

Probabilistic Physics of Failure

Approach to Reliability

Scrivener Publishing
100 Cummings Center, Suite 541J
Beverly, MA 01915-6106

Performability Engineering Series
Series Editors: Krishna B. Misra (kbmisra@gmail.com)
and John Andrews (John.Andrews@nottingham.ac.uk)

Scope: A true performance of a product, or system, or service must be judged over the entire life cycle activities connected with design, manufacture, use and disposal in relation to the economics of maximization of dependability, and minimizing its impact on the environment. The concept of performability allows us to take a holistic assessment of performance and provides an aggregate attribute that reflects an entire engineering effort of a product, system, or service designer in achieving dependability and sustainability. Performance should not just be indicative of achieving quality, reliability, maintainability and safety for a product, system, or service, but achieving sustainability as well. The conventional perspective of dependability ignores the environmental impact considerations that accompany the development of products, systems, and services. However, any industrial activity in creating a product, system, or service is always associated with certain environmental impacts that follow at each phase of development. These considerations have become all the more necessary in the 21st century as the world resources continue to become scarce and the cost of materials and energy keep rising. It is not difficult to visualize that by employing the strategy of dematerialization, minimum energy and minimum waste, while maximizing the yield and developing economically viable and safe processes (clean production and clean technologies), we will create minimal adverse effect on the environment during production and disposal at the end of the life. This is basically the goal of performability engineering.

It may be observed that the above-mentioned performance attributes are interrelated and should not be considered in isolation for optimization of performance. Each book in the series should endeavor to include most, if not all, of the attributes of this web of interrelationship and have the objective to help create optimal and sustainable products, systems, and services.

Publishers at Scrivener
Martin Scrivener (martin@scrivenerpublishing.com)
Phillip Carmical (pcarmical@scrivenerpublishing.com)

Probabilistic Physics of Failure Approach to Reliability

**Modeling, Accelerated Testing, Prognosis
and Reliability Assessment**

Mohammad Modarres

Center for Risk and Reliability, University of Maryland, College Park, U.S.A.

Mehdi Amiri

Department of Mechanical Engineering, George Mason University,
Fairfax, U.S.A.

Christopher Jackson

Henry Samueli School of Engineering and Applied Science, University of
California, Los Angeles, U.S.A.



WILEY

This edition first published 2017 by John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA and Scrivener Publishing LLC, 100 Cummings Center, Suite 541J, Beverly, MA 01915, USA
© 2015 and 2017 Center for Risk and Reliability (CRR), University of Maryland, U.S.A.
For more information about Scrivener publications please visit www.scriverpublishing.com.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

Edition History

First published 2015; reissued in 2017

Wiley Global Headquarters

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials, or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read.

Library of Congress Cataloging-in-Publication Data

ISBN 978-1-119-38863-0

Cover image: Russell Richardson
Cover design by Russell Richardson

Set in size of 11 pt Garamond by Exeter Premedia Services Private Ltd., Chennai, India

Printed in

10 9 8 7 6 5 4 3 2 1

Contents

Preface	ix
1 Overview of Probabilistic Physics-of-Failure Approach to Reliability	1
1.1 Introduction	1
1.2 Overview of Physics-of-Failure Modeling	2
1.3 Important Forms of PoF Models	4
1.4 PPoF Approach to Life Assessment	6
1.5 Accelerated Testing in PPoF Model Development	8
1.6 Organization of the Book	10
References	11
2 Summary of Mechanisms of Failure and Associated PoF Models	13
2.1 Introduction	13
2.2 Fatigue	15
2.2.1 Life-Stress	17
2.2.1.1 The S-N Diagram	18
2.2.1.2 Mean Stress Effects	19
2.2.1.3 Combined Loading	21
2.2.2 Strain-Life	23
2.2.2.1 Monotonic Stress-Strain Behavior	23
2.2.2.2 Cyclic Stress-Strain Behavior	26
2.2.2.3 Strain-Life Relationship	27
2.2.2.4 Mean Stress Effects	29
2.2.3 Variable Amplitude Loading	33
2.2.3.1 Non-Linear Damage Models	37
2.2.4 Notch Effect	37
2.2.4.1 Life-Stress	37
2.2.4.2 Strain-Life	41
2.2.5 Two-Stage Approach to Fatigue Life Estimation	47
2.2.6 Fracture Mechanics	47
2.2.6.1 Stress Intensity Factor	49
2.2.6.2 Region I	49
2.2.6.3 Region II	49
2.2.6.4 Region III	51
2.2.6.5 Fracture Mechanics Approach with Notch Effect	53
2.2.7 Factors Influencing Fatigue Failure	56
2.2.7.1 Size Effect	56
2.2.7.2 Frequency Effect	57
2.2.7.3 Environmental and External Effects	58
2.2.7.4 Miscellaneous Factors	59

2.3	Wear	60
2.3.1	General Form of Wear Equations	63
2.3.2	Sliding Wear	64
2.3.3	Abrasive Wear	66
2.3.4	Impact Wear	67
2.3.5	Rolling Wear	68
2.3.6	Life Models for Bearings	69
2.3.7	Life Models for Seals	73
2.3.8	Wear of Lubricated Contacts	78
2.3.9	Lubricated Wear and Lubricant Life	79
2.4	Creep	81
2.4.1	Larson-Miller Theory	82
2.4.2	Manson-Haferd Theory	83
2.4.3	Creep Under Uniaxial State of Stress	85
2.4.4	Cumulative Creep Prediction	89
2.5	Corrosion	90
2.5.1	Models for Prediction of Corrosion Rate and Service Life	92
	References	97
3	Types of Accelerated Testing and Modeling Concepts	101
3.1	Introduction	101
3.2	Types of Accelerated Testing – Qualitative and Quantitative	101
3.3	Qualitative Accelerated Tests	102
3.3.1	Environmental Stress Testing	103
3.3.2	Burn-In Testing	103
3.3.3	Environmental Stress Screening	104
3.3.4	Highly Accelerated Life Testing	104
3.3.4.1	Summary of HALT Process	105
3.3.5	Highly Accelerated Stress Screening	106
3.4	Quantitative Accelerated Tests	107
3.4.1	Modeling Degradation Associated with Various Failure Mechanisms	108
3.4.1.1	Stress-Strength Model	108
3.4.1.2	Damage-Endurance Model	109
3.4.1.3	Performance-Requirement Model	110
3.4.2	Forms of Degradation and Performance Models	113
	References	115
4	Analysis of Accelerated Life Testing Data and Physics-Based Reliability Model Development	117
4.1	Introduction	117
4.2	Accelerated Life Data Analysis Methods	117
4.3	Basics of ALT Data Analysis	117
4.4	Types of Collected Accelerated Life Test Data	118
4.5	Life-stress Models	119
4.6	Probability Plotting Method for ALT Model Estimation	124
4.6.1	Life-Stress Model by Regression	124
4.6.2	Summary of Plotting Method for Analyzing ALT Data	130
4.7	Maximum Likelihood Estimation Approach to ALT Data Analysis	131
4.8	Confidence Intervals for MLE	134

4.9	MLE Approach to Estimating Parameters of Common Distributions	136
4.9.1	Exponential Life Distribution	136
4.9.2	Weibull Life Distribution	137
4.9.3	Lognormal Life Distribution	137
4.10	MLE-Based Parameter Estimation for Different Life-Stress Models	139
4.10.1	The Exponential Life-Stress Model	139
4.10.2	Exponential Life-Stress Model with Weibull Life Distribution	140
4.10.3	Exponential Life-Stress Model with Lognormal Life Distribution	146
4.10.4	The Eyring Life-Stress Model	148
4.10.5	The Eyring-Weibull Model	149
4.10.6	The Eyring-Lognormal Model	151
4.10.7	Power Life-Stress Model	153
4.10.8	Power Life-Stress with Weibull Life Model	153
4.10.9	Power Life-Stress with Lognormal Model	155
4.10.10	Dual-Stress Exponential Life-Stress Model	159
4.10.11	Dual-Stress Exponential Life-Stress Model with Weibull Life Distribution	160
4.10.12	Dual-Stress Exponential Life-Stress Model with Lognormal Life Distribution	161
4.10.13	Power-Exponential Life-Stress Model	163
4.10.14	Power-Exponential Life-Stress Model Weibull Life Distribution	164
4.10.15	Power-Exponential Life-Stress Model Lognormal Life Distribution	165
4.11	Proportional Hazards (PH) Model	168
4.11.1	The Parametric PH Model, with an Example	169
4.12	Bayesian Estimation Approach to ALT Model Parameter Estimation	171
4.12.1	Prior Information for Bayesian Estimation	172
4.12.2	A Bayesian Estimation ALT Data Analysis Example	173
4.13	Determining Stress Dependencies	175
4.13.1	Confidence Bounds	176
4.14	Summary of the ALT Steps and Common Problems in Practice	178
4.15	Time Varying Stress Tests	179
4.16	Step-Stress Analysis and Model Development	182
4.16.1	Plotting Method for Step-Stress Data Analysis	182
4.16.2	Maximum Likelihood Estimation Method for Step-Stress Data Analysis	192
4.16.3	Bayesian Inference Method for Step-Stress Data Analysis	196
	References	201
5	Analysis of Accelerated Degradation Data and Reliability Model Development	203
5.1	Introduction	203
5.2	Degradation Models	205
5.2.1	Simple Degradation Model Without Variation	205
5.2.2	Consideration of the Variation in Degradation Model and Failure Time	206
5.2.3	General Degradation Path Model	207
5.2.4	Approximate Accelerated Life Degradation Analysis	208
5.2.5	Maximum Likelihood Approach to Estimating Acceleration Degradation Model Parameters	209
5.2.6	Bayesian Estimation of ADT Model Parameters	219
	References	231

6 Accelerated Test Planning	233
6.1 Introduction	233
6.2 Issues to Consider Prior to Accelerated Testing	233
6.3 Planning for Accelerated Life Tests	237
6.3.1 Steps for Accelerated Life Tests	237
6.3.2 Optimal Design of Accelerated Life Test	239
6.4 Planning for Accelerated Degradation Tests	246
References	250
7 Accounting for Uncertainties and Model Validation	251
7.1 Introduction	251
7.2 Uncertainties in Evidence	251
7.2.1 Classical Error: Uncertainty in the Physical Process	251
7.2.2 Berkson Error: Uncertainty in the Observation Process	256
7.2.2.1 Systematic Uncertainties	258
7.2.2.2 Stochastic Uncertainties	258
7.2.2.3 Relationship between Berkson and Classical Errors	258
7.3 PPoF Model Uncertainties, Errors, and Validation	259
7.4 Applications of Model Validation in ADT	263
References	268
Index	269

Preface

This book is result of the compilation of class notes from several years of teaching a graduate course on physics-of-failure and accelerated testing to the graduate students pursuing Master of Science, Master of Engineering and PhD degrees in Reliability Engineering at the University of Maryland. The book provides probabilistic and highly technical approaches to the physics-of-failure and mechanistic based reliability prediction and assessment. It relies on various methods and techniques published in the open literature regarding the development and practice of physics-of-failure analysis, accelerated life testing and accelerated degradation testing. The authors first discuss the overall concepts, objectives and framework for accelerated life assessment through the use of formal probabilistic physics-of-failure models. They review important failure mechanisms to demonstrate the process of examining and developing appropriate physics and mechanistic models that describe the degradation and failure phenomena in accompanying accelerated testing and accelerated degradation testing methods, including step-stress testing. The book presents data analysis methods to evaluate the probabilistic physics-of-failure models based on the observed data obtained from accelerated reliability tests. Further, it discusses the steps and methods of probabilistic life assessment and integrity of structures, components and systems based on the probabilistic physics-of-failure models. Since the book is intended for graduate-level students and for highly trained reliability engineers, it provides supplementary solved examples to clarify complex technical topics within each chapter. Some of these examples are benefitted directly or with some modifications from other sources, including Bannantine, et al. (1997), Collins (1993), Stephens, et al. (2003), Meeker & Escobar (1998), Nelson (2004), and Dowling (1998), which are referenced extensively in the book. Although qualitative accelerated tests such as the Highly Accelerated Life Test (HALT) and Environmental Stress Screening (ESS) have been briefly reviewed, the book is mainly about the quantitative methods in probabilistic physics-based and accelerated testing life assessment of structures, components and systems. A companion website under the auspices of the Center for Risk and Reliability at the University of Maryland (www.crr.umd.edu) provides downloadable support files for additional information and computational tools in form of MATLAB, R and OpenBUGS scripts to perform some of the more involved computational analyses discussed in the book. These files will be updated and conformed to the most recent versions of these tools. The companion website also includes a section on testing equipment and resources needed for accelerated testing. This book benefitted from contributions of many students who enrolled in the accelerated testing courses over many years at the University of Maryland. Particularly, inputs and solved example from Wendell Fox, Jonathan DeJesus, Reuel Smith, Reza Azarkhail, Andrew Bradshaw, and Taotao Zhou have been significant and are much appreciated.

Mohammad Modarres
Mehdi Amiri
Christopher Jackson

Chapter 1: Overview of Probabilistic Physics-of-Failure Approach to Reliability

1.1. INTRODUCTION

To address risk and reliability challenges in design, manufacturing and operation, reliability engineering has gone through a number of transformations over the past few decades. Reliability methods have progressively become more realistic by incorporating both data and information from real causes, and the modeling of failure phenomena. The evolution of reliability modeling from constant hazard rates to more representative life distributions (such as the Weibull and lognormal) was the first step towards better addressing wear-out and aging failure mechanisms in structures, systems and components. This trend was followed by the use of physics and mechanistic principles, as well as thermodynamic laws. Accelerated testing has borrowed concepts from materials degradation and fracture mechanics, through which the aggregate effects of operational and environmental conditions were formally accounted for in the life models.

The formal consideration of physics and mechanistic methods in reliability engineering is referred to as a “physics of failure” (PoF) approach. The PoF approach is a science-based means to reliability engineering and prediction as well as prognosis and health management, in contrast to the traditional statistical approach that relies on historical data. It uses physics-based modeling and simulations to assess design and reliability. The approach can be used to evaluate and predict system performance while reducing subjectivities in reliability assessments by modeling failure mechanisms such as fatigue, fracture, wear, and corrosion. The PoF approach is a comprehensive representation of wear out and aging, and is capable of bringing relevant physical factors into the life assessment and reliability models of the structures, components and systems.

The development of PoF models is still typically based on limited information. The uncertainties associated with this limitation have led to a Probabilistic Physics of Failure (PPoF) approach that formally addresses and incorporates uncertainties about the PoF models and their outputs.

Physics and mechanistic-based failure models can be categorized into three core frameworks: stress-strength, damage-endurance, and performance-requirements. In all these PoF modeling frameworks, metrics representing failure-inducing agents such as applied loads and environmental attack properties should be identified. Mechanical, thermal, electrical, chemical, and radiation-induced forces can cause stresses on an item. The passage of time drives the accumulation of damage. Both load and time may either be analyzed deterministically (e.g., identifying and studying the sources of stresses) or probabilistically (e.g., treating stress variation as a random variable). Substantial uncertainties associated with failure-inducing agents can emanate from environmental and operating (use) conditions and from the emergence of failure mechanisms that were not considered or well understood at the time of design.

Because of cost and time limitations, great emphasis has been placed on capturing reliability information from field data with minimal effort. As a result, design and assessment methodologies that address failures mechanistically have emerged as popular and powerful cost saving techniques. Accelerated life testing (ALT), an approach to mechanistic modeling of wear-out, damage process and failure, is a direct outcome of this movement. Unlike the reliability models developed on the basis of field data that suffer from wide variation in operating conditions and practices, reliability models based on PoF, developed using accelerated life or degradation tests, take into account operational conditions (applied stresses) that permit flexibility in applied stresses, leading to more relevant models.

2 PROBABILISTIC PHYSICS OF FAILURE APPROACH TO RELIABILITY

Before performing the accelerated test, a stress agent, which could be an aggregate effect of a single or multiple physical and operational conditions, should be identified. The next step involves accelerating this stress agent and applying it to samples of the structure, system or component in a test environment. Models of failure, damage and degradation are developed by using accelerated test data for a more flexible and representative description of the damage, failure phenomena, performance and life as compared to the traditional probabilistic approach to failure modeling.

Failure interdependency can also be a critical factor in reliability modeling of mechanical systems and components. In the study of system behavior, there are situations in which progressive failure of one component may activate or accelerate other failure mechanisms or the failure of other components. There are usually many links between different components by means of their properties and common environmental conditions. The PoF approach properly incorporates these interdependencies in complex structures, systems and components.

1.2. OVERVIEW OF PHYSICS-OF-FAILURE MODELING

Physics of failure modeling initially evolved out of examination of fatigue and fracture of materials. Reliability work related to fatigue and fracture of materials showed significant progress through the 1950s and early 1960s. In 1957, George R. Irwin proved that the fracture of materials was due to plastic deformation at the crack tip and generalized Griffith's Theory (Irwin 1957) that described the relationship between applied nominal stress and crack length at fracture. Between 1955 and 1963, Waloddi Weibull produced several publications related to modeling of fatigue and creep mechanisms that also discussed evaluating associated data (Weibull 1959). In 1961, Weibull published a book on materials and fatigue testing while working as a consultant for the U.S. Air Force Materials Laboratory (Weibull 1961). Building on Irwin's work on stress intensity factor, Paris et al. (Paris, Gomez and Anderson 1961) introduced methods for predicting the rate of fatigue crack growth.

Given this background in mechanistic-based life models (particularly to assess fatigue and fracture failures), Rome Air Development Center (RADC—the predecessor to the U.S. Air Force Rome Laboratory) introduced a PoF program in 1961 to address the growing complexity of military equipment and the consequent increase in number of failures observed. In 1962, researchers from Bell Labs published a paper on "High Stress Aging to Failure of Semiconductor Devices" that justified using the kinetic theory's interpretation of the Arrhenius equation: a simple yet accurate formula for the temperature dependence of the reaction rate constant as a basis for assessment of temperature-induced aging of semiconductor devices (Dodson and Howard 1961). Later, the RADC and Armor Research Foundation of the Illinois Institute of Technology (now IIT Research Institute) organized the first PoF symposium in electronics in Chicago in September 1962. This symposium laid the groundwork for future research and development activities related to PoF by RADC and several other organizations. Numerous original papers and ideas introducing and explaining the PoF concepts and methods were presented in these symposia, which continue today in IEEE International Reliability Physics symposia and Reliability and Maintainability Symposia (RAMS).

The PoF approach to reliability utilized scientific knowledge of damage and degradation processes and the load profile applied to an item, its structure, material properties and environmental conditions to identify potential failure mechanisms that individually or in combination lead to the item's failure. The PoF models would then be used to assess reliability, expended life and remaining life. Using PoF diminishes the need for enormous amounts of use-level life data and uses smaller sets of accelerated test data and other relevant data to present a more representative model. The PoF approach employs the available well-developed knowledge about the mechanisms of failure. The PoF models show how and why items fail, reducing the need for large (and expensive) quantities of life data.

The most critical step in a PoF approach is to understand failure mechanisms (such as corrosion or fatigue) in order to appropriately model degradation and time that a failure occurs. Accelerated life testing based on PoF models is an approach that can reduce long and costly life testing. In this approach, one seeks to relate the fundamental physical and chemical properties of materials to reliability metrics (such as degradation, life or cycles-to-failure). To eliminate (or reduce) the occurrence of failures, one must eliminate (or reduce) their root causes. To do that, one must also understand the physics of the material and failure mechanisms involved (Vaccaro 1962). Sometimes it is impossible to build several identical units or prototypes for reliability testing. Cases in point include large-scale systems (like buildings and space vehicles), one-of-a-kind or highly expensive systems, and units that must work properly at the first time. In these cases, performance and field data are not available, and a PoF approach to degradation and life assessment is the most appropriate. As such, the PoF approach is particularly useful in the design stage when there are limited prototypes or test facilities. Finally, the PoF approach has great utility when dealing with highly reliable units, when there is very little failure data to analyze.

PoF techniques can be used to interpret and extrapolate field data for failure prediction for in-service components. This field data might include parameters that are related to traditional physical measures but can only be used as a loose model for failure prediction. A good example of this is vibration of a bearing. The vibration is suggestive of a flaw, but since the flaw itself cannot be tracked, the vibration can be used to estimate failure. This obviously presents potential for high uncertainty because changes in the model cannot be easily detected; for example, there is no assurance that if a new flaw develops it would expedite the failure, so that it can be detected and incorporated into the model. However, this method at least presents some means of tracking the degradation of the component. This is useful for maintenance practitioners, as it provides a means of failure estimation when traditional methods cannot be used due to the lack of measurable PoF model parameters.

There is no single methodology for performing PoF-based reliability analysis. Chapter 2 discusses in more detail the steps involved in developing a PoF model of an item. If an item involves multiple subassemblies (parts and components), each subject to different failure mechanisms, then the combined effect of applicable failure mechanisms should be modeled. Figure 1.1 depicts the structural and dynamic hierarchy of PoF analysis elements for a multi-component system. The lowest level in this hierarchy is inter- and intra-environmental factors. The intra-environmental factors refer to conditions resultant from unit operation itself. This includes, for example, heat dissipation or vibration caused by an imbalanced rotating shaft. The inter-environmental factors are those imposed externally from its design boundary. Examples include relative humidity and prevalence of dust particles. There may be a causal chain among inter- and intra-environmental factors such that one may lead to another or vice versa in a synergistic manner. For example, a low temperature may cause condensation, leading to accelerated corrosion.

All environmental factors potentially lead to various forms of stress. For example, high temperature (as either an inter- or intra-environment factor) leads to thermal expansion, and (if the unit is confined) can cause mechanical stresses. Such stress agents are key actors in activating or accelerating degradation through corresponding failure mechanisms. While one failure mechanism may also accelerate another (such as corrosion accelerating fatigue), failure mechanisms can also produce new stresses. For example, wear in a journal bearing can cause vibration-induced fatigue. The top part of the hierarchy in Figure 1.1, known as the structural hierarchy, depicts the formal organization and topology of the system showing the functional and support relationships among parts, components and the whole systems. On the other hand, the lower part of the figure, the systems dynamics hierarchy, shows the underlying processes (failure mechanisms) and conditions that lead to the occurrence or acceleration of such mechanisms.

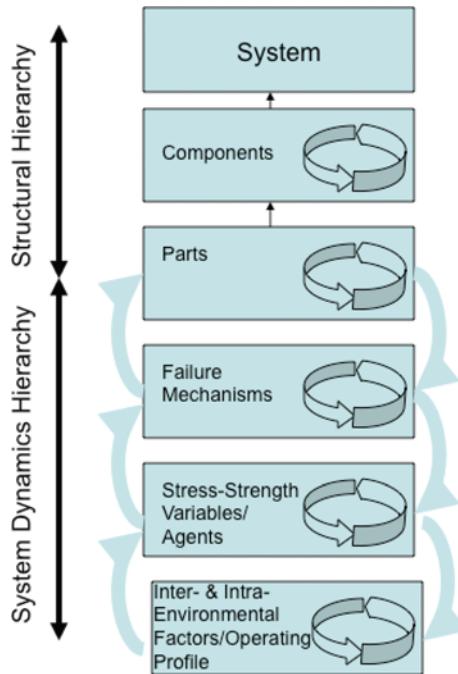


Figure 1.1: System hierarchy used in PoF analysis

1.3. IMPORTANT FORMS OF POF MODELS

As noted above, there are three possible PoF modeling frameworks subject to the nature of underlying failure and degradation mechanism. Each is described briefly below.

Stress-Strength Model. In this model, the item (e.g., a structure, system or component) fails if the applied stresses caused by design, operation and the external environment exceed its strength (see Figure 1.2). This failure model may depend on environmental conditions, applied operating loads and the occurrence of critical events, rather than the passage of time or cycles. Stress and strength are treated as a random variable encompassing variability in all conditions. Two examples of this model include a steel bar under a mean tensile stress lower than its yielding point but which will be randomly subjected to load that exceeds the yielding point over time.

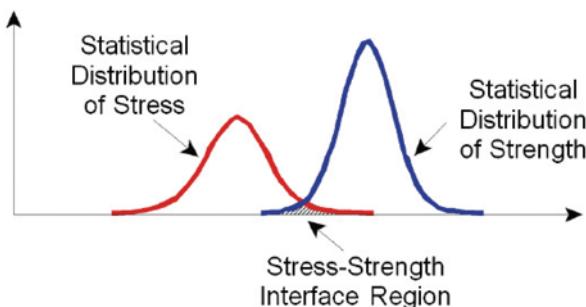


Figure 1.2: Stress-strength modeling

The second is a transistor with a mean voltage applied across the emitter-collector remaining below a failed level but which may randomly exceed the limit. In the case of the steel bar, the likelihood of failure is estimated from the probability that the stress random variable exceeds the strength random variable, which is obtained from a convolution of the two respective distributions.

Damage-Endurance Model. This model differs from the stress-strength model in that the *stress* (load) causes degradation in the form of irreversible cumulative damage through, for example, corrosion, wear, embrittlement, creep, or fatigue. The stress (load) aggregate drives the cumulative damage metric. Cumulative damage may not degrade performance; however, the item fails when the cumulative damage exceeds its endurance limit. For example, a crack grows on a structure until it reaches a critical length beyond which the growth will be catastrophically rapid. Accumulated damage does not disappear when the *stresses* are removed, although sometimes treatments such as annealing can repair cumulative damage. Variables representing damage and endurance may be treated as random and represented by probability density functions to capture distributions of initial damage, model parameter uncertainties, and model errors. Therefore, at any time or cycle (see Figure 1.3) the likelihood of failure may be represented by the exceedance of the damage distribution from the endurance probability density functions. If endurance is not a random variable and remains constant, then the distribution of the time to failure may be obtained when cumulative damage values randomly exceed the constant value of the endurance (see Figure 1.3). The distribution of the time-to-failure shown in Figure 1.3 is based on the assumption of a constant endurance limit around the median of the distribution of the endurance. Clearly, at a given time or cycle, N , the probability that the damage distribution exceeds endurance level (or distribution of endurance), would be equal to the probability that the random variable, time to failure (as represented by the time to failure distribution) is lower than N .

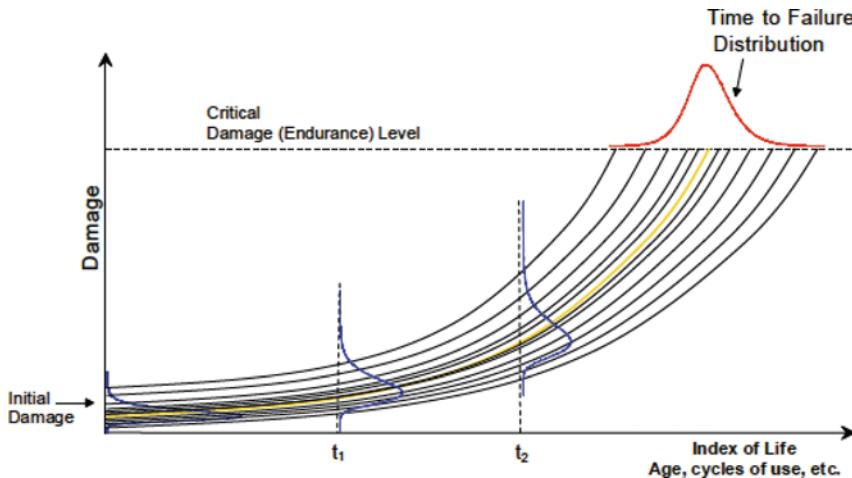


Figure 1.3: Damage-endurance model

Performance-Requirements Model. In this modeling approach, a system performance characteristic (such as system output capability, efficiency or availability) is satisfactory if it remains within acceptable tolerance limits. Examples include rotating machinery efficiency and printer print quality (such as one that is based on a level of efficiency or output at the pump head). Systems start with a positive margin of performance that cumulatively and irreversibly degrades due to the underlying failure mechanisms. These mechanisms cause degradation and damage until performance falls below the minimum requirement level (i.e. fails). As the stress applied to the unit increases the rate of performance degradation, the time to failure (the point at which the system reaches minimum or acceptable performance limit) is reduced. The concept is depicted in Figure 1.4.

6 PROBABILISTIC PHYSICS OF FAILURE APPROACH TO RELIABILITY

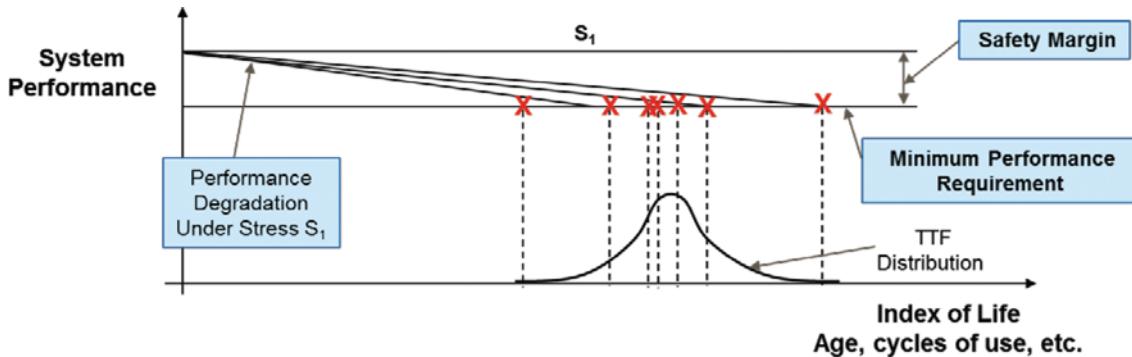


Figure 1.4: Performance-requirement model

1.4. PPOF APPROACH TO LIFE ASSESSMENT

Due to the inevitable stochastic variations of the many factors involved in degradation and failure processes described by PoF models, probabilistic physics-of-failure (PPoF) models can be used to formally account for the uncertainties and model errors. Factor variations include uncertainties in environmental and operational stresses, mission profile and variability in materials properties, and stress agents. The earliest effort in PPoF modeling was by Haggag, et al. (Haggag, McMahon, Hess, Cheng, Lee, & Lyding, 2000) who presented a PPoF approach to reliability assurance of high-performance chips by considering common defect activation energy distribution. Hall and Strutt (Hall & Strutt, 2003) have presented PPoF models for component reliabilities by considering parameter and model uncertainties. Azarkhail and Modarres (Azarkhail & Modarres, 2007) have presented a Bayesian framework for uncertainty management in physics-based reliability models. Matik and Sruk (Matik & Sruk, 2008) highlighted the need for PoF to be probabilistic in order to include inevitable variations of variables involved in processes contributing to the occurrence of failures in the analysis. Finally, Chatterjee and Modarres (Chatterjee & Modarres, 2012) have presented PPoF modeling of integrated steam generators in small modular reactors. Although substantial research has been done on PPoF modeling for reliability assessment, more research in this area is necessary.

The element of a PPoF model that assesses time-to-failure of a component (such as a ball bearing under fatigue-wear degradation mechanism) is illustrated in Figure 1.5. The lowest element in this figure shows the inter- and intra-environmental factors that produce the stresses that cause degradation and failure. The next three elements are drawn from the PoF models that describe the degradation. The final element (top level) is the probabilistic life assessment that formally accounts for parameter uncertainties in the PoF models and model errors. In Figure 1.5, a probabilistic approach (such as Bayesian inference) is shown to characterize the corresponding PoF model uncertainties. The arrows in Figure 1.5 show the direction of influences, such as how external ambient temperature may affect viscosity. Usually the direction of influences is upward (i.e., sequential causal relationships), but it is possible to have some influences going downward, causing a circular synergy among variables. For example, certain operating conditions, such as high internal temperature generated by poor lubrication during operation, lowers lubricant viscosity, which in turn increases the friction that further exacerbates the high internal temperature.

There are two basic types of uncertainties that can be described by a PPoF model of failure mechanisms: aleatory and epistemic uncertainty. Aleatory uncertainty is the inherent randomness of the phenomena that the model attempts to predict. This type of uncertainty is intrinsic and cannot be reduced. Examples of aleatory uncertainty include random environmental variations, random vibration in stress amplitude and certain material properties such size and density of flaws. Epistemic uncertainty is about

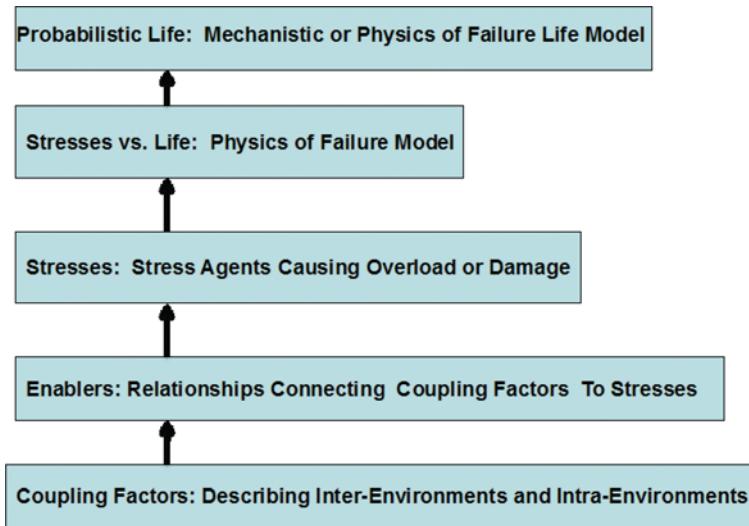


Figure 1.5: System hierarchy in probabilistic-mechanistic reliability life model

our lack of knowledge and consists of incomplete description of the modeled phenomena, measurement errors, and a lack of sufficiently accurate measurements to fully capture the phenomena. Incorporating additional PoF model data and information reduces this type of uncertainty: as such this uncertainty is reducible, whereas aleatory uncertainty is not. Since there can be uncertainties associated with failure-inducing agents (i.e., stresses), model parameters and the model itself, the prediction of failures is inherently a probabilistic problem requiring PPoF models for applicable primary failure mechanisms (Modarres, Kaminskiy, & Krivtsov, 2017). Each failure, damage or degradation mechanism should have its own PPoF model. All applicable PPoF models applied to an item need to be combined to find the overall degradation process. Methods for combining multiple PPoF models include the use of the weakest link approach, which assumes that one of such degradation mechanisms causes damage that will exceed the endurance limit before the other applicable mechanisms.

PPoF models are formulated considering all the variables that can initiate and propagate degradation in the item under study. As part of this process, one should identify important degradation causing variables such as (for a rotating tube example) normal loads, displacement amplitudes, and material properties contacting surfaces. In this example, tube degradations may be measured in terms of volume of material lost and then correlated with the stress variables. Experimental degradation data from accelerated testing would be needed to determine the PoF-based correlation between degradation and causal stress variables. In some cases, well-established correlations from the literature are used. In developing PoF models, other important variables (such as geometry) may also be considered. The next step is to characterize all forms of uncertainties associated with the PoF models and data, and estimate model parameters. This step converts the PoF models into PPoF models. A suitable regression approach should be developed to formally characterize all uncertainties. Bayesian regression is a powerful technique for estimating probability distributions of model parameters. In the tube example above, one requires experimental degradation data under prevailing environments experiencing operational conditions corresponding to each degradation mechanism.

Other factors that can lead to uncertainties in the likelihood of failures (such as manufacturing methods and material properties) also need to be considered. Most of these uncertainties should be accounted for when evaluating the stress agents acting on the tubes. Flow-induced vibratory stresses and thermal stresses that propagate tube degradations are two such examples for the tube example above. Each failure mechanism has specific stress agents that cause degradation. Agents like fatigue stress are alternating stresses, whereas stress agents of stress-corrosion-cracking (SSC) involve constant tensile

8 PROBABILISTIC PHYSICS OF FAILURE APPROACH TO RELIABILITY

stresses. In order to determine the stress agents, a detailed finite element analysis of the tube geometry and material properties is required, with prevailing operating conditions applied. The input parameters for the finite element analysis, (e.g., geometry, material properties) need to be entered probabilistically—not deterministically—with corresponding stresses estimated as probability distributions.

A Monte Carlo simulation approach complements PPoF models by propagating all the associated uncertainties (such as those associated with the model, its parameters, and initial material flaws) to estimate probability distribution of the unit failure or amount of damage as a function of time under the prevailing stresses. Monte Carlo simulation is the leading method for simulating systems with many coupled input variables. Appropriate failure criteria need to be defined for each failure mechanism considering the operability. For example, a failure criterion for a normal operating condition-induced fatigue mechanism can be defined as the through-wall cracks reaching the wall thickness of tubes.

1.5. ACCELERATED TESTING IN PPOF MODEL DEVELOPMENT

To develop the PoF models and estimate their parameters and model uncertainties, it is imperative to rely on failure and degradation (damage) versus time data. These data can be obtained from life and degradation testing or from valid field data. Many of today's structures, systems and components are capable of operating under benign environmental stresses for thousands of hours without failure. This makes normal life (non-accelerated) testing of such equipment difficult and costly. Field data in many cases are scarce, and even when they are available it is hard to judge their uniformity and accuracy. Alternatively, accelerated testing provides far more quickly a better understanding of equipment life and degradation processes, and generates data for development of PoF and PPoF models. As such, generating degradation data in the shortest possible time can be achieved by relying on formal accelerated testing methods.

