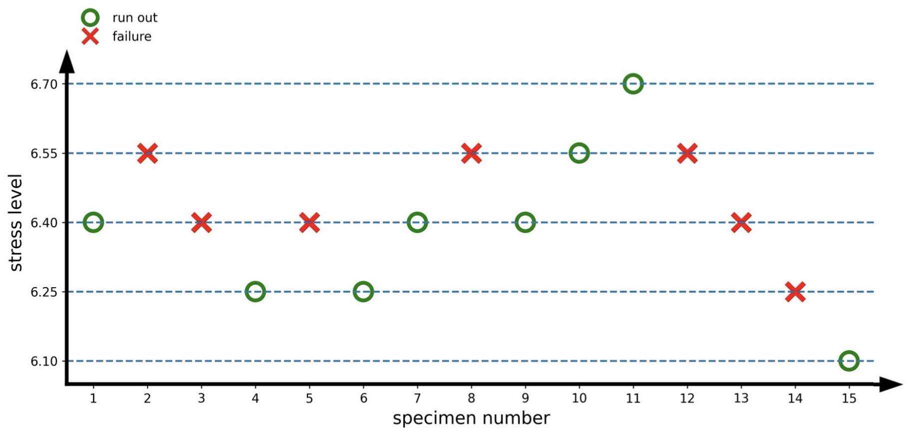


Figure 6.1: Staircase dataset used to fit the considered models.

models, the classical Dixon–Mood method, the bivariate normal model, and the

Random Fatigue Limit model, are estimated using the same observed information. Any differences in the obtained estimates are therefore attributable only to the different structural and probabilistic assumptions underlying the models, and not to a different informational content of the data.

The objective of this chapter is to compare the different approaches in terms of the estimation of the parameter B_{50}, the width of the associated confidence intervals, the stability of the parameters in the presence of small sample sizes, and, for the parametric models, the shape of the induced S–N curve.

Particular attention is given to the role of the threshold parameter γ, considered as fixed, as a free parameter in the Stromeyer formulation, and as a stochastic variable in the Random Fatigue Limit model, in order to assess the impact of model complexity on inferential uncertainty.

6.1 Dixon and Mood method

The classical Dixon–Mood method, described in detail in Chapter 2 within the ISO 12107 procedure, is here applied to the experimental dataset in order to obtain a reference estimate of the parameter B_{50} and of the standard deviation of the fatigue strength at the fixed life N_0.

In the analyzed sample, 7 failures and 8 runouts are observed. The parameter D, used to verify the validity of the standard deviation estimate according to ISO 12107, is equal to $D = 0.4898$, which is higher than the acceptability threshold of 0.3. The estimate can therefore be considered formally valid according to the standard criteria.

The obtained results are the following:

- $B_{50} = 6.368$ kN,
- $\hat{\sigma} = 0.126$ kN.

Based on these estimates, the one-sided lower tolerance limit for 90% coverage and 95% confidence level, calculated with factor $k = 2.755$ and degrees of freedom $v = 6$, is equal to

$$B_{90} = 6.021 \text{ kN}.$$

For completeness, a 95% confidence interval for the estimate of B_{50} was also computed, assuming normality and using the t distribution with $\nu = C - 1 = 6$ degrees of freedom. The result is:

$$CI_{95\%}(B_{50}) = [6.251,\ 6.485] \text{ kN}.$$

It can be observed that the width of the confidence interval is relatively small despite the limited sample size ($n = 15$, $C = 7$), while the ISO lower tolerance limit is significantly more conservative, reflecting its different design purpose compared to inference on the mean of the distribution.

6.2 Bivariate Normal Model

The bivariate normal model introduced in Chapter 4 is applied to the considered staircase dataset. The formulation is based on a Stromeyer-type relationship, in which the link between fatigue life and stress level is expressed through the transformation

$$Z = \log(S - \gamma),$$

assuming that the pair $(\log N, Z)$ follows a bivariate normal distribution. Consequently, conditionally on the stress level, fatigue life satisfies a relationship of the form

$$\log N = \beta_0 + \beta_1 \log(S - \gamma) + \varepsilon,$$

with ε normally distributed with zero mean and constant variance. The details of the parametrization and of the transformations adopted for optimization are reported in Chapter 4.

Different model specifications were considered, distinguishing between formulations where both variables are on a logarithmic scale (log–log) and mixed configurations, where only one of the two variables is transformed. Model selection was performed by comparing the Akaike Information Criterion (AIC).
Among the considered alternatives, the log–log specification showed the lowest AIC value and was therefore selected for the subsequent analysis. In the following, the results for this configuration are reported in the case where the threshold parameter γ is estimated as a free parameter.

The numerical optimization converges correctly and provides the following estimates: $\hat{\mu}_N = 15.633$, $\hat{\sigma}_N = 1.588$, $\hat{\mu}_Z = -1.899$, $\hat{\sigma}_Z = 0.620$, with correlation coefficient $\hat{\rho} = -0.315$. The negative correlation estimate is coherent with the inverse S–N relationship, reflecting shorter lives at higher stress levels. The threshold parameter is estimated as $\hat{\gamma} = 6.100$ kN, while the maximum log-likelihood value is $-\ell(\hat{\theta}) = 17.655$.

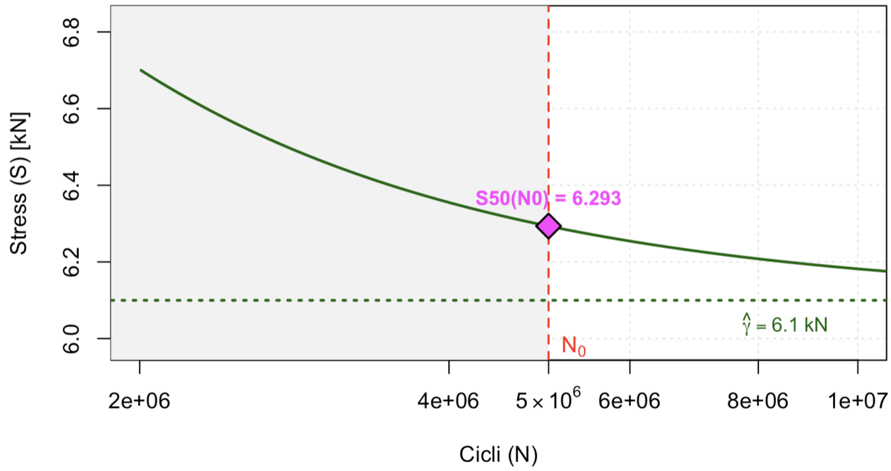


Figure 6.2: Estimated median S–N curve and the corresponding $S_{50}(N_0)$

Figure 6.2 shows the estimated median S–N curve together with the experimental data and the horizontal asymptote $\hat{\gamma}$. The median curve is obtained by considering the median of the conditional distribution $\log N \mid S$. In the following, $S_{50}(N_0)$ denotes the 50th percentile of the stress distribution at the fixed life N_0, induced by the bivariate model and defined as in Chapter 4.2.1.4. At N_0, the estimated median stress is

$$\hat{S}_{50}(N_0) = 6.293 \text{ kN}.$$

The uncertainty associated with the estimate of $S_{50}(N_0)$ was assessed using a likelihood-based approach by constructing a confidence interval based on the likelihood-ratio statistic.
For each candidate value of S_{50}, the maximum of the log-likelihood was computed under the corresponding constraint, while re-optimizing the remaining model parameters. The 95% confidence interval is therefore defined as the set of values

satisfying

$$2\Big(\ell(\hat{\theta}) - \ell(\theta_{S_{50}})\Big) \leq \chi^2_{1,0.95},$$

in accordance with Wilks' asymptotic approximation.

The resulting 95% likelihood-ratio confidence interval is

$$CI^{\mathrm{LR}}_{95\%}\Big(S_{50}(N_0)\Big) = [6.194,\ 7.106]\ \mathrm{kN}.$$

The asymmetry and width of the interval reflect the weak curvature of the likelihood surface along the directions associated with the threshold parameter γ, which is estimated close to the minimum observed stress level. Consequently, alternative parameter combinations yield statistically almost equivalent fits but lead to noticeably different values of $S_{50}(N_0)$, highlighting the limited informativeness of the available staircase sample.

6.2.1 Bivariate normal model with fixed fatigue limit

Given the small sample size, $n = 15$, estimating all six parameters of the bivariate model simultaneously places a considerable burden on the data. With such limited information, the resulting estimates may become unstable and strongly influenced by model flexibility.

For this reason, an alternative formulation was considered in which the threshold parameter γ was treated as fixed. This choice reduces the number of free parameters and allows a more stable estimation of the remaining components of the model.

In this configuration, maximum likelihood estimation is restricted to the remaining five parameters of the bivariate model, $\boldsymbol{\theta}_\gamma = (\mu_N, \mu_Z, \sigma_N, \sigma_Z, \rho)$, while the probabilistic structure remains unchanged with respect to the formulation with free γ.
The joint distribution of $(\log N, Z)$, with $Z = \log(S - \gamma)$, is still assumed to be bivariate normal in the presence of right censoring at the fixed life N_0, and the likelihood function retains the same form.

The analysis explored values of γ within the interval $[0.8\, S_{\min},\ 0.99\, S_{\min}]$. This range was selected in light of the free-γ estimation, where the estimate $\hat{\gamma}$ is located close to the minimum observed stress level $S_{\min}$. This suggests that the information contained in the staircase sample is concentrated in the lower region of the stress domain and that significantly lower threshold values would be poorly consistent with the physical meaning of a theoretical fatigue limit.

The Akaike Information Criterion,

$$\mathrm{AIC} = 2k - 2\ell(\hat{\theta}_\gamma), \qquad k = 5,$$

varies only marginally across the considered interval. This behavior indicates that the sample does not have sufficient discriminating power to clearly prefer one plausible threshold value over another.

From the point of view of point estimates, the median stress at the fixed life, $\hat{S}_{50}(N_0)$, is relatively stable as γ varies and lies in the interval [6.50, 6.58] kN. These values are systematically higher than the estimate obtained in the model with free γ.

The difference between the five- and six-parameter models is related to the different degree of statistical flexibility. In the free γ case, the model has one additional degree of freedom that allows the asymptote position to adjust in order to maximize the likelihood of the sample, producing a slightly more conservative estimate of $S_{50}(N_0)$. In contrast, fixing γ imposes a more rigid structure on the S–N curve: the asymptote position is constrained a priori, and the other parameters must compensate for this restriction, resulting in an upward shift of the curve at life N_0.