In essence, accelerated testing gathers more reliability and life information in a shorter span of time by utilizing a more severe test environment than what would otherwise be experienced under normal use conditions. Accelerated testing increases the stressors that are known to dominate the causes of failure of a system in order to test it in a compressed timeframe. Importantly, accelerated testing ensures that failure modes and mechanisms that would not be encountered under normal use are not inadvertently introduced in the test. The trajectory of cumulative degradation shown in Figure 1.3 and Figure 1.4 would be shifted to the left under an accelerated testing regime, as failures occur faster than under normal (or use) operating conditions.

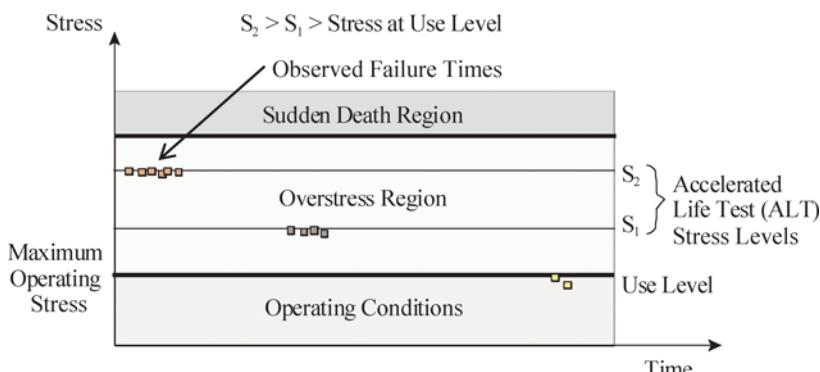


Figure 1.6: Conceptual acceleration of stress agents at two overstress conditions and corresponding data points generated from ALT

Accelerated tests are used to develop PPoF models, which in turn can be used to estimate and predict the equipment life or degradation and damage under normal operating conditions. This step in PPoF analysis underlines the importance of formally characterizing all the uncertainties in the PPoF models so as to reflect such uncertainties in the predicted life from such models.

Figure 1.6 illustrates the stress regions of a conceptual accelerated test that generate several failure data points at two stress levels in the “overstress” region. These data are then used to develop the PoF model that best describes them, including the associated uncertainties to extrapolate the resulting PPoF models (associated with each quantile of the life) to the “use” stress level to estimate the corresponding life distribution (see Figure 1.7).

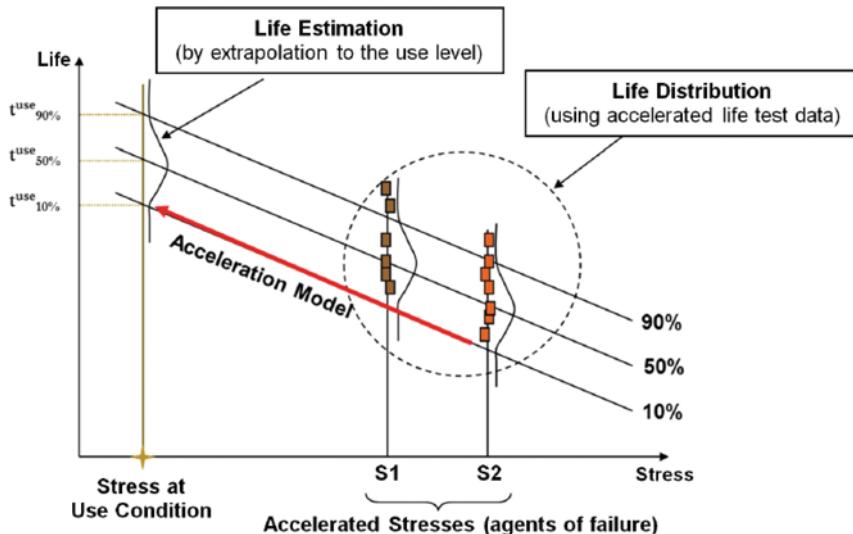


Figure 1.7: Conceptual PPoF model development and extrapolation in ALT

There are three fundamental approaches to accelerated testing: (1) field-testing of the unit under higher use frequency or higher operating stresses (loads); (2) laboratory testing of the unit, prototype, material samples or coupons under higher stress or higher use frequency; and (3) computer-based simulated acceleration using previously validated PPoF models. Acceleration of the stress variable is achieved either singly or in combination. Examples include:

- More frequent power cycling
- Higher vibration levels
- Higher humidity
- More severe temperature cycling
- Higher temperatures
- Higher load amplitudes

In addition to developing PoF and PPoF models, there are other motivations for accelerated testing to assure reliability. These motivations stem from, for example:

- The need to identify design failures. Results from accelerated tests can help eliminate or reduce design failures through redesign (e.g. intrinsic redundancy).
- The need for immediate verification of lifetime statistics (rather than waiting for field data). This is especially important for cases involving the prediction of performance of highly reliable

10 PROBABILISTIC PHYSICS OF FAILURE APPROACH TO RELIABILITY

products where access to normal use failure data is not feasible or practical.

- The need for shorter life tests due to the fast technological evolution of products.
- The need to assess and demonstrate component reliability in the design stage.
- The need to certify components and detect failure modes so that they can be corrected.
- The need to compare different manufacturers and vendors in a condensed timeframe. Accelerated testing can aid in choosing designs, components, suppliers, rated operating conditions, and test procedures.
- The need to determine an appropriate service policy in a condensed timeframe. Accelerated testing can aid in determining the most appropriate approach to inspecting, servicing and replacing units.

There are two basic categories of accelerated tests: quantitative tests and qualitative tests. The former commonly refers to Accelerated Life Tests (ALT) and Accelerated Degradation Tests (ADT). The latter is characterized by tests that aim to enhance the reliability of the item during design and operation.

Quantitative tests, are conducted on items (structures, systems, components) and manufacturing processes. They can take a few weeks to a few months to complete. ALT is fundamentally based on the assumption that the unit under test will exhibit the same behavior under a shorter time frame (at a high stress) as it would in a longer time frame at use stress conditions. Hence, there are several important planning considerations when conducting ALT tests so that this assumption remains valid.

Qualitative accelerated tests are designed to find failures linked to design or manufacturing without providing any life or damage characteristics associated with the items. By accelerating failures of structures, components or systems, these tests can determine the robustness of the unit in its useful life. When a failure occurs during a qualitative accelerated test, one needs to determine the root cause of the failure and judge whether the failure mode and mechanism observed would occur under normal use conditions. The most common type of qualitative tests is known as Highly Accelerated Life Testing (HALT). Being a qualitative accelerated test, it is important to note that HALT is not a life test: its purpose is not to determine life characteristics. Rather, it is a test designed to promote the occurrence of failure modes (mechanical or electronic) that will occur during the life of the product under normal use conditions. HALT provides valuable information to determine design weaknesses as well as the product's upper and lower destruct limits.

Another example of a qualitative accelerated test is known as Highly Accelerated Stress Screening (HASS). HASS tests are applied during the manufacturing phase and are used to screen marginal and defective units. HASS can expose infant mortality failures and other latent defects that would otherwise occur when the unit is being used. Whereas HALT is applied during the design phase to iron out potential design issues, HASS screens out defects (that can be fed back to designers) associated with the manufacturing process. HASS generally builds upon HALT but typically uses lower stresses.

1.6. ORGANIZATION OF THE BOOK

This book can be divided into three parts. Part one (chapters 1 and 2) introduces the PPoF approach more formally and provides a comprehensive coverage of key failure mechanisms, including their corresponding PoF models. Part two (chapters 3 to 6) focus on accelerated testing. Types of accelerated tests, accelerated degradation modeling, test planning, data gathering, PPoF model development, characterization and analysis methods will be covered. Part three (chapter 7) focuses on uncertainties in physics-based models in reliability and prognosis and health management.

REFERENCES

- Azarkhail, M., & Modarres, M. (2007). A novel Bayesian framework for uncertainty management in physics-based reliability models. ASME International Mechanical Engineering Congress and Exposition. Seattle.
- Chatterjee, K., & Modarres, M. (2012). A probabilistic physics-of-failure approach to prediction of steam generator tube rupture frequency. Nuclear Science and Engineering, 170(2), 136-150.
- Dodson, G., & Howard, B. (1961). High stress aging to failure of semiconductor devices. Seventh National Symposium on Reliability and Quality Control. Philadelphia.
- Dowling, N. (1998). Mechanical Behaviour of Materials. Upper Saddle River: Prentice Hall.
- Haggag, A., McMahon, W., Hess, K., Cheng, K., Lee, J., & Lyding, J. (2000). A probabilistic-physics-of-failure/short-time-test approach to reliability assurance for high-performance chips: Models for deep-submicron transistors and optical interconnects. Proceedings of IEEE Integrated Reliability Workshop, (pp. 179-182).
- Hall, P., & Strutt, J. (2003). Probabilistic physics-of-failure models for component reliabilities using Monte Carlo simulation and Weibull analysis: a parametric study. Reliability Engineering & System Safety, 80(3), 233-242.
- Irwin, G. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. Journal of Applied Mechanics, 24, 361-364.
- Levin, M., & Kalal, T. (2003). Improving Product Reliability. Chichester: John Wiley and Sons.
- Matik, Z., & Sruk, V. (2008). The physics-of-failure approach in reliability engineering. IEEE International Conference on Information Technology Interfaces, (pp. 745-750).
- Meeker, W., & Escobar, L. (1998). Statistical Methods for Reliability Data. New Jersey: John Wiley and Sons.
- Modarres, M., Kaminskiy, M., & Krivtsov, V. (2017). Reliability engineering and risk analysis: A practical guide 3rd Ed. Boca Raton: CRC Press.
- Paris, P., Gomez, M., & Anderson, W. (1961). A rational analytic theory of fatigue. The Trend in Engineering, 13, 9-14.
- Vaccaro, J. (1962). Reliability and physics of failure program at RADC. First Annual Symposium on the Physics of Failure in Electronics, (pp. 4-10).
- Weibull, W. (1959). Statistical evaluation of data from fatigue and creep rupture tests, Part I: fundamental concepts and general methods. Sweden: Wright Air Development Center.
- Weibull, W. (1961). Fatigue testing an analysis of results. London: Pergamon Press.

Chapter 2: Summary of Mechanisms of Failure and Associated PoF Models

2.1. INTRODUCTION

Chapter 1 provided an overview and the motivation behind accelerated testing. As discussed, one of the key elements required in order to successfully conduct an accelerated life test is determining the most suitable accelerated life model (which is contingent on the most suitable PoF model), upon which the test can be based. The selection of the most appropriate accelerated life model is directly influenced by the mechanism of failure that is being considered for the applicable material. For this reason, the study of the PoF is pertinent in any accelerated life testing analysis. This chapter will discuss important failure mechanisms that illustrate the process of deriving the relationship between life and applied stresses from the physics of the failure of the mechanism. Some of the mechanical, thermal, electrochemical and electrical failure mechanisms include:

- Fatigue cracking
- Creep
- Pitting corrosion
- General corrosion
- Crevice corrosion
- Erosion
- Wear
- Radiation embrittlement
- Hydrogen embrittlement
- Elecromigration
- Conductive filament formation
- Thermally-induced fatigue: cyclic creep-fatigue
- Fatigue induced by vibration
- Fracture induced by shock & drop
- Fretting-corrosion in connectors induced by vibration

Also, a few synergistic failure mechanisms are:

- Creep-fatigue
- Corrosion-fatigue
- Fretting-wear

Classification of material degradation by its root cause is the first step in any plan to protect a specific material in a specific situation. A classification of the known forms of material degradation as physical, chemical or biological phenomena is illustrated schematically in Figure 2.1 as a Venn diagram.

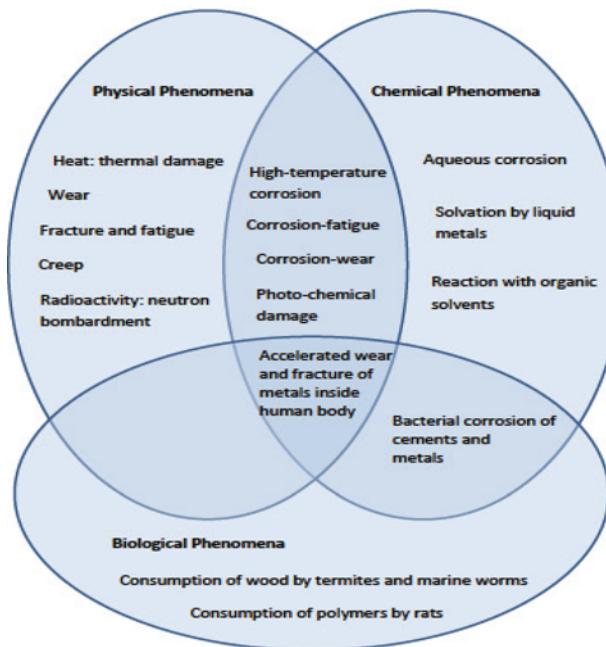


Figure 2.1: Classification of material degradation

For all materials (including electronics) there are three basic categories of material degradation: those that are (1) physical, (2) chemical and (3) biological in origin. A physical origin is one that involves the effect of force, heat or radiation. A chemical origin involves the destructive reactions between material and chemicals that contact it. A biological origin includes all interactions between life forms and engineering materials.

Physical and chemical material degradation (such as thermal damage or destructive chemical reactions) coexist with combined forms of material degradation (such as corrosive-wear). Environmental conditions strongly affect material degradation, meaning that any material degradation problem depends on locality. For example, a component operating in a corrosive environment is susceptible to corrosion, and a component that is inside of a living organism is more subject to biological damage. As will be shown in later chapters, there is a wide range of possible interactions or synergy between degradation processes, and considerable care is needed to predict which of these interactions are significant to any given situation.

An example of interaction between degradation processes is corrosive-wear, which is mechanical wear accelerated by chemical damage to the worn material. Biological phenomena, which can cause this corrosion, are either found in a pure form (such as organisms eating artificial materials) or in more complex relationships (such as bacterial corrosion where bacterial waste products are destructive to materials).

Damage that depends on just one phenomenon is easier to recognize than multi-phenomena damage such as those occurring in corrosive-wear. Damage measures are usually ineffective when dealing with multi-phenomena. In this book, single-phenomenon material degradation is described before more complex forms of degradation are discussed.

The PoF approach to ALT aims to use the science of physics to create analytical models to describe the failure of components and devices. In essence, the PoF methodology explains why and when a

particular component or material fails. From Chapter 1, the PoF modeling process can be summarized as follows:

- Specify the item or system, its operating limits, all its pertinent characteristics and operating requirements (i.e. what it is supposed to do).
- Define the item's design operating environment and profile.
- Use the profile to determine the item's static or dynamic mechanical, thermal, electrical and chemical loads (stress or damage agents).
- Identify "hot spots" exposed to the highest stress or load and the associated failure mechanisms that become activated.
- Determine materials, construction, manufacturing and assembly processes.
- Develop a PoF model that correlates loads applied to amount or rate of degradation.
- Use generic data or test data (primarily from accelerated tests) to estimate the parameters of the PoF model, including any parameter uncertainties and model error.
- Determine a level of degradation beyond which the item fails to operate or endure more damage.
- Using the PoF model and the endurance limit, estimate the time or cycles of operation that a failure event is expected to occur under various conditions, including uncertainties associated with such estimation.
- Perform computer-based simulation to estimate expected life or remaining life of the item.

In Table 2.1 a summary of some important failure mechanisms and their associated driving factors is presented. Wear and corrosion failures account for a significant proportion of those experienced in the field; however, corresponding quantitative PoF models are yet to be developed into mature and effective design tools. For more detailed information on wear and corrosion, refer to Chapter 17 of *Failure of Materials in Mechanical Design* (Collins, 1993). There are of course many other failure modes and mechanisms that are applicable to materials and therefore relevant to accelerated testing. The focus of this chapter will be on the physical and chemical degradation mechanisms that are commonly observed in machinery components and structures. These include fatigue, corrosion, wear, and creep.

2.2. FATIGUE

Fatigue is a degradation phenomenon emanating from the repeated loading of forces that cause component damage (and failure). The fatigue life of a component can be divided into (1) the life required to initiate a crack and (2) the life required for that crack to propagate to fracture. Fatigue is an important aspect of ALT, as many hardware components are subjected to some form of it. This chapter covers three methods for fatigue analysis: (1) the life-stress approach formulated between 1850-1870; (2) the strain-life approach formulated in the 1960s; and (3) the fracture mechanics approach also formulated in the 1960s. The first two methods concern the segment of life up to crack initiation and the third method with the crack propagation life. Each method can be used to perform fatigue analysis. Nonetheless, understanding all three enables the engineer to choose the most appropriate analysis method. A discussion on a hybrid approach combining the strain-life and fracture mechanics methods will also be presented. In addition, the effect of notches in the material requires a variation in the analysis, which will also be covered in this section.

16 PROBABILISTIC PHYSICS OF FAILURE APPROACH TO RELIABILITY

Table 2.1: Summary of common failure modes (Collins (1993), Arunajadai, Stone and Turner (2002) and Khonsari and Amiri (2012))

Failure Mechanism	Failure inducing agents	Sub-categories	Examples
Fatigue	Fluctuating stress/strain Assisted by environment: temperature, humidity, oxidation, corrosion	High-cycle fatigue Low-cycle fatigue Thermal fatigue Impact fatigue Surface fatigue Fretting fatigue Corrosion fatigue Creep fatigue	Airframes, pipelines, bridges, railroad structures, rotating shafts, turbine blades, pumps, bolts, gears, hip joint, welded structures, solder in electronic devices
Corrosion	Chemical or electrochemical reaction with environment Assisted by stress, deformation, abrasion, wear	Direct chemical attack Galvanic corrosion Uniform corrosion Pitting corrosion Erosion corrosion Crevice corrosion Intergranular corrosion Stress corrosion Biological corrosion Hydrogen damage Corrosion fatigue Dealloying corrosion	Pressure vessels, boiler tubes, pumps, compressors, bridges, crude oil storage tanks, airframes, marine structures, bolts, medical devices
Wear	Relative motion between mating surfaces Plastic deformation Assisted by environment: humidity, oxidation, temperature, corrosion	Adhesive wear Abrasive wear Fretting wear Corrosive wear Fatigue wear Impact wear Deformation wear	Pipe bends, seals, bearings, gears, disks and tapes, piston rings, nuclear machinery, drills, pump impellers, human teeth and joints
Creep	Plastic deformation due to stress and elevated temperature Assisted by fluctuating stress/strain, corrosion	Creep fatigue Creep corrosion	Boiler super-heaters, petro-chemical furnaces, reactor vessel components, gas turbine blades, aeroengines
Fracture	Static stress Elastic deformation (brittle material) Plastic deformation (ductile material)	Brittle fracture Ductile fracture	Ship structures, welded joints, composite structures, bridge structures, tanks
Impact	Dynamic or suddenly applied load Elastic/plastic deformation	Impact fracture Impact deformation Impact fatigue Impact wear Impact fretting	Aircrafts, composite structures, armor systems, combat vehicles, seat belt receptacles

Table 2.1: Summary of common failure modes (Collins (1993), Arunajadai, Stone and Turner (2002) and Khonsari and Amiri (2012)) (Continued)

Failure Mechanism	Failure inducing agents	Sub-categories	Examples
Brinelling	Static contact stress Local yielding Vibration of two stationary contacting bodies	Brinelling False brinelling	Idle rolling components, bearings, induction motors
Thermal Shock	Pronounced thermal gradient		Turbine engines, propulsion systems
Buckling	High and/or point compressive or torsional load Assisted by high temperature	Buckling Creep buckling	Bars, tubes, columns, engine shafts
Yielding	Plastic deformation		Ductile materials subjected to tensile stress, helicopter engine shaft and components

2.2.1. LIFE-STRESS

The life-stress method for fatigue analysis is the simplest. It involves counting the number of cycles to failure of a component subjected to fatigue based on the applied stress. This method is purely quantitative and has no specific consideration for the PoF of the phenomenon. The life-stress method has been shown to model failure of long-life components subjected to constant amplitude stresses well. There is also ample reliability data for many materials, making the life-stress method widely applicable.

The life-stress approach does not consider the true stress-strain behavior of the material being analyzed and treats all strains as elastic (Bannantine, et al., 1997). For this reason, the life-stress approach should only be used when the plastic strains of the material under load are small. For high cycle fatigue (where cycles to failure are greater than about 10^4 cycles), the life-stress approach is typically suitable.

In fatigue analysis, it is important to know the various measures of stress and the terminology associated with these measures. Figure 2.2 illustrates constant amplitude cyclic loading.

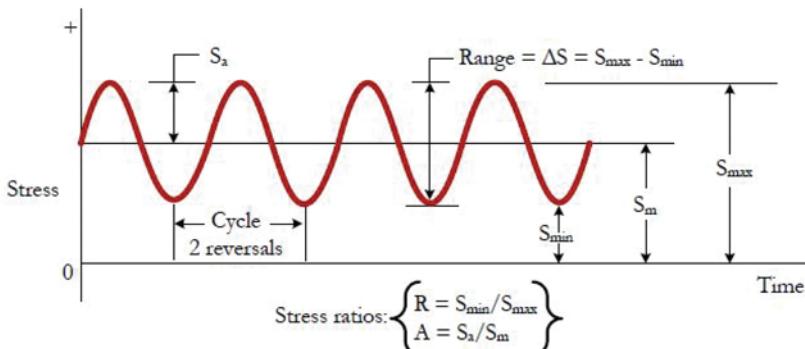


Figure 2.2: Nomenclature for constant amplitude cyclic loading

From Figure 2.2, the following is inferred:

- Stress Range: $\Delta S = S_{max} - S_{min}$
- Stress Amplitude: $S_a = \frac{\Delta S}{2}$
- Mean Stress: $S_m = S_{max} - S_a = S_{min} + S_a$
- Stress Ratio: $R = \frac{S_{min}}{S_{max}}$
- Amplitude Ratio: $A = \frac{S_a}{S_m}$

Note that $R=0$ (zero to maximum) and $R=-1$ (fully reversed) are two common reference test conditions used for obtaining fatigue properties.

2.2.1.1. The S-N Diagram

The S-N diagram is the basis for the life-stress approach method. It is a plot of the applied stress, S , versus the number of cycles to failure, N for fully reversed loading. An example of such a diagram is shown below in Figure 2.3. A number of methods exist for generating the required data for such an S-N diagram. S-N test data are generally plotted on a log-log graph with the S-N line representing the mean of the data (Bannantine, et al., 1997). Such an idealized S-N diagram is shown in Figure 2.3.

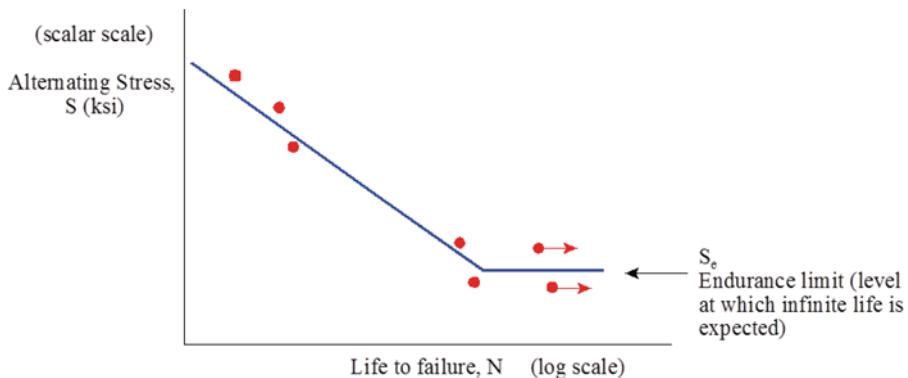


Figure 2.3: Idealized S-N diagram

As shown in Figure 2.3, the endurance limit S_e , is the level at which infinite life is expected, or the level at which the stress is low enough to consider that the damage caused by cyclic loading at that stress level will not cause failure. If the stress applied to the material is below this endurance limit, the material is said to have an infinite life. This is a basic assumption of the life-stress method, and although materials such as steel and titanium display this characteristic, there are some constraints in the assumption. Therefore, the assumption should only be applied in benign environmental conditions. The concept of an endurance limit is used in infinite life and safe stress designs. However, care must be taken in using the assumption of an endurance limit. Periodic overloads, corrosive environments, or high temperatures may invalidate it. The endurance limit is not based on the physics of the material: there are various factors that influence it, including surface finish, temperature, stress concentration, notch sensitivity, size, and environment.

Returning to the idealized S-N diagram in Figure 2.3, a log-log straight line S-N relationship results in the following (Stephens, Fatemi, Stephens, & Fuchs, 2001):

$$S_a \text{ or } S_{N_f} = A(N_f)^B \quad (2.1)$$

where S_a is the applied alternating stress and S_{N_f} is the fully reversed fatigue strength at N_f cycles. A and B are constants to be determined, with A being the value of S_{N_f} at one cycle, and B being the slope of the log-log S-N curve.

Therefore, the linear relationship on a log-log scale can be written as:

$$\log(S_{N_f}) = \log(A) + B \cdot \log(N_f) \quad (2.2)$$

Some other important relationships worth noting with regard to the life-stress method are as follows:

- Endurance limit related to hardness:

$$S_e \approx 0.25 \times \text{Brinnell Hardness Number (BHN)} \quad \text{for } BHN \leq 2800 \text{ MPa (400 ksi)}$$

if $BHN > 2800 \text{ MPa}, S_e \approx 700 \text{ MPa}$

- Endurance limit related to ultimate strength (point at which stress produces fracture):

$$S_e \approx 0.5 \times S_u \quad \text{for } S_u \leq 1400 \text{ MPa (200 ksi)}$$

if $S_u > 1400 \text{ MPa}, S_e \approx 700 \text{ MPa}$

2.2.1.2. Mean Stress Effects

The mean stress, S_m , can have a significant influence on the fatigue behavior. This is shown below in Figure 2.4 where alternating stress is plotted against the number of cycles to failure for three different mean stresses. In general, tensile mean stresses have a detrimental effect on the material's life, and compressive stresses have a beneficial effect. This intuitively makes sense, as cracks don't propagate under compression.

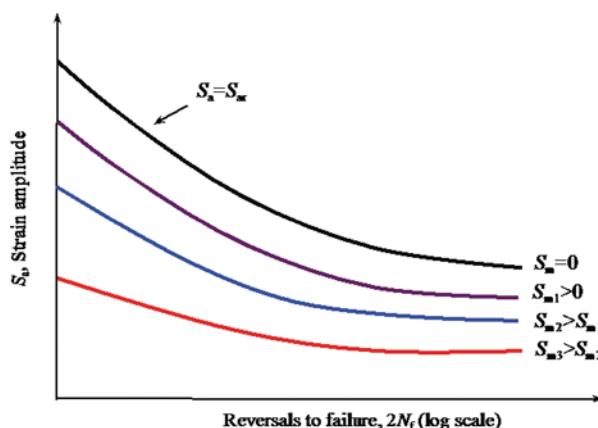


Figure 2.4: Effect of mean stress on fatigue life

The effect of mean stress thus results in various correction relationships as follows:

$$\begin{aligned}
 \frac{S_a}{S_{ar}} + \frac{S_m}{S_y} &= 1 \text{ (Soderberg relationship, } S_y \text{ denotes yield stress)} \\
 \frac{S_a}{S_{ar}} + \frac{S_m}{S_u} &= 1 \text{ (Modified Goodman relationship)} \\
 \frac{S_a}{S_{ar}} + \frac{S_m}{\sigma_f} &= 1 \text{ (Morrow relationship, } \sigma_f \text{ denotes true fracture} \\
 &\quad \text{strength)} \\
 \frac{S_a}{S_{ar}} + \left(\frac{S_m}{S_u}\right)^2 &= 1 \text{ (Gerber relationship)} \\
 \frac{S_a}{S_{ar}} &= \left(\frac{1-R}{2}\right)^{1/2} \text{ (Smith-Watson-Topper relationship, } R \text{ denotes} \\
 &\quad \text{stress ratio)} \\
 \frac{S_a}{S_{ar}} &= \left(\frac{1-R}{2}\right)^{1-\gamma} \text{ (Walker relationship, } \gamma \text{ is the fitting constant)}
 \end{aligned} \tag{2.3}$$

In the equations above, S_{ar} is the equivalent stress amplitude for a non-zero mean stress that can be used in an S-N diagram that assumes zero mean stress and fully reversed loading. These relationships are commonly used in fatigue design and analysis and are modified for notches, size, surface finish, environmental effects and finite life (Stephens, et al., 2001). Each one of these models has advantages and drawbacks in particular applications. For example, the Gerber relationship is simple to use and works better than the Goodman model for ductile materials and low applied stress. However, it is inaccurate for compressive mean stresses. The Smith-Watson-Topper relationship is recommended for general use (Dowling, 2004). A summary of applicability of these relationships is given in Khonsari and Amiri (2012).

Example 2.1: S-N Approach to Fatigue Analysis

Consider the example where a steel component is subjected to a cyclic stress with a maximum value of 760 MPa and a minimum of 70 MPa. The component has ultimate strength S_u of 1,030 MPa and endurance limit S_e of 410 MPa. A fully reversed stress at 1,000 cycles of 760 MPa is applied. In this example, the Goodman relationship is used to determine the life of the component.

Solution

First, we determine the stress amplitude and mean stress as follows:

$$\begin{aligned}
 S_a &= \frac{S_{max} - S_{min}}{2} = \frac{760 - 70}{2} = 345 \text{ MPa} \\
 S_m &= \frac{S_{max} + S_{min}}{2} = \frac{760 + 70}{2} = 415 \text{ MPa}
 \end{aligned}$$

Using the modified Goodman relationship presented in the previous section, one can determine S_{ar} as follows:

$$\begin{aligned}
 \frac{345}{S_{ar}} + \frac{415}{1030} &= 1 \\
 S_{ar} &\cong 578 \text{ MPa}
 \end{aligned}$$

The value for S_{ar} can now be used with the S-N diagram (Figure 2.5) to estimate the life of the component, N_f . (Recall that the S-N diagram represents fully reversed loading). When 578 MPa is used, the resulting life estimate is 22,113 cycles. This problem could also be solved using other relationships such as the Gerber or Morrow mean stress relationships. Each would provide a slightly different life estimate. This example is a modified version from Bannantine, et al. (1997).

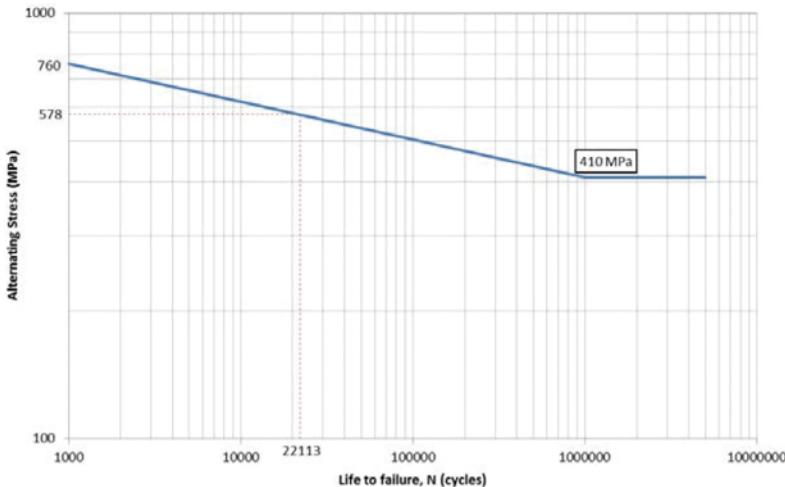


Figure 2.5: S-N diagram for example problem

2.2.1.3. Combined Loading

In engineering applications, components and structures commonly experience combined states of stress, such as combinations of torsion, axial and bending loads. For example, a thin-walled cylindrical pressure vessel experiences a circumferential (hoop) and longitudinal stress as shown in Figure 2.6. A shaft or axle may experience a combined effect of torsion and bending load. The multiaxial fatigue methodologies can be classified into five groups (You and Lee, 1996): (1) empirical formulas and modifications of the Coffin-Manson equation; (2) stress or strain invariants methods; (3) space averages of stress or strain; (4) critical plane theory; and (5) energy based methods. You and Lee (1996) provide a comprehensive review of these multiaxial fatigue methods with their validity regions and limitations. After the pioneering work of Stanfield (1935), the critical plane theory has gained considerable attention because of its effectiveness and broad application range (Karolczuk and Macha, 2005). In the critical plane approach, fatigue life is calculated on planes on which the combined effect of normal and shear stresses is maximum. These planes would be the fatigue fracture planes and are named “critical” by Stulen and Cummings (1959). Critical planes are not necessarily principal planes. The principal plane is the plane in Mohr’s circle where maximum and minimum principal stresses occur.

A simple multiaxial fatigue stress analysis is based on the *effective stress amplitude* and *effective mean stress*. It assumes that the multiaxial fatigue life can be evaluated based on the effective stress. For any convenient coordinate system (as shown in Figure 2.7), the amplitude and mean effective stress can be calculated from the amplitudes and means of the stress components:

$$\bar{\sigma}_a = \frac{1}{\sqrt{2}} \sqrt{(\sigma_{xa} - \sigma_{ya})^2 + (\sigma_{ya} - \sigma_{za})^2 + (\sigma_{za} - \sigma_{xa})^2 + 6(\tau_{xya} + \tau_{yza} + \tau_{zxa})^2} \quad (2.4)$$

$$\bar{\sigma}_m = \sigma_{xm} + \sigma_{ym} + \sigma_{zm}$$

The subscripts a and m refer to amplitude and mean of particular stress components, and the bar sign indicates effective stress. Having determined the effective amplitude and mean stresses, the fatigue life can be calculated from the life-stress approach in section 2.2.1.1.

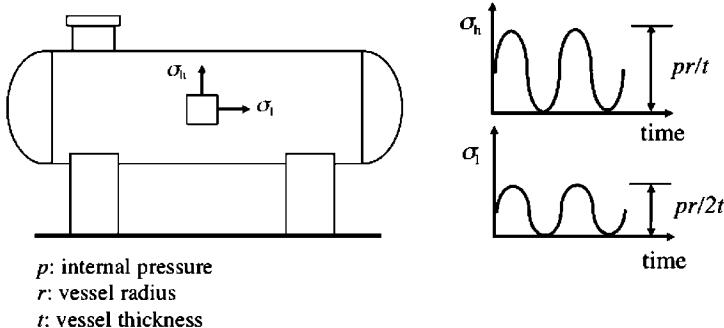


Figure 2.6: Biaxial stress state in thin-walled pressure vessel

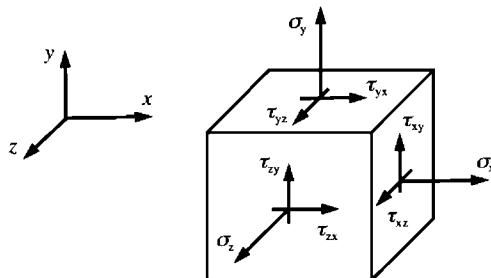


Figure 2.7: State of stress at a point represented by stress element

Example 2.2: Multiaxial Stress Analysis

A pressure vessel with a 1.4 m diameter and a 3.6 m length has a wall thickness of 16 mm made of AISI 1040 HR steel ($S_y = 227$ MPa and $S_u = 415$ MPa). The vessel is subjected to fluctuation of 10 cycles per minute. In each cycle, the pressure fluctuates between 2 and 6 MPa. What is the expected life of the pressure vessel (assuming the vessel is thin-walled)? The values of A and B for life-stress formulation in section 2.2.1.1 are 886 and -0.14 , respectively.

Solution

First, determine the amplitude and the mean pressure as follows:

$$p_a = \frac{p_{max} - p_{min}}{2} = \frac{6 - 2}{2} = 2 \text{ MPa}$$

$$p_m = \frac{p_{max} + p_{min}}{2} = \frac{6 + 2}{2} = 4 \text{ MPa}$$

Now, calculate stress amplitudes in longitudinal and circumferential (hoop) directions:

$$\sigma_{al} = \frac{p_a r}{2t} = \frac{2 \times (1.4/2)}{2 \times 0.016} = 43.8 \text{ MPa}$$

$$\sigma_{ah} = \frac{p_a r}{t} = \frac{2 \times (1.4/2)}{0.016} = 87.5 \text{ MPa}$$

Mean stresses in longitudinal and circumferential (hoop) directions are:

$$\sigma_{ml} = \frac{p_m r}{2t} = \frac{4 \times (1.4/2)}{2 \times 0.016} = 87.5 \text{ MPa}$$

$$\sigma_{mh} = \frac{p_m r}{t} = \frac{4 \times (1.4/2)}{0.016} = 175 \text{ MPa}$$

The effective stress amplitude and the effective mean stress are:

$$\bar{\sigma}_a = \frac{1}{\sqrt{2}} \sqrt{(87.5 - 43.8)^2 + (87.5 - 0)^2 + (0 - 43.8)^2} = 75.8 \text{ MPa}$$

$$\bar{\sigma}_m = 87.5 + 175 = 262.5 \text{ MPa}$$

Using the modified Goodman relationship presented in section 2.2.1.2, one can determine σ_{ar} , the fully reversed stress level corresponding to the same life as that obtained with the stress conditions S_a and S_m , as follows:

$$\sigma_{ar} = \frac{\bar{\sigma}_a}{1 - \frac{\bar{\sigma}_m}{S_u}} = \frac{75.8}{1 - \frac{262.5}{415}} = 206.3 \text{ MPa}$$

The value for σ_{ar} can now be used as stress amplitude in the life-stress equation presented in section 2.2.1.1 to determine the life:

$$\sigma_{ar} = A(N_f)^B$$

$$N_f = \left(\frac{\sigma_{ar}}{A} \right)^{\frac{1}{B}} = \left(\frac{206.3}{886} \right)^{\frac{1}{-0.14}} = 3.32 \times 10^4 \text{ cycles} \approx 2.3 \text{ days}$$

2.2.2. STRAIN-LIFE

In the strain-life method for fatigue analysis, both plastic and elastic deformation of material is measured. As discussed in section 2.2.1, the life-stress method does not consider plastic strain. At long lives, where plastic strain is negligible and stresses and strains are linearly related, the strain-life and life-stress methods are essentially equivalent. The strain-life method assumes that failure occurs when a crack, typically 1 mm in length, is established. One of the key advantages of the strain-life method is that the response of material in hot spots (such as notches) is often dependent on the strain or deformation factor. Unlike the life-stress method, the strain-life method is suitable for analyzing low cycle fatigue. However, it does not account for crack growth and is often used when considering crack initiation. For some applications, the mere existence of a crack is an overly conservative criterion for failure. As a result, the fracture mechanics method, discussed in section 2.2.2.1, is used to take into account both crack initiation and crack propagation in fatigue design and analysis.

The following sections will discuss some basic principles associated with the strain-life approach to fatigue analysis. They will cover relevant information pertaining to the monotonic stress-strain behavior, cyclic stress-strain behavior, and finally, the strain-life relationship.

2.2.2.1. Monotonic Stress-Strain Behavior

Applying a monotonic stress that induces strain is often used to obtain design parameters in engineering structures and components subjected to static loading. To understand the stress-strain behavior under monotonic load, it is important to define a number of parameters. Consider the cylindrical specimen shown in Figure 2.8. The *engineering stress* is the average normal stress in the specimen and is given by:

$$s = \frac{P}{A_0} \quad (2.5)$$

where P is the applied load and A_0 is the original unloaded cross-sectional area of the specimen. The *engineering strain* is the average linear strain along the specimen length and is given by:

$$e = \frac{l - l_0}{l_0} \quad (2.6)$$

where l_0 is the original unstrained specimen length and l is the strained length.

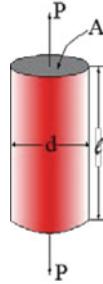


Figure 2.8: Cylindrical specimen under monotonic load

The *true stress* σ , is the actual stress based on the actual (instantaneous) area, A , corresponding at every instant in time. Thus true stress is given by:

$$\sigma = \frac{P}{A} \quad (2.7)$$

The *true strain* is calculated by integral length from the original length to the instantaneous length:

$$\varepsilon = \int_{l_0}^l \frac{dl}{l} = \ln\left(\frac{l}{l_0}\right) \quad (2.8)$$

The *engineering* term does not take into account how strain reduces the cross sectional area of the specimen as it is stretched. This stretching will eventually cause necking in the specimen when the ultimate tensile strength S_u is reached. Until necking occurs, the relationship between true stress and strain and engineering stress and engineering strain can be given by the following relationships:

$$\varepsilon = \ln(1 + e)$$

$$\sigma = s(1 + e) \quad (2.9)$$

The true fracture strength of the material, σ_f , is given by the stress experienced by the material at final fracture:

$$\sigma_f = \frac{P_f}{A_f} \quad (2.10)$$

where P_f and A_f are respectively the load and cross sectional area at fracture. The true fracture strain is the true strain at final fracture and is given by:

$$\varepsilon_f = \ln\left(\frac{A_0}{A_f}\right) \quad (2.11)$$

A plot of both the engineering and true stress against engineering and true strain is shown in Figure 2.9. The difference between the engineering and true stress-strain curves comes from the fact that stress and strain are defined differently as explained above.

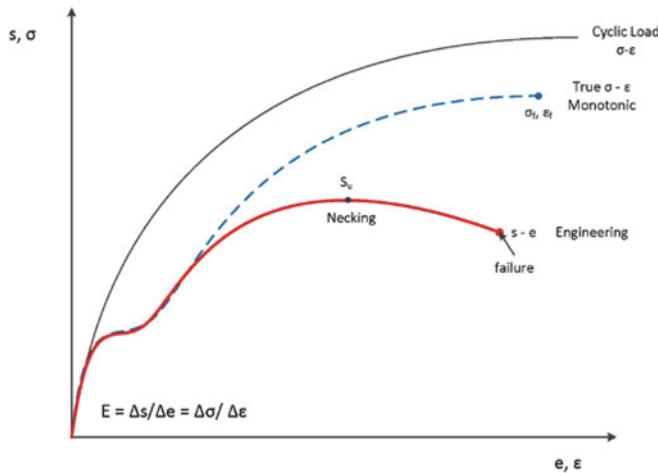


Figure 2.9: Typical monotonic and cyclic load stress-strain curves

The modulus of elasticity, E , represents the slope of the stress-strain curve in elastic region (the initial linear part of the curve). Elastic strain is given by Hooke's Law:

$$\varepsilon_e = \frac{\sigma}{E} \quad (2.12)$$

The plastic region of the stress-strain curve can often be sufficiently described by an empirically derived power equation as follows:

$$\sigma = K(\varepsilon_p)^n \quad (2.13)$$

where K is the strength coefficient and n is the strain hardening exponent.

The total true strain is the sum of the elastic and plastic strains as follows:

$$\varepsilon_{tot} = \varepsilon_e + \varepsilon_p \quad (2.14)$$

Therefore, the monotonic stress-strain relationship is given by:

$$\varepsilon_{tot} = \frac{\sigma}{E_{elastic}} + \left(\frac{\sigma}{K_{plastic}} \right)^{\frac{1}{n}} \quad (2.15)$$

2.2.2.2. Cyclic Stress-Strain Behavior

If the loading process described in section 2.2.1.1, and shown again in Figure 2.10(a), is cycled through a reversing process whereby the specimen is unloaded after yielding, the stress-strain relationship will follow a slope equivalent to the elastic modulus as shown in Figure 2.10(b). Massing's hypothesis states that materials behave similarly in tension and compression. Therefore, the hysteresis curve (shown in Figure 2.10(c)) can be obtained by doubling the values of the stress strain curve in Figure 2.10(a). This is further illustrated in Figure 2.11. Here, the loading process is continued in order to result in a stabilized hysteresis loop.

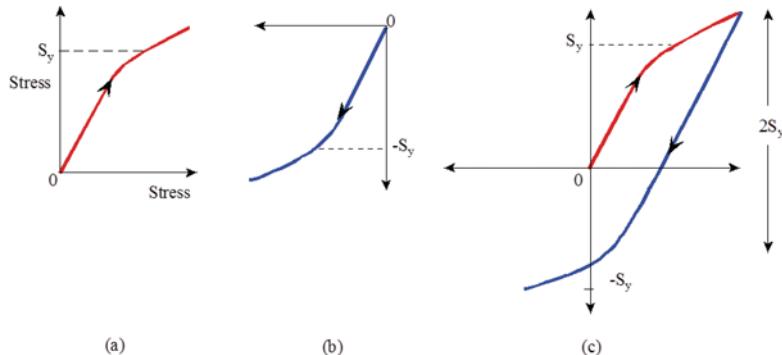


Figure 2.10: Cyclic stress-strain behavior

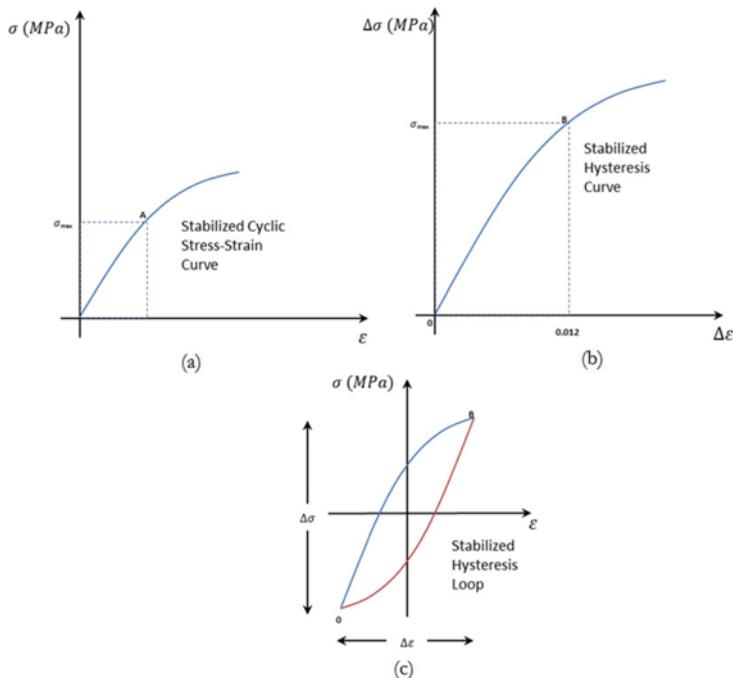


Figure 2.11: Massing's hypothesis showing how stabilized hysteresis loop is obtained

The cyclic stress-strain curve equation is the same as for the monotonic case, except the power law constants are updated to reflect the cyclic coefficient and exponent, as follows:

$$\varepsilon_{tot} = \underbrace{\frac{\sigma}{E}_{elastic}} + \underbrace{\left(\frac{\sigma}{K'}\right)^{\frac{1}{n'}}}_{plastic} \quad (2.16)$$

where K' is the cyclic strength coefficient, and n' is the cyclic hardening exponent.

From Massing's hypothesis, the hysteresis curve equation is given by:

$$\Delta\varepsilon = \underbrace{\frac{\Delta\sigma}{E}_{elastic}} + 2 \underbrace{\left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}}}_{plastic} \quad (2.17)$$

Note that the above equation is only valid for materials that exhibit symmetric behavior in tension and compression. Materials that do not exhibit the same response under tension and compression should not have the strain-life methodology applied to them. This symmetric behavior of material is known as the Bauschinger effect. The absolute value of the yield stress in compression is lowered by the previous application of tension load (Bannantine, et al., 1997). Figure 2.12 illustrates the Bauschinger effect.

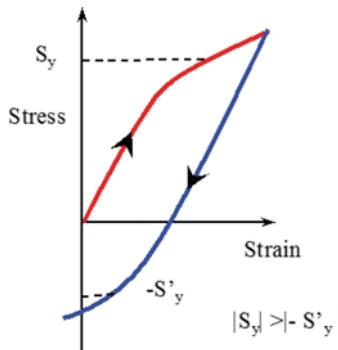


Figure 2.12: Bauschinger effect

2.2.2.3. Strain-Life Relationship

Both the elastic and plastic components of the cyclic strain can be modeled by an empirical power relationship as follows:

$$\varepsilon_e = \frac{\sigma_a}{E} = \frac{\sigma_f'}{E} (2N_f)^b \quad (2.18)$$

where σ_f' is the fatigue strength coefficient (y-intercept) and b is the fatigue strength exponent (slope). Traditionally, $2N_f$ instead of N_f is used in the formulation of strain-life approach; therefore, we follow the traditional notation.

$$\varepsilon_p = \varepsilon_f' (2N_f)^c \quad (2.19)$$

where ε'_f is the fatigue ductility coefficient (y-intercept) and c is the fatigue ductility exponent (slope). This is known as the Coffin-Manson relationship, proposed in the early 1960s to represent the relationship between plastic strain and life of a component (Manson, 1953; Coffin, 1954).

Thus, the strain-life curve can be obtained by combining the elastic and plastic components, as follows:

$$\varepsilon_{tot} = \frac{\sigma'_f}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c \quad (2.20)$$

A typical strain-life curve (in red) is illustrated in Figure 2.13. The elastic and plastic components influence the shape of the curve, which can be clearly seen. The transition life, $2N_t$, is the life corresponding to the intercept between the elastic and plastic strain regression lines, and the point at which failure transitions from being dominated by plastic strain to being dominated by elastic strain. The algebraic relationship for the transition life is:

$$2N_t = \left(\frac{\varepsilon'_f E}{\sigma'_f} \right)^{\frac{1}{b-c}} \quad (2.21)$$

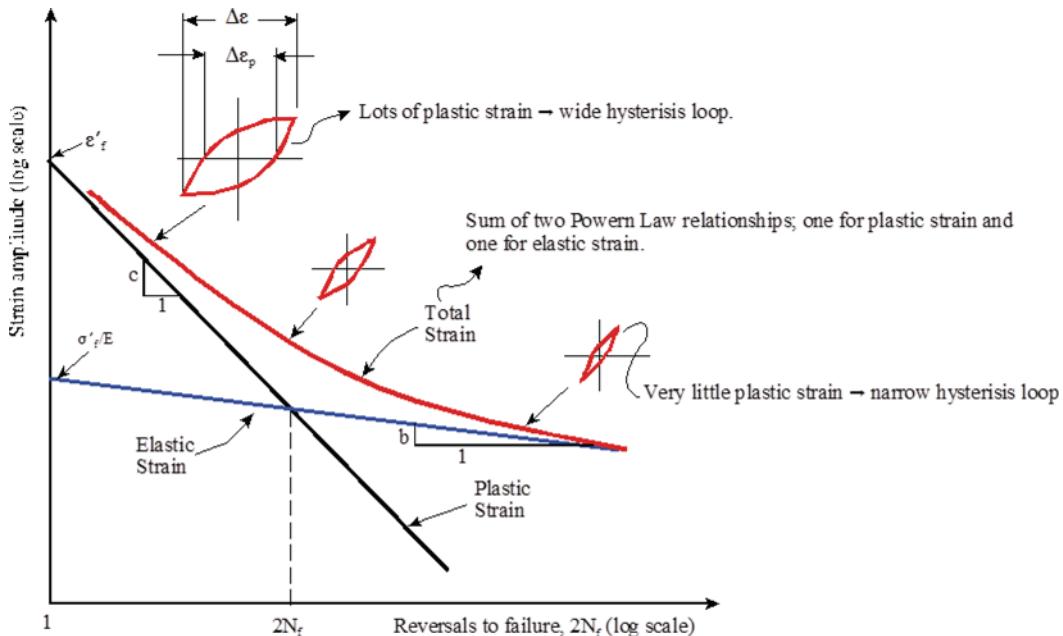


Figure 2.13: Typical strain-life curve

Elastic strain has a greater influence on the curve for fatigue life above the transition life, and plastic strain has a greater influence below the transition life. Therefore, it is often convenient to use the transition life as the boundary between low and high cycle fatigue. As mentioned previously, this chart also illustrates generally that plastic strain has negligible effect at higher cycles and the strain-life and life-stress approaches essentially arrive at the same result. It should be noted that the strain life graph plots the strain against reversals to failure as opposed to cycles to failure. Two reversals equal one cycle.

2.2.2.4. Mean Stress Effects

In terms of the strain-life relationship, three corrections are commonly used to account for the effects of mean stress. As mentioned in section 2.2.1.2, tensile mean stresses generally reduce fatigue life, while compressive mean stresses generally extend life. The three common corrections used are as follows:

- Morrow Mean Stress Correction

$$\varepsilon_{tot} = \frac{\Delta\varepsilon}{2} = \frac{\sigma'_f - \sigma_m}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c \quad (2.22)$$

where the mean stress σ_m is positive for tensile stresses and negative for compressive stresses. Morrow's mean stress correction is shown graphically in Figure 2.14, which is consistent with the observation that mean stress has a significant effect at long fatigue life with low plastic strains and negligible effect on short fatigue life.

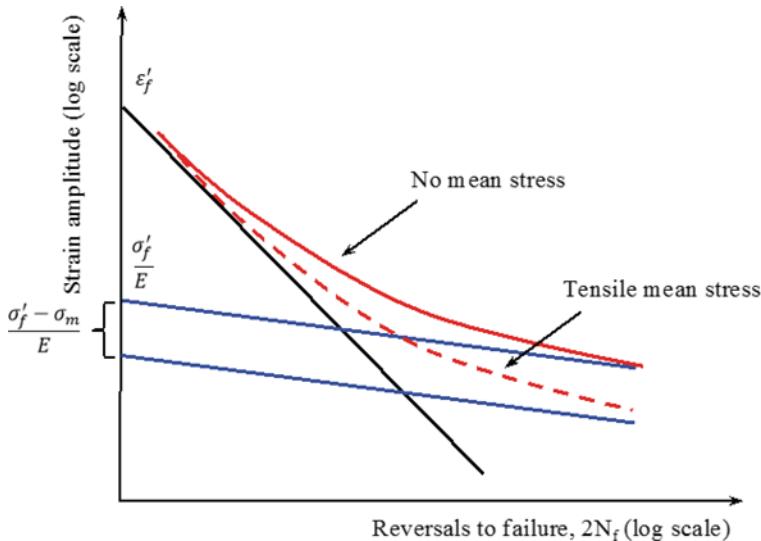


Figure 2.14: Morrow's mean stress correction to the strain-life curve for a tensile mean stress (Bannantine, et al., 1997)

- Modified Morrow Mean Stress Correction. This method corrects for the Morrow method with the assumption that the ratio of elastic to plastic strain is dependent on mean stress. Thus, the elastic and plastic terms of the strain-life equation are modified to maintain the independence of the elastic-plastic strain ratio from stress.

$$\varepsilon_{tot} = \frac{\Delta\varepsilon}{2} = \frac{\sigma'_f - \sigma_m}{E} (2N_f)^b + \varepsilon'_f \left(\frac{\sigma'_f - \sigma_m}{\sigma'_f} \right)^{c/b} (2N_f)^c \quad (2.23)$$

- Smith-Watson-Topper Mean Stress Correction. Often called the SWT parameter, this method is based on strain-life test data obtained with various mean stresses:

$$\sigma_{max} \frac{\Delta\varepsilon}{2} = \frac{(\sigma'_f)^2}{E} (2N_f)^{2b} + \sigma'_f \varepsilon'_f (2N_f)^{b+c} \quad (2.24)$$

where $\sigma_{max} = \sigma_m + \sigma_a$ and $\Delta\varepsilon$ is the alternating strain. Note that the SWT correction method assumes that no fatigue damage will occur when the maximum stress is equal to or

less than zero (i.e. compressive), which is not always true. For this reason, the Morrow or Modified Morrow correction method should be used for loading sequences that are predominantly compressive, and the SWT method should be used for those that are predominantly tensile.

Example 2.3: Strain-Life Approach to Fatigue Analysis

Consider a test specimen with the following material properties:

- Modulus of elasticity, $E = 210$ GPa
- Cyclic strain hardening component, $n' = 0.2$
- Cyclic strength coefficient, $K' = 1.2$ GPa

The specimen is subjected to a fully reversed cyclic strain with a strain range, $\Delta\varepsilon$, of 0.05. The stress-strain response of the material needs to be determined.

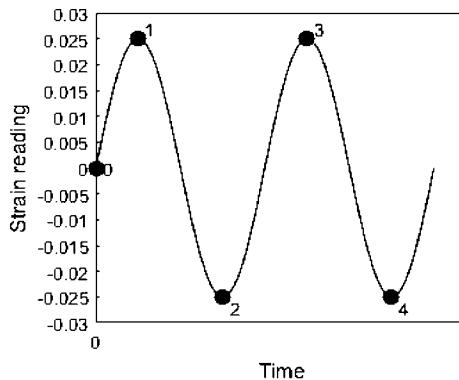


Figure 2.15: Strain history

Solution

The strain history for the specimen is shown in Figure 2.15. On the initial application of strain (point 1), the material response follows the cyclic stress strain curve equation, as follows (see section 2.2.2.2):

$$\varepsilon_1 = \frac{\sigma_1}{E} + \left(\frac{\sigma_1}{K'} \right)^{\frac{1}{n'}}$$

Substituting given values (with a strain value of 0.02) gives:

$$0.025 = \frac{\sigma_1}{210,000 \text{ MPa}} + \left(\frac{\sigma_1}{1,200 \text{ MPa}} \right)^{\frac{1}{0.2}}$$

Solving for σ_1 by iteration yields:

$$\sigma_1 \cong 561 \text{ MPa}$$

The cyclic stress-strain curve is used only for the initial application of strain. On all subsequent strain reversals, the material response is modeled using the hysteresis curve equation:

$$\Delta\varepsilon = \underbrace{\frac{\Delta\sigma}{E}_{elastic}}_{\text{elastic}} + 2 \underbrace{\left(\frac{\Delta\sigma}{2K'} \right)^{\frac{1}{n'}}}_{\text{plastic}}$$

Substituting the material properties into the above equation with a strain range $\Delta\varepsilon = 0.05$ yields:

$$0.05 = \frac{\Delta\sigma}{210,000} + 2 \left(\frac{\Delta\sigma}{2 \times 1,200} \right)^{\frac{1}{0.2}}$$

Solving the above for $\Delta\sigma$ through iteration yields:

$$\Delta\sigma = 1,122 \text{ MPa}$$

The stress and strain values corresponding to point 2 in Figure 2.15 can now be determined by subtracting the changes in stress and strain ($\Delta\sigma, \Delta\varepsilon$) from the values at point 1 (σ_1, ε_1).

$$\varepsilon_2 = \varepsilon_1 - \Delta\varepsilon = 0.025 - 0.05 = -0.025$$

$$\sigma_2 = \sigma_1 - \Delta\sigma = (561 - 1,122) \text{ MPa} = -561 \text{ MPa}$$

The stress and strain values corresponding to point 3 are similarly determined using the hysteresis curve equation.

$$\varepsilon_3 = \varepsilon_2 - \Delta\varepsilon = -0.025 - (-0.05) = 0.025$$

$$\sigma_3 = \sigma_2 - \Delta\sigma = (-561 - (-1,122)) \text{ MPa} = 561 \text{ MPa}$$

The solution for point 3 shows that the material response returns (as expected) to point 1. Therefore, the material response forms a closed stabilized hysteresis loop and all successive strain cycles would follow this loop. This stabilized stress-strain response is illustrated in Figure 2.16.

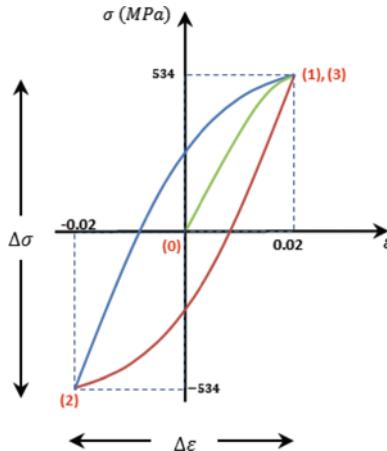


Figure 2.16: Stress-strain response for given example

Example 2.4: Mean Stress Correction Method

A test specimen of the material given in Example 2.3 has the following material properties:

- Fatigue strength coefficient, $\sigma'_f = 1 \text{ GPa}$
- Fatigue ductility coefficient, $\varepsilon'_f = 1.1$
- Fatigue strength exponent, $b = -0.1$
- Fatigue ductility exponent, $c = -0.6$

The fatigue life of the specimen needs to be estimated based on:

- It being subjected to the strain range plotted in Figure 2.17a; and
- It being subjected to strain range plotted in Figure 2.17b (using the SWT mean stress correction model).

Solution

- As the mean stress is zero, the strain-life prediction model presented in section 2.2.2 is used:

$$\varepsilon_{tot} = \frac{\sigma'_f}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c$$

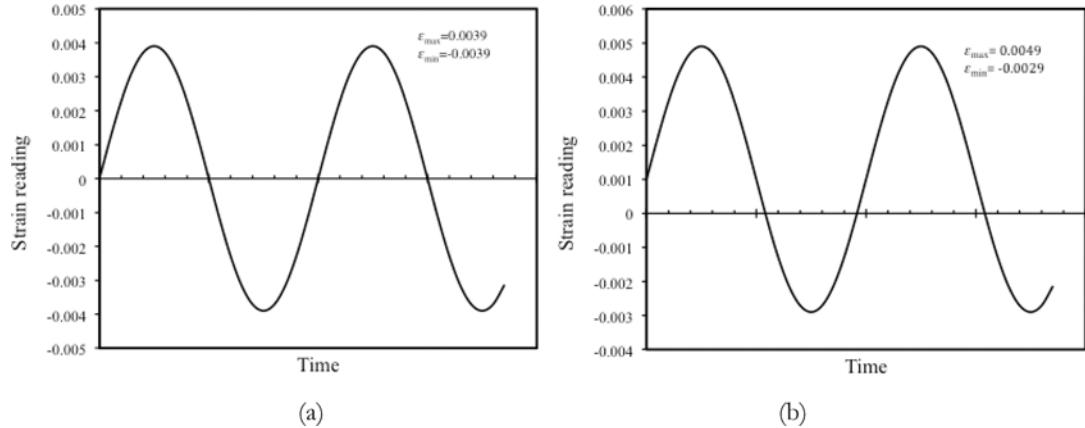


Figure 2.17: Strain reading: a) zero mean stress, b) tensile mean stress

Substituting the material properties into the above equation with a strain amplitude $\varepsilon_{tot} = \Delta\varepsilon/2 = 0.0039$ yields:

$$0.0039 = \frac{1000}{210,000} (2N_f)^{-0.1} + 1.1(2N_f)^{-0.6}$$

Solving the above for $2N_f$ through iteration yields:

$$2N_f = 31,293 \text{ reversals}$$

- On the initial application of maximum strain, the material response follows the cyclic stress strain curve equation (as presented in section 2.2.2.2) as follows:

$$\varepsilon_{max} = \frac{\sigma_{max}}{E} + \left(\frac{\sigma_{max}}{K'}\right)^{\frac{1}{n'}}$$

Substituting given values (with a strain value of 0.0049) gives:

$$0.0049 = \frac{\sigma_{max}}{210,000} + \left(\frac{\sigma_{max}}{1200}\right)^{\frac{1}{0.2}}$$

Solving for σ_{max} by iteration yields:

$$\sigma_{max} \cong 378 \text{ MPa}$$

Minimum stress can be evaluated in a similar way:

$$\sigma_{min} \cong -321 \text{ MPa}$$

Mean stress is:

$$\sigma_0 = -321 + [378 - (-321)]/2 = 28.5 \text{ MPa}$$

The strain range is evaluated as:

$$\frac{\Delta\varepsilon}{2} = \frac{\varepsilon_{max} - \varepsilon_{min}}{2} = \frac{0.0049 - (-0.0029)}{2} = 0.0039$$

Using Morrow's mean stress correction model presented as:

$$\frac{\Delta\varepsilon}{2} = \frac{\sigma'_f - \sigma_0}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c$$

and substituting the material properties into the above equation yields:

$$0.0039 = \frac{(1000 - 28.5)}{210,000} (2N_f)^{-0.1} + 1.1(2N_f)^{-0.6}$$

Solving the above for $2N_f$ through iteration yields:

$$2N_f = 30,305 \text{ reversals}$$

Using modified Morrow's mean stress correction model gives:

$$2N_f = 29,031 \text{ reversals}$$

Using the SWT mean stress correction model amends the estimated life to:

$$2N_f = 27,538 \text{ reversals}$$

It can be seen that the mean stress reduces the fatigue life.

2.2.3. VARIABLE AMPLITUDE LOADING

The discussion provided in the previous sections dealt with cases where the stresses involved are constant amplitude. However, most practical applications deal with variable amplitude loads. For example, an aircraft experiences a spectrum of variable amplitude loading emanating from a diverse array of operating modes from taxiing to landing. In fatigue analysis, methods for damage summation and cycle counting (based on loading history) are required.

Figure 2.18 illustrates a variable amplitude loading consisting of "blocks" of constant amplitude load cycles. Many variable amplitude loading scenarios would be much more complicated than this; however, it conveniently sets the stage for explaining the concepts required for analyzing variable amplitude loading.

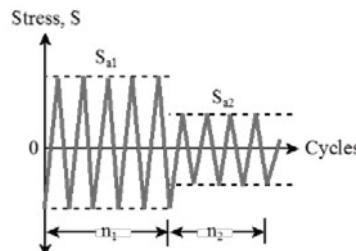


Figure 2.18: Variable amplitude loading consisting of constant amplitude "blocks"

This is because most of the methods used to analyze variable amplitude loading involve taking the variable amplitude loading history and extracting constant amplitude blocks to sum the damage accumulated by the item under test. The most common theory used to support fatigue analysis involving variable amplitude loading is the Palmgren-Miner linear damage rule. Consider the S-N plot in Figure 2.19.

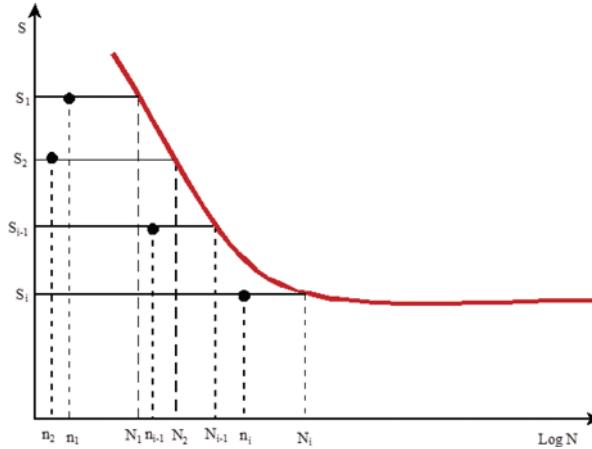


Figure 2.19: S-N plot showing spectrum of loading where n_i cycles are accrued at each of the different corresponding stress levels S_i , and the N_i are cycles to failure at each S_i .

The basic assumptions of the linear damage rule are as follows:

- Given an S-N plot, operation at a constant stress amplitude, S , will produce complete damage or failure in N cycles.
- Operation at a stress level less than this value will produce a smaller *damage fraction*.
- Operation over a spectrum of different stress levels results in a damage fraction for each of the different stress levels in the spectrum. Failure is then predicted when these damage fractions sum to unity. The linear damage theory assumes that the damage fraction at any stress level is linearly proportional to the ratio of the number of cycles to the number of cycles that would produce failure at that stress level. Therefore, the theory predicts that failure will occur if:

$$\frac{n_1}{N_{f1}} + \frac{n_2}{N_{f2}} + \dots = \sum \frac{n_i}{N_{fi}} \geq 1 \quad (2.25)$$

where n_i is the number of cycles of operation at stress level of S_i , and N_{fi} is the number of cycles to failure at constant stress level S_i .

The major disadvantage of the linear damage theory is that significant influences are not accounted for, potentially affecting failure prediction. Specifically, linear damage theory assumes the order of application of various stress levels has no influence, with damage accumulating at the same rate at any given stress level regardless of the past (known as the “sequence effect”). In reality, the sequence effect may have a significant influence, with the damage rate a function of prior cycle stress history. Nonetheless, the linear damage theory is often used due to its simplicity and the convenience it provides in calculating an initial failure prediction. For example, when a sequence of variable amplitude loading is repeated, it is convenient to analyze each sequence and then multiply the effect of the sequence for each repetition. Cycle ratios are summed over each sequence repetition, which will predict failure (in terms of repetitions) when they multiply to unity (Dowling (1998), Sec. 9.9.1):

$$B_f \left[\sum \frac{n_i}{N_{f_i}} \right]_{\text{one rep.}} = 1 \quad (2.26)$$

where B_f is the number of repetitions to failure.

Each load level contributes a proportion of damage as determined by:

$$D_i = B_f \times \frac{n_i}{N_{f_i}} \quad (2.27)$$

In determining the number of repetitions using the above equations, one must be consistent with the units used. Note the units of one repetition (e.g. 5 cycle repetitions, 10 hour repetition) and the units of cycles (e.g. hours, revolutions)

There are some non-linear damage theories that have been proposed which attempt to overcome some of the pitfalls of the linear damage theory discussed above. Section 2.2.3.1 discusses some of the non-linear damage models.

Example 2.5

Consider a modified version of the test of ball bearings discussed in www.efatigue.com that showed the following expected life at 2 variable amplitude loads:

- 3×10^8 cycles under 2 kN load
- 4×10^7 cycles under 3 kN load

The variable amplitude load applied in the test is assumed to be 2 kN 70% of the time and 3 kN 30% of the time. To calculate the number of cycles that the bearing is expected to last, Palmgren Miner's rule is used to determine an expression for the damage accumulated as a result of the given variable amplitude loads. The total applied cycles to failure is n , meaning that the number of cycles at the 2 kN and 3 kN loads are $n_1 = 0.7n$ and $n_2 = 0.3n$ respectively. The total damage accumulated is given by:

$$\frac{0.7n}{3 \times 10^8} + \frac{0.3n}{4 \times 10^7} = 9.83 \times 10^{-9}n$$

The number of cycles expected for failure to occur is then given by Equation (2.26) and therefore:

$$n = \frac{1}{9.83 \times 10^{-9}} = 1 \times 10^8 \text{ cycles}$$

Example 2.6

A smooth specimen made from material with $\sigma'_f = 1,500$ MPa and $b = -0.1$ is subjected to a stress test of repeated blocks of axial stress. Each stress block consists of 200 reversals of -550 MPa to 550 MPa, followed by 1000 reversals of 0 to 690 MPa, followed by 100 reversals of -690 to 0 MPa. Determine the number of blocks, B_f , of this load that can be applied before expecting a crack initiation failure.

Solution

The number of reversals to fail for the stress levels of each of the three stress loading blocks as follows need to be determined.

- (1) Stress loading block of -550 to 550 MPa:

Using the relationship $\sigma_a = \sigma_f' (2N_f)^b$, one obtains from the material properties:

$$550 = 1,500(2N_f)^{-0.1}$$

Solving for $2N_f$ yields:

$$2N_f = 22,766 \text{ reversals}$$

- (2) Stress loading block of 0 to 690 MPa:

Here, the mean stress effects need to be determined to calculate the effective stress at this loading block. Using Morrow's mean stress correction relationship,

$$\frac{\sigma_a}{\sigma_{ar}} + \frac{\sigma_m}{\sigma_f} = 1$$

where

- $\sigma_a = \frac{\Delta\sigma}{2} = 345 \text{ MPa}$
- $\sigma_m = \frac{\sigma_{min} + \sigma_{max}}{2} = 345 \text{ MPa}$
- $\sigma_f = 1,500 \text{ MPa}$ (given)

then:

$$\frac{\sigma_a}{\sigma_{ar}} + \frac{\sigma_m}{\sigma_f} = \frac{345}{\sigma_{ar}} + \frac{345}{1,500} = 1$$

$$\text{Solving for } \sigma_{ar} = \frac{\sigma_a}{1 - \frac{\sigma_m}{\sigma_f}} = 448 \text{ MPa}$$

Therefore, the effective stress at this stress loading block is 448 MPa. Applying the same relationship as in part 1, we get:

$$448 = 1,500(2N_f)^{-0.1}$$

Solving for $2N_f$ yields:

$$2N_f = 176,860 \text{ reversals}$$

- (3) Stress loading block of -690 to 0 MPa:

Again, one needs to account for mean stress effects, as in part (2). This time,

- $\sigma_a = \frac{\Delta\sigma}{2} = 345 \text{ MPa}$
- $\sigma_m = \frac{\sigma_{min} + \sigma_{max}}{2} = -345 \text{ MPa}$
- $\sigma_f = 1,500 \text{ MPa}$ (given)

$$\text{Solving for } \sigma_{ar} = \frac{\sigma_a}{1 + \frac{\sigma_m}{\sigma_f}} = 281 \text{ MPa}$$

Therefore,

$$281 = 1,500(2N_f)^{-0.1}$$

Solving for $2N_f$ yields:

$$2N_f = 19,132,545 \text{ reversals}$$

Now that the reversals to failure based on a continuous application of each stress loading block has been determined, one can determine the number of blocks to failure using Palmgren-Miner's rule as follows:

$$B_f \left[\sum \frac{n_i}{N_{f_i}} \right]_{\text{one rep.}} = 1$$

Therefore,

$$B_f \left[\frac{200}{22,766} + \frac{1,000}{176,860} + \frac{100}{19,132,545} \right] = 1$$

$$B_f \cong 69 \text{ repetitions}$$

2.2.3.1. Non-Linear Damage Models

Non-linear damage models attempt to account for the non-linearity of fatigue damage that the linear models cannot. One such example is known as the Marco-Starkey non-linear damage rule. The theory introduces an exponential term, m (a function of the stress level), to augment the Palmgren-Miner rule, as follows:

$$D_i = \left(\frac{n_i}{N_i} \right)^m ; 0 < m \leq 1 \quad (2.28)$$

Note that with $m = 1.0$, the above rule becomes equivalent to Miner's rule. As explained previously, a specimen subjected to any sequence of completely reversed sinusoidal stresses will fail when D reaches 1. This nonlinear method reasonably correlates to observed material behavior, particularly in high temperature applications where creep interacts with fatigue (Bannantine, et al., 1997). However, the Palmgren-Miner and other nonlinear rules do not incorporate many of the factors that complicate damage estimation in complex variable amplitude loading scenarios.