For each fixed value of the threshold parameter γ, the uncertainty on $S_{50}(N_0)$ was assessed through profile likelihood. Since $S_{50}(N_0)$ is a smooth function of the parameter vector $\boldsymbol{\theta}_\gamma = (\mu_N, \mu_Z, \sigma_N, \sigma_Z, \rho)$, the log-likelihood was profiled with respect to this derived quantity.

Specifically, for each candidate value s of $S_{50}(N_0)$, the log-likelihood was maximized over the parameter space under the constraint that the model implies $S_{50}(N_0) = s$. According to Wilks' theorem the 95% confidence interval was therefore defined as the set of values s such that

$$2\left\{\ell(\hat{\boldsymbol{\theta}}_\gamma) - \ell_p(s)\right\} \leq \chi^2_{1,0.95},$$

where $\ell_p(s)$ denotes the profile log-likelihood of $S_{50}(N_0)$. Across the explored range of threshold values, the confidence intervals for $S_{50}(N_0)$ are relatively wide and largely overlapping, typically lying approximately between 6.1 and 6.8 kN. This behavior confirms that, although fixing γ reduces model flexibility, the available staircase data

γ	$\gamma/S_{\min}$	$\hat{S}_{50}(N_0)$	$\mathrm{CI}_{95\%}$	AIC
4.88	0.8	6.517	[6.12,6.89]	45.92
5.49	0.9	6.527	[6.02, 7.33]	45.85
5.795	0.95	6.505	[6.12, 6.90]	45.77
5.917	0.97	6.492	[6.09, 6.89]	45.70
6.039	0.99	6.459	[6.06, 6.86]	45.54

Table 6.1: Estimation of $S_{50}(N_0)$, likelihood-ratio 95% confidence intervals and the AIC for different fixed values of γ.

remain weakly informative with respect to the median fatigue strength at the fixed life N_0.

Table 6.1 reports the estimated values of $\hat{S}_{50}(N_0)$, the corresponding likelihood-ratio 95% confidence intervals and the AIC values for the different fixed choices of the threshold parameter γ.

Overall, the results highlight two main aspects. On the one hand, the point estimate of $S_{50}(N_0)$ is relatively robust to moderate variations of the threshold parameter. On the other hand, the wide and overlapping confidence intervals reveal substantial inferential uncertainty, driven primarily by the limited sample size and the binary nature of the staircase information. The uncertainty on the threshold position directly affects the inference on $S_{50}(N_0)$, suggesting caution in assuming a value of γ without strong physical justification or support from a larger sample.

6.2.1.1 Profile likelihood of the fatigue threshold γ

The identification of the threshold parameter γ is investigated through profile likelihood, adopting a likelihood-based approach consistent with the framework of Pascual and Meeker [15] for censored fatigue models. The profile likelihood provides a natural tool for assessing the identifiability of a structural parameter and for constructing confidence regions based on Wilks' theorem.

Consider the log-likelihood of the bivariate normal model (Eq. (4.8)), based on the Stromeyer-type relationship $\log N = \beta_0 + \beta_1 \log(S - \gamma) + \varepsilon$, with ε normally distributed with zero mean and constant variance, and with right-censoring at the fixed life N_0. In terms of the adopted parametrization, the log-likelihood can be written as

$$\ell(\boldsymbol{\theta}, \gamma), \qquad \boldsymbol{\theta} = (\mu_N, \mu_Z, \sigma_N, \sigma_Z, \rho).$$

The profile log-likelihood of γ is defined as

$$\ell_p(\gamma) = \max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}, \gamma),$$

that is, the maximum log-likelihood obtained by fixing γ and estimating the remaining five parameters by maximum likelihood, yielding the profile $\ell_p(\gamma)$. It was evaluated on the grid $\gamma \in [0.95\, S_{\min},\ 0.99\, S_{\min}]$, corresponding to the interval $[5.795,\ 6.039]$ kN.

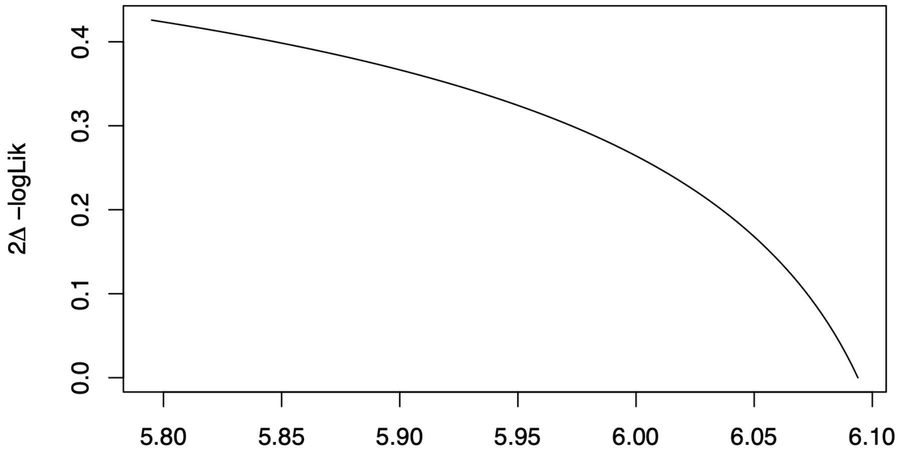


Figure 6.3: Profile likelihood of the threshold parameter γ.

Figure 6.3 shows that the increase in the statistic $2\Delta(\gamma)$ remains modest over the explored range of the threshold parameter. Even for variations of γ of the order of a few tenths of kN, the likelihood ratio statistic remains well below the critical value $\chi^2_{1,0.95}$. This indicates that a relatively wide range of γ values provides statistically comparable fits, suggesting limited precision in the estimation of the asymptotic stress level within the available staircase sample.

Under regularity conditions, Wilks' theorem provides the asymptotic approximation

$$2\Delta(\gamma) = 2\Big(\ell_p(\hat{\gamma}) - \ell_p(\gamma)\Big) \ \dot{\sim}\ \chi^2_1.$$

Accordingly, the 95% confidence region for the threshold parameter is defined as

$$\Gamma_{0.95} = \Big\{\gamma : 2\Big(\ell_p(\hat{\gamma}) - \ell_p(\gamma)\Big) \le \chi^2_{1,0.95}\Big\}.$$

Within this entire range, the likelihood-ratio statistic $2\Delta(\gamma)$ remains below the critical value $\chi^2_{1,0.95}$. Consequently, all the investigated values of γ belong to the 95% profile-likelihood confidence region, and the resulting interval coincides with the whole explored domain.
This behavior indicates that, within the practically relevant region close to $S_{\min}$ suggested by the six-parameter fit, the available staircase data do not provide sufficient curvature in the likelihood to discriminate sharply between alternative threshold positions.

For each admissible value of γ in this region, the corresponding median fatigue strength at the fixed life N_0 was computed by conditional maximum likelihood estimation of the remaining five parameters. Over the entire plausible set of γ, the induced variation is $S_{50}(N_0) \in [6.505,\ 6.577]$ kN.

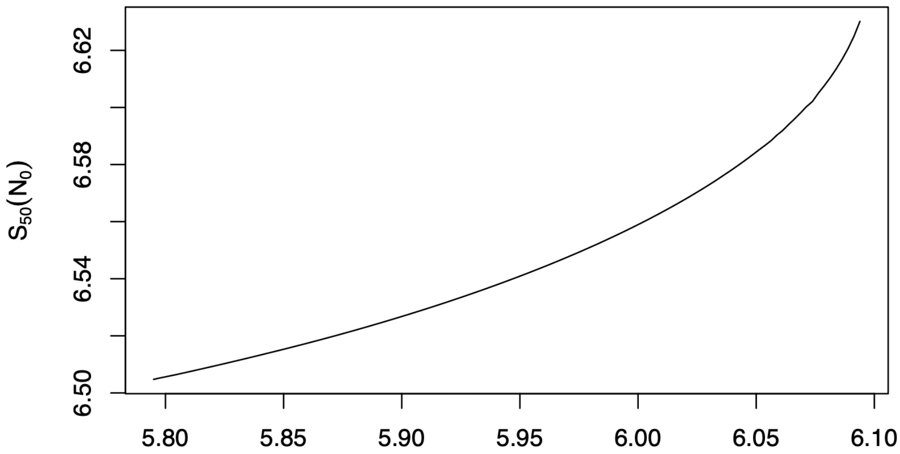


Figure 6.4: Variation of the estimated $S_{50}(N_0)$ as a function of the threshold parameter γ along the likelihood profile.

As shown in Figure 6.4, the relationship between $S_{50}(N_0)$ and γ is monotone increasing and mildly nonlinear: larger threshold values shift the median S–N curve upward in the stress domain at the fixed life N_0. The induced variation of $S_{50}(N_0)$ over the 95% plausible region of γ remains comparatively limited, suggesting that, although the precise location of the threshold is subject to moderate uncertainty, the median strength estimate at N_0 is relatively stable with respect to admissible variations of the asymptote.

6.3 Random Fatigue Limit model

In order to evaluate an alternative structural hypothesis with respect to the model with deterministic threshold, the Random Fatigue Limit model with immunity component was applied. In this formulation, the fatigue limit is not treated as a fixed parameter common to all specimens, but as a random variable varying from unit to unit, thereby introducing an additional source of structural heterogeneity, as detailed in Chapter 5. It is assumed that

$$V = \log \Gamma \sim \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2),$$

where $\Gamma = (\gamma_1, \ldots, \gamma_n)$ while, conditionally on the individual limit γ_i and on the applied stress, the fatigue life on the logarithmic scale follows a Stromeyer-type relationship. Whenever $\gamma_i \geq S_i$, the specimen is considered immune to failure. The model therefore explicitly incorporates a probability mass associated with specimens that are structurally non-susceptible to failure at the applied stress level, while runout observations are treated as right-censored at the fixed life N_0.

Maximum likelihood estimation converges correctly; however, the parameter associated with the distribution of the fatigue limit is estimated in such a way that the median limit,

$$\Gamma_{50} = \exp(\hat{\mu}_\Gamma),$$

assumes a value on the order of 10^{-8} kN, thus negligible compared to the stress levels observed in the sample (ranging between 6 and 7 kN). This result indicates that the estimated distribution of the fatigue limit is concentrated at values far below the informative region of the experimental data. In this case, since the estimated fatigue limit is approximately zero, the model effectively reduces to a Basquin-type S–N relation.