2.2.4. NOTCH EFFECT

The previous sections covered the life-stress and strain-life approaches used in fatigue analysis. In practice, fatigue failures usually occur at stress concentrations or notches in the specimen under load. This section will cover the use of previously discussed methods of fatigue analysis to account for notch effect. In section 1.6.6, the fracture mechanics approach to notched specimens is discussed.

2.2.4.1. Life-stress

Almost all machined components and structural members contain some form of geometric discontinuities that result in stress concentrations. Examples include holes, grooves and fillets - all of which elevate local stress. The local maximum stress, σ_{max} , is greater than the nominal stress, S being applied to the specimen. Nominal stress is defined as the applied load divided by the nominal cross sectional area (area as if there is no notch). Maximum stress, on the other hand, is defined as the applied load divided by the reduced cross sectional area. For elastic members, the ratio of these stresses is known as the elastic stress concentration factor K_t and is given by:

$$K_t = \frac{\sigma_{max}}{S} \quad (2.29)$$

This factor is the ratio of the maximum local stress to the specimen's nominal stress. K_t solely depends on geometry of the specimen and the mode of loading (such as axial or in plane bending). Many fatigue textbooks list typical examples of elastic stress concentration curves that contain useful data for assisting in fatigue analysis of notched members. Some of these are listed below:

1. Fundamentals of Fatigue Analysis (Bannantine, et al., 1997) – Section 4.2, Figure 4.1, p. 125.
2. Failure of Materials in Mechanical Design (Collins, 1993) – Section 12.3, Figures 12.3 – 12.8, pp. 418-425.
3. Metal Fatigue in Engineering (Stephens, et al., 2001) – Section 7.1, Figures 7.3 & 7.4, pp. 191-192.

In addition to the textbooks above, there are many online sources such as www.efatigue.com which provide these data. The efatigue.com “fatigue calculator” tool is an online tool that allows engineers to quickly and easily conduct a fatigue or durability analysis for common basic fatigue scenarios. Simple step-by-step run-through examples are provided in this website to demonstrate the tool’s functionality.

In the Life-Stress approach, the effect of notches is accounted for by the *fatigue notch factor*, K_f (also known as the fatigue stress concentration factor). K_f relates the unnotched fatigue strength (which is the endurance limit for ferrous materials) to the notched fatigue strength of the member under analysis, as given by:

$$K_f = \frac{S_e^{(unnotched)}}{S_e^{(notched)}} \quad (2.30)$$

In almost all cases, the K_f is less than the elastic stress concentration factor K_t . Unlike K_t , K_f depends on the type of material and the notch size. The notch sensitivity factor, q , accounts for these effects as follows:

$$q = \frac{K_f - 1}{K_t - 1} \quad (2.31)$$

Notch sensitivity factor is a measure of the reduction in strength of a metal caused by the presence of a notch. The value of q ranges from 0 (implying that there is no notch effect) to 1 (implying full theoretical effect, $K_f = K_t$). A number of analytical relationships for determining q has been proposed based on empirical data. The two most common relationships are the Peterson and Neuber relationships, described below. The Peterson Equation for notch factor sensitivity is:

$$q = \frac{1}{1 + \frac{a}{r}} \quad (2.32)$$

where

- r = notch root radius
- a = material constant. This constant depends on the material strength and ductility and is obtained experimentally from long life fatigue tests on both notched and unnotched specimens. For relatively high strength steels, a is given by (Dowling, 1998):

$$a = 0.025 \left(\frac{2,070 \text{ MPa}}{S_u} \right)^{1.8} \text{ mm} \quad (S_u \geq 550 \text{ MPa})$$

where S_u is the ultimate strength. The Neuber Equation for notch factor sensitivity is:

$$q = \frac{1}{1 + \sqrt{\frac{\rho}{r}}} \quad (2.33)$$

where

- r = notch root radius
- ρ = material constant (which relates to the grain size of the material). For steels, ρ is given by (Dowling, 1998):

$$\log_{10} \rho = -\frac{S_u - 134 \text{ MPa}}{586} \quad (\text{mm}) \quad (S_u \leq 1,520 \text{ MPa})$$

Another use of the fatigue notch factor K_f is to correct the fatigue strength for the notched specimen, as follows:

$$S_{ar} = \frac{\sigma_{ar}}{K_f} \quad (2.34)$$

The fatigue notch factor for stresses corresponding to lives of 1000 cycles is K'_f . One can correct the S-N curve to account for a notched member using K'_f as shown in Figure 2.20. This method is known as the modified Juvinal approach.

Example 2.7

A notched steel component consists of a bar of width $W = 100$ mm, thickness $t = 5$ mm and two semi-circular edge notches with radii $r = 10$ mm (see Figure 2.21). This gives the plate a width at the reduced section of 80 mm. Determine the life of the component subjected to a fully reversed load with amplitude of $F = 150$ kN. The steel has ultimate strength S_u of 790 MPa and material constant a of 1.45×10^{-4} m. The ratio of fatigue stress concentration factor at 1,000 cycles is 0.3.

Solution

First, the theoretical stress concentration factor K_t is determined. From the given dimensions, this can be achieved by using the efatigue online tool. This results in a calculated stress concentration factor of $K_t = 2.41$. (see <https://www.efatigue.com/constantamplitude/stressconcentration>).

Next, the notch sensitivity factor, q , is determined using the following relationship (discussed in the previous section):

$$q = \frac{1}{1 + \frac{a}{r}} = \frac{1}{1 + \frac{1.45 \times 10^{-4}}{0.01}} = 0.9857$$

The fatigue stress concentration factor is given by:

$$K_f = 1 + q(K_t - 1) = 1 + 0.9857(2.41 - 1) = 2.39$$

Recall the ultimate strength value for the material is $S_u = 790$ MPa. The endurance limit is related to the ultimate strength by:

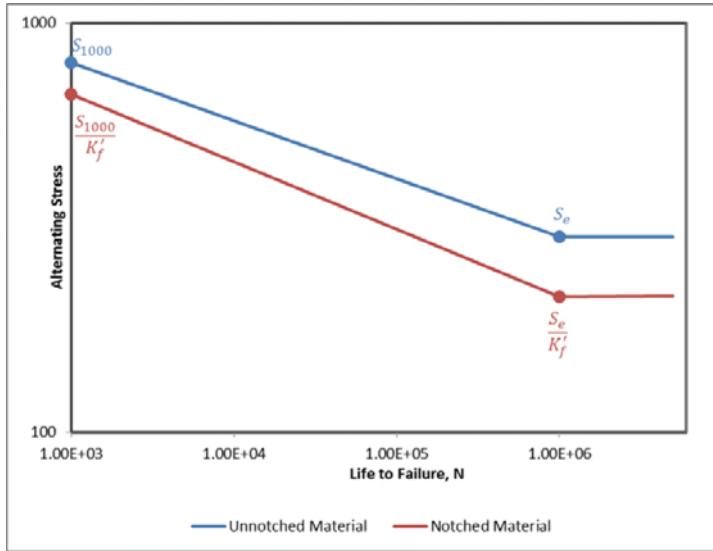


Figure 2.20: Juvinal method for modification for the S-N curve for notched components

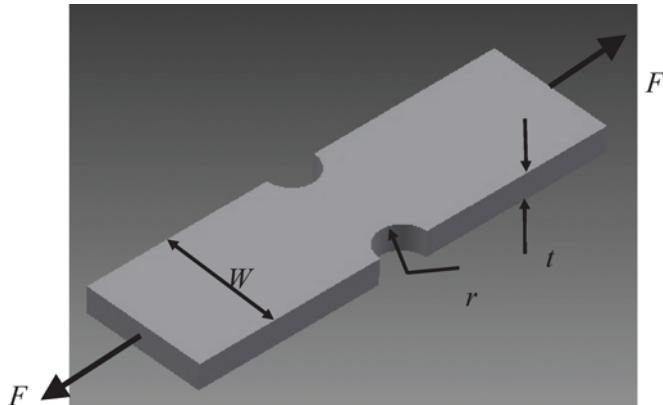


Figure 2.21: Notched bar configuration for Example 2.7

$$S_e = \frac{S_u}{2} = 395 \text{ MPa}$$

The alternating stress level corresponding to a life of 1,000 cycles, S_{1000} , is:

$$S_{1000} = 0.9 \times S_u = 711 \text{ MPa}$$

The fatigue stress concentration factor at 1,000 cycles, K_f' , is determined using the relationship (Bannantine, et al., 1997):

$$\frac{K_f' - 1}{K_f - 1} = 0.3$$

Therefore,

$$K_f' = (2.39 - 1)(0.3) + 1 = 1.417$$

Next, the design S-N curve is constructed by locating the following points on log-log coordinates.

Life (cycles)	Stress (MPa)
1×10^6	S_e/K_f
1×10^3	S_{1000}/K'_f
1	$S_u + 345$ (MPa)

$$\text{At } 10^6 \text{ cycles: Stress} = \frac{S_e}{K_f} = \frac{395 \text{ MPa}}{2.39} \approx 165 \text{ MPa}$$

$$\text{At } 10^3 \text{ cycles: Stress} = \frac{S_{1000}}{K'_f} = \frac{711 \text{ MPa}}{1.417} \approx 502 \text{ MPa}$$

$$\text{At 1 cycle: Stress} = S_u + 345 = 1,135 \text{ MPa}$$

The S-N curve is extended into the range 1 to 1,000 cycles for comparison purposes only and is not recommended for general design purposes. The estimate at one cycle is the true fracture strength, σ_f . For steels, this is approximately the ultimate strength plus 345 MPa (50 ksi) [See Chapt 1 reference Bannantine, et al. (1997)], sec 4.2). These points are plotted on log-log coordinates as shown in Figure 2.22. The net section stress (nominal stress) in the component is determined as follows:

$$S_{net} = \frac{P_a}{A_{net}} = \frac{150 \text{ kN}/10^6}{0.005 \times 0.08 \text{ m}^2} = 375 \text{ MPa}$$

The number of cycles to failure, N_f , is determined by locating the life corresponding to the net section stress on the design S-N curve. The life corresponding to 375 MPa on the S-N curve is obtained by solving the following line of best fit equation for N_f :

$$375 \text{ MPa} = 1527.3(N_f)^{-0.161}$$

Therefore:

$$N_f \approx 6140 \text{ cycles}$$

2.2.4.2. Strain-Life

An understanding of the local stress-strain behavior at the root of a notch is required in order to use the strain-life fatigue method of analysis. There are a number of methods for determining local stress-strain behavior, such as:

1. Strain gauge measurements
2. Finite element analysis
3. Methods that relate local stresses and strains

Of the three methods above, the third is the simplest (in terms of time and cost). This method involves using the stress concentration factor, K_t , or the fatigue notch factor, K_f . It relates the stress-strain field in the immediate vicinity of a notch to the remote stresses and strains determined from fatigue tests of smooth specimens. Consider Figure 2.23, which shows a specimen under tensile stress where the maximum stress concentrations are at the edges of a hole. If conditions at both the nominal and local regions remain purely elastic, then the theoretical elastic stress

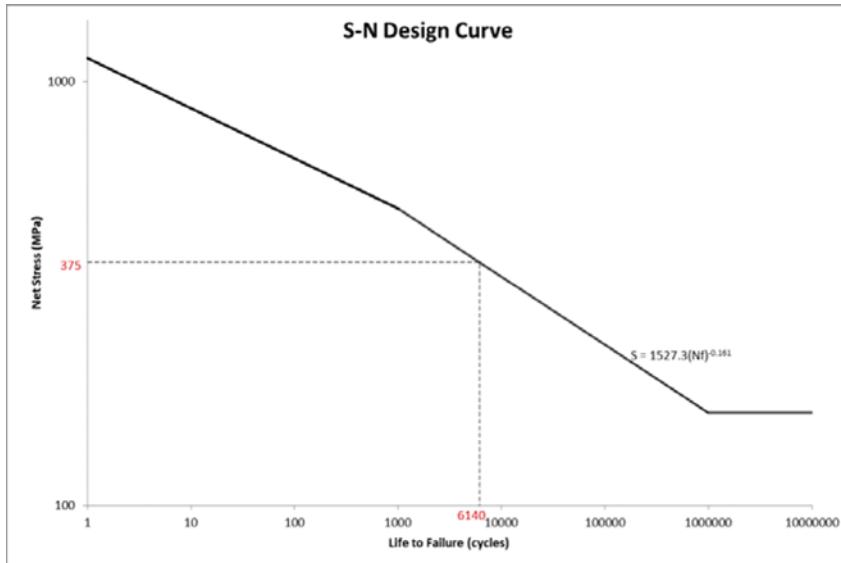


Figure 2.22: S-N design curve for Example 2.7

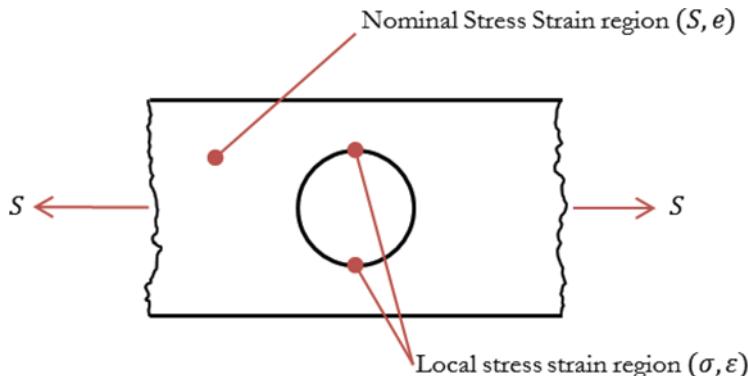


Figure 2.23: Nominal and local stress-strain regions

concentration factor, K_t , is equal to both the local stress concentration factor, $K_\sigma = \frac{\sigma}{S}$, and the local strain concentration factor $K_\varepsilon = \frac{\varepsilon}{e}$. That is,

$$K_t = K_\sigma = K_\varepsilon \quad (2.35)$$

However, once the material yields at the tip of the notch, the stress concentration factor K_σ decreases with respect to K_t . For plastic deformation, Neuber proposed that the elastic stress concentration factor K_t can be approximated by the geometric mean of the stress and strain concentration factors K_σ and K_ε . Neuber's Rule is thus given by:

$$K_t = \sqrt{K_\sigma \times K_\varepsilon} \quad (2.36)$$

Note that the use of K_t in the application of Neuber's rule gives conservative life estimates for component failure. A modified approach using fatigue notch factor K_f is recommended for life predictions corresponding to component failure, as follows:

$$K_f = \sqrt{K_\sigma \times K_\varepsilon} = \sqrt{\frac{\sigma}{S} \times \frac{\varepsilon}{e}} \quad (2.37)$$

Rearranging the relationship above yields:

$$K_f^2 \times S \times e = \sigma \varepsilon \quad (2.38)$$

Combining the above equation with the cyclic stress strain equation presented in section 2.2.2.2 yields:

$$K_f^2 \times S \times e = \sigma \left[\underbrace{\frac{\sigma}{E}_{elastic}}_{\text{elastic}} + \underbrace{\left(\frac{\sigma}{K'} \right)^{\frac{1}{n'}}}_{\text{plastic}} \right] \quad (2.39)$$

To account for load reversals, Neuber's rule has been extended as follows:

$$\begin{aligned} K_f^2 \times \Delta S \times \Delta e &= \Delta \sigma \Delta \varepsilon \\ K_f^2 \times \Delta S \times \Delta e &= \Delta \sigma \left[\frac{\Delta \sigma}{E} + \left(\frac{\Delta \sigma}{K'} \right)^{\frac{1}{n'}} \right] \end{aligned} \quad (2.40)$$

where

- Δe = nominal strain range
- $\Delta \varepsilon$ = local strain range
- ΔS = nominal stress range
- $\Delta \sigma$ = local stress range

The Neuber method is commonly used to estimate notch root elastic-plastic stresses and strains on the basis of far-field elastic nominal stresses and strains. For fully elastic behavior in the nominal region, nominal stresses and strains are related by:

$$e = \frac{S}{E} \quad (2.41)$$

Given this relationship, Neuber's rule can be rewritten as:

$$\frac{(K_f S)^2}{E} = \sigma \varepsilon \quad (2.42)$$

The true stress and strain can be approximated by the nominal stress and strain for limited yielding in the nominal region:

$$e = \frac{S}{E} + \left(\frac{S}{K'} \right)^{\frac{1}{n'}} \quad (2.43)$$

From above, Neuber's rule can be expressed as:

$$K_f^2 \times S \times \left[\frac{S}{E} + \left(\frac{S}{K'} \right)^{\frac{1}{n'}} \right] = \sigma \varepsilon \quad (2.44)$$

Example 2.8

Redo Example 2.7, using Neuber analysis, given the following material properties:

Modulus of elasticity, $E = 210$ GPa

$S_y = 690$ MPa

$K' = 1,060$ MPa

$n' = 0.14$

$\sigma'_f = 1,160$ MPa

$b = -0.081$

$\varepsilon'_f = 1.1$

$c = -0.65$

Solution

Strain life fatigue concepts are needed to determine the expected number of cycles to induce a small crack of approximately 1 mm in length. The first step is to determine the notch root stress-strain behavior using a Neuber analysis. Since

$$K_t S_{net} = 2.41 \times 375 = 903.75 \text{ MPa} > S_y (= 690 \text{ MPa})$$

the notch root behavior is inelastic. For monotonic load:

$$\sigma \varepsilon = \frac{(K_f S)^2}{E} = \frac{(2.39 \times 375)^2}{210,000} = 3.83$$

and plastic strain:

$$\varepsilon = \frac{\sigma}{E} + \left(\frac{\sigma}{K'} \right)^{1/n'}$$

Therefore,

$$\begin{aligned} \frac{\sigma}{E} + \left(\frac{\sigma}{K'} \right)^{1/n'} &= \frac{3.83}{\sigma} \\ \frac{3.83}{\sigma} &= \frac{\sigma}{210,000} + \left(\frac{\sigma}{1,060} \right)^{1/0.14} \end{aligned}$$

Solving for σ yields:

$$\sigma = 507 \text{ MPa}$$

And substituting σ to determine ε :

$$\varepsilon = \frac{3.83}{507} = 0.00755$$

Therefore, this analysis implies that at the root notch the maximum stress of 507 MPa corresponds to a maximum strain of 0.00755. At the compression point of -375 MPa (when cyclic load commences in hysteresis curve):

$$\Delta\varepsilon\Delta\sigma = \frac{(K_f\Delta S)^2}{E} = \frac{(2.39 \times 750)^2}{210,000} = 15.3$$

and:

$$\frac{\Delta\varepsilon}{2} = \frac{\Delta\sigma}{2E} + \left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}}$$

Therefore,

$$\frac{15.3}{2\Delta\sigma} = \frac{\Delta\sigma}{2 \times 210,000} + \left(\frac{\Delta\sigma}{2 \times 1,060}\right)^{1/0.14}$$

Solving for $\Delta\sigma$ and $\Delta\varepsilon$ yields:

$$\Delta\sigma = 1,013.5 \text{ MPa}$$

$$\Delta\varepsilon = 0.0151$$

Therefore, at the notch, the minimum stress is $\sigma = 507 - 1,013.5 = -506.5$ MPa and the minimum strain is $\varepsilon = 0.00755 - 0.0151 = -0.00755$.

With the notch root strain known, the fatigue life N_f can be determined using the strain life equation (SWT model) as follows:

$$\sigma_{max} \frac{\Delta\varepsilon}{2} = \frac{(\sigma'_f)^2}{E} (2N_f)^{2b} + \sigma'_f \varepsilon'_f (2N_f)^{b+c}$$

$$507 \times \frac{0.0151}{2} = \frac{(1,160)^2}{210,000} (2N_f)^{2(-0.081)} + 1,160 \times 1.1 (2N_f)^{-(0.081+0.65)}$$

Solving numerically, one obtains:

$$2N_f = 5,830 \text{ reversals}$$

Therefore,

$$N_f = 2,915 \text{ cycles}$$

Example 2.9

The fatigue failure of a steel bar is obtained by bending it back and forth 1,000 times daily for a period of 20 years. We wish to accelerate fatigue failure in a new steel bar of the same type by placing a notch in the bar. If the fatigue notch concentration factor is $K_t = K_f = 10x \text{ cm}^{-1}$, where x is the notch depth in centimeters, and the properties of the steel bar are:

$$b = -0.1$$

$$c = -0.5$$

$$\sigma'_f = 1,400 \text{ MPa}$$

$$E = 200,000 \text{ MPa}$$

$$\varepsilon'_f = 0.5$$

$$\begin{aligned}K &= 1,500 \text{ MPa} \\n &= 0.193\end{aligned}$$

Determine the minimum notch depth, x , for the new steel bar to fail in of two weeks using the same bending rate.

Solution

The number of cycles to failure for the original bar is given by:

$$N_f = 1,000 \frac{\text{cycles}}{\text{day}} \times 365 \frac{\text{days}}{\text{year}} \times 20 \text{ years} = 7,300,000 \text{ cycles}$$

Using the strain-life relationship, it is reasonable to assume that the original bar only experienced elastic strains (since this is a high cycle fatigue problem). Therefore, one can apply the original stress σ_1 as:

$$\begin{aligned}\text{Original Stress} &= \sigma_1 = \sigma'_f (2N_f)^b \\&= 1,400 \text{ MPa} \times (2 \times 7,300,000)^{-0.1} \\&= 269 \text{ MPa}\end{aligned}$$

Therefore, an applied stress 269 MPa was needed at each cycle.

For the new bar, applying a notch would introduce plastic strain from the bending. Using the strain-life model,

$$\varepsilon_a = \frac{\sigma'_f}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c$$

The cycles to failure corresponding to two weeks of cyclic strains at 1,000 cycles/day is 14,000 cycles. Therefore, the stress applied to cause fatigue failure at 14,000 cycles needs to be determined using the above relationship.

$$\varepsilon_{a2} = \frac{1,400 \text{ MPa}}{200,000 \text{ MPa}} (2 \times 14,000)^{-0.1} + 0.5 \times (2 \times 14,000)^{-0.5} = 0.0055$$

From the Ramberg-Osgood relationship,

$$\varepsilon_{a2} = \frac{\sigma_2}{E} + \left(\frac{\sigma_2}{K} \right)^{1/n}$$

$$0.0055 = \frac{\sigma_2}{200,000 \text{ MPa}} + \left(\frac{\sigma_2}{1,500 \text{ MPa}} \right)^{1/0.193}$$

$$\sigma_2 = 490 \text{ MPa}$$

From Neuber's Law (notch effects):

$$\varepsilon_2 \sigma_2 = \frac{(K_f \Delta S)^2}{E}$$

$$0.0055 \times 490 \text{ MPa} = \frac{(10x \text{ cm}^{-1} \times 269 \text{ MPa})^2}{200,000 \text{ MPa}}$$

Solving for x yields:

$$x = 0.23 \text{ cm} = 2.7 \text{ mm}$$

Note that this solution utilizes the strain-life approach to fatigue life prediction, which only takes into account crack initiation, not crack propagation. In reality, one may need to take into account the propagation life of the crack to more accurately determine of fatigue life. The next section discusses the fracture mechanics approach to fatigue analysis of smooth and notch members.

2.2.5. TWO-STAGE APPROACH TO FATIGUE LIFE ESTIMATION

The strain-life methodology (discussed above) and fracture mechanics methodology (discussed below) may be combined to determine a more accurate life prediction. This is because the life of a structure can often be separated into two stages as follows (Stephens, et al., 2001):

- Stage 1: life to the crack formation of the order of 1 mm (strain-life)
- Stage 2: life from crack formation to failure (fracture mechanics).

In fatigue design, engineers must determine the extent of life contribution that each stage provides. Such determination is based on factors such as stress level, load history, residual stresses, material properties, and geometry. There are situations where stage 2 may be negligible in comparison with stage 1. These include low amplitude loading of simple geometric structure (such as a small shaft) with little or no stress raisers (notches). Alternately, life prediction may be calculated more appropriately using the fracture mechanics approach (stage 2). For example, there may be a requirement to present safety margins for a particular design with inherent small cracks. Nonetheless, there are valid reasons for combining these stages for fatigue life predictions. The existence of stress raisers (such as notches) is one example of where the two-stage or hybrid approach has utility. In essence, whenever any applicable factors such as geometry and material properties change stress distribution across the material, the hybrid approach is likely to provide a more accurate life prediction.

2.2.6. FRACTURE MECHANICS

The fracture mechanics approach to fatigue analysis involves calculating fatigue crack growth rate and subsequent cycles to failure of the component under analysis. The fatigue life of a component is assumed to be comprised of an initiation and a propagation stage. In many materials, a significant amount of useful life remains after crack initiation, and therefore failure should not simply correspond with crack initiation. Figure 2.24 shows a comparison of the stress vs. life curve for both crack initiation and crack propagation. One must apply some initial assumptions (if not already known) as to what constitutes crack initiation to perform crack propagation analysis. The failure mechanics approach then evaluates the strength of the structure or component, given that this initial crack is already present, and calculates life based on the growth rate of the crack (Stephens, et al., 2001).

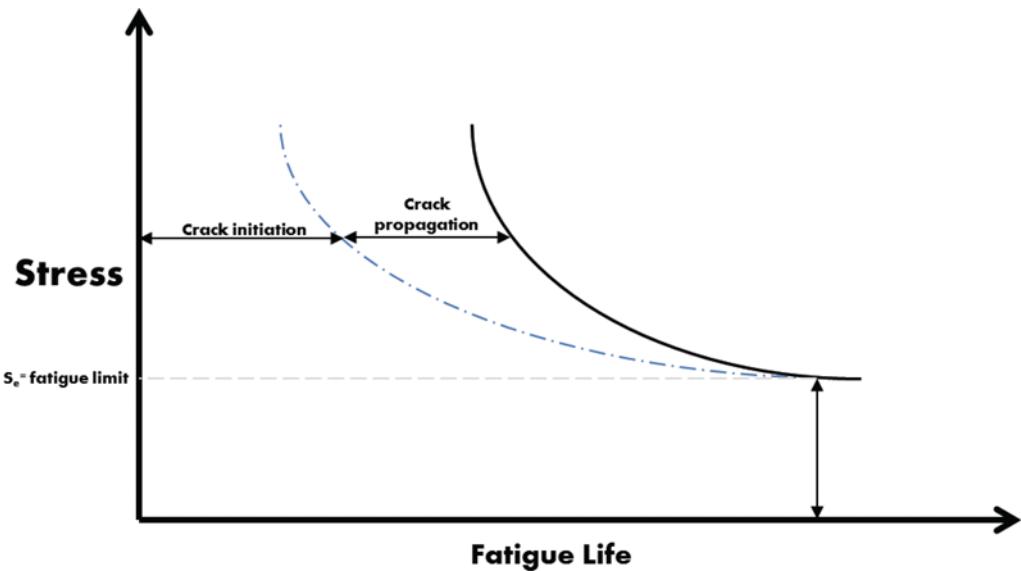


Figure 2.24: Crack initiation and propagation portions of fatigue life

There is no universally accepted definition for the size of crack initiation. For metals, several definitions for the size of the crack initiation have been reported. For example, Kujawski and Ellyin (1992) used crack length ranging from grain diameter to about 50-100 μm , depending on the material and physical scale of interest. Crack length of 0.5 mm was considered for initiation size in structural welds by Martin and Wirsching (1991). A value of 1 mm was used for En7A steel (Kaynak et al., 1996). Murtaza and Akid (1995) considered 0.12 mm as crack initiation length for BS250A53 steel, and a value of 0.051 mm was used for carbon steel (Majumdar, et al., 1993).

For a cracked specimen subjected to a constant amplitude cyclic load with a remote stress range given by:

$$\Delta\sigma = \sigma_{max} - \sigma_{min} \quad (2.45)$$

A typical plot of the crack length a versus the corresponding number of cycles N is shown in Figure 2.25. The majority of component life comprises the period where crack length is relatively small (compared to final crack length). This plot also shows crack growth rate as the slope of the curve, i.e. da/dN .

This section covers Linear Elastic Fracture Mechanics (LEFM), which is based on the theory of elasticity to bodies containing cracks or defects (i.e. small displacements and linear stress versus strain relationships) (Bannantine, et al. 1997). For the LEFM method, the initial crack size is often assumed to be about 1 mm. LEFM is used to relate stress magnitude near crack tip to:

1. Remote stresses
2. Crack size and shape
3. Materials properties

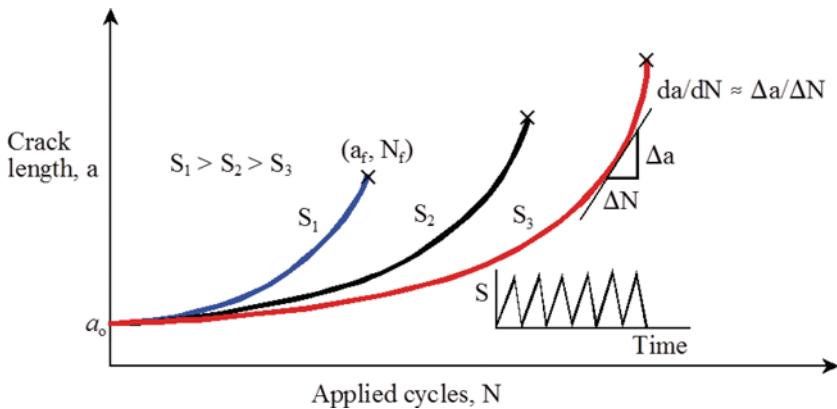


Figure 2.25: Fatigue crack length vs. applied cycles (x marks the fracture)

2.2.6.1. Stress Intensity Factor

Given a propagating crack, the stress intensity factor, K , represents the scale of the crack tip local stress magnitude. There are several elements that influence this factor, including crack size, crack shape, and geometric boundaries. The general form of the equation for stress intensity factor is given by:

$$K = f(g)\sigma\sqrt{\pi a} \quad (2.46)$$

where

σ = remote stress applied to component

a = crack length

$f(g)$ = correction factor which is influenced by material properties of the test specimen and crack geometry. This is covered in detail in many fatigue analysis textbooks (e.g. section 3.2.4 of Bannantine, et al. (1997)).

The stress intensity factor range is defined as:

$$\Delta K = K_{max} - K_{min} \quad (2.47)$$

This quantity is extremely important in the LEFM method of fatigue analysis, as it provides the initial indication of the stress environment at the crack tip. If one is to plot a typical graph of $\log da/dN$ against $\log \Delta K$, the plot illustrated in Figure 2.26 is generated. Note here that the resulting curve is divided into three regions: I, II and III, each of which will be discussed in the following subsections.

2.2.6.2. Region I

In this region the value for ΔK is low, which implies that crack propagation is very slow. Figure 2.26 shows a quantity $\Delta K_{threshold}$, which represents the stress intensity value below which the crack does not grow. In the design of components, it is therefore desirable to have a stress intensity factor that is below ΔK_{th} at use conditions.

2.2.6.3. Region II

Region II is the mid region of stress intensities for the crack growth rate vs. stress intensity range plot shown in Figure 2.26. As shown in this plot, the slope of the curve is approximately linear. The most widely accepted equation that fits this region II curve is the Paris equation, proposed in the 1960s:

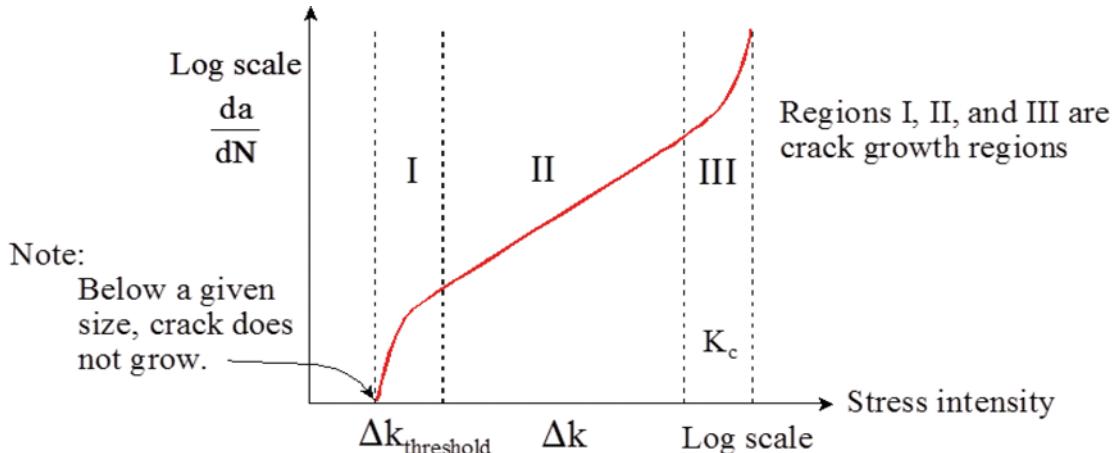


Figure 2.26: Typical curve for crack growth rate vs range of stress intensity factor

$$\frac{da}{dN} = C(\Delta K)^m \quad (2.48)$$

where C and m are material constants, found in many material data references.

Integrating each side of the equation with respect to dN , one can derive the expression for cycles to failure N_f as follows:

$$N_f = \int_{a_i}^{a_f} \frac{da}{C(\Delta K)^m} \quad (2.49)$$

where a_i is the initial crack length and a_f is the final (critical) crack length.

Since K is a function of correction factor, $f(g)$, which in turn, depends on the crack length a , the above integral is usually numerically solved. There are common approximations used for $f(g)$ (such as 1.12 for infinite plate with a single edge crack or 1.0 for an internal crack) to simplify calculations. If $f(g)$ is assumed constant, then substituting the equation (2.45) in the above gives:

$$N_f = \int_{a_i}^{a_f} \frac{da}{C(f(g)\Delta\sigma\sqrt{\pi a})^m} \quad (2.50)$$

and for $m \neq 2$,

$$N_f = \frac{a_f^{(-m/2)+1} - a_i^{(-m/2)+1}}{\left(-\frac{m}{2} + 1\right) C(\Delta\sigma)^m \pi^{m/2} [f(g)]^m} \quad (2.51)$$

To solve the above equation, the final crack size a_f must first be determined. Since

$$K = f(g)\sigma\sqrt{\pi a} \quad (2.52)$$

Then we can derive final crack size a_f as:

$$a_f = \frac{1}{\pi} \left(\frac{K_c}{f(g)\sigma_{max}} \right)^2 \quad (2.53)$$

K_c is the “fracture toughness”: the critical value of K in which a crack extends in a rapid (unstable) manner without an increase in load or applied energy (i.e. tearing effect). This value depends on the material, strain rate, environment, thickness, and to a lesser extent crack length. The minimum value of fracture toughness is known as “plain strain fracture toughness”: K_{Ic} . In this notation, the subscript I refers to the fact that these fractures occur almost entirely by the mode I crack opening. Mode II and mode III of fatigue fracturing are not discussed here. Interested readers may refer to Bannantine, et al. (1997) for discussions on mode II and III fracture. Note that fatigue life is largely dependent on the initial size of the crack as opposed to the final size. This is illustrated in Figure 2.27. As a result, fatigue life is less sensitive to fracture toughness beyond cases where a very hard material is subjected to very large stresses (i.e. the difference between initial and final crack size is very small).

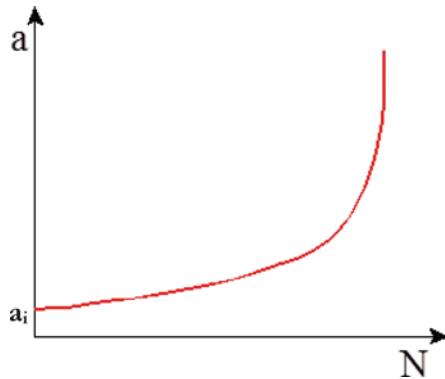


Figure 2.27: Effect of initial crack size on fatigue life

2.2.6.4. Region III

In this region, rapid, unstable crack growth occurs. Here, the crack growth rate accelerates as the maximum stress intensity factor approaches the fracture toughness of the material. Little fatigue life is involved in this region, and thus in many practical applications, is ignored since it has little influence in estimating total crack propagation life. Transition from region II to region III depends on yield strength of the material, stress intensity factor, and stress ratio. A common representation of the behavior in this region was formulated by Forman and is given by:

$$\frac{da}{dN} = \frac{C(\Delta K)^m}{(1 - R)K_c - \Delta K} \quad (2.54)$$

where R is the stress ratio defined as $\frac{\sigma_{min}}{\sigma_{max}}$.

The applied stress ratio R can have a significant impact on the crack growth rate. Generally, the higher the stress ratio, the higher the crack growth rate. Stress ratio in turn is dependent on material properties. The Forman equation above is only applicable when R is positive (i.e. tensile load). For negative values of R (i.e. compressive loads), there is no significant change in growth rate when compared to the case of $R=0$. This is intuitive, as compressive loads should generally “close the gap,” thereby prohibiting

crack growth. It has been shown, however, that for certain materials, higher crack growth rates can exist for compressive loadings (compared to $R = 0$ loadings).

Example 2.10

A large plate containing a center crack of length $2a = 2$ mm is used in seawater environment (see Figure 2.28). The following laboratory data was obtained for this material in salt spray:

$$\frac{da}{dN} = A(\Delta K)^3$$

$$K_{IC} = 66 \text{ MPa}\sqrt{\text{m}}$$

$$A = 3.81 \times \frac{10^{-12} \text{ m}}{\text{cycle}}$$

$$K = \sigma\sqrt{\pi a}$$

$$f(g) = 1.0$$

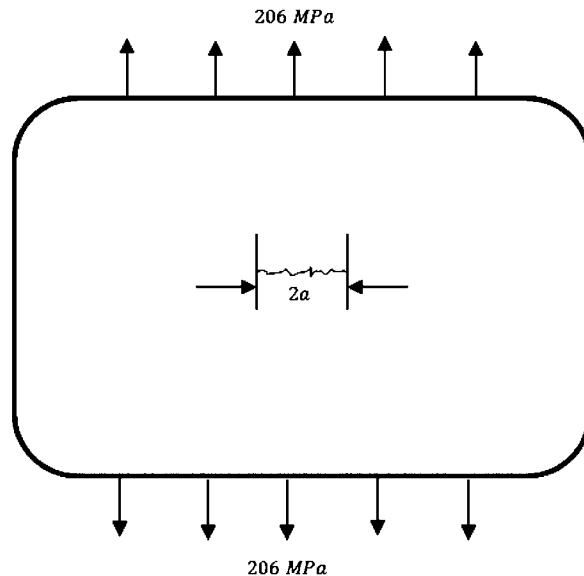


Figure 2.28: Diagram for Example 2.10

This plate is subjected to a cyclic stress of 206 MPa with $R = 0$. The time to failure needs to be calculated.

Solution:

The following relationship is known:

$$\frac{da}{dN} = A(\Delta K)^3$$

Substituting: $\Delta K = \Delta\sigma\sqrt{\pi a}$ yields:

$$\frac{da}{dN} = A(\Delta\sigma\sqrt{\pi a})^3$$

Rearranging gives the following relationship:

$$dN = \frac{da}{A(\Delta\sigma)^3 \pi^{3/2} a^{3/2}}$$

Integrating each side of the equation yields:

$$\int_0^{N_f} dN = \frac{1}{A \cdot \Delta\sigma^3 \cdot \pi^{3/2}} \int_{a_0}^{a_f} \frac{da}{a^{3/2}}$$

The crack length is determined through integrating the above expression from the initial crack size, a_0 , to the final crack size a_f .

$$N_f = \frac{2}{A \cdot \Delta\sigma^3 \cdot \pi^{3/2}} \left(\frac{1}{\sqrt{a_0}} - \frac{1}{\sqrt{a_f}} \right)$$

One can determine a_f from the relationship

$$a_f = \frac{1}{\pi} \left(\frac{K_{Ic}}{f(g) \sigma_{max}} \right)^2$$

$$a_f = \frac{1}{\pi} \left(\frac{66 \text{ MPa}\sqrt{\text{m}}}{1 \times 206 \text{ MPa}} \right)^2 = 32.67 \text{ mm}$$

$$a_0 = 1 \text{ mm (initial flaw size)}$$

Therefore,

$$N_f = \frac{2}{3.81 \times 10^{-12} \times 206^3 \times \pi^{3/2}} \left(\frac{1}{\sqrt{0.001}} - \frac{1}{\sqrt{0.03267}} \right)$$

$$N_f \cong 2.8 \times 10^5 \text{ cycles}$$

Note: an edge crack length a , and a center crack length $2a$, have the same initial crack length a .

2.2.6.5. Fracture Mechanics Approach with Notch Effect

The use of LEFM methods for notched components involves considering the notch stress-strain field. Inherent in these methods when evaluating notched members is the calculation of the transition crack length. The transition crack length, denoted as l_t , is the length corresponding to the distance taken from the edge of the notch to the nearest point in the specimen where local stress approaches bulk (nominal) stress.

From Bannantine, et al. (1997), the relationship between l_t , notch depth D , stress intensity factor K_t is given by:

$$l_t \left[\frac{1.12 K_t}{f(g)} \right]^2 = D \pm l_t \quad (2.55)$$

Rearranging the above equation yields:

$$l_t = \frac{D}{\left[\frac{1.12K_t}{f(g)}\right]^2 \pm 1} \quad (2.56)$$

Note the \pm in the above equation allows for alternate notch scenarios. For shallow notches, such as the one depicted in Figure 2.29, the crack length is considered to be equal to the notch depth plus the free surface crack length. For deep notches, such as the one depicted in Figure 2.30, the crack length is considered to be the depth of the notch minus the length of the free surface crack.

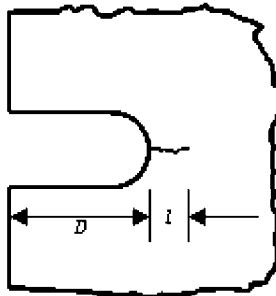


Figure 2.29: A shallow notch

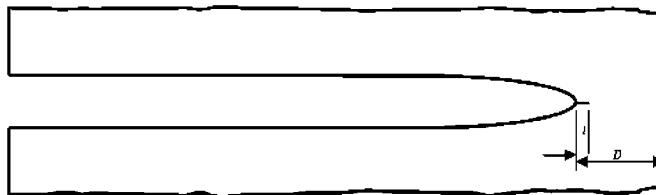


Figure 2.30: A deep notch

When the physical crack length, l , exceeds the value of the transition crack length, the crack is assumed to be outside the notch stress field, and the crack growth may then be calculated using the fracture mechanics techniques discussed in section 2.2.6, with the crack length equal to $l + D$.

Example 2.11

Using the fracture mechanics approach, the number of cycles required to grow the crack in Example 2.7 and Example 2.8 to fracture needs to be determined. An initial crack size of 1 mm is assumed.

$$A = 6.91 \times \frac{10^{-12}\text{m}}{\text{cycle}} \text{ and } K_C = 66 \text{ MPa}\sqrt{\text{m}}$$

Solution

First, the transition crack length l_t is determined using:

$$l_t = \frac{D}{[1.12K_t/f(g)]^2 - 1}$$

In this analysis, the geometry correction factor $f(g)$ is approximately 1.12: the free edge correction factor. The transition crack length becomes:

$$l_t = \frac{D}{K_t^2 - 1}$$

where

$$D = \text{depth of elliptical notch} = 10 \text{ mm}$$

$$K_t = 2.41$$

Substituting into the above expression yields:

$$l_t = \frac{10}{2.41^2 - 1} = 2.08 \text{ mm} > 1 \text{ mm} \text{ (Absolute crack length)}$$

Therefore, a crack length of 1 mm represents a “short” crack, meaning crack growth is affected by the notch’s local stress distribution. Therefore, analysis should be conducted in two parts:

1. Propagation life from crack length of 1 mm to l_t
2. Propagation life from crack length of $l_t + D$ to $a_f + D$

For part 1

Using Paris’ equation, N_f for crack growth exponent $n \neq 2$ is:

$$N_f = \frac{a_f^{(-n/2)+1} - a_i^{(-n/2)+1}}{\left(-\frac{n}{2} + 1\right) A(\Delta S)^n \pi^{n/2} [f(g)]^n}$$

where $A = 6.91 \times 10^{-12}$ and $n = 3$ (for ferretic steels)

Therefore, the propagation life for the crack is given by:

$$N_f = \frac{0.00208^{-0.5} - 0.001^{-0.5}}{(-0.5)6.91 \times 10^{-12}(375)^3 \pi^{1.5} 1.12^3}$$

$$N_{f1} = 6803 \text{ cycles}$$

For part 2

$$a_f = \frac{1}{\pi} \left(\frac{K_c}{S_{max} f(g)} \right)^2$$

where

$$S_{max} = 375 \text{ MPa}$$

$$K_c = 66 \text{ MPa}\sqrt{\text{m}} \text{ (given)}$$

Therefore,

$$a_f + D = \frac{1}{\pi} \left(\frac{66}{375 \times 1.12} \right)^2 + 10 \text{ mm} = 17.86 \text{ mm}$$

$$N_f = \frac{0.01786^{-0.5} - 0.01208^{-0.5}}{(-0.5)6.91 \times 10^{-12}(375)^3 \pi^{1.5} 1.12^3}$$

$$N_{f2} = 1134 \text{ cycles}$$

Total propagation time is: $N_f = N_{f1} + N_{f2} = 6803 + 1134 = 7937$ cycles.

2.2.7. FACTORS INFLUENCING FATIGUE FAILURE

Previous sections in this chapter discussed some loading parameter effects on fatigue life. The influence of mean stress, combined mode loading, variable amplitude and notch sensitivity were studied in detail. In this section, other factors that significantly influence fatigue life and should be considered in component and structure design are reviewed. These include the effects of size (also known as the geometry effect), frequency, and environment (humidity, temperature, electric field, magnetic field and corrosive medium).

2.2.7.1. Size Effect

Larger structural members tend to produce more locations susceptible to fatigue failure compared to smaller members. This is because fatigue is a “weakest link” process (Dowling, 1998). Fatigue degradation may initiate at locations where stresses, geometry, defects and material properties combine to form optimum conditions for fatigue crack initiation and growth (Berge, 1985). Dowling (1998) states that the decrease of stress with depth is less abrupt in larger cross sections compared to smaller cross sections. Consequently, the larger volume of material experiences high stress. The size effect becomes more important in the fatigue life predictions of welds (where pre-existing flaws increase with the size of the weld). The size effect is important in both crack initiation and crack growth. There are several studies available that study size effect in fatigue life prediction. Shigley and Mischke (2001) present the size factor, K_b , evaluated using 133 sets of data points from bending and torsion tests as follows:

$$\begin{aligned} K_b &= 1.24d^{0.107} & 2.79 \leq d \leq 51 \text{ mm} \\ K_b &= 0.859 - 0.000837d & 51 \leq d \leq 254 \text{ mm} \\ K_b &= 1, \text{for axial loading} & \end{aligned} \quad (2.57)$$

The above models are applicable for members with circular cross sections. For other shapes, an effective dimension d is introduced by Shigley and Mischke (2001) particularly Section 7-6. For example, consider two circular specimens undergoing torsion fatigue load. One specimen has diameter of $d_1=3$ mm and the other $d_2=35$ mm. According to the equation above, the ratio of size factors is:

$$\frac{K_{b1}}{K_{b2}} = \left(\frac{d_1}{d_2}\right)^{-0.107} = \left(\frac{3}{35}\right)^{-0.107} = 1.3 \quad (2.58)$$

This means the endurance limit decreases by a factor of 1.3 when the diameter of the specimen increases from 3 mm to 35 mm. The above formulae relate to crack initiation. However, several studies investigate the size effect on just crack growth, and both crack initiation and growth. For example, Berge (1985) studied the effect of plate thickness on the fatigue strength of transverse fillet welds in axial loading. He tested C-Mn steel plates with thicknesses ranging from 12.5 to 80 mm. Results of the study showed that the fatigue strength of plate with thickness 80 mm was 40% less than that of plate with 12.5 mm thickness. Berge (1985) proposed a power law model:

$$\frac{N}{N_0} = \left(\frac{t_0}{t}\right)^{3/4} \quad (2.59)$$

where N_0 refers to fatigue life for a reference plate thickness t_0 .

Kim and Shim (2000) studied the effect of plate thickness on fatigue crack growth rate of aluminum 7075-T6 under constant amplitude load. They presented a modified version of the $da/dN-\Delta K$ equation by introducing variable Z that accounts for thickness:

$$\frac{da}{dN} = Z \cdot f(\Delta K, R) \quad (2.60)$$

where f is a function of stress intensity range, ΔK , and stress ratio, R . Kim and Shim (2000) assumed a normal distribution for the logarithm of Z as:

$$\gamma = 1 - \phi \left[\frac{\log Z_\gamma - \mu_{\log Z}}{\sigma_{\log Z}} \right] \quad (2.61)$$

where $\sigma_{\log Z}$ is the standard deviation of the logarithm of Z and $\mu_{\log Z}$ is the mean of the logarithm of Z ($\mu_{\log Z} = 0$). Their results showed that the variance of the logarithm of Z can be used to estimate the variation in crack growth rates due to the thickness effect. The variance of the logarithm of Z increases with decreasing thickness via a power law correlation:

$$\sigma_{\log Z}^2 = \alpha B^\beta \quad (2.62)$$

where α and β are material constants, and B is the plate thickness. The values of $\alpha=0.0143$ and $\beta=-0.792$ were obtained for aluminum 7075-T6 (Kim and Shim, 2000).

Size effect also becomes important in viscoelastic materials such as polymers that produce internal heat during cyclic loading. The internal heating gives rise to the temperature of the sample, which, in turn, decreases the fatigue life. Since these materials have poor thermal conductivity, thinner specimens have better fatigue performance (1989).

2.2.7.2. Frequency Effect

Fatigue tests often require a long test period, especially for high-cycle fatigue failure ($>10^7$ cycle). A cost-effective fatigue test can involve increasing fatigue frequency to apply a large number of cycles in a more reasonable period of time. Not only does an understanding of the frequency effect shed light on fatigue failure mechanisms, but it also informs the design of structures and accelerated tests.

In this section, the effect of frequency on room temperature fatigue life is discussed without considering influence of environmental factors. Morrow (1965) conducted a series of axial tension-compression low-cycle fatigue tests at low frequency (order of 0.017 Hz to 0.17 Hz) and concluded that change of frequency within this range does not influence damping due to plastic strain energy. This is because the magnitude of the plastic strain and plastic strain energy is reasonably independent of test frequency at temperatures well below the creep range.

Gucer and Capa (1970) studied the effect of frequency (0.003 Hz, 0.017 Hz, 0.17 Hz and 1.7 Hz) on the strain behavior and crack growth rate for low-cycle fatigue of fully reversed bending of steel specimens. For high-cycle fatigue, they reported a small frequency effect within the range of 80 to 170 Hz. For frequencies above this range, a continuous increase in fatigue strength is observed for high-cycle fatigue. In low-cycle fatigue (lives below 10^4 or 10^5), the frequency effect becomes significant. Their results show that as frequency increases, specimens undergo more elastic deformation and less plastic deformation. This effect will increase fatigue life, as failure is predominantly controlled by plastic deformation in low-cycle fatigue. The crack growth rate was also reported to increase as the frequency decreases. Figure 2.31a shows crack growth for different frequencies in low-cycle fatigue. In terms of fatigue life, results of the work of Gucer and Capa (1970) show a 50% decrease in fatigue

life as frequency decreases from 1.7 Hz to 0.003 Hz. The effect of frequency on fatigue life is also plotted in Figure 2.31b.

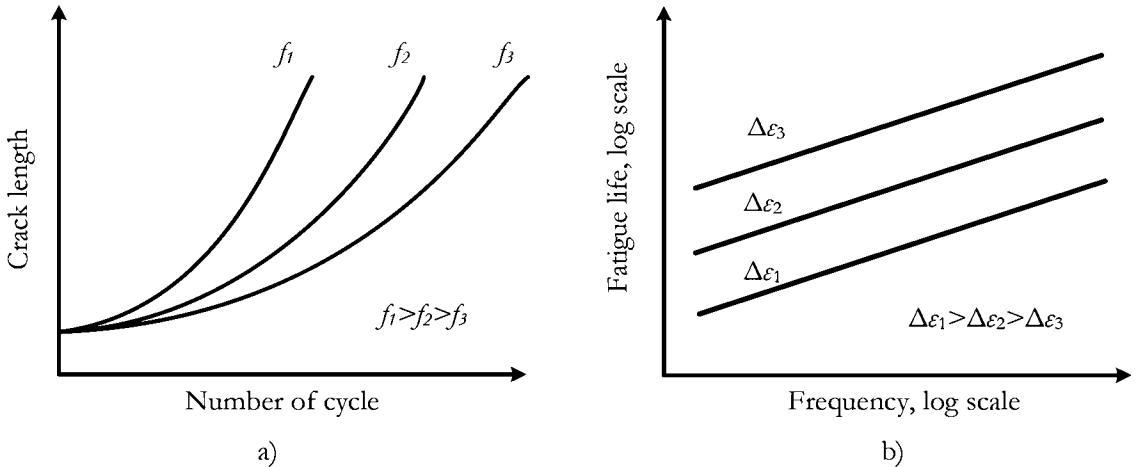


Figure 2.31: Effect of frequency, f , on low-cycle fatigue life, a) crack growth b) fatigue life

2.2.7.3. Environmental and External Effects

Environmental effects (such as chemical and thermal) can adversely influence fatigue life. This can occur in both crack initiation and propagation by contributing to pit initiation, surface cracks and dissolution of crack tip material. The effects of temperature are included in the subsequent discussion of creep and corrosion fatigue.

Figure 2.32 shows a reduction in fatigue life due to environmental effects through corresponding SN curves.

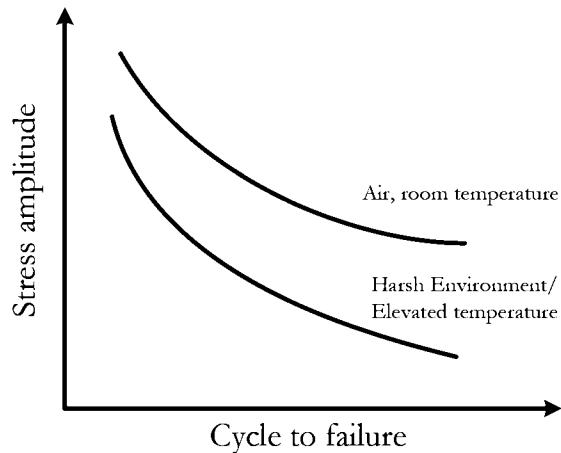


Figure 2.32: Fatigue life reduction due to harsh environment and/or elevated temperature

Khonsari and Amiri (2012) studied external influences of fatigue life such as electric current, alternating magnetic field and surface cooling. These factors can significantly enhance life by decelerating crack initiation and propagation processes. For example, Tobushi et al. (1997) studied the influence of air and water on rotating-bending fatigue life of a shape-memory alloy wire. Their experimental results showed that fatigue life improves in water for low-cycle fatigue ($<10^4$ cycles). Water maintained

constant wire temperature, in turn maintaining yield stress that extended fatigue life. However, for high-cycle fatigue ($>10^5$ cycles), fatigue life is shorter in water than that in air, mainly due to corrosion.

Hirano et al., (2003) investigated the effects of water flow rate on fatigue life of carbon steel in a simulated light-water reactor (LWR) environment. They tested carbon steel at 289 °C for various dissolved oxygen contents (DO) at strain rates of 0.4, 0.01, and 0.001 percent per second (%/s). Their experimental results showed that at the strain rate of 0.01 %/s, the fatigue life increased with increasing flow rate under all DO conditions. Specifically, they reported that the fatigue life at a 7 m/s flow rate

Table 2.2: Some values of wear coefficient (from Engel (1993))

Pair of contact	k (10 ⁻³)
Zinc on zinc	53
Mild steel on mild steel	15
Platinum on platinum	13
Copper on copper	11
Stainless steel pair	7
Silver pair	4
Copper on mild steel	0.5
Platinum on mild steel	0.5
Platinum on silver	0.33

was about three times longer than that at a 0.3 m/s flow rate. This increase in fatigue life was attributed to increases in the crack initiation life and small-crack propagation life.

Amiri and Khonsari (2013) investigated the effect of surface cooling in atmospheric air on the fatigue life of two materials: stainless steel 304L and steel 4145. They conducted fatigue tests with surface cooling and without surface cooling. Significant improvement in fatigue life was observed for both materials when the surfaces of the specimens were cooled. Fatigue life for stainless steel 304L at stress amplitude of 495 MPa improved by about 100%, and about 1000% at 450 MPa. No physical explanation for the observed results is presented in the study.

There are several studies on the effect of external elements such as electric current pulsation and magnetic field on the fatigue life of metals (Karpenko et al., 1976; Abd E1 Latif, 1979; Bezborodko, 1984; Cao, 1989; Conrad et al., 1991). Conrad et al. (1991) investigated the effect of electric current pulses on 99.9 % pure copper fatigue life. Their results showed a positive effect on both low and high-cycle fatigue life from electric current through increasing the number of cycles required for micro-crack initiation and decreasing inter-granular cracking tendencies. This phenomenon was attributed to the increase in slip homogenization, which involves decreased slip band spacing and width.

Another external factor influencing fatigue life is magnetic field. Yong et al. (1993) studied this effect on A3 steel specimens loaded at 25 Hz frequency and stress ratio of 0.01. Their results show that by applying an alternating magnetic field, the fatigue life substantially increased, with an average increase of 269% reported by Yong et al. (1993).

2.2.7.4. Miscellaneous Factors

In addition to the factors mentioned above, there are several other factors that need to be accounted for in fatigue life assessment. These factors include surface condition factor (to account for ground, machined, hot-rolled or as-forged surfaces), load factor (to account for torsion, axial, bending) and miscellaneous-effects factor. Therefore, the endurance limit, S'_e , can be corrected due to modifying factors as:

$$S_e = k_a k_b k_c k_d k_e S'_e \quad (2.63)$$

where

k_a = surface condition modifying factor

k_b = size modifying factor

k_c = load modifying factor

k_d = temperature modifying factor

k_e = miscellaneous effects modifying factor

Interested readers may refer to Section 7-6 of the “Mechanical Engineering Design” book by Shigley and Mischke (2001), which provides a comprehensive understanding of endurance limit modifying factors.

2.3. WEAR

Wear is a phenomenon related to the gradual loss of material from the surface of two solid bodies due to friction. These two bodies are either in direct contact (dry sliding) or separated by a layer of lubricant. In this section, wear mechanisms and classification are discussed. Factors that accelerate wear are outlined in detail, and wear (as a form of degradation in bearings, clutches, brakes and seals) is studied.

Perhaps the first notation of friction dates back to 1508 when the Italian mathematician, engineer and artist Leonardo da Vinci proposed the coefficient of friction as the ratio of friction force to normal force. This concept was later developed by the French physicists Guillaume Amontons in 1699 and Charles-Augustin de Coulomb in 1785, who distinguished the difference between static and kinetic friction. More recently, numerous PoF-based wear models have replaced traditional empirical relationships. The most practical and easy-to-use is Archard's wear law (1953), which relates wear volume, W , to the hardness of wearing material, H , applied normal load, N , and sliding distance L :

$$W = kNL/H \quad (2.64)$$

where k is a dimensionless wear coefficient. Values of the wear coefficient for some pairs of contacting materials are given in Table 2.2.

Friction and wear are not material properties, but rather responses of a tribo-system (Kato, 2000). In general, wear mechanisms can be classified into three groups: mechanical, chemical and thermal. Each group can be further classified into sub-categories. For example, mechanical wear can be categorized into abrasive, adhesive, flow and fatigue wear. Figure 2.33 shows a schematic of wear modes (taken from Kato (2000) with permission). As inferred from its name, mechanical wear is caused by deformation and fracture of mating surfaces. Chemical wear is mainly caused by a chemical reaction in the tribofilm. Tribofilms are dynamic structures that form at the interface during frictional sliding. These films play a significant role in friction control. Thermal wear is the local interface melting from excessive frictional heat (Kato, 2000). In most practical applications, friction involves a complex interaction of factors including mechanical, chemical and thermal mechanisms. A comprehensive study of a tribosystem requires complete understanding of friction mechanisms and factors affecting wear rate. In this section, common wear mechanisms and influencing factors are discussed.

There is no wear model derived from fundamental principles for general and practical use. Useful equations are empirical and suitable for very special materials and conditions. Meng and Ludema (1995) analyzed over 180 wear equations with over 100 variables used in these equations. They found no way to unify a group of equations suitable for general use. However, Ludema (1981) categorized these

equations into three groups based on typical wear parameters. Table 2.3 shows influencing parameters used in development of wear models (Ludena, 1981).

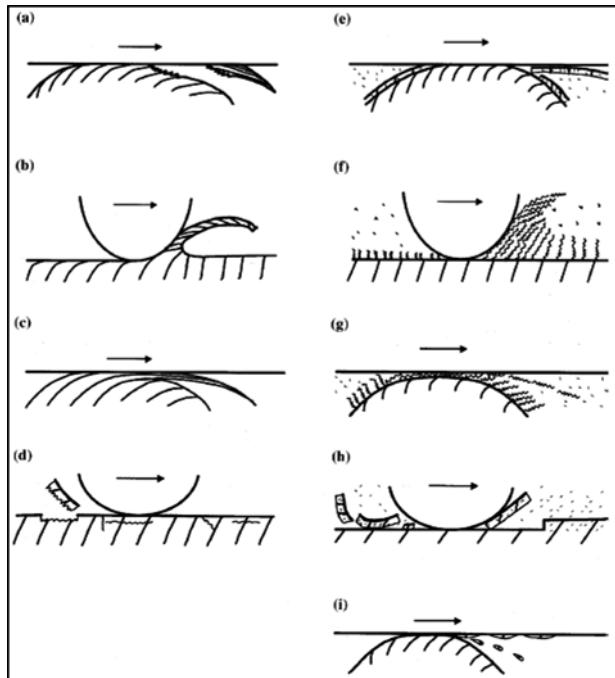


Figure 2.33: Different modes of wear: (a) adhesive wear by adhesive shear and transfer (b) abrasive wear by microcutting of ductile bulk surface (c) flow wear by accumulated plastic shear flow (d) fatigue wear by crack initiation and propagation (e) corrosive wear by shear fracture of ductile tribofilm (f) corrosive wear by shaving of soft tribofilm; (g) corrosive wear by accumulated plastic shear flow of soft tribofilm (h) corrosive wear by delamination of brittle tribofilm (i) melt wear by local melting and transfer or scattering (Kato, 2000; with permission)

Table 2.3: Parameters used in development of wear models

Material parameters

- 1) Hardness, cold and hard; 2) Ductility; 3) Fracture toughness; 4) Strength; 5) Work hardenability; 6) Modulus of elasticity; 7) Material morphology; 8) Type and thickness of surface film; 9) Thermal properties

Operational parameters

- 1) Surface topography; 2) Applied load; 3) Contact geometry; 4) Sliding speed, rolling speed; 5) Coefficient of friction

Environmental parameters

- 1) Lubricant properties; 2) Debris and dirt; 3) Rigidity of supporting structures; 4) Working temperature; 5) Clearance, alignment and fit; 6) Similar or dissimilar pair of contact; 7) Continuous, stop-start, reciprocating

Most adhesive wear models deal with steady-state wear with a time-independent and constant rate of wear generation. However, experiments show that wear volume versus sliding distance has transient as well as steady-state wear regimes (Yang, 2004) as shown in Figure 2.34. The transient wear regime is known as the “running-in” or “breaking-in” period, where the rate of wear volume is high and decreases as the steady-state is approached. During the running-in process, two mating surfaces

improve in conformity, topography and frictional compatibility. The high wear rate during this period is due to the rough asperity contact. As the more prominent asperities are lost or flattened, the wear rate decreases. The duration of the running-in period depends on the surface roughness, material properties and operating conditions. Once the running-in is completed, the steady-state regime commences.

Yang (2004) presents a model encompassing both transient and steady-state wear regimes. For the steady-state regime, Archard's wear law is adopted. For the transient regime, it is assumed that the rate of wear volume per unit sliding distance is a function of the volume of material at the interface:

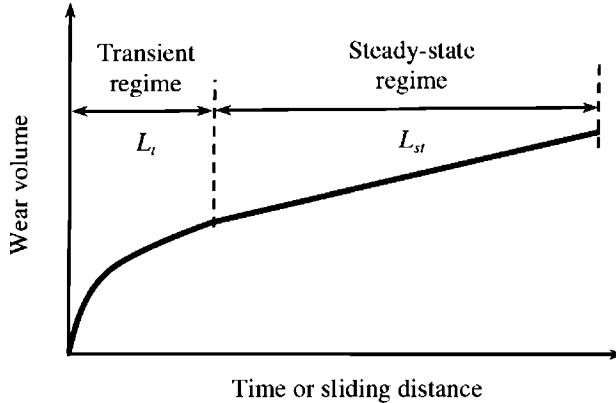


Figure 2.34: Transient and steady-state wear regime for adhesive wear mechanism

$$\frac{dW}{dL} = -BW \quad (2.65)$$

or

$$\frac{dW}{W} = -B dL \quad (2.66)$$

The negative sign indicates decrease in the rate of wear volume by increasing sliding distance. Considering the available material volume of $W = W_0$ at $L = 0$, the solution to the equation above is:

$$W = W_0(1 - e^{-BL}) \quad L < L_t \quad (2.67)$$

For the steady-state regime ($L_t < L$), the Archard's law yields:

$$W = k \frac{NL}{H} \quad (2.68)$$

The transient distance, L_t , can be obtained by differentiating transient and steady-state wear equations and equating the obtained equations as:

$$W_0 B e^{-BL_t} = k \frac{N}{H} \quad (2.69)$$

Solving the above equation for L_t yields:

$$L_t = \frac{1}{B} \ln \left(\frac{W_0 BH}{kN} \right) \quad (2.70)$$

2.3.1. GENERAL FORM OF WEAR EQUATIONS

Contact mechanics-based equations are generally based on the contacting surface topography to calculate the local region of contact (Meng and Ludema, 1995). They assume that a material property, usually Young's modulus E or hardness H , is important in the wear process. An example of a contact mechanics-based wear equation is the previously presented Archard's wear law. Archard (1953) put the ideas of contact pressure, sliding speed, and real area of contact of plastically deformed bodies into one simple equation, which is of paramount importance in practice.

Examples of empirically derived equations are provided by Barwell (1957) when expressing wear rate as one of the three equations:

$$\begin{aligned} W &= \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \\ W &= \alpha t \\ W &= \beta e^{\alpha t} \end{aligned} \quad (2.71)$$

where α is a constant, t is time and β is defined as "some characteristic of initial surfaces."

These equations relate to different regimes of wear over time (or equivalently sliding distance) as shown in Figure 2.34 with an additional regime after the steady-state regime where the rate of wear volume exponentially increases over time. This effect can cause seizure of the contacting bodies. Empirical work of Rhee (1970) suggested that wear volume is a function of applied load, N ; sliding speed, V ; and time, t :

$$\Delta W = K N^a V^b t^c \quad (2.72)$$

where ΔW is the weight loss of soft material and K , a , b , and c are constants obtained from experiment.

This type of equation, which is a multiplication of some testing parameters, is common in tribosystem analysis. Meng and Ludema (1995) state that empirical exponents are commonly assumed to be independent of each other but are rarely proven to be so. This type of equation is useful to plan accelerated wear tests, since each parameter is explicitly expressed in the equation independent of other parameters.

Ludema (1996) presented a generalized equation for the life of cutting tool in a lathe, T , that depends on the cutting speed, V , depth of cut, d , feed rate, f in the form of:

$$V T^n f^a d^b = C \quad (2.73)$$

where n , a , b , and C are taken from experiments.

This equation is not a wear equation, but rather a life model for a cutting tool with T the cutting time after which the tool is not useful. This equation is applicable over a wide range of cutting speeds; however, it is valid only over the range of cutting conditions.

Wear is generally measured in three ways:

- Volume of material loss (unit of mm³)
- Mass of material loss (g)
- Wear depth (mm)

It is possible to convert wear volume to wear using the density of the wearing material.

There are two types of empirical wear models used for design: one relates the wear life to the operating conditions and the other provides a relationship between cumulative wear or wear rate and operating conditions. Bayer (2004) provides a summary of engineering wear models from both types. The forms of relationships for various wear situations are similar; however, the coefficients and exponents may vary for different materials, speeds, environments, and lubrication. Some of the models are presented below. For a comprehensive discussion on the applicability of the models, readers may refer to chapter 2 of Bayer (2004).

2.3.2. SLIDING WEAR

One of the most common forms of wear is sliding wear. Using the adhesive wear concept discussed in the previous section, Archard's law can be applied to model the amount of wear or wear rate. Other models, such as those developed for running-in and nonlinear relationships between wear and load, were also discussed in the preceding section. Considering Archard's wear equation and assuming a conformal contact of constant area A , wear can be written as:

$$W = Ah = k \frac{NL}{H} \quad (2.74)$$

where h is the wear depth. Therefore,

$$\begin{aligned} h &= k \left(\frac{N}{A} \right) \frac{L}{H} \\ \frac{dh}{dt} &= \left(\frac{k}{H} \right) \left(\frac{N}{A} \right) \left(\frac{dL}{dt} \right) = \frac{k}{H} (PV) \end{aligned} \quad (2.75)$$

where P is the pressure, $\left(\frac{N}{A} \right)$.

This formulation is known as the PV factor approach, relating wear to the operating conditions of pressure and velocity. In wear experimentation based on the PV factor, the wear coefficient, k , is obtained for different PV values. After a limit value of PV , known as the PV limit, the wear regime changes from mild to severe wear. In a severe wear regime, the wear coefficient depends on the PV factor as shown in Figure 2.35. The PV factor approach is useful in wear analysis in situations where the contact area does not change. This is the case in journal and thrust bearings (Bayer, 2004).

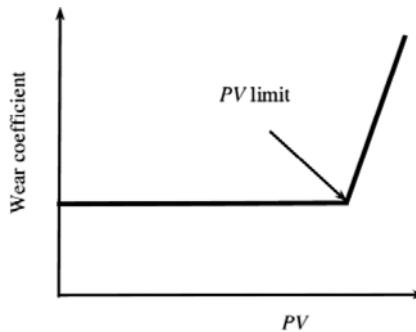


Figure 2.35: Change of wear coefficient with PV (Bayer, 2004)

Example 2.12: Adhesive Wear

A pin-on-disk friction test is set up as shown in Figure 2.36. Disk and pin materials are steel 4140 and 70-30 brass, respectively. Steel 4140 hardness is 2840 MPa, and 70-30 brass hardness is 390.5 MPa. The test setup is equipped with a linear variable differential transformer (LVDT) so that the vertical displacement of the pin-disk configuration can be accurately measured. The pin diameter is 3 mm and is placed 15 mm away from the disk center. The disk is rotating at 10 rpm. The wear coefficient is $k=4.3\times 10^{-4}$. Given the load on the pin is 2 kN, the time taken to reach 1 mm vertical displacement needs to be determined.

Solution

Since the hardness of steel 4140 is much greater than that for the 70-30 brass, it is assumed that it is only the brass that wears. Using Archard's wear law,

$$W = k \frac{NL}{H}$$

If both sides are divided by time, t , and $V = L/t$ and $\dot{W} = W/t$ substituted, one obtains:

$$\dot{W} = k \frac{NV}{H}$$

In this equation, V is the sliding velocity and can be calculated from rotating speed as:

$$V = 2\pi \times \text{radius} \times \text{rotating velocity} = 2\pi \times 0.015 \times \left(\frac{10}{60}\right) = 0.0157 \text{ m/s}$$

Therefore,

$$\dot{W} = 4.3 \times 10^{-4} \frac{2 \times 10^3 \times 0.0157}{390.5 \times 10^6} = 3.457 \times 10^{-11} \frac{\text{m}^3}{\text{s}} = 0.03457 \frac{\text{mm}^3}{\text{s}}$$

Assuming the real area of contact equals the nominal area of contact,

$$W = \text{area of contact} \times \text{removed length} = \pi \left(\frac{3}{2}\right)^2 \times 1 = 7.069 \text{ mm}^3$$

Therefore, to remove 7.069 mm³ of Brass, the elapsed time is:

$$t = \frac{7.069}{0.03457} = 204.5 \text{ sec}$$

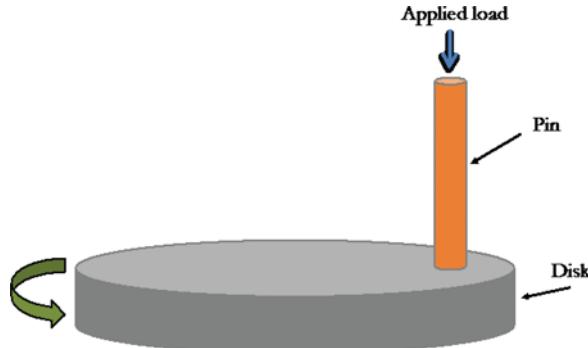


Figure 2.36: Pin-on-disk tribosystem

2.3.3. ABRASIVE WEAR

Similar to adhesive wear, in abrasive wear, a wear coefficient k_{abr} can be defined to characterize the wear process as in Archard's law. Figure 2.37 shows a hard-conical slider removing (plowing) a groove of cross sectional area $r \cdot h$. For a sliding distance of x , the removed volume is (Engel, 1993):

$$W = r \times h \times x = N \times x \times \tan(\theta)/\pi H \quad (2.76)$$

The term $\tan(\theta)/\pi$ can be replaced by k_{abr} to result in an equation similar to Archard's law for adhesive wear:

$$W = k_{abr} \frac{Nx}{H} \quad (2.77)$$

A generic form of the above equation can be used for abrasive wear as:

$$W = k_{abr} H^m N^n L^k \quad (2.78)$$

where the exponents are functions of contacting abrasive and abraded materials and operating conditions and are determined from experiment.