The marginal probability of failure by N_0 was computed by integrating with respect to the distribution of the individual fatigue limit, yielding

$$\hat{S}_{50}(N_0) = 6.399 \text{ kN},$$

with quantiles

$$S_{10} = 5.610 \text{ kN}, \qquad S_{90} = 7.300 \text{ kN}.$$

The 95% confidence interval for $S_{50}(N_0)$, obtained through a likelihood-based approach, is

$$CI_{95\%}\Big(S_{50}(N_0)\Big) = [6.106,\ 7.083] \text{ kN}.$$

Overall, the RFL model does not provide evidence of a well-identified positive fatigue limit in the analyzed sample. The collapse of the estimated median limit toward values close to zero suggests that the information contained in the data is insufficient to support the existence of a structural threshold distinct from the intrinsic dispersion of fatigue life. In this configuration, the inferential behavior of the model is dominated by the variability of the conditional life distribution, while the immunity component remains formally present but not substantially supported by the empirical evidence.

The estimate of $S_{50}(N_0)$ is slightly higher than the B_{50} value obtained with the Dixon–Mood method, although the two estimates remain close.

6.3.1 Bayesian Random Fatigue Limit model

The Random Fatigue Limit model with immunity was also estimated in a Bayesian framework using JAGS.

The probabilistic specification coincides with the frequentist RFL formulation. The fatigue limit is treated as a specimen-specific latent variable, with $V = \log \Gamma \sim \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2)$, while, conditionally on (S_i, γ_i), fatigue life follows a Stromeyer-type relationship: $\log N_i \mid (S_i, \gamma_i) \sim \mathcal{N}\Big(\beta_0 + \beta_1 \log(S_i - \gamma_i), \sigma^2\Big), \quad \text{for } \gamma_i < S_i$. If $\gamma_i \geq S_i$, the specimen is considered immune to failure. Runout observations are handled as right-censored at the fixed life N_0.

Posterior inference was obtained from three parallel MCMC chains. Convergence diagnostics based on the Gelman–Rubin statistic indicate satisfactory mixing for all monitored parameters, with potential scale reduction factors close to 1, in particular all below 1.06. No evidence of lack of convergence was detected.

The posterior distribution of the median fatigue strength at the fixed life N_0 yields a posterior median estimate

$$S_{50}(N_0) = 6.469 \text{ kN},$$

with a 95% credible interval

$$CrI_{95\%}\Big(S_{50}(N_0)\Big) = [5.134,\ 7.968]\ \text{kN}.$$

This interval is substantially wider than those obtained under the frequentist formulations, reflecting the full propagation of uncertainty on both the structural parameters and the latent fatigue limits.

The posterior distribution of the fatigue limit parameter provides additional insight. The median of μ_Γ corresponds to a fatigue limit of approximately 5.5 kN, while the 95% credible region for Γ spans from values close to zero up to stress levels comparable with those applied experimentally. Such a broad posterior distribution indicates that the data do not contain sufficient information to precisely locate a positive fatigue limit within the observed stress range.

Unlike the frequentist RFL formulation, which drives the estimated fatigue limit toward negligible values, the Bayesian approach does not collapse the threshold to a boundary solution. Instead, it expresses the structural uncertainty through a very wide posterior distribution. In practical terms, with $n = 15$ staircase observations and binary failure/runout information only, the existence and position of a distinct fatigue limit cannot be reliably inferred. The uncertainty on the threshold propagates directly to $S_{50}(N_0)$, resulting in the widest uncertainty interval among the considered models.

6.4 Comparative discussion

The different approaches considered in this chapter provide structurally distinct representations of fatigue behavior at the fixed life $N_0 = 5 \times 10^6$ cycles. Since all methods are fitted to the same staircase dataset, $n = 15$, binary failure/runout outcomes, stress range approximately 6.1–6.7 kN, any differences in the results arise exclusively from modeling assumptions and from the way uncertainty is quantified.

From the perspective of central estimation, the inferred median fatigue strength at N_0 is reasonably consistent across models. The Dixon–Mood procedure yields $B_{50} = 6.368$ kN, while the model-based estimates of $S_{50}(N_0)$ range from 6.293 kN in the bivariate model with free γ to approximately 6.50–6.58 kN when γ is fixed within a plausible interval below $S_{\min}$. The Random Fatigue Limit model produces

$\hat{S}_{50}(N_0) = 6.399$ kN in the frequentist formulation and 6.469 kN as posterior median in the Bayesian implementation. Overall, central estimates differ by only a few tenths of kN, indicating that the data support a relatively stable point estimate of median fatigue strength at the investigated life.

More substantial differences emerge in the associated uncertainty. The Dixon–Mood approach produces the narrowest interval; however, this should not be interpreted as evidence of greater informativeness of the data. The Dixon–Mood estimator operates within the ISO staircase framework and estimates fatigue strength locally at N_0 without introducing an explicit S–N functional form, slope parameter, or threshold component, with accuracy depending on the staircase step size being on the same order as the standard deviation of the fatigue-strength distribution. In contrast, the parametric models embed additional structural elements, including slope terms and, crucially, a deterministic or random fatigue-limit parameter. With only 15 binary observations, these additional degrees of freedom are weakly identified, and the corresponding uncertainty is therefore propagated into wider confidence or credible intervals. The larger intervals observed in the parametric approaches thus reflect the cost of structural modeling in a small-sample setting rather than numerical instability.

The treatment of the threshold parameter represents the main structural distinction among models. In the bivariate formulation, γ is weakly identified and is driven close to $S_{\min}$ when estimated freely; the profile likelihood exhibits limited curvature, confirming practical non-identifiability. Fixing γ can be interpreted as a form of regularization, which stabilizes inference but induces a systematic upward shift in $S_{50}(N_0)$. In the frequentist Random Fatigue Limit model, the estimated fatigue-limit distribution collapses toward negligible values, suggesting that the data do not support the existence of a positive threshold distinct from intrinsic life variability. In the Bayesian RFL formulation, instead of collapsing to a boundary solution, the posterior distribution of Γ becomes extremely diffuse, spanning values close to zero up to stress levels comparable with those experimentally applied. This structural uncertainty propagates directly to $S_{50}(N_0)$ and results in the widest uncertainty interval among the considered models.

Taken together, the results indicate that although the central estimates of median fatigue strength lie within a relatively narrow numerical band (approximately 6.29–

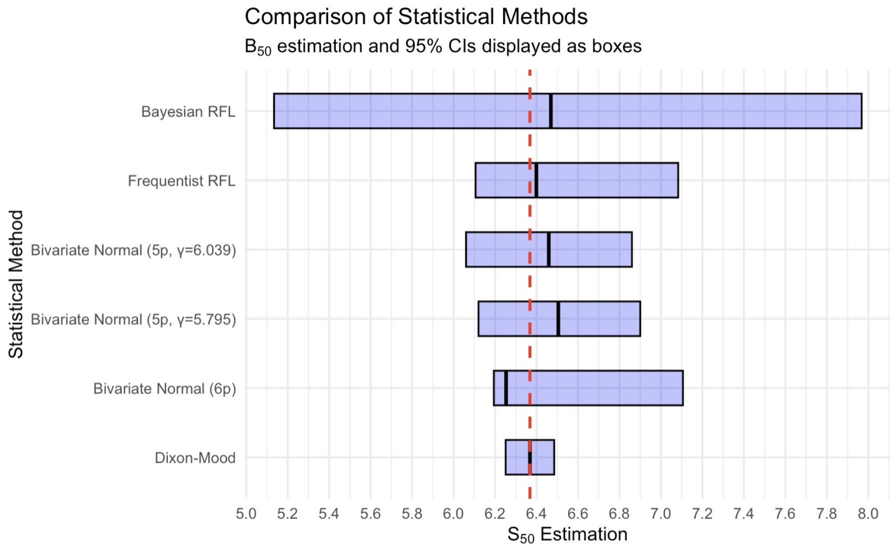


Figure 6.5: Comparison of the estimated $S_{50}(N_0)$ values and the corresponding 95% confidence intervals obtained using the different statistical methods. Each box represents the interval estimate, while the horizontal line indicates the corresponding point estimate. The red dashed line represents the value obtained using the ISO 12107 Dixon–Mood method.

6.57 kN), the associated uncertainty remains substantial once structural components are introduced. The instability in identifying the threshold or fatigue-limit parameter and the sensitivity of interval widths to modeling assumptions both suggest that the staircase dataset is weakly informative. With only $n = 15$ specimens, binary failure/runout information, and a limited stress range, the available data do not allow robust identification of asymptotic S–N behavior or of a distinct fatigue limit. The proximity of the point estimates should not be interpreted as evidence of strong identification, but rather as a consequence of the limited information contained in the available staircase data. The substantial width of the associated intervals indicates a considerable level of structural and inferential uncertainty. Additional experimental information, such as a wider stress range, a larger sample size, or tests conducted at multiple fixed lives, would therefore be required to reliably support richer parametric formulations.

Chapter 7

Conclusions and Future Work

The work presented in this thesis originates from an industrial problem related to fatigue testing of high-strength fasteners. In current practice, fatigue strength is commonly estimated according to the ISO 12107 standard using the staircase method and the Dixon–Mood procedure, from which tolerance limits such as the B_{50} value are derived. Industrial experience, however, suggests that these estimates may be overly conservative.

Using experimental staircase data from fatigue tests on high-strength bolts, kindly provided by Prof. Carlo Rosso, this study investigates alternative model-based approaches for estimating fatigue strength parameters at a fixed life $N_0 = 5 \times 10^6$ cycles. Starting from the classical ISO 12107 Dixon–Mood procedure, progressively more flexible statistical formulations are considered, including a bivariate normal Stromeyer-type model and a Random Fatigue Limit (RFL) model with immunity, estimated under both frequentist and Bayesian frameworks.

The comparative analysis highlights that, although different structural assumptions lead to slightly different point estimates of the median fatigue strength, the dominant feature across all models is the substantial inferential uncertainty. The estimated values of $S_{50}(N_0)$ range approximately between 6.25 and 6.47 kN depending on the formulation, but the associated confidence and credible intervals are wide and strongly influenced by modeling choices.