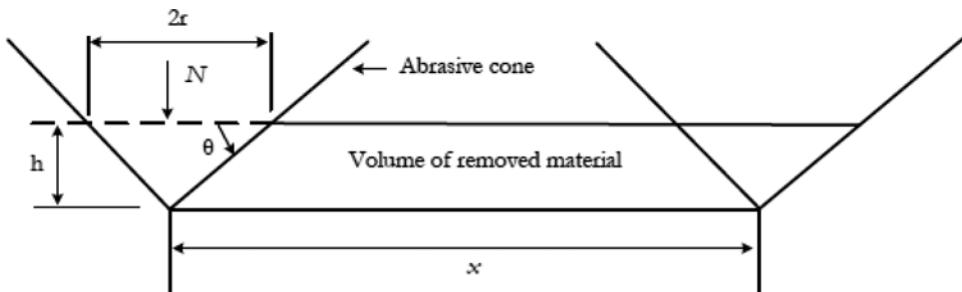


Figure 2.37: Schematic of a hard cone plows a groove in a soft material (Adopted from Engel, 1993)

2.3.4. IMPACT WEAR

Percussive mechanical wear occurs between two impacting components through a mechanism referred to as impact wear. It occurs in many aerospace, transportation, manufacturing, and business machines such as printers (Wear Control Handbook, edited by Peterson and Winer, 1980). An empirical model of impact wear suggests that the following equation accurately models the phenomenon (Bayer, 2004):

$$W = k_{imp} V^n N_{imp} \quad (2.79)$$

where k_{imp} is the wear coefficient in impact wear, V the impact velocity, and N_{imp} the number of impacts.

In the above equation n depends on the tribosystem and needs to be determined from experimentation. Its value is within the range $2 < n < 3$.

There are other impact wear equations that predict wear life cycle. For example, Bayer (2004) presents a Zero Wear Model for compound impact situations where both impact and sliding are causing wear. Zero wear refers to the wear being negligible from an engineering point of view with surfaces considered intact. In a zero-wear regime, the depth of wear is below the half roughness of the original surface. The number of cycles to zero wear, N_0 , can be calculated using:

$$N_0 = \frac{2000}{1 + \beta} \left(\frac{\gamma \sigma_y}{\sigma_0} \right)^9$$

where σ_y is the tensile yield strength of the wearing material, σ_0 the peak contact stress during impact, γ system wear coefficient and β the contribution factor of sliding wear with respect to impact wear.

For pure impact damage, $\beta = 0$. If the number of impacting cycles exceeds that of for zero wear, N_0 , there is a transition in wear regime from zero wear to measurable wear as shown in Figure 2.38. The system wear coefficient γ is found to be 1.1 for several steels.

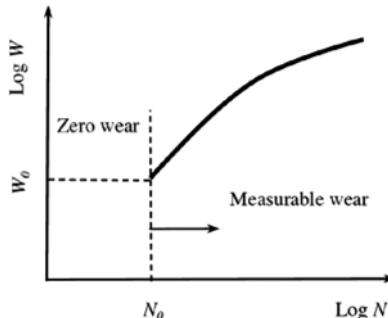


Figure 2.38: Impact wear regimes (adapted from Wear Control Handbook, 1980)

Example 2.13: Impact Wear

The zero-wear limit of steel needs to be determined with the peak contact stress of $\sigma_0 = 450$ MPa and yield strength of $\sigma_y = 550$ MPa.

Solution

Since there is only impact (no sliding), $\beta = 0$. Using the equation above for zero wear life cycle yields:

$$N_0 = \frac{2000}{1 + \beta} \left(\frac{\gamma \sigma_y}{\sigma_0} \right)^9 = \frac{2000}{1 + 0} \left(\frac{1.1 \times 550}{450} \right)^9 = 28702 \text{ cycles}$$

2.3.5. ROLLING WEAR

Similar to sliding and impact wear, a model for rolling wear is developed as (Bayer, 2004):

$$W = K N^m N_r \quad (2.80)$$

where N is load, N_r number of revolutions, K the wear coefficient, and m needs to be evaluated from experiment.

However, there are models for rolling wear life based on the observation of macro-surface damage such as cracks, spalls or pits. The rolling elements can be in pure rolling contact or rolling with sliding (such as slip). Figure 2.39 shows a schematic of two parallel cylinders in rolling-sliding contact. Bayer (2004) presents a form of equation for both rolling and rolling-sliding wear that relates the maximum contact stress, σ , to the number of revolutions for initiation of surface damage, N_r :

$$N_r = C \sigma^m \quad (2.81)$$

where C and m are constants and are determined from experiment.

This equation is similar to the life-stress equation (S-N curve) in fatigue, which indicates the damage mechanism in this type of wear is dominated by fatigue wear. Model parameters C and m depend on the materials of mating surfaces, surface conditions and the slide-to-roll ratio. The slide-to-roll ratio has a significant influence on the model parameters and, therefore, on the life of the rolling elements. This ratio is defined as the difference in surface speed between two contacting surfaces divided by the velocity sum:

$$\frac{R_1 \omega_1 - R_2 \omega_2}{R_1 \omega_1 + R_2 \omega_2} \quad (2.82)$$

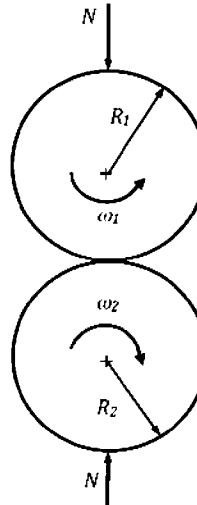


Figure 2.39: Two parallel cylinders in rolling-sliding contact

Figure 2.40 qualitatively shows the effect of slide-to-roll ratio on the wear life of the rolling elements. It can be seen that the slope of the curves may also change as the slide-to-roll ratio changes.

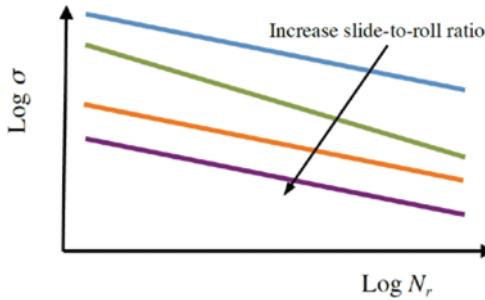


Figure 2.40: Effect of slide-to-roll ratio on life curves

2.3.6. LIFE MODELS FOR BEARINGS

Rolling bearings are used in a variety of applications. They are generally grouped into ball bearings and roller bearings. Each group can be further divided into radial (journal) bearings, axial (thrust) bearings, and a combination of these two. Bearings are used in different sizes and geometries based on application. Generally, they consist of a rolling element (ball or cylinder), inner race, outer race, shoulders, and a cage to keep the roller in position. Bearings have several different failure mechanisms. Common failure mechanisms are listed in Table 2.4.

Fatigue properties of the bearing materials are determining factors in the life prediction of rolling element bearings (Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994). It is not practical to predict the fatigue life of an individual bearing via a precise relationship between life, load, and design characteristics. This is due to the large number of design parameters that need to be considered in relation to the sensitivity of the operating environment.

Table 2.4: Typical bearing failure mechanisms (adopted from the Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994)

Failure mechanism	Description
Spalling	Subsurface chipping or breaking by loading exceeding design load
Fatigue or peeling	Surface cracking or peeling due to poor lubrication or surface damage
Smeared	Surface damage from unlubricated sliding contact in bearing
Brinelling	Indentation of a rolling element under excessive load or impact
Fretting	Caused by an improper fit between the bearing and the shaft
Scuffing	Caused by bearing exposure to an excessive load for a long time
Corrosion	Caused by the chemical reaction between the acids in the lubricants and the base metals in the Babbitt

L_{10} is the standard load carrying capacity where 90% of all bearings have a life greater than 10^6 cycles. Note that L_{10} indicates life of the bearing. Therefore, correction factors are used to adjust the L_{10} life to the actual conditions being projected. For the same set of reference conditions, the life in million cycles is given as (Bayer, 2004):

$$L_{10} = \left(\frac{C}{P}\right)^n \quad (2.83)$$

In the above equation, C is the basic dynamic load rating, P is the equivalent radial load, and constant n is 3 for ball bearings and 3.3 for roller bearings. For bearings subjected to combined axial (thrust) and radial forces, the equivalent radial load, P , is:

$$P = XF_r + YF_a \quad (2.84)$$

where factors X and Y depend on the bearing type and design. Their values for different bearings are given in the Wear Control Handbook (1980), page 706.

For non-standard conditions, several empirical modifying factors are introduced to determine the actual bearing life as (Bayer, 2004):

$$L = D \times E \times F \times G \times H \times L_{10} \quad (2.85)$$

where

D : multiplying factor for materials

E : multiplying factor for material processing and finishing

F : multiplying factor for lubrication

G : multiplying factor for speed

H : multiplying factor for alignment

The basic dynamic load rating, C , can be obtained from the standard formulae or the relevant manufacturing catalog. For example, dynamic load ratings of radial ball bearings are given by following formulae:

$$\begin{aligned} C &= f_c(i \cos \alpha)^{0.7} Z^{0.67} D_w^{1.8}, & D_w \leq 25.4 \text{ mm} \\ C &= 3.647 f_c(i \cos \alpha)^{0.7} Z^{0.67} D_w^{1.4}, & D_w > 25.4 \text{ mm} \end{aligned} \quad (2.86)$$

where i is the number of rows of ball bearings, Z is the number of balls per row, α is the angle between ball and its race (see Figure 2.41) and D_w is the diameter of the balls. The load capacity factor, f_c , is a function of $D_w \cos(\alpha) / d_m$, and is given in for radial ball bearings. It is to be noted that d_m is the pitch diameter as shown in Figure 2.41. As mentioned above, the multiplying factors of bearing life prediction models need to be determined experimentally. However, there are empirical formulae available for some of the factors such as load, lubricant, and water contamination. The multiplying factor for load L_{10} is:

$$\begin{aligned} L_{10} &= \left(\frac{C}{P}\right)^3 \quad \text{For ball bearings} \\ L_{10} &= \left(\frac{C}{P}\right)^{3.3} \quad \text{For roller bearings} \end{aligned} \quad (2.87)$$

The multiplying factor for lubrication depends on the bearing lubricant viscosity (Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994):

$$F = \left(\frac{\nu_0}{\nu_L}\right)^{0.54} \quad (2.88)$$

where ν_0 is the viscosity of the lubricant in standard testing condition and ν_L is the viscosity of the lubricant used in actual application.

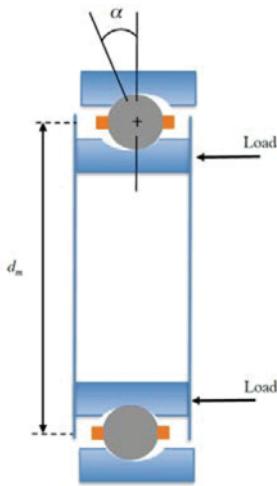


Figure 2.41: Schematic of a ball bearing with a contact angle between ball and its race

In addition to the damage factors (such as static overload, fatigue, wear, corrosion, lubricant failure, and overheating), water contamination can also be detrimental to bearing fatigue life (Armstrong, et al., 1977). A water contamination multiplying factor can account for fatigue life reduction due to water leakage into the lubrication (Hindhede, et al., 1983):

Table 2.5: Values of f_c for radial and angular ball bearings (adopted from Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994)

$D_w \cos(\alpha) / d_m$	Single row radial contact; single and double row angular contact (mm)	Double row radial contact (mm)
0.05	46.7	44.2
0.06	49.1	46.5
0.07	51.1	48.4
0.08	52.8	50.0
0.09	54.3	51.4
0.10	55.5	52.6
0.12	57.5	54.5
0.14	58.8	55.7
0.16	59.6	56.5
0.18	59.9	56.8
0.20	59.9	56.8
0.22	59.6	56.5
0.24	59.0	55.9
0.26	58.2	55.1
0.28	57.1	54.1
0.30	56.0	53.0
0.32	54.6	51.8
0.34	53.2	50.4
0.36	51.7	48.9
0.38	50.0	47.4
0.40	48.4	45.8

$$C_w = 1.04 + 1.03CW - 0.065CW^2 \quad (2.89)$$

where C_w is the percentage of water in the lubricant.

Example 2.14: Fatigue Life of Ball Bearing

A double row angular contact ball bearing has 12 balls (with diameter of 4.97 mm) and a 30 mm pitch diameter. The contact angle is 15 degrees. The bearing is designed to carry 250 N thrust load. The expected life of the bearing (in hours) needs to be determined at rotational speed of 50,000 rpm.

Solution

First, the load capacity factor, f_c , needs to be calculated.

$$\frac{D_w \cos \alpha}{d_m} = \frac{4.97 \cos(15^\circ)}{30} = 0.16$$

From, $f_c = 56.5 \text{ mm}$. Therefore:

$$\begin{aligned} C &= f_c (i \cos \alpha)^{0.7} Z^{0.67} D_w^{1.8} \\ &= 56.5 [2 \cos(15^\circ)]^{0.7} (12)^{0.67} (4.97)^{1.8} \\ &= 2411 \text{ N} \end{aligned}$$

From the L_{10} equation for ball bearings:

$$\begin{aligned} L_{10} &= \left(\frac{C}{P}\right)^3 \\ &= \left(\frac{2411}{250}\right)^3 \\ &= 897 \text{ million cycles} \\ &= \frac{897 \times 10^6}{50,000 \times 60} = 299 \text{ hr} \end{aligned}$$

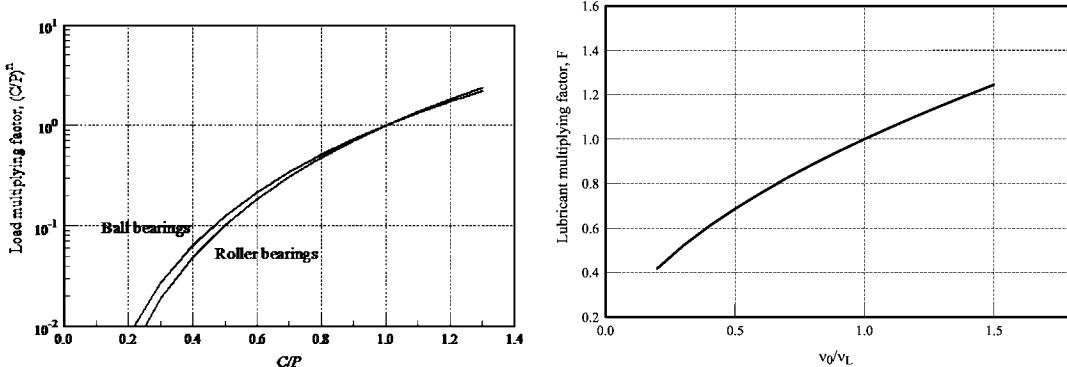


Figure 2.42: Bearing life multiplying factors, a) load factor b) lubricant factor

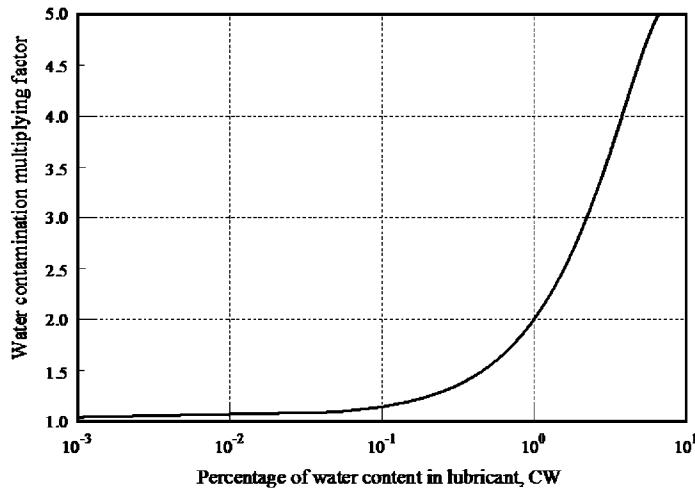


Figure 2.42: c) water contamination factor

2.3.7. LIFE MODELS FOR SEALS

Seals' primary failure mechanism is wear. They are primarily used to restrict lubricant flow from one side of the seal to the other (i.e. to prevent leakage). Figure 2.43 shows a schematic of static and dynamic seals. In dynamic seals, direct contact between the seal surface and the rotating part causes friction, wear, and frictional heating. Dynamic seal reliability and parameters that affect seal life are discussed below. For a comprehensive review of static seal life models, refer to the Handbook of Reliability Prediction Procedures for Mechanical Equipment (1994), Chapter 3.

Abrasive particles in the fluid cause seal wear during operation. Hard particles can become embedded in soft elastomeric and metal surfaces, leading to abrasion of the harder mating surfaces forming the seal, ultimately resulting in leakage. Several different types of wear are responsible for seal failure. Common seal wear mechanisms are listed in Table 2.6.

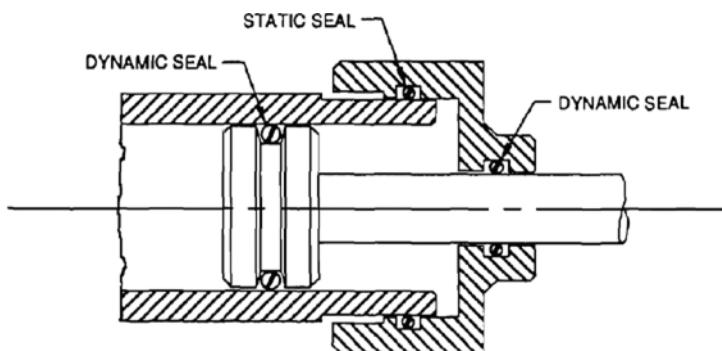


Figure 2.43: Static and dynamic seals (source: Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994)

Some other failure mechanisms beyond those in Table 2.6 can accelerate seal failure. These mechanisms include dynamic instability due to misalignment, contaminant embrittlement, fluid/seal incompatibility, and thermal degradation.

Table 2.6: Typical bearing failure mechanisms (adopted from Wear Control Handbook, 1980)

Wear type	Description
Adhesive	Dominant wear type in seals caused by solid contact wear during startup, shutdown, and operating perturbation. Adhesive wear is mild in seals. The PV criterion is used in design to identify the limit of mild wear transition.
Abrasive	Determines the service life of an entire system like an aircraft engine; It is mainly caused by particles coming to the seal from operating environment, from within the sealed system, from mechanical components, etc.
Corrosive	Commonly observed in seals that are exposed to corrosive working fluid, the frictional heat generated by sliding promotes chemical reactions, which are further accelerated by contact pressure. The rate of chemical reactions doubles for every 10 °C increase in temperature.
Pitting or fatigue	Pitting is usually associated with fatigue. Other causes include oxidative erosion on carbons, carbon blistering, and thermal stress cracks.
Impact	Caused by dynamic instability of seal moving normal to the interface. Very high vibration and acceleration forces can be generated.
Fretting	Observed commonly on the secondary sealing surfaces; fretting can increase total seal leakage significantly in 200 hr in a piston ring secondary seal in a gas seal.

The *PV* test procedure (explained in section 2.3.2) is popularly used to evaluate seal face material performance and reliability, giving a measure of adhesive wear. In a seal *PV* test, face and working fluid pressure are two important parameters. By multiplying these parameters, the *PV* factor can be expressed as (Wear Control Handbook, 1980):

$$PV = [\Delta P(b - k) + P_{sp}]V_m \quad (2.90)$$

where PV = pressure \times velocity

ΔP = pressure differential across seal face

b = seal balance, the ratio of hydraulic closing area to seal face area

k = pressure gradient factor

P_{sp} = mechanical spring pressure

V_m = fluid velocity at the seal mean face diameter

The pressure gradient factor k is assumed to be:

- 0.5 for water-base solutions
- 0.3 for liquids such as light hydrocarbons
- 0.7 for lubricating oils

In a *PV* test, frictional power is detrimental to seal life and should be considered in reliability analysis. Frictional power increases seal face material temperature. Excessive heat causes thermal distortions on the seal face and gaps, which can increase the leakage rate (Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994). Heat also causes material change that can significantly increase the seal wear rate. The frictional power can be calculated from:

$$Q_s = \mu_f(PV)A \quad (2.91)$$

where Q_s = frictional power (heat) input to the seal

PV = pressure \times velocity

μ_f = coefficient of friction

A = apparent area of seal face

The coefficients of friction are obtained from original design tests with water as the lubricant, and are reported in Table 2.7 for some commonly used materials. It is, however, suggested in the Wear Control Handbook (1980) that the values could be 25 to 50% when oils are the lubricant. This is due to the additional drag force caused by oil viscosity. Conversely, introducing lubrication grooves, hydropads, and surface texture on the face of one of the sealing rings can reduce the coefficient of friction. There is no universally accepted PV test procedure. To maximize seal reliability, one should consider tradeoffs when selecting seal material. For example, some materials, such as solid silicon carbide, have better abrasion and thermal shock resistance than carbon-graphite based material. However, solid silicon carbide generates five times more heat than carbon-graphite materials. In order to compensate for excessive heat generation and the associated seal face temperature rise, a cooling flow should be supplied to maintain the film thickness on the dynamic seal faces to prevent increased wear due to higher surface temperatures.

Since adhesive wear is the predominant mechanism of seal wear, Archard's wear equation can be used:

$$W = k \frac{NL}{H} \quad (2.92)$$

Table 2.7: Coefficient of friction for various seal materials (adopted from Wear Control Handbook, 1980)

Sliding Materials		Coefficient of friction, μ_f
Rotating	Stationary	
Carbon-Graphite (Resin Filled)	Cast Iron	0.07
	Ceramic	0.07
	Tungsten Carbide	0.07
	Silicon Carbide	0.02
	Silicon Carbide, Converted Carbon	0.015
Silicon Carbide	Tungsten Carbide	0.02
Silicon Carbide, Converted Carbon	Silicon Carbide, Converted Carbon	0.05
Silicon Carbide	Silicon Carbide	0.02
Tungsten Carbide	Tungsten Carbide	0.08

This equation can be modified to express wear coefficient k in terms of PV factor. To do so, the sliding distance, L , can be expressed as velocity \times time ($L = Vt$), load N can be expressed in terms of pressure ($N = PA$), and linear wear h can be written as $h = W/A$. Therefore,

$$k = \frac{hH}{tPV} \quad (2.93)$$

The effect of different testing conditions and materials on wear rate is established through a series of PV tests. Figure 2.44 shows a schematic of wear rate vs. PV factor for a typical seal life of two years.

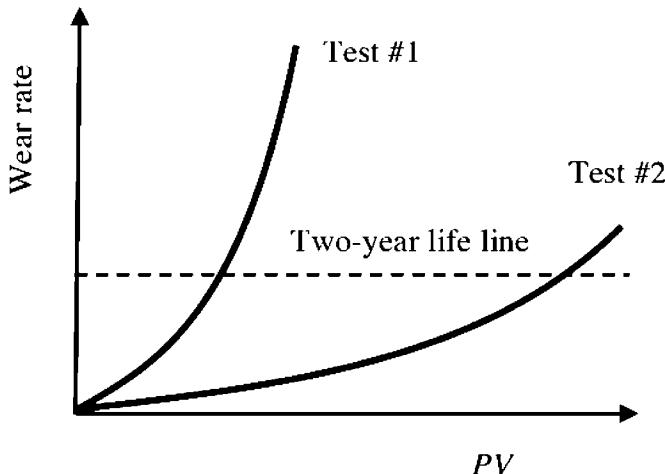


Figure 2.44: Wear rate vs. PV factor in seal tests (Adopted form Wear Control Handbook, 1980)

Similarly, to reliability considerations for bearings discussed previous sections, several modifying factors can be introduced to account for non-standard testing conditions (Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994):

$$L = C_Q \times C_H \times C_F \times C_V \times C_T \times C_N \times C_{PV} \times L_{base} \quad (2.94)$$

where

L = actual failure rate of a seal in failures/million hr

L_{base} = base failure rate of a seal in failures/million hr

C_Q = Multiplying factor for the effect of allowable leakage

C_H = Multiplying factor for the effect of contact stress and seal hardness

C_F = Multiplying factor for the effect of surface finish

C_V = Multiplying factor for the effect of fluid viscosity

C_t = Multiplying factor for the effect of temperature

C_N = Multiplying factor for the effect of contaminants

C_{PV} = Multiplying factor for the effect of the pressure-velocity coefficient

C_Q is determined as follows:

$$\begin{aligned} \text{For leakage, } Q_f \text{ (cm}^3/\text{min}) \leq 0.49, \quad C_Q = 4.1 - (4.82 \times Q_f) \\ \text{For leakage, } Q_f \text{ (cm}^3/\text{min}) > 0.49, \quad C_Q = 0.90/Q_f \end{aligned} \quad (2.95)$$

C_H is determined as follows:

$$C_H = \left(\frac{M}{0.55C} \right)^{4.3} \quad (2.96)$$

where

M = Meyer hardness, MPa

C = Contact pressure, MPa

C_F is determined as follows:

$$\begin{aligned} \text{For surface finish, } f \text{ (microns)} \leq 0.38, \quad C_F = 0.25 \\ \text{For surface finish, } f \text{ (microns)} > 0.38, \quad C_F = 1.21f^{1.65} \end{aligned} \quad (2.97)$$

C_T is determined as follows:

$$C_T = \frac{1}{t^2} \quad (2.98)$$

where

$$t = \frac{T_R - T_0}{10}$$

T_R = Rated temperature of seal, °C

T_0 = Operating temperature of seal, °C

C_{PV} is determined as follows:

$$C_{PV} = \frac{PV_{op}}{PV_{ds}} \quad (2.99)$$

where

PV_{op} = PV factor for the original design

PV_{ds} = PV factor for actual seal operation

The multiplying factors for C_v and C_N are given in Table 2.8 and Table 2.9, respectively.

Table 2.8: Fluid viscosity multiplying factor (adopted from Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994)

FLUID	C _v						
	Fluid Temperature, °C						
	-46	-18	10	38	66	93	121
MIL-H-83282	0.6	0.7	0.8	0.9	1	2	3
MIL-H-5606	0.7	0.8	0.85	0.9	1	2	...
Aircraft Phosphate Esters	...	0.8	0.85	1	1	2	...
Industrial Phosphate Esters	0.7	0.8	0.85	0.9	...
Water Glycol	...	0.7	0.8	0.8	0.9
SAE 10 Oil	0.8	0.8	1	1	2
SAE 60 Oil	0.7	0.7	0.8	0.85	1

Table 2.9: Contaminant multiplying factor (adopted from Handbook of Reliability Prediction Procedures for Mechanical Equipment, 1994)

Typical Quantities of Particles Produced by Hydraulic Components	Particle Material	Number Particles Under 10 Micron Per Hour Per Rated GPM (N ₁₀)
Piston Pump	Steel	1.7
Gear Pump	Steel	1.9
Vane Pump	Steel	0.6
Cylinder	Steel	0.8
Sliding Action Valve	Steel	0.04
Hose	Rubber	0.13

2.3.8. WEAR OF LUBRICATED CONTACTS

In most engineering applications, contacting bodies are separated (partially or completely) by a layer of lubricant. Therefore, friction and wear characteristics of the tribosystem are greatly influenced by the thickness, mechanical and chemical properties of this lubricant layer. Lubrication of sliding components such as bearings, gears, and shafts is of great interest in designing for reliability, although it adds complexity to component design. Once lubricant is degraded it cannot separate two parts efficiently, which results in increasing direct solid contact and ultimately causing failure by seizure. One important lubricant property in determining its performance is viscosity, μ . Figure 2.45 shows two plates with area A , separated by a layer of lubricant of thickness h . The relative velocity of two plates is U , and the force required to keep this velocity constant is F . Viscosity is a measure of fluid resistance against motion and is defined as:

$$\mu = \frac{Fh}{AU} \quad (2.100)$$

The lubricant film thickness plays an important role in the generation of frictional force between contacting surfaces. A plot of coefficient of friction μ_f vs. the Sommerfeld number ($\mu U/N$) exhibits

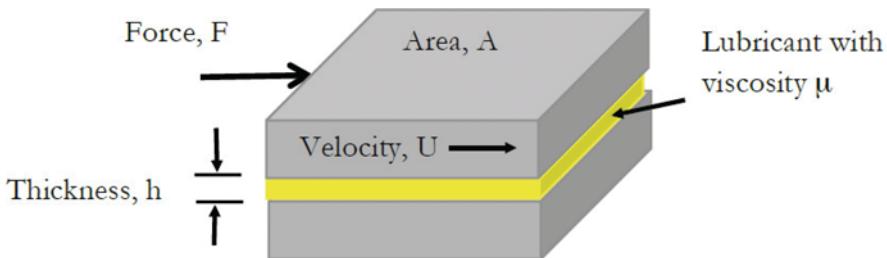


Figure 2.45: Schematic for the definition of viscosity

a minimum point in friction for a variety of testing conditions. This curve, known as the Stribeck curve (Figure 2.46), encompasses three distinct regions: boundary lubrication, mixed lubrication and hydrodynamic (or full film) lubrication. In the boundary lubrication region, the average height of the asperities is larger than the lubricant film thickness. The coefficient of friction in boundary lubrication region can reach values typical of dry sliding contact. In the mixed lubrication region, the lubricant layer separates surfaces. However, there is still asperity contact from mating surfaces. As the film thickness increases, surfaces completely separate from each other. The coefficient of friction increases due to increase in fluid friction, which is related to increase in viscous dissipation.

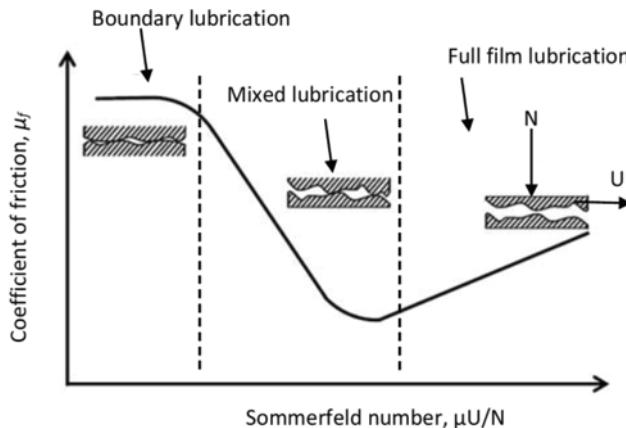


Figure 2.46: Stribeck curve

2.3.9. LUBRICATED WEAR AND LUBRICANT LIFE

Archard's wear law quantifies adhesive wear volume for non-lubricated metal contacts:

$$W = k \frac{NL}{H} \quad (2.101)$$

For lubricated contact, Archard's wear law is used with a modified wear coefficient ($K = \alpha k$) that accounts for effect of lubrication on wear mitigation (Rowe, 1966):

$$W = (\alpha k) \frac{NL}{H} = K \frac{NL}{H} \quad (2.102)$$

In this equation, α is the fraction of true area of metal on metal contact. The values of lubricated wear coefficient, K , are tabulated throughout the literature for different test conditions and lubricants (such as page 143 of the Wear Control Handbook). Regardless of test conditions (that include applied load,

speed, and contacting materials), lubricant chemical composition significantly affects the wear coefficient. Rowe (in Wear Control Handbook, 1980) states that lubricants with a higher energy of desorption or higher heat of adsorption are more effective in reducing wear. Many lubricant additives are excessively reactive, which, in turn, promotes chemical wear in addition to adhesive wear. Therefore, an effective lubricant with additive should minimize both adhesive and chemical wear.

One important and determining factor in lubricated wear is the surface temperature in a sliding contact. Temperature rise can significantly decrease the life and efficiency of the lubricant. For example, chemical reaction rate doubles by each 10°C increase in temperature, which results in accelerated aging of the oil. Oil reacts chemically with oxygen to increase acidity and viscosity, darken its color, and create varnish-like surface deposits (Booser, 1997). Standard laboratory tests for evaluating oxidation life of oils, such as ASTM D943, ALCOR deposition test, and Federal Test Method Standard 791, show that the oxidation life of lubricating oils decreases by half for every 10°C increase in the range of 100°C to 150°C (Khonsari and Booser, 2001). If L denotes the oxidation life in hours, the following Arrhenius type equation can be used to estimate oil life (Khonsari and Booser, 2001):

$$\log_{10} L = k_1 + \frac{4750}{T} \quad (2.103)$$

where T is temperature ($T = {}^{\circ}\text{C} + 273$). Typical values of k_1 for some industrial lubricants are given in Table 2.10.

Table 2.10: Typical values for k_1 in oxidation life prediction of industrial mineral oils (adopted from Booser, 1997)

Oil type	k_1	Maximum temperature at the indicated life, ${}^{\circ}\text{C}$			
		1 hour	100 hour	1000 hour	10,000 hour
Uninhabited	-10.64	135	103	75	51
EP gear lubricant	-10.31	147	113	84	59
Hydraulic	-8.76	214	168	131	99
Turbine	-8.45	230	182	142	106
Heavily refined, hydrocracked	-8.05	252	200	157	121

Example 2.15: Oxidation Life of Lubricating Oils

What is the useful lifetime of hydraulic and turbine oils at 135°C ? Also plot the useful lifetime for temperature ranging from 100 to 400°C .

Solution

From Table 2.10, $k_1 = -8.76$ for hydraulic oil, gives oxidation life of:

$$\begin{aligned} \log_{10} L &= -8.76 + \frac{4750}{(135 + 273)} = 2.88 \\ L &= 10^{2.88} = 759 \text{ hr} \end{aligned}$$

Similar calculation for turbine oil results in $L = 1549$ hr. Figure 2.47 shows life expectancy for hydraulic and turbine engines as a function of temperature in the range of 100 to 400°C .

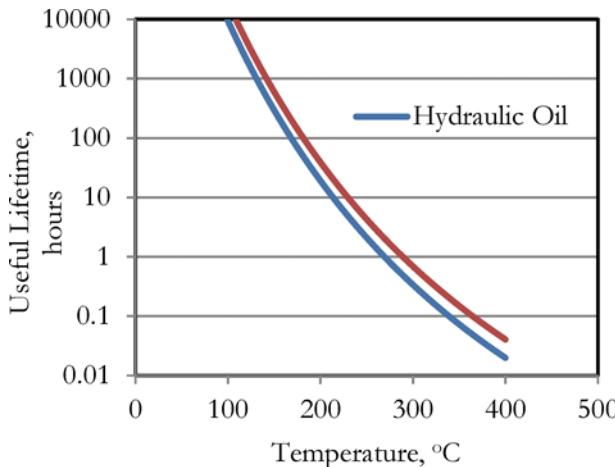


Figure 2.47: Expected useful life of hydraulic and turbine oils as a function of temperature

2.4. CREEP

The previous sections described the common methods used for fatigue analysis. These sections are an important exercise in appreciating the science behind the life-stress models and how they are developed or selected for the purposes of accelerated life testing. The use of fatigue as an example was done for its simplicity. However, there are many other failure mechanisms that are applicable to accelerated life tests. The next to be discussed in detail is creep. Creep is defined as the progressive accumulation of plastic strain in a specimen under stress at elevated temperature over a period of time (Collins, 1993). Creep failure occurs when the accumulated creep strain results in a deformation of the system or component part that exceeds the design limits. It is a predominant failure mechanism that is applicable to all types of materials in many engineering applications, especially those involving high temperature, such as steam turbines in power plants, jet and rocket engines, and nuclear reactors. Some other mundane examples include failure of light bulb filaments and gradual loosening of bolts.

A typical creep test involves subjecting a specimen to a constant load or stress while maintaining a constant temperature. Strain (or deformation) is then measured and plotted as a function of elapsed time (Callister Jr, 2007). Figure 2.48 shows a schematic of the typical constant load creep behavior of materials. As shown, when load is applied to the material, there is an instantaneous deformation, which is mostly elastic. There are three regions in this creep curve. Region I is the primary, or transient creep region, characterized by a decreasing creep rate (i.e. decreasing slope of the curve as time increases). This region suggests that the material is experiencing an increase in creep resistance, also known as strain hardening, where deformation requires more effort as the material is strained. A steady state creep characterizes region II, where the rate of creep is constant (i.e., the creep curve is linear). This region is often the longest stage of creep. The linear nature of the rate of creep that happens in this region is due to the balance between competing processes of strain hardening and recovery, whereby the material becomes softer and retains its ability to experience deformation. Finally, region 3 of the curve is known as tertiary creep, and is characterized by an accelerated rate of creep and ultimate failure/rupture, resulting from microstructural changes, such as grain boundary separation, formation of cracks, cavities, and voids.