In particular, the threshold parameter γ proves to be weakly identifiable in the considered dataset. In the bivariate model with free γ, the estimate is driven close to the minimum observed stress level, and the profile likelihood reveals very

limited curvature in that direction, indicating practical non-identifiability. Fixing γ stabilizes the estimation of the remaining parameters but induces a systematic shift in $S_{50}(N_0)$, illustrating the sensitivity of the results to structural constraints imposed in a small-sample setting.
The Random Fatigue Limit model further confirms the lack of information on the existence of a distinct fatigue limit. In the frequentist implementation, the estimated distribution of the fatigue limit collapses toward negligible values, effectively removing evidence of a positive asymptotic threshold. In the Bayesian formulation, instead of collapsing to a boundary solution, the posterior distribution of Γ becomes extremely diffuse, spanning from values close to zero up to stress levels comparable to those tested experimentally. These two behaviors are not contradictory but represent different inferential expressions of the same structural limitation: the data do not contain sufficient information to identify a fatigue limit separate from intrinsic life variability.

This limitation is intrinsic to the statistical structure of staircase experiments. In the staircase procedure, the random quantity is not the applied stress S, but the binary indicator of whether the fatigue life exceeds a fixed threshold N_0. Consequently, the appropriate likelihood is constructed from the conditional distribution $N \mid S$. With only one fixed lifetime threshold and a limited stress range, this likelihood inherently contains limited information about the asymptotic behavior of the S–N curve and about the existence or location of a fatigue limit. The weak identifiability observed for γ or for the random fatigue limit Γ is therefore not a numerical artifact, but a structural consequence of the experimental design.

Overall, with only $n = 15$ staircase observations, binary failure/runout information, and a relatively narrow stress range, approximately 6.1 – 6.7 kN, the dataset is weakly informative. The data do not allow robust discrimination between deterministic and random threshold formulations, nor do they support precise inference on the median fatigue strength.

Future developments should therefore focus primarily on enhancing the informational richness of the data. Increasing the sample size would directly improve parameter stability. Even more fundamentally, performing staircase experiments at multiple fixed lives $N_0, N_1, \ldots$ would substantially enrich the likelihood structure and allow a clearer separation between slope effects and threshold effects.

From a modeling perspective, richer datasets would also enable the development of unified hierarchical formulations combining the bivariate Stromeyer-type approach with a specimen-specific random fatigue limit. Such a fully integrated bivariate random-threshold model could simultaneously capture slope, asymptotic behavior, and inter-specimen heterogeneity, but its identifiability requires substantially more informative experimental designs.

In conclusion, model-based approaches offer a flexible and coherent framework for fatigue parameter estimation. However, in small-sample staircase settings, their practical performance is fundamentally constrained by the limited information available. Reliable identification of a fatigue limit, deterministic or random, cannot rely solely on single-life staircase data, but requires experimental designs specifically tailored to capture the full finite-sample and asymptotic S–N behavior.

Appendix A

Dixon & Mood – R code

This appendix reports the R notebook used to perform the staircase analysis. The procedure includes:

- Data import and preprocessing;
- Construction of the test sequence table;
- Identification of the rare event and computation of intermediate parameters;
- Validity check according to Dixon & Mood through parameter D;
- Estimation of $\hat{\mu}_y$ (B50) and $\hat{\sigma}_y$;
- Calculation of the lower tolerance limit.

The required R libraries are first loaded. These packages are used for data manipulation, reading the Excel dataset, and performing statistical computations.

```
library(readxl)
library(dplyr)
library(tidyr)
library(tolerance)
```

Listing A.1: Loading required libraries

The dataset is imported from an Excel file and the relevant variables are renamed. The stress step d used in the staircase procedure is also defined.

```r
path <- "~/Desktop/staircase/staircase.xlsx"

imported_data <- read_excel(path, sheet = "staircase 1", skip = 5)
data <- imported_data %>%
  rename(
    stress_level = "Fa [kN]",
    outcome = "end of test criterion"
  )

# Define stress step (Delta F)
stress_step <- 0.15
Ntarget <- 5.00e+06
```

The following code generates the table used to visualize the fatigue test outcomes. Each specimen is identified by an index and its result is encoded symbolically: `X` for failure and `O` for run-out. The data are then reshaped into a wide format and ordered in descending stress level to reproduce the standard layout of staircase test results.

```r
data_for_table <- data %>%
  mutate(
    specimen_n = row_number(),
    symbol = ifelse(outcome == "failed", "X", "O")
  ) %>%
  select(specimen_n, stress_level, symbol)

table_output <- data_for_table %>%
  pivot_wider(
    names_from = specimen_n,
    values_from = symbol,
    values_fill = ""
  ) %>%
  arrange(desc(stress_level))
```

The total number of failures and run-outs is computed. The rare event is defined as the least frequent outcome.

```r
total_counts <- table(data$outcome)

# Identify the least frequent outcome (rare event)
rare_event <- names(total_counts)[which.min(total_counts)]

# Number of rare events
n <- min(total_counts)
```

The data are grouped by stress level and the frequency f_i of rare events at each level is computed.

S_0 is defined as the lowest stress level at which at least one rare event occurs. The level index is then defined as:

$$i = \frac{S - S_0}{d}$$

```r
data_analysis <- data %>%
  filter(outcome == rare_event) %>%
  group_by(stress_level) %>%
  summarise(fi = n()) %>%
  arrange(stress_level)

# Base level S0
S0 <- min(data_analysis$stress_level)

# Assign level index i
data_analysis <- data_analysis %>%
  mutate(i = round((stress_level - S0) / stress_step))
```

The following parameters are defined:

$$C = \sum f_i, \quad A = \sum (i f_i), \quad B = \sum (i^2 f_i)$$

The normalized sample variance parameter is:

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

```
d <- stress_step

C <- sum(data_analysis$fi)
A <- sum(data_analysis$i * data_analysis$fi)
B <- sum(data_analysis$i^2 * data_analysis$fi)

if (C < 2) {
  stop("Error: Fewer than two rare events. Standard deviation cannot be
     estimated.")
}

D <- (B * C - A^2) / C^2
```

According to the adopted procedure, the estimate of the standard deviation is considered reliable if $D > 0.3$.

```
cat("\n--- CHECK DIXON AND MOOD (Estimate Validity) ---\n")

if (D > 0.3) {
  cat(paste("Result: Valid test. (D =", round(D, 4), "is > 0.3)\n"))
  cat("The estimate of the standard deviation is reliable.\n")
} else {
  cat(paste("Result: NOT valid test. (D =", round(D, 4), "is <= 0.3)\n"))
  cat("The estimate of the standard deviation is not reliable.\n")
  cat("The step size d may be too large compared to the actual standard
     deviation.\n")
}
```

The mean estimate (B50) is computed as:

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

The standard deviation is estimated as:

$$\hat{\sigma}_y = 1.62\, d\, (D + 0.029)$$

```r
if (rare_event == "survived") {
  mu_factor <- 0.5
} else {
  mu_factor <- -0.5
}

# Mean estimate (B50)
mean_B50 <- S0 + d * (A / C + mu_factor)

# Estimated standard deviation
std_dev <- 1.62 * d * (D + 0.029)
```

The lower bound of the fatigue strength for a failure probability $\mathbb{P}$ and confidence level $1 - \alpha$ is calculated as:

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

```r
# Degrees of freedom (v = n - 1)
total_n <- nrow(data)
v <- n - 1

failure_probability_P <- 0.10
confidence_level <- 0.95
P_coverage <- 1 - failure_probability_P
alpha_value <- 1 - confidence_level

# One-sided tolerance factor
k_factor <- tolerance::K.factor(n, alpha = alpha_value, P = P_coverage)
lower_limit <- mean_B50 - k_factor * std_dev
```

Appendix B

Bivariate model – R code

B.0.1 Gamma as free parameter

This appendix reports the R code used to estimate the bivariate model for staircase data in which the relationship between S and N is defined by the equation 4.3 where fatigue-limit parameter γ is treated as an unknown quantity and estimated jointly with the remaining parameters of the model.
The code implements the likelihood formulation described in the chapter 4 and applies numerical optimization to estimate the model parameters using the staircase fatigue test data provided by Prof. Rosso.
The procedure consists of the following main steps:

- loading the required R libraries;
- importing and preprocessing the staircase dataset;
- defining the censored likelihood function for the bivariate model;
- numerical maximization of the likelihood and parameter reconstruction;
- computation of derived quantities such as the S_{50} at the target life N_0;
- graphical representation of the estimated median S–N curve.

The required R libraries are first loaded. These packages are used for data manipulation, reading the Excel dataset, and performing statistical computations.

```
library(stats)
library(readxl)
library(dplyr)
library(tidyr)
library(tolerance)
set.seed(42)
```

The experimental data from the staircase fatigue test are then imported from an Excel file. The relevant variables are renamed for convenience, and the quantities required for the statistical model are computed. In particular, the logarithm of the number of cycles is calculated and a binary indicator variable is defined to distinguish failures from run-outs.

```
n <- 15
path <- "~/Desktop/staircase/staircase.xlsx"
imported_data<- read_excel(path, sheet = "staircase 1",skip=5)
data <- imported_data %>%
  rename(
    S = "Fa [kN]",
    outcome = "Test termination criterion",
    N="CTest cycles",
  )

N=data$N
stress_step<-0.15
N0<-5.00e+06
S = data$S
logN <- log(N)
logN0 <- log(N0)

# Indicator: 1 = failure, 0 = run-out
x_i <- ifelse(data$outcome=="failure", 1, 0)

data <- data.frame(S = S, logN = logN, x = x_i)
```

The following code implements the censored negative log-likelihood for the bivariate model in terms of the variables $\log N$ and $Z = \log(S - \gamma)$, assuming a joint normal model. The parameter vector contains six unknown quantities: μ_N, μ_Z, σ_N, σ_Z, ρ, and the fatigue-limit parameter γ.

Failures contribute through the conditional normal density of $\log N$, while run-outs are treated as right-censored observations and contribute through the survival probability $\mathbb{P}(\log N > \log N_{\text{obs}})$.