Figure 2.48 also shows the effect of both temperature and stress on creep characteristics. With increasing stress or temperature, there is (1) an increase in instantaneous strain at the time of stress application, (2) an increase in the steady state creep rate, and (3) a decrease in rupture life.

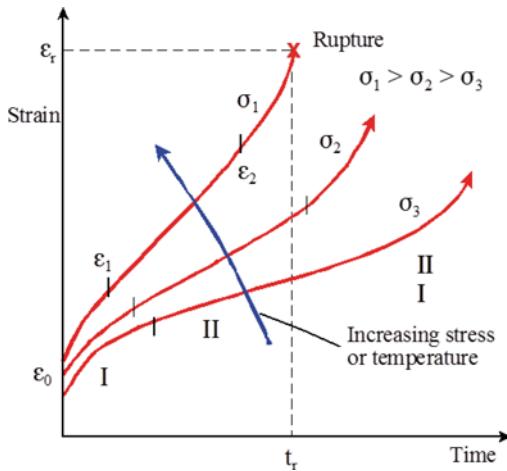


Figure 2.48: Schematic of creep curves for different stress levels or temperatures

The most reliable creep data are those obtained from tests that duplicate actual loadings and temperatures at normal use conditions. Unfortunately, and for practical reasons, it is often not feasible to wait for this type of data to be used as input into the design of components, due to time constraints. Therefore, there has been significant effort invested in devising accelerated creep tests that provide accurate and reliable predictions of long-term creep behavior. One of the ways this is done is via thermal acceleration. The thermal acceleration method involves testing a component for creep behavior at temperatures much higher than the actual use level temperatures applicable to the item under test. Figure 2.49 shows a plot of such a test, stress vs. time, for a family of constant temperatures where the creep strain produced is constant for the whole plot. Given the following plot, the data can then be extrapolated to the design life. The point at which the design temperature curve intersects the design life indicates the appropriate design stress level.

Several accelerated life theories have been proposed to correlate the results of short-time elevated temperature tests with long-term service performance at more moderate temperatures. The more accurate and useful of these proposals to date are the Larson-Miller theory and the Manson-Haferd theory.

2.4.1. LARSON-MILLER THEORY

The Larson-Miller theory postulates that for each combination of material and stress level, there exists a unique value of a parameter P that is related to temperature and time by the equation:

$$P = (\theta + 460)(C + \log_{10} t) \quad (2.104)$$

where

P = Larson-Miller parameter, constant for a given material and stress level

θ = temperature ($^{\circ}$ F)

C = constant, usually assumed to be 20

t = time in hours to rupture or to reach a specified value of creep strain

By using the Larson-Miller relationship, it is a simple matter of finding a short-term combination of temperature and time that is equivalent to any desired long-term service requirement. For instance, for any given material at a specified stress level, the test conditions listed in Table 2.11 should be equivalent to the operating conditions (Collins, 1993). The Larson-Miller parameter has been shown

to be effective for a wide variety of materials in predicting long term creep behavior and stress rupture performance.

Table 2.11: Equivalent conditions based on Larson-Miller parameter

Operating Condition	Equivalent Test Condition
10,000 hours at 1000°F	13 hours at 1200°F
1,000 hours at 1200°F	12 hours at 1350°F
1,000 hours at 1350°F	12 hours at 1500°F
1,000 hours at 300°F	2.2 hours at 400°F

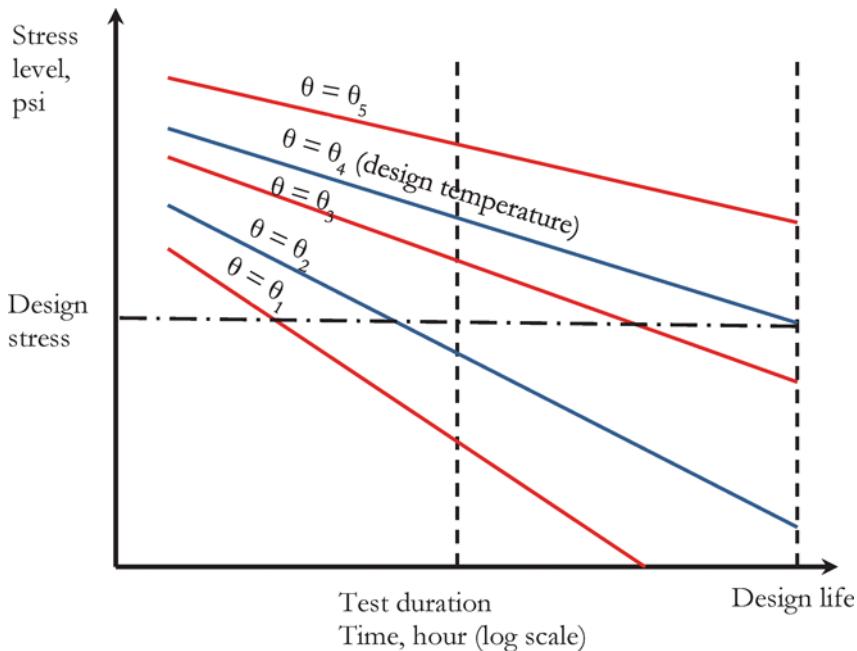


Figure 2.49: Thermal acceleration method for creep testing, Collins (1993)

2.4.2. MANSON-HAFERD THEORY

The Manson-Haferd theory postulates that for a given material and stress level, there exists a unique value of a parameter P' that is related to temperature and time by the equation:

$$P' = \frac{\theta - \theta_a}{\log_{10} t - \log_{10} t_a} \quad (2.105)$$

where

- P' = Manson-Haferd parameter, constant for a given material and stress level
- θ = temperature (°F)
- t = time in hours to rupture or to reach a specified value of creep strain
- θ_a, t_a = material constants

Note that the Manson-Haferd parameter introduces material constants. Once these constants have been evaluated, the above equation (like the L-M parameter) has been shown to give good agreement with experimental results. Table 2.12 shows Manson-Haferd constants for various materials.

Table 2.12: Various constants for the Manson-Haferd equation

Material	Creep or Rupture	θ_a	$\log_{10} t_a$
25-20 stainless steel	Rupture	100	14
18-8 stainless steel	Rupture	100	15
S-590 alloy	Rupture	0	21
DM steel	Rupture	100	22
Inconel X	Rupture	100	24
Nimonic 80	Rupture	100	17
Nimonic 80	0.2 percent plastic strain	100	17
Nimonic 80	0.1 percent plastic strain	100	17

Example 2.16

Consider the creep rupture data from a series of tests on Cr-Mo-V steel shown in Table 2.13.

- (a) Using data in Table 2.13, plot the creep rupture curves for the various temperatures on a single plot of stress versus log-time.
- (b) Using the data for a stress level of 70 ksi at a temperature of 900 °F, predict the time to rupture at the same stress level if the temperature is 1100 °F, using the Larson-Miller parameter.
- (c) Compare predictions from part (b) with the data from part (a).

Solution

Using MS-Excel, the plot of the data provided in Table 2.13 is shown in Figure 2.50.

- (b) From the problem statement, we are given:

$$\theta = 900^{\circ}\text{F}$$

$$C = 20$$

$$t = 9878 \text{ hrs}$$

Therefore, using the Larson-Miller parameter yields:

$$\begin{aligned} P &= (\theta + 460)(C + \log_{10} t) \\ &= (900 + 460)(20 + \log_{10} 9878) = 32632.75 \end{aligned}$$

For $P = 32632.75$, at a temperature of 1100 °F, we can solve the L-M equation for t to find the time to rupture at the same stress level of 70 ksi, as follows:

$$32632.75 = (1100 + 460)(20 + \log_{10} t)$$

Solving the above for t yields:

$$t \approx 8.3 \text{ hrs}$$

- (c) Note from Table 2.13 that at 1100 °F, at a stress level of 70 ksi, an instance of stress rupture of 1 hr was recorded which differs significantly from the Larson-Miller estimate calculated at part (b). However, from the plot at part (a), it appears that the data point at 70 ksi and 1100 °F may

potentially be an outlier due to the lack of fit with the other data at this temperature. The fitted line shown in Figure 2.50 for 1100°F actually considers the first point as an outlier, and estimates a rupture failure time at 70 ksi of between 7 and 8 hours, which agrees more closely with the estimate in part (b).

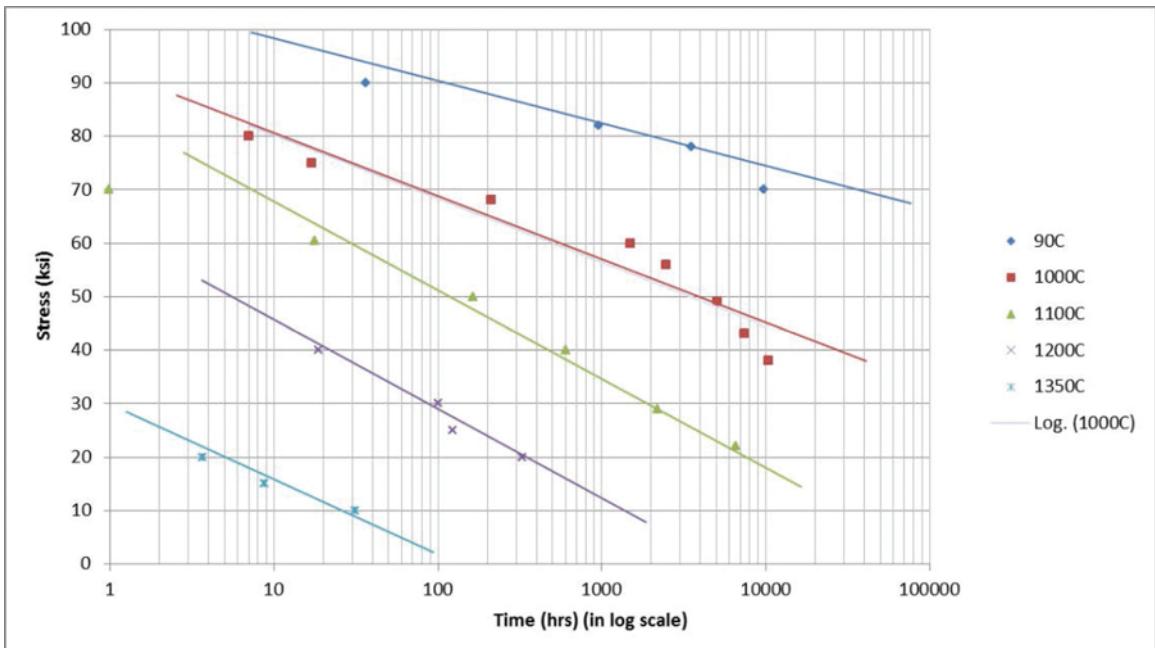


Figure 2.50: Creep rupture curves for Cr-Mo-V steel

2.4.3. CREEP UNDER UNIAXIAL STATE OF STRESS

For metallic materials, most creep tests are conducted in uniaxial tension. On the other hand, uniaxial compression tests are more appropriate for brittle materials, as they provide a better measure of the intrinsic creep properties since there is no stress amplification and crack propagation in tensile loads (Callister Jr, 2007). Uniaxial creep and stress rupture tests of 100 hours (4 days), 1000 hours (42 days), and 10,000 hours (420 days) are common, with longer duration testing of 100,000 hours (11.5 years) being performed in a few cases. Many relationships have been proposed to relate stress, strain, time, and temperature in the creep process. From the experimental creep strain versus time data, it can be observed that the data are close to linear for a wide variety of materials when plotted on log-strain versus log-time coordinates. Such a plot is shown below in Figure 2.51 for three different materials (Collins, 1993). An equation describing this type of behavior is described by:

$$\delta = At^a \quad (2.106)$$

where

$$\delta = \text{true creep strain}$$

$$t = \text{time}$$

$$A, a = \text{empirical constants}$$

Table 2.13: Creep rupture data for Cr-Mo-V steel

Test Temperature (°F)	Stress (ksi)	Rupture Time (hrs)
900	90	37
900	82	975
900	78	3581
900	70	9878
1000	80	7
1000	75	17
1000	68	213
1000	60	1493
1000	56	2491
1000	49	5108
1000	43	7390
1000	38	10447
1100	70	1
1100	60.5	18
1100	50	167
1100	40	615
1100	29	2220
1100	22	6637
1200	40	19
1200	30	102
1200	25	125
1200	20	331
1350	20	3.7
1350	15	8.9
1350	10	31.8

Differentiating the above with respect to time and setting $aA = b$ and $(1 - a) = n$, gives:

$$\dot{\delta} = bt^{-n} \quad (2.107)$$

Here, $-n$ represents the slope of the curve in Figure 2.51 and b represents the intersect. This equation represents a variety of different types of creep strain versus time curves, depending on the magnitude of the exponent n . If n is zero, the behavior is termed *constant creep rate*. This type of creep behavior is most commonly found at high temperatures. If n is 1, the behavior is termed *logarithmic creep*, which is a type of behavior displayed by rubber, glass, and certain types of concrete, as well as metals at lower temperatures. If the exponent lies between 0 and 1, the behavior is termed *parabolic creep*. This type of creep behavior occurs at intermediate and high temperatures.

The influence of stress level on creep rate can often be represented by the empirical expression:

$$\dot{\delta} = B\sigma^N \quad (2.108)$$

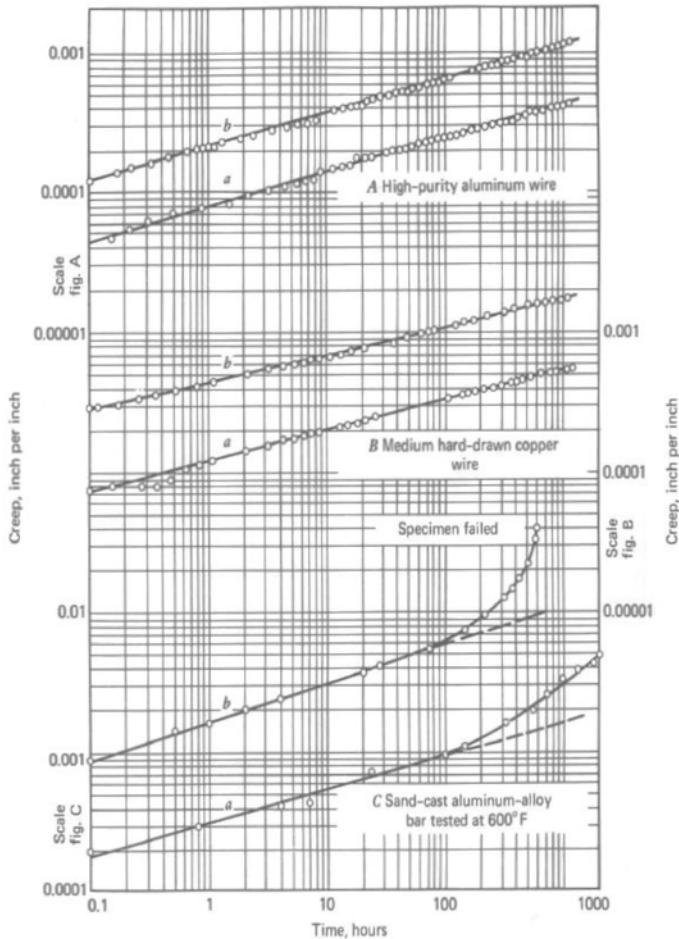


Figure 2.51: Creep curves for three materials plotted on log-log coordinates (From Collins, 1993, with permission-original from Sturm, et. al, 1936)

where B and N are constants. Assuming stress σ is independent of time, the above equation may be integrated to yield the true creep strain as follows:

$$\delta = Bt\sigma^N + C' \quad (2.109)$$

If the constant C' is small compared with $Bt\sigma^N$ (which is often the case), the result is known as the log-log stress-time creep law, given by:

$$\delta = Bt\sigma^N \quad (2.110)$$

With the above equation, one may determine the stress required at a specified temperature to hold creep deformation within specified limits.

Table 2.14: Constants for log-log stress-time creep law

Material	Temperature	$B, (\text{in}^2/\text{lb})^N$ per day	N
1030 steel	750°F	48×10^{-38}	6.9
1040 steel	750°F	16×10^{-46}	8.6
2Ni-0.8Cr-0.4Mo steel	850°F	10×10^{-20}	3.0
12Cr steel	850°F	10×10^{-27}	4.4
12Cr-3W-0.4Mn steel	1020°F	15×10^{-16}	1.9

Example 2.17

A cylindrical pressure vessel of Ni-Cr-Mo steel has closed ends and a diameter of 10 inches. The tank is to operate continuously for 5 years at a temperature of 850 °F under an internal pressure of 10 ksi. If a safety factor of 1.2 is desired, and permanent deformation must not exceed 3 percent, what should the minimum thickness of the cylindrical pressure vessel wall be made?

Solution

From above, we can determine the failure stress σ_f as follows:

$$\sigma_f = \left[\frac{\delta}{Bt} \right]^{1/N}$$

Or

$$\sigma_f = \left[\frac{(0.03 \times 10)/10}{(10 \times 10^{-20})(5 \times 365)} \right]^{1/3} \cong 54780 \text{ psi}$$

where the constant B and N are obtained from

Table 2.14. The design stress based on desired safety factor is then given by:

$$\sigma_d = \frac{\sigma_f}{n} = \frac{54780}{1.2} = 45650 \text{ psi}$$

The critical stress applicable to this cylindrical vessel is the tangential stress:

$$\sigma_d = \frac{pD}{2t}$$

where

p = internal tank pressure

D = diameter of cylinder

t = thickness of cylinder wall

Therefore,

$$45650 = \frac{pD}{2t} = \frac{10000 \times 10}{2 \times t}$$

Solving for t , the required minimum cylinder vessel thickness is:

$$t = 1.095 \text{ inches}$$

2.4.4. CUMULATIVE CREEP PREDICTION

There is currently no universally accepted method for estimating creep strain accumulation as a function of time at different temperatures and stress levels. A linear hypothesis suggested by E.L. Robinson is one of the simplest methods proposed. It follows that if a design limit of creep strain δ_D is specified, it is predicted that this creep strain will be reached when (Collins, 1993):

$$\sum_{i=1}^k \frac{t_i}{L_i} = 1 \quad (2.111)$$

where

t_i = time to exposure at the i 'th combination of stress level and temperature.

L_i = time required to produce creep strain δ_D if entire exposure were held constant at the i 'th combination of stress level and temperature.

Example 2.18

A sand-cast aluminum alloy that has been creep tested at 600 °F at a stress level of 10 ksi results in the data shown as curve *b* in Figure 2.51. A hollow cylindrical support bracket of this material has a 2.0-inch outside diameter, a 0.175-inch wall thickness, and is 5.0 inches long. The bracket is subjected to a direct tensile load of 5 tons. Each ambient temperature cycle is 400 °F for 1000 hours, then 600 °F for 3 hours, and then 800 °F for 0.5 minute.

- (a) If the design requires that the bracket must not elongate more than 0.025 inches, how many cycles would you predict could be survived by the bracket?
- (b) If the bracket were internally pressurized at 1750 psi using a clever end-sealing arrangement so that the pressure produced no axial component of stress in the bracket wall, would you predict that the bracket would achieve a longer or shorter life before failure, according to the criterion of part (a)?

Solution

Part A

As given, one ambient temperature cycle is:

$$t_A = 1000 \text{ hours at } 400 \text{ °F}$$

$$t_B = 3 \text{ hours at } 600 \text{ °F}$$

$$t_C = 30 \text{ seconds at } 800 \text{ °F}$$

The limiting design elongation is given as:

$$\Delta L = 0.025 \text{ inch}$$

Hence the limiting design strain is:

$$\delta_D = \frac{\Delta L}{L} = \frac{0.025}{5} = 0.005 \frac{\text{in}}{\text{in}}$$

Using the Larson-Miller parameter, as discussed in Section 2.4.1, to convert t_A and t_C to 600 °F equivalent times of exposure,

$$P_A = (400 + 460)(20 + \log_{10} 1000) = 19780$$

Hence, for 600 °F operation, the equivalent time t_{A-eq} may be calculated from the L-M equation as:

$$19780 = (600 + 460)(20 + \log_{10} t_{A-eq})$$

Therefore,

$$t_{A-eq} \cong 0.04575 \text{ hr} = 165 \text{ seconds at } 600 \text{ °F}$$

likewise,

$$P_C = (800 + 460)(20 + \log_{10} 0.0083) \cong 22580$$

Therefore,

$$22580 = (600 + 460)(20 + \log_{10} t_{C-eq})$$

solving for t_{C-eq} yields:

$$t_{C-eq} \cong 20.04 \text{ hrs at } 600 \text{ °F}$$

Reading from Figure 2.51, curve b using $\delta_D = 0.005 \text{ in/in}$ and noting that 5 short tons = 10,000 pounds, the equivalent stress of the test data, then:

$$L_{0.005} = 50 \text{ hours}$$

Next, using the Robinson hypothesis, and setting N equal to the number of repeated temperature cycles required to produce the limiting design strain δ_D ,

$$N \left[\frac{0.04575}{50} + \frac{3}{50} + \frac{20.04}{50} \right] = 1$$

Therefore,

$$N = 2.16 \text{ cycles}$$

to produce the limiting elongation.

Part B

From part A, we obtained a life of 2.16 cycles with each cycle being 1003.0083 hours. Therefore, the expected life based on the design criteria is 2166.5 hours = 90.27 days.

Now, if we have a hollow cylindrical bracket that is internally pressurized at 1750 psi, the critical stress in this case would be the tangential stress created by the internal pressure.

$$\sigma_d = \frac{pD}{2t} = \frac{1750 \times 2}{2 \times 0.175} = 10000 \text{ psi}$$

which is greater than the applied tensile load of 5 tons. Therefore, it is expected that the bracket would achieve a shorter life before failure. To make this work, some other design criteria need to be considered, such as type of material used and thickness of the cylinder wall.

2.5. CORROSION

In this section, another important degradation mechanism that is commonly observed in practice, corrosion, will be discussed. In order to better understand corrosion, some basic concepts will be

introduced, and different corrosion types and factors that influence its rate will be reviewed. Also, corrosion testing methods and techniques for diagnosing corrosion failure are included.

Corrosion has been under scientific research for almost 150 years. The electrochemistry of corrosion is well studied in a long list of publications. This section will not cover the fundamentals of electrochemical kinetics of corrosion, but rather its association with structural health management. Corrosion is the degradation of a material (usually a metal) due to chemical or electrochemical reaction with its environment. Degradation is usually in the form of change in properties and/or material loss from the surface as the metal is dissolved and leaves the surface in the form of ions or other kind of corrosion product. According to a study by the National Association of Corrosion Engineers, NACE, the cost of corrosion exceeded \$1 trillion in the United States in 2013, approximately 6.1% of GDP. The study revealed that, although corrosion management has improved over the past several decades, the U.S. must find more and better ways to encourage, support, and implement optimal corrosion control practices.

Corrosion mechanisms and corrosion damage are related to the electrochemistry and thermodynamics of the corrosion process. This is because almost all corrosion processes of metals involve two or more electrochemical reactions with transfer of electron charge in aqueous solution. Electrochemical reactions require anodes and cathodes in electrical contact and an ionic conduction path through an electrolyte (Davis, 2000). However, the focus of this section is not on the electrochemistry and thermodynamics of corrosion; rather, we study prediction models for corrosion rates and the factors affecting its severity. Corrosion rate is one of the most important input parameters in models for corrosion-induced damage prediction. To have an accurate service life assessment of structures with corrosion damage, the corrosion rate needs to be accurately assessed. Corrosion-life prediction models basically predict time to different forms of corrosion damage such as crack covering, loss of metal cross section area, and loss of stiffness (Liu and Weyers, 1998; El Maaddawy and Soudki, 2007; Torres-Acosta et al., 2007; Zhang et al., 2009). Corrosion damage prediction, thus, relies on the field variables chosen to describe the anticipated degradation (e.g., pitting cracks, loss of stiffness). Therefore, it seems necessary for a modeler to first study different forms of corrosion damage and identify their underlying causes and effects. In what follows, we briefly discuss some commonly observed forms of corrosion damage. After that, we discuss available models for assessment of corrosion rate and its influencing factors.

Corrosion damage can be divided into several categories based on the appearance of the corrosion damage or the mechanism of attack. Jones (1996) has classified corrosion damage into nine distinct forms as shown in Figure 2.52 (taken from Jones (1996) with permission) and listed as follows:

- Uniform or general corrosion
- Galvanic corrosion
- Crevice corrosion
- Pitting corrosion
- Environmentally assisted cracking, including stress corrosion cracking, fatigue corrosion and hydrogen-induced cracking
- Hydrogen damage
- Intergranular corrosion
- Dealloying
- Erosion-corrosion, including fretting-corrosion and cavitation-corrosion

Uniform or general corrosion: As inferred from its name, uniform corrosion results in a uniform removal of metal over the entire exposed surface or over a large area. This is the most commonly observed form of corrosion; however, it is not a great concern from technical viewpoint because it can be accurately assessed via comparatively simple immersion tests (Davis, 2000). Uniform corrosion damage can be measured through weight loss and reduction of metal thickness over time.

Galvanic corrosion: This form of corrosion occurs when two dissimilar metals or alloys are electrically coupled (galvanic coupling) in a corrosive electrolyte. The metal that is less resistive against corrosion corrodes while the other is preferentially protected from corrosion. The potential difference between the metals in galvanic coupling affects the corrosion rate.

Crevice corrosion: Narrow gaps between two metals or one metal and one nonmetal cause regions of stagnant solutions, and attack the material. This form of attack is defined as crevice corrosion. Examples can be listed as crevice corrosion damage occurring at bolt holes, gaskets, riveted seams, threaded joints, lap joints, washers, and O-rings (Davis, 2000).

Pitting corrosion: Pitting is a form of extremely localized corrosion damage that results in holes on the surface which may be small or large in diameters, deep or shallow, undercut, close to each other or isolated. Deep pits can penetrate through the wall of the component and cause leaking. According to Davis (2000), pitting corrosion is one of the most difficult forms of corrosion damage to detect because it can cause only a small mass loss and, if small in size, it can be easily overlooked.

Stress corrosion cracking: This is a form of cracking damage in presence of static tensile stress and corrosive environment. An example of this form of damage is pressure vessels containing corrosive media. This form of cracking includes crack nucleation and propagation. The crack propagation phase includes three stages: 1) rapid increase in crack growth rate; 2) slow crack growth rate (crack growth rate is independent of stress intensity factor); and 3), rapid growth rate before catastrophic failure (Davis, 2000).

Corrosion fatigue: This form of corrosion occurs when metal is under cyclic mechanical load or deformation while its surface is in contact with corrosive fluid. It is a synergistic effect of plastic deformation and anodic dissolution current density in the surface layer of metal. The presence of a corrosive environment reduces the number of cycles to failure than would be the case in corrosion free fatigue damage. Examples are boiler tubes, airframe structures operating in sea environment, and medical implants.

Erosion corrosion: This form of damage is a combination of corrosion damage accelerated or increased by the high flow velocity of a corrosive fluid or by mechanical wear. The high velocity fluid basically erodes the protective corrosion product film, resulting in accelerated corrosion damage. Therefore, the combination of erosion and corrosion damage is more severe than one of them alone and is generally recognized by the appearance of grooves, waves and rounded holes on the surface of metal. Examples of erosion corrosion damage are piping systems, pumps, impellers, and turbine blades.

2.5.1. MODELS FOR PREDICTION OF CORROSION RATE AND SERVICE LIFE

In general, the available models for prediction of corrosion degradation can be divided into two groups: 1) approaches that assess corrosion rate (expressed as the corrosion current density, i_{corr}) and 2) approaches that assess corrosion-induced damage (e.g., pitting cracks, loss of stiffness, etc). Otieno et al., (2011) present a review of available prediction models for reinforced concrete (RC) structures. In this section, however, both groups will be covered, as we present available models and factors that influence severity and rate of corrosion. Some of the corrosion rate prediction models (i_{corr}) are explained in detail as follows.

Stern-Gary's model: This model is applicable to the linear regions of polarization in which slight changes in current applied to corroding metal in an ionic solution cause corresponding changes in potential of metals. Stern and Gary (1957) experimentally showed that for a simple corroding system the corrosion rate is inversely related to the polarization resistance, R_p as:

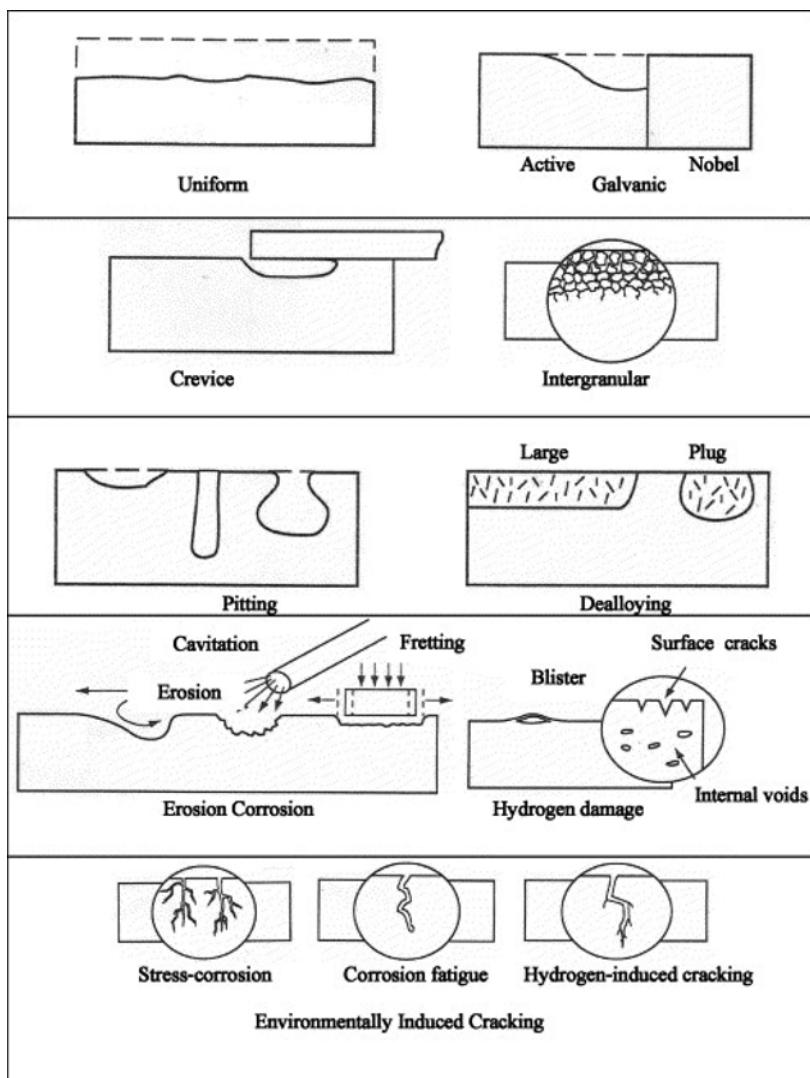


Figure 2.52: Various forms of corrosion (Jones, 1996; with permission)

$$i_{corr} = \frac{B}{R_p} \quad (2.112)$$

where i_{corr} = corrosion current density in A/cm^2 ; B = Stern-Gary constant in V; and R_p = polarization resistance in $\Omega \cdot \text{cm}^2$. The value of B is determined from the particular electrochemical cell and generally ranges from 13 to 52 mV depending on the system: a value of 26 mV is commonly used for steels embedded in concrete for active corrosion (Andrade and Gonzalez, 1978). The Stern-Gary model has been criticized for its simplicity; however, from a practical viewpoint, it is still widely used in the field of corrosion science.

Yalcyn and Ergun's model: This model was developed to study the effect of chloride and acetate ions on i_{corr} from accelerated corrosion testing (Yalcyn and Ergun, 1987).

$$i_{corr} = i_0 e^{-Ct} \quad (2.113)$$

where i_{corr} = corrosion rate at time t ; i_0 = initial corrosion rate; and C = corrosion constant that depends on the degree of concrete pore saturation. Yalcyn and Ergun's model has some shortcomings, including that it is based on accelerated corrosion tests and is not proved for natural corrosion tests; it was developed for uncracked concretes and may not be applicable to cracked concretes; and factors affecting i_{corr} , such as temperature, cover depth, and concrete resistivity, are not included in the model.

Liu and Weyers' model: As mentioned above, the corrosion rate can be influenced by factors such as changes of temperature, chloride content, resistance of concrete, and corrosion period. Liu and Weyers performed a statistical analysis on experimental results obtained from a five-year accelerated corrosion test on 44 uncracked bridge deck slabs with 2927 corrosion rate measurements from seven series of chloride contaminated specimens. The corrosion rates were correlated with concrete ohmic resistance, temperature, chloride content and exposure time by using a non-linear regression model as follows:

$$i_{corr} = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 CL) - \frac{3034}{T} - 0.000105 R_c + 2.32 t^{-0.215} \right) [\mu\text{A}/\text{cm}^2] \quad (2.114)$$

where CL is the free chloride content (kg/m^3); T is temperature at the depth of steel surface (K); R_c is the ohmic resistance of the cover concrete (Ω); and t is time from initiation of corrosion (years).

Example 2.19

Consider a block of concrete reinforced with steel bars. Find the corrosion rate using Liu and Weyers' model for the following conditions:

- a) with chloride content of 1.8 (kg/m^3) and ohmic resistance of 1500 (Ω), compare the corrosion rate after 1 year for environments with temperatures of 5 °C and 30 °C.
- b) with ohmic resistance of 1500 (Ω) and temperature of 13 °C, compare the corrosion rate after 1 year for environments with chloride content of 1 (kg/m^3) and 4 (kg/m^3).
- c) with chloride content of 1.8 (kg/m^3), ohmic resistance of 1500 (Ω), and temperature of 13 °C, compare the corrosion rate after 1 year and 5 years.

Solution

Using the Liu and Weyers' model as given above:

- a) $CL = 1.8 (\text{kg}/\text{m}^3)$, $R_c = 1500 (\Omega)$, $t = 1 (\text{year})$, $T_1 = 5 + 273.15 = 278.15 (\text{K})$:

$$(i_{corr})_1 = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 \times 1.8) - \frac{3034}{278.15} - 0.000105 \times 1500 + 2.32 \times 1^{-0.215} \right) = 1.257 [\mu\text{A}/\text{cm}^2]$$

$$T_1 = 30 + 273.15 = 303.15 (\text{K}):$$

$$(i_{corr})_2 = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 \times 1.8) - \frac{3034}{303.15} - 0.000105 \times 1500 + 2.32 \times 1^{-0.215} \right) = 3.09 [\mu\text{A}/\text{cm}^2]$$

$$\frac{(i_{corr})_2}{(i_{corr})_1} = \frac{3.09}{1.257} = 2.46$$

- b) $R_c = 1500 (\Omega)$, $T = 13 + 273.15 = 286.15 (\text{K})$, $t = 1 (\text{year})$
 $(CL)_1 = 1 (\text{kg/m}^3)$:

$$(i_{corr})_1 = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 \times 1) - \frac{3034}{286.15} - 0.000105 \times 1500 + 2.32 \times 1^{-0.215} \right) = 1.185 [\mu\text{A}/\text{cm}^2]$$

$(CL)_1 = 4 (\text{kg/m}^3)$:

$$(i_{corr})_2 = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 \times 4) - \frac{3034}{286.15} - 0.000105 \times 1500 + 2.32 \times 1^{-0.215} \right) = 2.793 [\mu\text{A}/\text{cm}^2]$$

$$\frac{(i_{corr})_2}{(i_{corr})_1} = \frac{2.793}{1.185} = 2.36$$

- c) $CL = 1.8 (\text{kg/m}^3)$, $R_c = 1500 (\Omega)$, $T = 13 + 273.15 = 286.15 (\text{K})$
 $t_1 = 1 (\text{year})$:

$$(i_{corr})_1 = 0.92 \exp \left(8.37 + 0.618 \ln(1.69 \times 1.8) - \frac{3034}{286.15} - 0.000105 \times 1500 + 2.32 \times 1^{-0.215} \right) = 1.705 [\mu\text{A}/\text{cm}^2]$$

$t_2 = 5 (\text{year})$:

$$(i_{corr})_2 = 0.92 \exp \left[8.37 + 0.618 \ln(1.69 \times 1.8) - \frac{3034}{286.15} - 0.000105 \times 1500 + 2.32 \times 5^{-0.215} \right] = 0.865 [\mu\text{A}/\text{cm}^2]$$

$$\frac{(i_{corr})_2}{(i_{corr})_1} = \frac{0.865}{1.705} = 0.507$$

Having determined the corrosion rate (i_{corr}) from the models discussed above or other available models, metal loss due to corrosion can be estimated. The volumetric loss V_{corr} (or mass loss, M_{corr}) of metal by unit of area and unit of time is expressed in mm/year or in kg/m² year in the case of mass loss. V_{corr} is obtained from the corrosion current (i_{corr}) through Faraday's law and the density of the metal. For uniform corrosion of steel, 1 $\mu\text{A}/\text{cm}^2$ is equivalent to a rate of metal loss of 0.0116 mm/year (Andrade and Alonso, 2004). The volumetric and mass loss of steel due to corrosion can be expressed as:

$$V_{corr} = 0.0116 i_{corr}$$

where V_{corr} is in mm/year and i_{corr} is in $\mu\text{A}/\text{cm}^2$. Mass loss is:

$$M_{corr} = 9.127 i_{corr}$$

where M_{corr} is in kg/m² year and i_{corr} is in $\mu\text{A}/\text{cm}^2$.