For numerical optimization, constrained parameters are expressed through an unconstrained parameter vector as described in Sections 4.2.1.2 and 4.2.1.3.

```
negloglik_gamma6 <- function(par, data) {
  muN <- par[1] # mean of logN
  muZ <- par[2] # mean of Z = log(S - gamma)
  log_sigN <- par[3] # log(sigma_N)
  log_sigZ <- par[4] # log(sigma_Z)
  atanh_rho <- par[5] # transformation for rho
  eta_gamma <- par[6] # unconstrained parameter for gamma

  # Transformations to enforce constraints
  sigN <- exp(log_sigN) # > 0
  sigZ <- exp(log_sigZ) # > 0
  rho <- tanh(atanh_rho) # in (-1, 1)

  S <- data$S
  logN <- data$logN
  x <- data$x # 1 = failed, 0 = run-out

  S_min <- min(S)

  # Since I want 0 < gamma < S_min, we use the logit function
  gamma <- S_min * plogis(eta_gamma)

  # Defining the new covariate: Z = log(S - gamma)
  Z <- log(S - gamma)

```

```
  # Conditional logN | Z ~ Normal
  mu_cond <- muN + rho * (sigN / sigZ) * (Z - muZ)
  sd_cond <- sigN * sqrt(1 - rho^2)

  # Log-likelihood:
  # - failed: normal density in logN
  # - survivors: P(logN > logN_obs)
  logL_failed <- dnorm(logN, mean = mu_cond, sd = sd_cond, log = TRUE)
  logL_surv <- pnorm(logN, mean = mu_cond, sd = sd_cond, lower.tail =
      FALSE, log.p = TRUE)

  # Log-likelihood with x (1 = failed, 0 = run-out)
  total_logL <- sum(x * logL_failed + (1 - x) * logL_surv)

  return(-total_logL) # for minimization with optim
}
```

To estimate the model parameters, suitable initial values are specified and numerical minimization of the negative log-likelihood is performed using `optim()`. Details on the initialization strategy and parameter transformations are provided in Sections 4.2.1.2 and 4.2.1.3.

After convergence, the estimated parameters are mapped back to their natural scale and used to compute derived quantities of interest, including the fatigue strength S_{50} corresponding to the target life N_0, defined in Section 4.2.1.4.

```
S_min=min(S)
S_max=max(S)
gamma_start <- 0.90 * S_min
Z_start <- log(S - gamma_start)

start_muN <- mean(logN)
start_muZ <- mean(Z_start)
start_log_sigN <- log(sd(logN))
start_log_sigZ <- log(sd(Z_start))
start_rho <- 0
start_eta_g <- qlogis(gamma_start / S_min)
```

```

start_par <- c(start_muN,
               start_muZ,
               start_log_sigN,
               start_log_sigZ,
               start_rho,
               start_eta_g)

fit6 <- optim(par = start_par,
              fn = negloglik_gamma6,
              data = data,
              method = "BFGS",
              control = list(maxit = 5000))

fit6$value # value of -log-likelihood
fit6$par # Parameter estimates

# Parameter reconstruction
par_hat <- fit6$par

muN_hat <- par_hat[1]
muZ_hat <- par_hat[2]
sigN_hat <- exp(par_hat[3])
sigZ_hat <- exp(par_hat[4])
rho_hat <- tanh(par_hat[5])
eta_g_hat <- par_hat[6]

S_min <- min(S)
gamma_hat <- S_min * plogis(eta_g_hat)

muS_hat <- gamma_hat + exp(muZ_hat + 0.5 * sigZ_hat^2)
VarS_hat <- exp(2*muZ_hat + sigZ_hat^2) * (exp(sigZ_hat^2) - 1)
sigS_hat <- sqrt(VarS_hat)

# find Z0 = E[Z | logN = logN0]
```

```
Z0 <- muZ_hat + rho_hat * (sigZ_hat / sigN_hat) * (logN0 - muN_hat)

# transform Z0 in S50 = gamma_hat + exp(Z0)
S50 <- gamma_hat + exp(Z0)

k1 <- length(fit6$par) # number of parameter estimates
negLL1 <- fit6$value # -log L_max
AIC1 <- 2 * k1 + 2 * negLL1
fit6$convergence
```

Finally, the estimated median S–N relationship is computed and displayed graphically. The curve represents the conditional median of $\log N$ given the stress level S, obtained through the transformation $Z = \log(S - \hat{\gamma})$.
The S_{50} corresponding to the target life N_0 is also shown on the estimated curve.

```
logN_med_S <- function(S_val) {
  Z_val <- log(S_val - gamma_hat)
  muN_hat + rho_hat * (sigN_hat / sigZ_hat) * (Z_val - muZ_hat)
}

S_seq <- seq(from = gamma_hat + 0.01,
             to = max(S),
             length.out = 200)

logN_seq <- logN_med_S(S_seq)
N_seq <- exp(logN_seq)

# S\-N Plot with logarithmic N\-axis
plot(N, S,
     log = "x",
     xlab = "N [cycles]",
     ylab = "S [MPa]",
     pch = ifelse(x_i == 1, 16, 1),
     col = ifelse(x_i == 1, "red", "blue"),
     main = "Curva S-N")

```

```
lines(N_seq, S_seq,
      col = "darkgreen",
      lwd = 2)

points(N0, S50, pch = 19, col = "darkgreen")
text(N0, S50,
     labels = "N0",
     pos = 3, cex = 0.8, col = "darkgreen")
```

B.0.2 Gamma as fixed parameter

This appendix reports the R code used to estimate the bivariate model for staircase data in which the relationship between S and N is defined by Eq. 4.3, while the fatigue-limit parameter γ is treated as fixed rather than estimated.

The probabilistic structure of the model remains unchanged with respect to the six-parameter formulation described in the previous subsection. However, the optimization is now performed over the reduced parameter vector

$$\boldsymbol{\theta}_\gamma = (\mu_N, \mu_Z, \sigma_N, \sigma_Z, \rho),$$

while the value of γ is specified a priori.

The data import, preprocessing steps, and graphical representation follow exactly the same procedure described in the previous subsection. Only the components that differ substantially from the six-parameter implementation are reported below.

The following code implements the censored negative log-likelihood for the bivariate model in terms of the variables $\log N$ and $Z = \log(S - \gamma)$, assuming a joint normal model with fixed fatigue-limit parameter.

The parameter vector therefore contains five unknown quantities: μ_N, μ_Z, σ_N, σ_Z, and ρ.

Failures contribute through the conditional normal density of $\log N$, while run-outs are treated as right-censored observations and contribute through the survival probability $\mathbb{P}(\log N > \log N_{\text{obs}})$.

For numerical optimization, constrained parameters are expressed through an unconstrained parameter vector as described in Sections 4.2.1.2 and 4.2.1.3.

```r
negloglik_gamma5 <- function(par, data, gamma) {

  muN <- par[1]   # mean of logN
  muZ <- par[2]   # mean of Z = log(S - gamma)
  log_sigN <- par[3]   # log(sigma_N)
  log_sigZ <- par[4]   # log(sigma_Z)
  atanh_rho <- par[5]   # transformation for rho

  # Transformations to enforce constraints
  sigN <- exp(log_sigN)  # > 0
  sigZ <- exp(log_sigZ)  # > 0
  rho <- tanh(atanh_rho) # in (-1, 1)

  S <- data$S
  logN <- data$logN
  x <- data$x   # 1 = failure, 0 = run-out

  # Transformed covariate
  Z <- log(S - gamma)

  # Conditional distribution logN | Z
  mu_cond <- muN + rho * (sigN / sigZ) * (Z - muZ)
  sd_cond <- sigN * sqrt(1 - rho^2)

  # Log-likelihood contributions
  logL_failed <- dnorm(logN, mean = mu_cond, sd = sd_cond, log = TRUE)
  logL_surv <- pnorm(logN,
                     mean = mu_cond,
                     sd = sd_cond,
                     lower.tail = FALSE,
                     log.p = TRUE)

  total_logL <- sum(x * logL_failed + (1 - x) * logL_surv)
  return(-total_logL)
```

```
}
```

To estimate the model parameters, a grid of candidate values for the fixed fatigue-limit parameter γ is considered.

For each value, the transformed variable $Z = \log(S - \gamma)$ is used to construct suitable starting values, and the corresponding five-parameter negative log-likelihood is minimized numerically using `optim()`. The resulting parameter estimates are then mapped back to their natural scale and used to compute derived quantities, including the median fatigue strength S_{50} at the target life N_0, defined as in Section 4.2.1.4.

For each fixed value of γ, the estimated median S–N curve is also computed and displayed graphically, together with the corresponding point S_{50} at N_0.

```
S_min <- min(S)
gamma_grid <- S_min * c(0.8, 0.9, 0.95, 0.97, 0.99)
res_list <- vector("list", length(gamma_grid))

for (j in seq_along(gamma_grid)) {

  fixed_gamma <- gamma_grid[j]

  # Starting values
  Z_start <- log(S - fixed_gamma)

  start_par5 <- c(
    start_muN = mean(logN),
    start_muZ = mean(Z_start),
    start_log_sigN = log(sd(logN)),
    start_log_sigZ = log(sd(Z_start)),
    start_rho = 0
  )

  fit5 <- optim(par = start_par5,
                fn = negloglik_gamma5,
                data = data,
                gamma = fixed_gamma,
```

```
                method = "BFGS",
                control = list(maxit = 5000))

negLL <- fit5$value  # value of -log-likelihood
par_hat <- fit5$par  # estimate parameters

k <- length(par_hat)  # number of estimated parameters
AIC <- 2 * k + 2 * negLL

# Parameter reconstruction
muN_hat <- par_hat[1]
muZ_hat <- par_hat[2]
sigN_hat <- exp(par_hat[3])
sigZ_hat <- exp(par_hat[4])
rho_hat <- tanh(par_hat[5])

muS_hat <- fixed_gamma + exp(muZ_hat + 0.5 * sigZ_hat^2)
VarS_hat <- exp(2 * muZ_hat + sigZ_hat^2) * (exp(sigZ_hat^2) - 1)
sigS_hat <- sqrt(VarS_hat)

# Median fatigue strength at N0
Z0 <- muZ_hat + rho_hat * (sigZ_hat / sigN_hat) * (logN0 - muN_hat)
S50 <- fixed_gamma + exp(Z0)

res_list[[j]] <- data.frame(
  gamma = fixed_gamma,
  muN_hat = muN_hat,
  muZ_hat = muZ_hat,
  sigN_hat = sigN_hat,
  sigZ_hat = sigZ_hat,
  rho_hat = rho_hat,
  negLL = negLL,
  AIC = AIC,
  convergence = fit5$convergence,
  muS_hat = muS_hat,
```

```
    VarS_hat = VarS_hat,
    sigS_hat = sigS_hat,
    S50_NO = S50
  )

  # Estimated median S-N curve
  logN_med_S <- function(S_val) {
    Z_val <- log(S_val - fixed_gamma)
    muN_hat + rho_hat * (sigN_hat / sigZ_hat) * (Z_val - muZ_hat)
  }

  S_seq <- seq(from = fixed_gamma + 0.01,
               to = max(S),
               length.out = 200)
  logN_seq <- logN_med_S(S_seq)
  N_seq <- exp(logN_seq)

  plot(N, S,
       log = "x",
       xlab = "N [cycles]",
       ylab = "S [MPa]",
       pch = ifelse(x_i == 1, 16, 1),
       col = ifelse(x_i == 1, "red", "blue"),
       main = paste0("S-N curve, fixed gamma = ", round(fixed_gamma, 3)))

  lines(N_seq, S_seq, col = "darkgreen", lwd = 2)
  abline(h = fixed_gamma, col = "darkgreen", lty = 3, lwd = 1)

  points(N0, S50, pch = 19, col = "magenta")
  text(N0, S50,
       labels = "N0",
       pos = 3,
       cex = 0.8,
       col = "magenta")
}
```

```

risultati <- do.call(rbind, res_list)
risultati
```

Appendix C

Random Fatigue Limit (RFL) model – R code

C.0.1 Frequentist version

This appendix reports the R code used to estimate the Random Fatigue Limit (RFL) model under a frequentist framework, as described in Chapter 5.
In this formulation, the fatigue limit is not treated as a deterministic material constant but as a specimen-specific random variable. The logarithm of the individual fatigue limit $V = \ln(\Gamma)$ is assumed to follow a normal distribution $V \sim \mathcal{N}(\mu_g, \sigma_g^2)$.
Fatigue life is expressed on the logarithmic scale $W = \ln(N)$, and is modeled through a Stromayer-type relationship(Eq. 4.3).
If the applied stress does not exceed the individual fatigue limit ($\Gamma_i \geq S_i$), the specimen is considered immune to failure and its fatigue life is theoretically infinite.
The procedure implemented in the code consists of the following steps:

- loading the required R libraries;
- importing and preprocessing the staircase fatigue data;
- defining auxiliary numerical functions used for likelihood evaluation;
- defining the negative log-likelihood of the RFL model;
- estimating model parameters via numerical maximization;

- computing the marginal probability of failure at the target life N_0;
- determining the stress levels corresponding to selected failure probabilities.

The required R libraries are first loaded. These packages are used for data import and manipulation.

```
library(readxl)
library(dplyr)
```

The staircase fatigue data provided by Prof. Rosso are then imported from the Excel file. Stress and cycle variables are converted to logarithmic scale and an indicator variable is introduced to distinguish failures from run-outs.

```
path <- "~/Desktop/staircase/staircase.xlsx"
sheet_name <- "staircase 1"
skip_rows <- 5

N0 <- 5.00e+06
logN0 <- log(N0)

# quadrature points
K_quad <- 81

data_imported <- read_excel(path, sheet = sheet_name, skip = skip_rows)
data <- data_imported %>%
  rename(
    S = "Fa [kN]",
    outcome = "End of test criterion",
    N = "Test cycles"
  )

S <- data$S
x <- log(S)
w <- log(data$N)

# 1 = censored (run-out), 0 = failure
```

```
is_cens <- ifelse(data$outcome == "survived", 1, 0)

# censor limits
c_lim <- w
c_lim[is_cens == 1] <- logN0 # runouts have liminf logN0
```

The likelihood of the RFL model involves integration with respect to the latent fatigue limit V. The presence of this latent variable leads to an integral in the likelihood that does not admit a closed-form solution. This integral is therefore approximated numerically using K quadrature nodes. Two auxiliary functions are therefore introduced.

The function `logmeanexp` computes the quantity

$$\log\left(\frac{1}{K}\sum_{k=1}^{K} e^{x_k}\right)$$

in a numerically stable way.

The function `truncnorm_quantiles_left` generates K nodes from the distribution

$$V \sim \mathcal{N}(\mu_g, \sigma_g^2)$$

truncated to the interval $(-\infty, x_i)$. These nodes are used to approximate the integral appearing in the likelihood function.

```
logmeanexp <- function(logv) {
  m <- max(logv)
  m + log(mean(exp(logv - m)))
}

# Quantiles of V ~ N(mu, sd^2) truncated to (-inf, upper)
# Returns points for conditional distribution V | V < upper
truncnorm_quantiles_left <- function(mu, sd, upper, K = 81, eps = 1e-12) {
  a <- (upper - mu) / sd
  F_a <- pnorm(a)
  if (!is.finite(F_a) || F_a < 1e-14) return(rep(NA_real_, K))
  u <- (seq_len(K) - 0.5) / K
  q <- qnorm(pmax(pmin(u * F_a, 1 - eps), eps))
```

```
  V <- mu + sd * q
  pmin(V, upper - 1e-10)
}
```

The function `rfl_hard_negloglik` implements the negative log-likelihood of the RFL model.

For each observation i, let $x_i = \ln(S_i)$. Under the model, the probability that the specimen is *active* (i.e. $\Gamma_i < S_i$) is

$$F_a = \Phi\left(\frac{x_i - \mu_g}{\sigma_g}\right) = P(V_i < x_i),$$

while the complementary probability

$$1 - F_a = P(V_i \geq x_i)$$

represents the *immune* mass for which failure cannot occur.

Conditional on being active, the likelihood involves integration with respect to the truncated distribution $V_i \mid V_i < x_i$. This integral is approximated numerically as the average over K quadrature nodes $\{v_{k,i}\}_{k=1}^{K}$.

The likelihood contribution differs depending on the observed outcome.

Failures contribute through the conditional normal density of $W_i = \ln(N_i)$ evaluated at w_i. Run-outs are treated as right-censored observations at censoring level $c_i = \ln(N_0)$ and contribute through the survival probability $P(W_i > c_i)$, including the immune mass $(1 - F_a)$.

```
rfl_hard_negloglik <- function(par, x, S, w, is_cens, c_lim, K = 81) {

  beta0 <- par[1]
  beta1 <- par[2]
  sigma <- exp(par[3])
  mu_g <- par[4]
  sig_g <- exp(par[5])

  if (!is.finite(sigma) || sigma <= 0)
    return(Inf)
  if (!is.finite(sig_g) || sig_g <= 0)
```

```
    return(Inf)
  if (!is.finite(beta1) || beta1 >= 0)
    return(Inf) # Stromayer slope must be negative

  n <- length(S)
  ll <- numeric(n)

  for (i in seq_len(n)) {
    xS <- x[i]
    a <- (xS - mu_g) / sig_g
    F_a <- pnorm(a) # P(Gamma < S)
    tail_a <- 1 - F_a # P(Gamma >= S)

    # If almost surely Gamma >= S: immune
    if (!is.finite(F_a) || F_a < 1e-14) {
      if (is_cens[i] == 1) {
        ll[i] <- log(pmax(tail_a, 1e-300))
      } else {
        return(Inf) # failure impossible
      }
      next
    }

    # Points for V | V < logS
    Vi <- truncnorm_quantiles_left(mu_g, sig_g, upper = xS, K = K)

    if (any(!is.finite(Vi)))
      return(Inf)

    gamma <- exp(Vi)
    d <- S[i] - gamma

    if (any(d <= 0))
      return(Inf)

```

```
    mu_w <- beta0 + beta1 * log(d)

    if (is_cens[i] == 0) {
      # failure: log( F_a * E_cond[f_W] )
      logf <- dnorm(w[i], mean = mu_w, sd = sigma, log = TRUE)
      ll[i] <- log(pmax(F_a, 1e-300)) + logmeanexp(logf)
    } else {
      # run-out: log( tail_a + F_a * E_cond[ P(W>c) ] )
      z <- (c_lim[i] - mu_w) / sigma
      surv <- pmax(1 - pnorm(z), 1e-300) # prob surviving beyond N0
      pr_runout <- tail_a + F_a * mean(surv)
      ll[i] <- log(pmax(pr_runout, 1e-300))
    }
  }

  nll <- -sum(ll)
  nll
}
```

An initial vector of parameters is specified to facilitate numerical optimization. The parameters $(\beta_0, \beta_1, \sigma)$ are initialized using a linear regression fitted on failure observations only. The fatigue limit parameters (μ_g, σ_g) are assigned plausible values consistent with the observed stress range.

```
idx_fail <- which(is_cens == 0)

if (length(idx_fail) >= 2) {
  fit0 <- lm(w[idx_fail] ~ log(S[idx_fail]))
  beta0_init <- as.numeric(coef(fit0)[1])
  beta1_init <- as.numeric(coef(fit0)[2])
  sigma_init <- max(0.3, sd(residuals(fit0)))
} else {
  beta0_init <- mean(w)
  beta1_init <- -1
  sigma_init <- 1
}
```

```

mu_g_init <- min(x) * 0.9
sig_g_init <- 0.2
par_init <- c(beta0_init, min(beta1_init, -1e-3), log(sigma_init),
              mu_g_init, log(sig_g_init))
print(par_init)
```

Model parameters are estimated by numerically minimizing the negative log-likelihood using the function `optim`. In the implementation, σ and σ_g are parameterized on the logarithmic scale in order to enforce positivity.

```
ctrl <- list(maxit = 8000)

set.seed(1)
opt1 <- optim(
  par = par_init,
  fn = rfl_hard_negloglik,
  x = x, S = S, w = w, is_cens = is_cens, c_lim = c_lim,
  K = K_quad,
  control = ctrl
)

par_hat <- opt1$par

beta0_hat <- par_hat[1]
beta1_hat <- par_hat[2]
sigma_hat <- exp(par_hat[3])
mu_g_hat <- par_hat[4]
sig_g_hat <- exp(par_hat[5])
```

Using the estimated parameters $\hat{\theta}$, the function `pfail_mle(Sval)` computes the marginal probability of failure at the target life N_0 for a given stress level S. This quantity corresponds to the marginal fatigue-life distribution $F_N(n \mid S)$ derived in Section 5.1 and given in Equation (5.6), evaluated at $n = N_0$.

In the theoretical formulation of the RFL model, the logarithm of the individual fatigue limit is defined as $V = \ln(\Gamma)$ and is assumed to follow a normal distribution

$$V \sim \mathcal{N}(\mu_\Gamma, \sigma_\Gamma^2).$$

In the R implementation, these parameters are denoted by `mu_g` and `sig_g`, respectively.

Let $x = \ln(S)$ and

$$F_a = \Phi\left(\frac{x - \hat{\mu}_\Gamma}{\hat{\sigma}_\Gamma}\right)$$

denote the probability that the specimen is *active*, i.e. that $\Gamma < S$. Conditional on $V < x$, the integral with respect to the latent fatigue limit is approximated using K quadrature nodes drawn from the truncated distribution

$$V \mid V < x.$$

For each node v_k, the conditional mean log-life is

$$\mu_k = \hat{\beta}_0 + \hat{\beta}_1 \ln(S - \exp(v_k)),$$

and the conditional probability of failure before N_0 is

$$P(W \leq \ln N_0 \mid v_k, S) = \Phi\left(\frac{\ln N_0 - \mu_k}{\hat{\sigma}}\right).$$

The marginal probability of failure is then obtained by averaging these probabilities over the quadrature nodes and multiplying by the active probability F_a.

```
pfail_mle <- function(Sval, target_logN = logN0, K = 81) {
  xS <- log(Sval)

  a <- (xS - mu_g_hat) / sig_g_hat
  F_a <- pnorm(a) # P(Gamma < S)
  if (!is.finite(F_a) || F_a < 1e-14) return(0)

  V <- truncnorm_quantiles_left(mu_g_hat, sig_g_hat, upper = xS, K = K)
  if (any(!is.finite(V))) return(NA_real_)

  gamma <- exp(V)
  d <- Sval - gamma
```

```
  if (any(d <= 0))
    return(NA_real_)

  mu_w <- beta0_hat + beta1_hat * log(d)

  F_a * mean(pnorm((target_logN - mu_w) / sigma_hat))
}
```

The function `find_Sp(p,S_lo,S_hi)` determines the stress level S_p such that

$$p_{\text{fail}}(S_p) = p.$$

The equation is solved numerically using the `uniroot` function. The search interval is initialized using the observed stress range and is automatically expanded if necessary until a sign change is detected.

Using this procedure, the stress levels S_{10}, S_{50} and S_{90} are obtained by solving $p_{\text{fail}}(S) = p$ for $p \in \{0.1, 0.5, 0.9\}$.

```
find_Sp <- function(p, S_lo, S_hi, K = 81) {
  f <- function(Sval) pfail_mle(Sval, K = K) - p

  flo <- f(S_lo); fhi <- f(S_hi)
  if (!is.finite(flo) || !is.finite(fhi))
    return(NA_real_)

  # widen bracket if needed
  if (flo > 0) {
    S_lo2 <- max(1e-9, S_lo * 0.7)
    flo2 <- f(S_lo2)
    if (is.finite(flo2) && flo2 <= 0)
      { S_lo <- S_lo2; flo <- flo2 }
  }
  if (fhi < 0) {
    S_hi2 <- S_hi * 1.3
    fhi2 <- f(S_hi2)
    if (is.finite(fhi2) && fhi2 >= 0)
      { S_hi <- S_hi2; fhi <- fhi2 }
```

```
  }

  flo <- f(S_lo); fhi <- f(S_hi)
  if (!is.finite(flo) || !is.finite(fhi) || flo * fhi > 0)
    return(NA_real_)

  uniroot(function(Sval) f(Sval), lower = S_lo, upper = S_hi)$root
}

S_lo0 <- min(S) * 0.8
S_hi0 <- max(S) * 1.3

S10 <- find_Sp(0.10, S_lo0, S_hi0, K = K_quad)
S50 <- find_Sp(0.50, S_lo0, S_hi0, K = K_quad)
S90 <- find_Sp(0.90, S_lo0, S_hi0, K = K_quad)
```

As a final verification step, the function `pfail_mle` is evaluated at the observed stress levels in the dataset. This allows us to check that the resulting failure probabilities lie in the interval $[0, 1]$ and increase monotonically with the applied stress.

```
Up <- sort(unique(S))
p_at_S <- sapply(Up, pfail_mle, K = K_quad)
print(data.frame(S = Up, p_fail = p_at_S))
```

C.0.2 Bayesian version

This appendix reports the R code used to estimate the Random Fatigue Limit (RFL) model under the Bayesian framework described in Chapter 5. The implementation follows the hierarchical formulation introduced in Section 5.2.1 and performs posterior inference using Markov Chain Monte Carlo (MCMC) sampling implemented through the `JAGS` software.

The procedure implemented in the code consists of the following main steps:

- loading the required R libraries;

- importing and preprocessing the staircase fatigue dataset;
- specifying the hierarchical RFL model in JAGS;
- defining the data structures and initial values for the MCMC chains;
- sampling from the joint posterior distribution of the model parameters;
- computing the posterior distribution of the median fatigue strength S_{50}.

The required R libraries are first loaded. These packages are used for data import, data manipulation, and Bayesian inference through MCMC sampling.

```
library(readxl)
library(dplyr)
library(rjags)
library(coda)
library(upndown)
```

The staircase fatigue dataset is then imported from the Excel file provided by Prof. Rosso. Stress and cycle variables are converted to numeric format and the logarithmic transformations required by the model are computed.

An indicator variable is introduced to distinguish failures from run-outs. Observed log-lives are stored for failure observations, while run-outs are represented as right-censored observations at the target life N_0.

```
path <- "~/Desktop/staircase/staircase.xlsx"
sheet_name <- "staircase 1"
skip_rows <- 5

N0 <- 5e6
logN0 <- log(N0)

dati <- read_excel(path, sheet = sheet_name, skip = skip_rows) %>%
  rename(S = "Fa [kN]", esito = "criterio fine prova", N = "Cicli prova")

S_obs <- dati$S
x <- log(S_obs)
```

```
x_mean <- mean(x)
min_x <- min(x)

# 1 = run-out, 0 = failure
is_cens <- ifelse(dati$esito == "sopravvissuto", 1, 0)

# observed log-life for failures, NA for run-outs
w_obs <- log(dati$N)
w <- w_obs
w[is_cens == 1] <- NA

cens_limit <- rep(logN0, length(x))
V_max_vec <- ifelse(is_cens == 0, x - 0.0001, 10)
```

The Random Fatigue Limit model is specified in JAGS according to the hierarchical structure described in Section 5.2.1. For each specimen i, the latent fatigue limit is represented by the variable $V_i = \log(\gamma_i)$ and is assumed to follow a normal distribution at the population level b.

A physical constraint is imposed to ensure that failures can occur only when the applied stress exceeds the individual fatigue limit. This constraint is implemented through truncation of the normal distribution of V_i.

Structural immunity is represented through a deterministic indicator variable

$$active_i = \mathbb{I}(x_i > V_i),$$

which determines whether a specimen is susceptible to fatigue failure or structurally immune at the applied stress level.

For susceptible specimens, fatigue life follows the Stromeyer-type stress–life relationship introduced in Chapter 5.

In order to improve numerical stability and reduce posterior dependence between regression parameters, the intercept is parameterized in centered form through $\beta_{0,\text{centered}}$, defined with respect to the mean stress level x_{mean}. The original intercept is then recovered deterministically as

$$\beta_0 = \beta_{0,\text{centered}} - \beta_1 x_{\text{mean}}.$$

Immune specimens are assigned a mean log-life well beyond the censoring threshold, in order to provide a computational representation of the theoretically infinite fatigue life implied by structural immunity.

Right censoring of run-out observations is implemented using the `dinterval` construction available in JAGS, which ensures that latent log-lives for censored observations exceed the censoring threshold $\log N_0$.

```
model_string <- "
model {
  for (i in 1:N) {

    V[i] ~ dnorm(mu_gamma, tau_gamma) T(, V_max[i])
    active[i] <- step(x[i] - V[i])
    diff[i] <- max(exp(x[i]) - exp(V[i]), 0.0001)
    mu_act[i] <- beta0_centered + beta1 * (log(diff[i]) - x_mean)
    mu_imm[i] <- logN0 + (kImm * sigma)
    muW[i] <- (active[i] * mu_act[i]) + ((1 - active[i]) * mu_imm[i])
    w[i] ~ dnorm(muW[i], tau)
    is_cens[i] ~ dinterval(w[i], c[i])
  }

  beta0 <- beta0_centered - beta1 * x_mean
  mu_gamma ~ dnorm(mu_g0, 0.1) T(, min_x)
  sigma_gamma ~ dunif(0.001, 0.5)
  tau_gamma <- 1/(sigma_gamma * sigma_gamma)
  beta0_centered ~ dnorm(logN0, 0.01)
  beta1 ~ dnorm(-2, 0.1) T(-100, 0)
  sigma ~ dunif(0, 2)
  tau <- 1/(sigma * sigma)
}
"
```

The data required by the JAGS model are organized in a list structure containing the observed stresses, log-lives, censoring indicators, and auxiliary quantities used in the hierarchical formulation.

```
jags_data <- list(
  N = length(x),
  x = x,
  x_mean = x_mean,
  w = w,
  is_cens = is_cens,
  c = cens_limit,
  V_max = V_max_vec,
  min_x = min_x,
  logN0 = logN0,
  mu_g0 = 0.9 * min(x),
  kImm = 10
)
```

Initial values for the model parameters are obtained from a preliminary linear regression fitted on the failure observations only. This strategy provides data-driven starting values for the regression parameters and improves the numerical stability of the MCMC procedure.

Three parallel chains are initialized using slightly perturbed values of the fatigue-limit parameters in order to facilitate convergence diagnostics. For reproducibility, each chain is initialized with a different random-number generator seed so that the MCMC sampling procedure can be exactly replicated.

```
idx_fail <- which(is_cens == 0)

if (length(idx_fail) >= 2) {
  fit0 <- lm(w_obs[idx_fail] ~ log(S_obs[idx_fail]))
  beta0_init <- as.numeric(coef(fit0)[1])
  beta1_init <- as.numeric(coef(fit0)[2])
  sigma_init <- max(0.3, sd(residuals(fit0)))
} else {
  beta0_init <- logN0
  beta1_init <- -2
  sigma_init <- 0.5
}
```

```
mu_g_init <- min(x) * 0.9
sig_g_init <- 0.2

inits_list <- list(
  list(
    beta0_centered = beta0_init + beta1_init * x_mean,
    beta1 = beta1_init,
    sigma = sigma_init,
    mu_gamma = mu_g_init - 0.05,
    sigma_gamma = sig_g_init,
    .RNG.name = "base::Wichmann-Hill",
    .RNG.seed = 123
  ),
  list(
    beta0_centered = beta0_init + beta1_init * x_mean,
    beta1 = beta1_init,
    sigma = sigma_init,
    mu_gamma = mu_g_init,
    sigma_gamma = sig_g_init,
    .RNG.name = "base::Wichmann-Hill",
    .RNG.seed = 456
  ),
  list(
    beta0_centered = beta0_init + beta1_init * x_mean,
    beta1 = beta1_init,
    sigma = sigma_init,
    mu_gamma = mu_g_init + 0.05,
    sigma_gamma = sig_g_init,
    .RNG.name = "base::Wichmann-Hill",
    .RNG.seed = 789
  )
)
```

Posterior sampling is performed using three independent MCMC chains. After an adaptation phase and a burn-in period, posterior samples are generated for the model parameters

$$(\beta_0, \beta_1, \sigma, \mu_\Gamma, \sigma_\Gamma).$$

Convergence of the chains is assessed through trace plots and Gelman–Rubin diagnostics.

```
jm <- jags.model(
  textConnection(model_string),
  data = jags_data,
  inits = inits_list,
  n.chains = 3,
  n.adapt = 2000
)

update(jm, 30000)
params <- c("beta0_centered","beta0","beta1","sigma","mu_gamma","sigma_gamma")
samples <- coda.samples(jm, variable.names = params, n.iter = 150000, thin = 50)

print(summary(samples))
print(gelman.diag(samples, multivariate = FALSE))

post_mat <- as.matrix(samples)
post_mat <- cbind(
post_mat,
beta0 = post_mat[, "beta0_centered"] - post_mat[, "beta1"] * x_mean
)
```

Posterior inference for the median fatigue strength S_{50} follows the definition introduced in Section 5.1.2. For each posterior draw of the model parameters, the marginal probability of failure $F_N(N_0 \mid S)$ is evaluated using the same quadrature approximation employed in the frequentist formulation.

The stress level S_{50} is defined as the solution of

$$F_N(N_0 \mid S) = 0.5,$$

and is obtained numerically using a one-dimensional root-finding algorithm. Repeating this procedure across posterior draws produces a posterior sample of S_{50} values. From this sample, the posterior median is computed.

```
K_quad <- 81

pfail_paperlike <- function(Sval, par, logN0, K = 81, eps = 1e-9) {

  if (!is.finite(Sval) || Sval <= 0)
    return(NA_real_)

  b0 <- par["beta0"]
  b1 <- par["beta1"]
  sig <- par["sigma"]
  mg <- par["mu_gamma"]
  sg <- par["sigma_gamma"]
  xS <- log(Sval)
  a <- (xS - mg) / sg
  Phi_a <- pnorm(a)

  if (!is.finite(Phi_a) || Phi_a < 1e-14)
    return(0)

  u <- (seq_len(K) - 0.5) / K
  p <- pmin(pmax(u * Phi_a, eps), 1 - eps)
  Vk <- mg + sg * qnorm(p)
  gammak <- exp(Vk)
  diff_val <- pmax(Sval - gammak, eps)
  mu_k <- b0 + b1 * log(diff_val)
  Phi_a * mean(pnorm((logN0 - mu_k) / sig))
}

S50_one_draw <- function(par, logN0, S_lo0, S_hi0, K = 81) {
```

```
  f <- function(S) pfail_paperlike(S, par, logN0, K) - 0.5
  grid <- seq(S_lo0, S_hi0, length.out = 200)
  fg <- suppressWarnings(sapply(grid, f))
  idx <- which(is.finite(fg) & fg >= 0)[1]

  if (is.na(idx) || idx == 1)
    return(NA_real_)

  tryCatch(
    uniroot(f, lower = grid[idx - 1], upper = grid[idx])$root,
    error = function(e) NA_real_
  )
}

set.seed(42)

ndraws <- min(800, nrow(post_mat))
idx_draw <- sample(seq_len(nrow(post_mat)), ndraws)

S_lo0 <- min(S_obs) * 0.5
S_hi0 <- max(S_obs) * 1.5

S50_draws <- sapply(idx_draw, function(i) {
  S50_one_draw(post_mat[i, ], logN0, S_lo0, S_hi0, K = K_quad)
})

S50_clean <- S50_draws[is.finite(S50_draws)]

cat("S50 median:", median(S50_clean), "\n")
```

Bibliography

[1] August Wöhler. "Ueber die Festigkeits-Versuche mit Eisen und Stahl". In: *Zeitschrift für Bauwesen* 20 (1870), pp. 73–106 (cit. on p. 1).

[2] *ISO 12107. Metallic Materials — Fatigue Testing — Statistical Planning and Analysis of Data.* International Organization for Standardization, Aug. 2012 (cit. on pp. 2, 7, 10, 32, 41).

[3] W. J. Dixon and A. M. Mood. "A Method for Obtaining and Analyzing Sensitivity Data". In: *Journal of the American Statistical Association* 43.241 (1948), pp. 109–126 (cit. on pp. 2, 3, 17, 20, 23, 26, 27, 29, 41).

[4] Kim R. W. Wallin. "Statistical Uncertainty in the Fatigue Threshold Staircase Test Method". In: *International Journal of Fatigue* 33.3 (2011), pp. 354–362 (cit. on pp. 4, 29, 38).

[5] J. J. Braam and S. van der Zwaag. "A Statistical Evaluation of the Staircase and the ArcSinVP Methods for Determining the Fatigue Limit". In: *Journal of Testing and Evaluation* 26.2 (Mar. 1998), pp. 125–131 (cit. on pp. 4, 29, 39).

[6] Christian Müller, Michael Wächter, Rainer Masendorf, and Alfons Esderts. "Accuracy of Fatigue Limits Estimated by the Staircase Method Using Different Evaluation Techniques". In: *International Journal of Fatigue* 100 (2017), pp. 296–307 (cit. on pp. 4, 29, 30, 32).

[7] C. R. A. Schneider and S. J. Maddox. *Best Practice Guide on Statistical Analysis of Fatigue Data.* Tech. rep. The Welding Institute, 2003 (cit. on p. 8).

[8] William Q. Meeker, Peng Liu, Luis A. Escobar, Wayne M. Falk, Francis G. Pascual, Yili Hong, and Balajee Ananthasayanam. "Modern Statistical Models

and Methods for Estimating Fatigue-Life and Fatigue-Strength Distributions from Experimental Data". In: *arXiv preprint* (2024) (cit. on pp. 10, 14).

[9] Ibrahim Burhan and Ho Sung Kim. "S-N Curve Models for Composite Materials Characterisation: An Evaluative Review". In: *Journal of Composites Science* 2.38 (2018) (cit. on p. 14).

[10] Francis G. Pascual and William Q. Meeker. "Estimating Fatigue Curves with the Random Fatigue-Limit Model". In: *Technometrics* 41.4 (1999), pp. 277–289 (cit. on p. 15).

[11] William Q. Meeker, Luis A. Escobar, Francis G. Pascual, Yili Hong, Peng Liu, Wayne M. Falk, and Balajee Ananthasayanam. "Modern Statistical Models and Methods for Estimating Fatigue-Life and Fatigue-Strength Distributions from Experimental Data". In: *Statistical Science* 41.1 (2026), pp. 1–27 (cit. on p. 15).

[12] R. Pollak, A. Palazotto, and T. Nicholas. "A Simulation-Based Investigation of the Staircase Method for Fatigue Strength Testing". In: *Mechanics of Materials* 38 (2006), pp. 1170–1181 (cit. on pp. 19, 39).

[13] T. Svensson, S. Loren, J. de Maré, and B. Wadman. "Statistical Models of the Fatigue Limit". In: *Fatigue & Fracture of Engineering Materials & Structures* 22.10 (1999), pp. 909–916 (cit. on p. 39).

[14] Peng Liu and William Q. Meeker. "A Robust Numerical Method for Nonlinear Regression". In: *arXiv preprint* (2024) (cit. on p. 59).

[15] Francis G. Pascual and William Q. Meeker. "Analysis of Fatigue Data with Runouts Based on a Model with Nonconstant Standard Deviation and a Fatigue Limit Parameter". In: *Technometrics* 38.4 (1996), pp. 358–367 (cit. on p. 67).

[16] C. Blanco and A. Martín Meizoso. *Sensitivity Analysis of the Staircase Method to Determine the Fatigue Limit*. Tech. rep. Donostia–San Sebastián, Spain: School of Engineering (TECNUN), University of Navarra and CEIT Materials Department.

[17] V. Roué, C. Doudard, S. Calloch, Q. Pujol d'Andrebo, F. Corpace, and C. Guévenoux. "Simulation-Based Investigation of the Reuse of Unbroken Specimens in a Staircase Procedure: Accuracy of the Determination of Fatigue Properties". In: *International Journal of Fatigue* 131 (2020), p. 105288.

[18] *Methods of Fatigue Testing — Guide to the Application of Statistics*. British Standards Institution, 1966.

[19] *Standard Method of Statistical Fatigue Testing*. Japan Society of Mechanical Engineers, 1981.

[20] *Metallic Products — Fatigue Tests — Statistical Treatment of Data*. Association Française de Normalisation, 1991.

[21] *Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ϵ-N) Fatigue Data*. ASTM International, 1991.

[22] W. Weibull. *Fatigue Testing and the Analysis of Results*. Pergamon Press, 1961.

[23] Samuel S. Shapiro. *How to Test Normality and Other Distributional Assumptions*. Milwaukee: American Society for Quality Control, 1986.

[24] John Neter, William Wasserman, and Michael H. Kutner. *Applied Linear Statistical Models*. Homewood, IL: Irwin, 1985.

[25] Norman R. Draper and Harry Smith. *Applied Regression Analysis*. 2nd. New York: Wiley, 1981.

[26] S. Nishijima. "Statistical Fatigue Properties of Some Heat-Treated Steels for Machine Structural Use". In: *ASTM Special Technical Publication* 744 (1981), pp. 75–88.

[27] K. A. Brownlee, J. L. Hodges Jr., and Murray Rosenblatt. "The Up-and-Down Method with Small Samples". In: *Journal of the American Statistical Association* 48.262 (1953), pp. 262–277.

[28] F. Bastenaire, G. Pomey, and P. Rabbe. "Étude statistique des durées de vie en fatigue et des courbes de Wöhler de cinq nuances d'acier". In: *Mémoires Scientifiques de la Revue de Métallurgie* 68 (1971), pp. 645–664.

[29] F. Bastenaire. “New Method for the Statistical Evaluation of Constant Stress Amplitude Fatigue Test Results”. In: *Probabilistic Aspects of Fatigue (ASTM STP 511)*. ASTM, 1972.

[30] C. E. Stromeyer. “The Determination of Fatigue Limits Under Alternating Stress Conditions”. In: *Proceedings of the Royal Society A* 90.620 (1914), pp. 411–425.

[31] Xing-Wang Sheng, Wei-Qi Zheng, and Ying Yang. “Tensile and High-Cycle Fatigue Performance of HRB500 High-Strength Steel Rebars Joined by Flash Butt Welding”. In: *Construction and Building Materials* (2020).

[32] Randall D. Pollak and Anthony N. Palazotto. “A Comparison of Maximum Likelihood Models for Fatigue Strength Characterization in Materials Exhibiting a Fatigue Limit”. In: *Probabilistic Engineering Mechanics* (2008).