Prediction of the service life of reinforced concrete structures is a complicated task because corrosion of steel embedded in concrete is influenced by several environmental factors that are not well known. However, several attempts have been made by researchers to develop prediction models for service life of RCs. Morinaga's model is among many that are discussed here. Morinaga (1990) developed an empirical model to predict the time from corrosion initiation to corrosion cracking when the concrete

cover cracks due to expansion by means of rust formation on rebar surface. The expression to estimate t_{corr} is:

$$t_{corr} = \frac{0.602 D \left(1 + 2 \frac{C}{D}\right)^{0.85}}{I_{corr}}$$

where t_{corr} = the time from corrosion initiation to corrosion cracking (in days); D = the steel bar diameter (mm); C = the clear concrete cover (mm); and I_{corr} = the corrosion rate ($\mu\text{A}/\text{cm}^2$ day). It should be mentioned that Morinaga's model does not account for the mechanical properties of concrete.

Example 2.20

Consider a block of concrete reinforced with steel bars with diameter of 16 mm and clear cover thickness of 20 mm. Find the time for corrosion cracking at the corrosion rate of $i_{corr}=100 (\mu\text{A}/\text{cm}^2)$.

Solution

Since i_{corr} is given in ($\mu\text{A}/\text{cm}^2$), we first convert it to I_{corr} ($\text{gr}/\text{cm}^2 \text{ day}$) via:

$$M_{corr} = 9.127 i_{corr} = 9.127 \times 100 = 912.7 \text{ (kg/m}^2 \text{ year)}$$

$$I_{corr} = M_{corr} \times \frac{1000}{10000 \times 365} = 912.7 \times \frac{1000}{10000 \times 365} = 0.25 \text{ (gr/cm}^2 \text{ day)}$$

Using Morinaga's model as given above,

$$D = 16 \text{ (mm)}, C = 20 \text{ (mm)}, I_{corr} = 0.25 \text{ (gr/cm}^2 \text{ day)}$$

$$t_{corr} = \frac{0.602 D \left(1 + 2 \frac{C}{D}\right)^{0.85}}{I_{corr}} = \frac{0.602 \times 16 \times \left(1 + 2 \times \frac{20}{16}\right)^{0.85}}{0.25} = 112 \text{ days}$$

REFERENCES

- Abd El Latif, A. K., Fatigue behaviour from application of AC., Eng. Fract. Mech. 12 (1979) 449-454.
- Amiri, M., and Khonsari, M.M., Effect of surface cooling on fatigue life improvement. J. Failure Anal. Preven. 13 (2013) 183-187.
- Andrade, C., and Gonzalez, J.A., Quantitative measurements of corrosion rate of reinforcing steels embedded in concrete using polarization resistance measurements. Werkstoff und Korrosion, 29 (1978) 515-519.
- Andrade, C., and Alonso, C., Test methods for on-site corrosion rate measurement of steel reinforcement in concrete by means of the polarization resistance method. Materials and Structures' / Matériaux et Constructions, 37 (2004) 623-643.
- Archard, J. F., Contact and rubbing of flat surfaces, J. Appl. Phys., 24 (1953) 981-988.
- Armstrong, E.L., W.R. Murphy and P.S. Wooding, "Evaluation of Water Accelerated Bearing Fatigue in Oil-Lubricated Ball Bearings," Lubrication Engineering. Vol. 34, No. 1, pp. 15-21 (1 Nov 1977).
- Bannantine, J., Comer, J., & Handrock, J. (1997). Fundamentals of Metal Fatigue Analysis. Prentice Hall.
- Barwell, F. T., Wear of metals, Wear, 1 (1957-1958) 317-332.
- Bayer, R. G., Engineering Design for Wear, Marcel Dekker, Inc. (2004) New York, NY, USA.
- Berge, S., On the effect of plate thickness in fatigue of welds. Engineering Fracture Mechanics. 21 (1985) 423–435.
- Bezborodko, L. G., Effect of alternating current on the strength of steel undergoing cyclic tensile-compressive and torsional loading, Inst. Mech., Acad. Sci. Ukrainian, (1884) 95-98.
- Booser, E.R., Life of oils and greases, Tribology Data Handbook, CRC Press, Boca Raton, FL, pp. 1018-1028, 1997.
- Callister Jr, W. (2007). Materials Science and Engineering - An Introduction (7th ed.). New York: John Wiley & Sons.
- Cao, W., Effect of electric current and field on the behavior of metallic materials, Ph.D. Thesis, North Carolina State University, 1989.
- Coffin, L.F. Jr., A study of the effect of cyclic thermal stresses on a ductile metal. Transactions of ASME, 76 (1954) 931-950.
- Conrad, H., White, J., Cao, W.D., Lu, X.P., Sprecher, A.F., Effect of electric current pulses on fatigue characteristics of polycrystalline copper, Mater. Sci. Eng. A 145 (1991) 1-12.
- Collins, J. (1993). Failure of Materials in Mechanical Design (2nd ed.). New York: John Wiley & Sons.
- Dowling, N. (1998). Mechanical Behaviour of Materials. Upper Saddle River: Prentice Hall.
- Engel, P. A., Failure models for mechanical wear modes & mechanisms, IEEE Transactions of Reliability, 42 (1993) 262-267.
- Gucer, D.E., Capa, M., Effect of loading frequency on the strain behavior and damage accumulation in low-cycle fatigue. AGARD Report No. 527 (1970).

- Handbook of Reliability Prediction Procedures for Mechanical Equipment, CARDEROCKDIV, NSWC-94/L07, March 1994.
- Hindhede, U., Zimmerman, J.R., Hopkins, R.B., Erisman, R.J., Hull, W.C., Lang, J.D., Machine Design Fundamentals, John Wiley & Sons, 1983.
- Hirano, A., Yamamoto, M., Sakaguchi, K., Shoji, T., ida, K., Effects of water flow rate on fatigue life of carbon steel in simulated LWR environment under low strain rate conditions. *J. Pres. Ves. Tech.* 125 (2003) 52-58.
- Karolczuk, A., Macha, E., A review of critical plane orientations in multiaxial fatigue failure criteria of metallic materials, *Int. J. Fract.* 134 (2005) 267-304.
- Karpenko, G.V., Kuzin, O.A., Tkachev, V.I., Rudenko, V.P., The effect of electric current on the low-cycle fatigue of steels, *Dokl. Akad. Nauk SSSR* 227 (1976) 85-86.
- Kato, K., Wear in relation to friction – a review, *Wear* 241 (2000) 151-157.
- Kaynak, C., Ankara, A., Baker, T.J., Effects of short cracks on fatigue life calculations. *Int. J. Fatigue*, 18 (1996) 25–31.
- Khonsari, M.M., Amiri, M., Introduction to Thermodynamics of Mechanical Fatigue. 2012, CRC Press, Taylor & Francis Group, Boca Raton, FL. (166 Pages).
- Khonsari, M.M., Booser, E.R., Applied tribology: bearing design and lubrication, CRC Press, Boca Raton, FL, pp. 477, 2001.
- Kim, J.K., Shim, D.S., The variation in fatigue crack growth due to the thickness effect. *International Journal of Fatigue* 22 (2000) 611–618.
- Kujawski, D., Ellyin, F., Crack initiation and total fatigue life of a carbon steel in vacuum and air. *J. Testing Eval.*, 20 (1992) 391–395.
- Liu, T., and Weyers, R.W., Modeling the dynamic corrosion process in chloride contaminated concrete structures, *Cement and Concrete Research*, 28 (1998) 365–379.
- Ludema, K. C., Selecting materials for wear resistance, Conference on wear of materials 1981, ASME, New York.
- Ludema, K. C., Friction, Wear, Lubrication: A Textbook in Tribology, CRC Press, Inc. Boca Raton, Florida, 1996.
- Majumdar, S., Chopra, O.K., Shack, W.J., Interim failure design curves for carbon, low-alloy, and austenitic stainless steels in LWR environments. In Proceedings of 20th WRSM, Vol. 3, NUREG/CP-0126, 3 March 1993.
- Manson, S.S., Behavior of Materials Under Conditions of Thermal Stress, Heat Transfer Symposium, University of Michigan Engineering Research Institute, (1953) 9–95.
- Martin, W.S., Wirsching, P.H., Fatigue crack initiation–propagation reliability model. *J. Mater. Civil Eng.*, ASCE, 3 (1991) 1–18.
- Morinaga, S., Prediction of service lives of reinforced concrete buildings based on the corrosion rate of reinforcing steel. Proceedings of Building Materials and Components, Brighton, UK, 7–9 November 1990) 5–16.
- Murtaza, G., Akid, R., Modelling short fatigue crack growth in a heat-treated low-alloy steel. *Int. J. Fatigue*, 17 (1995) 207–214.

- Rhee, S.K., Wear equation for polymers sliding against metal surfaces, *Wear*, 16 (1970) 431-445.
- Rowe, C.N., Some aspects of the heat of adsorption in the function of a boundary lubricant, *ASLE Trans.* 9 (1966) 110-111.
- Shigley, J.E., Mischke, C.R., *Mechanical engineering design*. McGraw Hill, 2001.
- Stern, M., and Geary A.L., Electrochemical Polarization No. 1: Theoretical Analysis of the Shape of Polarization Curve, *Journal of Electrochemical Society*, 104 (1957) 56-63.
- Stephens, R., Fatemi, A., Stephens, R., & Fuchs, H. (2001). *Metal Fatigue in Engineering*. New York: John Wiley and Sons.
- Stulen F.B. and Cummings, H.N., A failure criterion for multiaxial fatigue stresses. *Proceedings of the ASTM* 54 (1954) 822–835.
- Sturm, R.G., Dument, C., and Howell, F.M. (1936). A Method of Analyzing Creep Data. *ASME Transactions*, 58: A62.
- Tobushi, H., Hachisuka, T., Yamada, S., Lin, P.H., Rotating-Bending Fatigue of a TiNi Shape-Memory Alloy Wire. *Mech Mater.* 26 (1997) 35-42.
- Yalcyn, H., and Ergun, M., The prediction of corrosion rates of reinforcing steels in concrete. *Cement and Concrete Research*, 26 (1996) 1593-1599.
- Yang, L. J., Prediction of net steady-state wear coefficient in an Al-Al₂O₃(P)/steel system with an integrated wear model, *Tribology Letters.*, 17 (2004) 105-118

Chapter 3: Types of Accelerated Testing and Modeling Concepts

3.1. INTRODUCTION

In this chapter we discuss the types of qualitative and quantitative Accelerated Testing (AT) methods used by practicing engineers and elaborate on the general techniques and models used in analyzing such accelerated test results. An AT provides diverse information about an item's performance and its failure mechanisms; depending on the objectives of the reliability analysis, the type of AT will differ. Qualitative tests are performed on small samples with the samples tested subjected to a single high level of stress, to multiple stresses or to a time-varying stress to observe and eliminate degradation and failure and to improve reliability and performance of the items without consideration of the amount of the resulting improvement. Quantitative Accelerated Life Testing (ALT), unlike the qualitative testing methods, consists of lower stress level tests to estimate the life characteristics (distribution) of the product, component or system under normal use conditions. As such, the type of AT chosen is governed by the types of the results desired. Both types of accelerated tests will allow the engineer to collect information to increase the reliability of the item. However, if one needs to estimate life characteristics (i.e., mean life, reliability growth, warranty returns) then a quantitative test would be more useful. In reality, if resources are available, it may be advisable to conduct both types: information gained from the qualitative tests can be used to help design the quantitative test.

Quantitative tests and their extended version, known as Accelerated Degradation Tests (ADT), are meant to develop life characteristics relating to the item and require developing life-stress models for estimating the reliability life measures. In the remaining chapters, this book will emphasize quantitative tests and the various analysis methods associated with those tests. However, qualitative tests are discussed at a detailed level in this chapter and later in Chapter 4 to provide readers with an appreciation of the spectrum of different AT methods used in reliability engineering.

3.2. TYPES OF ACCELERATED TESTING – QUALITATIVE AND QUANTITATIVE

Both qualitative and quantitative tests identify and evaluate failure modes and failure mechanisms that cause degradation and damage. While qualitative tests can be done quickly in matters of a few hours to a few days, the quantitative ALTs can take from a few weeks to a few months to complete. An underlying assumption in quantitative ALT is that the unit under stress will exhibit the same behavior when subjected to a high stress for a short period that it would when under use conditions for longer periods of time. Hence, there are several important planning considerations when conducting ALT tests so that this assumption remains valid.

Regardless of the type of AT, there are two ways to accelerate degradation caused by failure mechanisms: **usage rate acceleration**, and **overstress acceleration**. The former works by accelerating the frequency of use while keeping the stress at use-level. For instance, increasing the frequency of switching a piece of equipment on and off will accelerate its rate of failure of products (toaster, lighting system, automobile engine). The overstress acceleration on the other hand, involves the application of relevant stresses that exceed the levels that the item encounters under normal use conditions. These data are then used to extrapolate performance characteristics at use conditions. Examples of these stresses include elevated temperature, humidity, voltage, load, pressure, or vibration. These can be applied singly or in selected combination. Examples of stresses include:

1. Temperature cycling – T_{\max} , T_{\min} , dwell, ramp-times
2. Sustained temperature – temperature and exposure time

3. Humidity – controlled, condensation
4. Corrosion – salt, corrosive gases (NaCl, Cl₂, etc.)
5. Power cycling – duty cycles, power dissipation
6. Electrical loads – voltage, current, current density (static and transient)
7. Electrical noise
8. Mechanical stress – tensile and compressive, static and cyclic
9. Mechanical bending – strain, static and cyclic
10. Random vibration – gravitational acceleration root mean square (G_{rms})
11. Harmonic vibration – G_{rms} and frequency
12. Mechanical shock
13. Combinations of above stresses

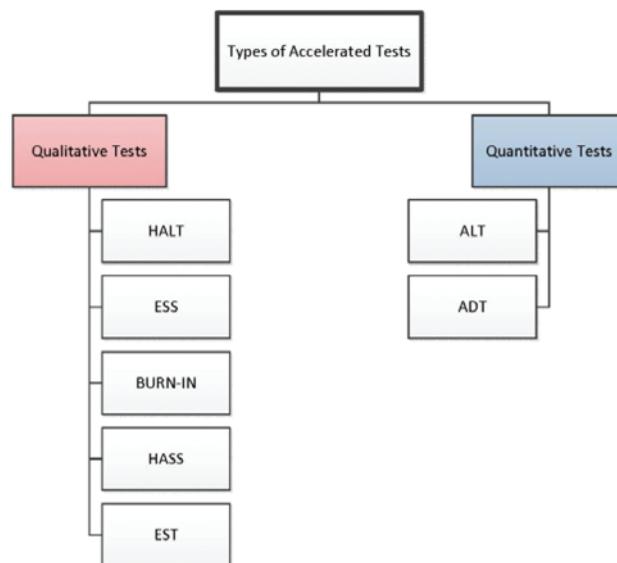


Figure 3.1: Types of accelerated testing

Figure 3.1 illustrates specific important qualitative and quantitative ATs that will be discussed in more detail in the following sections.

3.3. QUALITATIVE ACCELERATED TESTS

Qualitative accelerated tests, sometimes referred to as Reliability Enhancement Tests (RETs), are designed to find failures linked to design, manufacturing and handling, without providing the life characteristics associated with the items. By accelerating the failures of such items at a very high stress level, RETs can determine the reliability robustness of the item during its normal operating condition (i.e., use condition). When a failure occurs during a qualitative AT, we need to determine the root causes of the failure and assess whether the failure mode would occur under normal use conditions. The most common type of RET is the Highly Accelerated Life Test (HALT). It is important to note

that the term HALT does not refer to a life test as such, as its purpose is not to determine life characteristics. Rather, it is a test designed to identify the principal failure modes and mechanisms of hardware units (mechanical or electronic) that will occur during the life of the item under common use operating conditions. Another example of a qualitative RET is known as Highly Accelerated Stress Screening (HASS). HASS tests are applied during the manufacturing phase and are used to prevent marginal and defective units from being released during the process (Levin & Kalal, 2003).

The following subsections outline in more detail the most common qualitative and quantitative ATs shown in Figure 3.1.

3.3.1. ENVIRONMENTAL STRESS TESTING

An Environmental Stress Test (EST) is designed to accelerate the failure of an item, motivated by the pressure to quickly develop new, highly reliable products. One example of an EST is known as STRIFE (life-stress) testing. STRIFE testing involves aggressively stressing an item's prototype to force failures, and usually involves only one or two units. A typical STRIFE test starts with a stress profile (e.g., temperature cycling) just below the designed operating specification. The amplitude of the cycling is increased continuously until the end of the test. When possible, use rate may also be increased.

When there is a failure in a STRIFE test, it is necessary to study the failure's root cause, and assess whether the failure could actually occur within the item's normal use. It is important to avoid possible costly efforts to eliminate high-stress induced failure modes that could never occur during actual use. Since the results of STRIFE testing are used to make changes on the product design and manufacturing process, it would be difficult and improper to use the test data to predict what will happen during normal use conditions. For instance, if a newly designed automobile engine were to run continuously at high RPM to simulate 50,000 miles of service, the main risk in interpreting results from such a test is that the assumed acceleration factors¹ may be significantly inaccurate. Hence it should be noted that different failure modes have different acceleration factors. For instance, an AT may accurately predict a wear mechanism, but not at all identify a corrosion mechanism. Similarly, humidity may accelerate corrosion mechanism, but could potentially reduce the wear rate.

3.3.2. BURN-IN TESTING

The most common reliability problem for manufacturers and consumers of electronic equipment has been early burn-in (infant mortality) failures. The hazard rate of most items exhibit the so-called bathtub curve, which shows that the average early instantaneous failure rate of an item is often higher than the failure rate of those items during their useful life. This is why it is sometimes preferable to buy a nearly new item, as early failures due to manufacturing and handling causing failure may have been removed. Theoretically, one could completely avoid infant mortality by operating the item before release for a considerable period (e.g., months). This is clearly impractical for many obvious reasons, involving accumulated damage, cost, and time-to-market. Therefore, if one can accelerate the item's early life through a burn-in test, then those items with manufacturing or material flaws may be removed and infant mortality hazard rate will be bypassed. Hence, the burn-in process is designed to accelerate operating life of the item so that early failures occur before the item is released. It is applied to components and final products as a final acceptance test. For components like integrated circuits, it is common to conduct burn-in at high humidity and temperature. For more complex items and systems, burn-in is necessary to avoid the use of high levels of the accelerating variables to avoid damaging

¹ Acceleration factor is the factor by which the life will be reduced under high stress condition.

sensitive units. The length of a burn-in test is based on the desired level of reliability and observed failures during the screening.

3.3.3. ENVIRONMENTAL STRESS SCREENING

Environmental Stress Screening (ESS) is a common form of a burn-in test. It provides a more economical and more effective means of removing defective units while testing units at the system or subsystem level. Because systems and subsystems cannot tolerate high levels of stress for long periods, ESS uses mild but more complicated application of stresses. This is done, for example, by replacing high levels of temperature and humidity at the component level with more moderate temperature cycling, physical vibration, and perhaps stressful operational regimes (e.g., running computer boards at higher than usual clock speeds). These tests are usually conducted while the item is in service and being monitored. The main difference between a burn-in test and ESS is that the latter induces multiple stresses on the product such as a combination of temperature and vibration cycling.

Some guidelines for ESS activities include:

- Stress conditions must be carefully chosen such that they can best identify manufacturing defects whilst minimizing the chance of damaging good units.
- Design screens should be used to provide feedback for improving product design or the manufacturing process, by reducing the frequency of manufacturing defects (or eliminating them).
- Information that would allow prediction of field reliability should be assessed. The typically complicated stress profiles and patterns make it difficult to use ESS data directly to predict field performance.

It is important to note that, like burn-in, ESS is an inspection/monitoring/screening method. Today, however, most manufacturers deliver reliable parts through modern quality assurance. Furthermore, inspection schemes such as burn-in and ESS tests tend to be expensive, and studies have shown that they are not necessarily very effective in removing early failures (Levin & Kalal, 2003). For instance, research indicated that burn-in at the component level tended to damage more good parts due to overstress than identifying faulty parts. By improving reliability through continuous improvement of the product design and the manufacturing process, it is often possible to reduce or eliminate reliance on screening tests. In some cases ESS is only applied on an audit basis to monitor production quality on an ongoing basis. ESS can be useful for electronic manufacturers having poor field failure experience. ESS should be tailored to meet specific product requirements by applying stresses that lead to the maximum number of failures without affecting the useful life of components. Of course, ESS should pay for itself by removing costly latent failures.

3.3.4. HIGHLY ACCELERATED LIFE TESTING

Highly Accelerated Life Testing (HALT) is a qualitative AT designed to identify field failures before the first product is shipped (Levin & Kalal, 2003). HALT applies very high stresses to an item (i.e., acceleration factors of up to 2500), while it is still in the design phase, revealing imperfections, design errors, and design nonconformities. When these design issues are identified, they can then be remedied through redesign. To verify that the redesign works and that no further design issues have been introduced by the change, the HALT process may be repeated. The end result is a final product that is free from the identified design defects, thereby presenting value through increased reliability. HALT can be used for:

- Sampling inspections

- Burn-in screening tests
- Pilot tests to obtain information required to plan a more extensive ALT (quantitative test)
- Obtaining information on the relevance of failure modes

Some of the benefits that HALT can provide include:

- Very quick observation of failures
- Improved reliability
- Reduced chance of failure and recalls
- Quick identification of possible weak areas during the design
- Determination of multiple failure modes and root causes
- Determination of functional operating limits and destruct limits
- Increasing stresses without needing to correlate to stresses applied in the use conditions

The ultimate intent of HALT is to drive the unit to failure, so that the failure can be formally investigated to determine its root cause. Once these failures are identified, they can be rectified through redesign. A less ambitious but important goal of the HALT is to find the weakest areas of the design in the shortest possible time, and fix them until the fundamental limit of the technology is reached. Finally, HALT is used to compare new designs to proven previous generation designs using the same accelerated test, thus providing an efficient benchmarking test.

The two most common types of stresses applied in the HALT chambers are vibration and temperature. Some other typical stresses applied in this test (in conjunction with vibration and temperature) include:

- Voltage margining (i.e., determining voltage-related safety margin)
- Clock frequency
- AC supply margining
- Humidity
- Load
- Pressure
- Power cycling
- Voltage sequencing

HALT procedures are highly experiential. The analyst decides which stresses to apply based on his or her judgment. However, to date there are no specific set of standards for carrying out the HALT process. Generally, the choice of stresses applied depends on many factors relating to the component's use profile and environment.

3.3.4.1. Summary of HALT Process

HALT does not attempt to simulate the field environment, but only seeks to find design and process flaws. The purpose of HALT is to determine failure modes and mechanisms, but not to demonstrate that a component meets any specified requirements. It may prove difficult to get both management and engineering personnel to accept that HALT results are real indicators of future field failures. HALT is not meant to determine reliability but to improve it without quantifying the amount or degree of improvement.

Key steps associated with HALT process include:

- Conduct a sequence of HALT starting first with a single environment, then again with combined environments for comparison.
- Stress the items to far beyond their design specifications, typically significantly greater than what is involved in traditional quantitative ALT tests, achieving acceleration factors as high as 2500.
- Conduct failure root cause analysis to inform appropriate design changes.
- Conduct further HALT for the purpose of verification. The design change is considered successful if no new failure modes are identified.

One weakness of the HALT process is that test environments may not be realistic and thus could produce questionable results. HALT can also be difficult to conduct for complex structures, with time dependent failure modes not being exposed. Finally, long-term damage mechanisms such as fatigue and corrosion failure mechanisms are difficult to identify through HALT. HALT results then can be used to:

- Inform subsequent redesign activities to improve reliability
- Identify manufacturing process problems needed to design a proper HASS test (to be discussed in Section 3.3.5)
- Identify of load/stress limit conditions
- Address whether identified problems can be reproduced in real-life failure modes and mechanisms.

3.3.5. HIGHLY ACCELERATED STRESS SCREENING

Once the design issues identified by HALT have been rectified through redesign, the designed item is ready for production and manufacturing. Since the manufacturing process itself can introduce failure mechanisms, it is necessary to ensure that such processes are sufficiently robust to avoid field failures. Highly Accelerated Stress Screening (HASS) is used to achieve this objective. While HASS is likened to HALT, it is solely related to the manufacturing process to manage any variations from the original design. HASS identifies manufacturing process related defects and weaknesses.

Both HALT and HASS use similar types of accelerated stress variables to identify failures. However, HASS involves subjecting the item to lower stress levels than those used for HALT. This is to minimize the probability of introducing defects and failure mechanisms while still identifying manufacturing process variations. HASS should not noticeably reduce an item's expected life when it goes into operation after the test. HASS is performed as part of the manufacturing process as a means of product screening and acceptance (Levin & Kalal, 2003).

HASS consists of two main elements: precipitation and detection screens. HASS begins with the former, which applies a stress level that is below the destruct limit and above the operating limit of the product under test. A suitable stress level of temperature is between 80 and 50 percent of the destruct limits, and for vibration, at about 50 per cent of the destruct limits. A Proof of Screen (PoS) activity is used to identify whether the chosen stress levels are too severe or ineffective. The PoS uses production units and executes the HASS profile repeatedly until the item fails. If, for instance, it takes 20 runs of the HASS screen profile to fail the item, then it is reasonable to estimate that 5 percent of the product is removed with each HASS screen. If, for instance, only four runs were required, then it can be assumed that 25 percent of the life of the product was removed at each run. PoS results form the basis of determining final HASS precipitation stress levels. After precipitation screening, a detection screen is conducted. The detection screen applies appropriate stress levels to detect the “bad”

products that have been sufficiently damaged during the precipitation test. Some of the advantages of HASS are:

- Reduced manufacturing defects
- Improved yield rates and quality of the manufacturing process
- Increased item quality
- Faster production and manufacturing
- Improved reliability

HASS is more involved than HALT, because one must conduct HALT first, and operational and overstress limits are somewhat subjective, making it difficult to optimize the effectiveness of the screening process.

Figure 3.2 shows the areas of the bathtub curve where HALT and HASS are applied during the design and manufacturing. As seen in the figure, HALT can extend the life and increase life and postpone arrival of the wear-out time. HASS and HALT can reduce the failure rate by removing or reducing the effects of failure mechanisms early in the life or during the item's life and wear-out. As such HALT and HASS can affect both the life and failure rate axes of the bathtub curve.

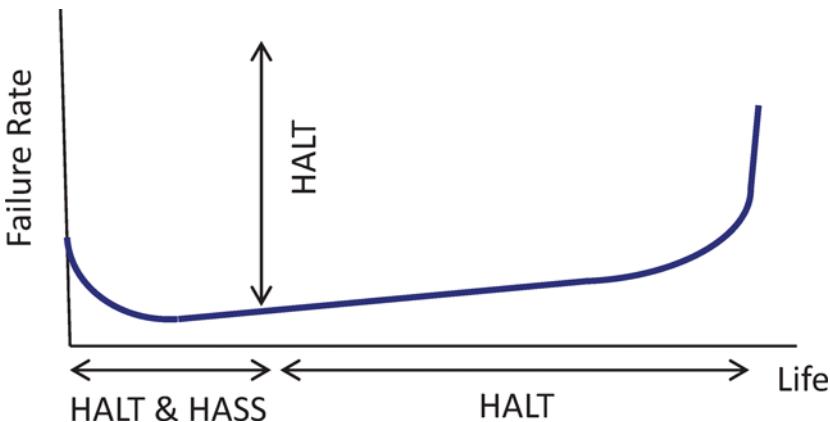


Figure 3.2: Areas of life where HALT and HASS are useful

3.4. QUANTITATIVE ACCELERATED TESTS

ALT is the leading type of quantitative AT. ALT is designed to estimate parameters of mechanistic life-stress models and quantify life characteristics at the use level, by accelerating failures of units by subjecting them to more severe levels of stresses. Unlike qualitative tests, ALT is not primarily intended to expose defects for subsequent redesign, but this can often be a secondary benefit. ALT involves the assumption that the item under test will exhibit the same degradation behavior in a shorter time frame at an elevated stress as it would in “at use” conditions. This concept is illustrated in Figure 3.3 representing the so-called life-stress relationship. The assumption will only be valid assuming either a single failure mechanism or multiple failure mechanisms that synergize in a well understood and controlled way. It is also assumed that competing failure mechanisms are not present. From these failure mechanisms we can determine (either physically or empirically) the corresponding life-stress model to be used to relate the life or life distribution of the item at the various stress levels, including the life distribution at the use stress level.

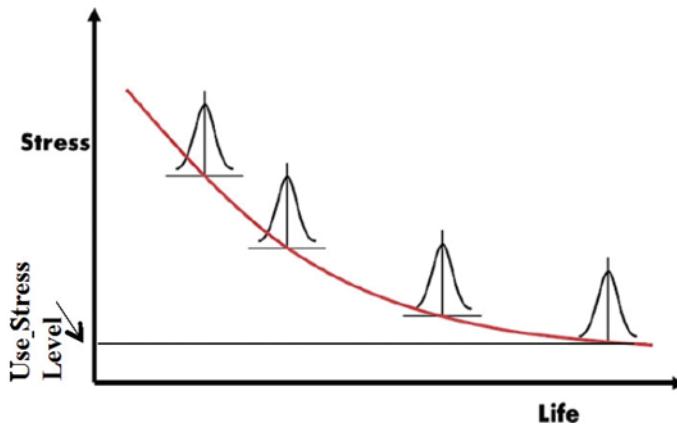


Figure 3.3: Stress vs. life relationship

3.4.1. MODELING DEGRADATION ASSOCIATED WITH VARIOUS FAILURE MECHANISMS

There are three types of PoF modeling based on applied failure mechanisms to an item of interest. These are the stress-strength model, damage-endurance model and performance-requirement model described below.

3.4.1.1. Stress-Strength Model

These types of failures occur when an item is subjected to a stress that exceeds its strength. The assumption behind the stress-strength model for reliability modeling is that so long as the stresses applied to a unit are below its strength, no damage will occur, and it remains as good as new. The overlap between the two distributions represents cases where a random stress value in the high tail of its distribution coincides with the item having a random strength on the weak tail of the strength distribution. Then it is possible that strength falls short of the applied stress and a failure will occur. This is illustrated by the graphical representation in Figure 1.2. This type of failure usually results from temporal random application of high stresses, or from insufficient strength due to poor design, manufacturing or maintenance.

The probability that no failure occurs is equal to the probability that the applied stress is less than the item's strength. That is,

$$R = \Pr(s > l) \quad (3.1)$$

where R is the reliability of the item, l is the applied stress (load), and s is the item's strength.

Examples of stress related failures include application of extra load to a pressure vessel, such as a compressed natural gas container. Engineers need to ensure that the strength of an item substantially exceeds the applied stress under possible stress conditions. Traditionally, in the deterministic design process, safety factors are used to determine the degree of separation (between the means) between these two distributions considering the spectrum of possible applied stresses. This is generally a good engineering principle, but failures occur despite these safety factors due to the overlap of the tails. On the other hand, safety factors that are too stringent result in over design, high cost, and sometimes-poor performance.

If the distribution of s (i.e., strength) and l (i.e., load or stress) can be expressed by the pdfs $f(s)$ and $g(l)$, then

$$R = \int_0^{\infty} g(l)df(s) \quad (3.2)$$

$$F = \int_0^{\infty} f(s) \left[\int_s^{\infty} g(l)dl \right] ds \quad (3.3)$$

According to this model the *safety margin* (*SM*) is defined as

$$SM = \frac{E(s) - E(l)}{\sqrt{var(s) + var(l)}} \quad (3.4)$$

where $E(s)$ and $E(l)$ are expected values of the random variables s and l , and $var(s)$ and $var(l)$ are variances of these random variables. The SM shows the relative difference between the mean values of stress and strength. The larger the *SM*, the more reliable the item will be. Use of (3.3) is a more objective way of measuring the safety margin of an item. It also allows for calculation of reliability and probability of failure as compared with the traditional deterministic approach using safety factors. However, good data on the variability of stress and strength are often not easily available. In these cases, engineering judgment can be used to obtain the distributions of stress and strength including engineering uncertainty.

The distribution of stress is highly influenced by the way the item is used and the internal and external operating environments. The design determines the strength distribution, and the degree of quality control in manufacturing primarily influences the strength variation.

It is easy to show that for a normally distributed s and l , the statistic SM also follows a normal distribution. As such, the reliability of the item, R , can be determined from

$$R = \Phi(SM) \quad (3.5)$$

where $\Phi(SM)$ is the cumulative standard normal distribution of reliability with $z = SM$.

3.4.1.2. Damage-Endurance Model

Although some stresses applied to an item may be below the level that item experiences permanent damage, such as applying monotonic tensile force to the point of exceeding the yield strength in metals, there are cases where applied stress brings about a small but accumulating amount of irreversible damage. The repeated application of these stresses results in the accumulation of damage, until the damage surpasses the endurance of the item. Figure 1.3 shows a depiction of this model with multiple traces of damage starting with an uncertain amount of initial damage (shown by a probability density function) and growing until it exceeds the endurance limit. Each damage accumulation trace produces one instance of time to failure. All such instances result in the time-to-failure distribution of the item as shown in Figure 1.3. At each instant of time, so long as the distribution of cumulative damage does not surpass the endurance limit, no failure will occur. However, when the cumulative damage exceeds the endurance level, a failure will be expected. Examples of application of the damage-endurance model include modeling fatigue, corrosion, wear and creep failure mechanisms, which were discussed in Chapter 2. Since most failure mechanisms exhibit cumulative damage behavior, the damage-endurance model is the most appropriate modeling approach in ALT.

For this model the probability of failure is calculated from the time-to-failure distribution, $f(t)$, which is also related to the distribution of the cumulate damage, $g(\delta)$, according to

$$g(\delta)d\delta = f(t)dt \quad (3.6)$$

Examples of damage include fatigue crack growth in a structure subject to cyclic loading. When the crack size reaches a critical size, the item fails. The critical size is the endurance limit of the item. Two examples are crack propagation and fracture (discussed in Chapter 2), and nuclear reactors where the pressure vessel is damaged by neutron irradiation causing embrittlement. The physical endurance of the pressure vessel reduces over time to a point that it can no longer withstand possible anticipated or unanticipated thermal and mechanical stresses. This would cause rupture of the pressure vessel. It should be noted that as time can affect both the damage level (e.g., increase it) and the endurance limit (e.g., decrease it). Figure 3.4 depicts exceedance of the cumulative damage distribution from endurance distribution, where these two distributions approach each other as a function of time.

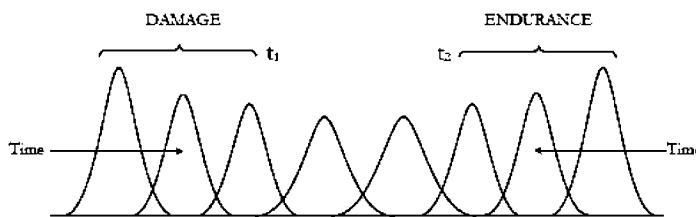


Figure 3.4: Impact of aging over time as described by damage-endurance modeling

From Figure 3.4, the problem is to establish a distribution of damage and endurance at an instant of time. Clearly, as we are unable to reduce all the uncertainties involved in measuring damage and endurance, they are often represented by distributions as shown in Figure 3.4. The mathematical concept determining the failure of an item at a given time is similar to the strength-strength modeling approach, and Equations (3.2) - (3.3) apply to the damage-endurance model at a given age (time) or cycle of operation. Clearly, the “stress” is now represented by “damage” and “strength” is represented by “endurance.” As such, Equations (3.3) and (3.6) would be used to calculate the probability of the item failure at each time, t , as the damage accumulates.

3.4.1.3. Performance-Requirement Model

This type of failure modeling is used for active components and systems and is characterized by a gradual decline in performance (efficiency, output yield, pressure head). When performance falls below the required level of performance, the system is considered to have failed. In items important to safety and integrity, the separation between the requirements of initial performance is defined as the safety margin. Figure 1.4 depicts traces of gradual decline in performance of an item over time until reaches the lowest requirement limit, beyond which the item is considered failed (even though its output may not have ceased). The initial performance can either be known exactly or similar to the damage-endurance example in Figure 3.4; it can start from an uncertain initial amount of performance expressed by a distribution function. The various traces are the results of uncertainties in the performance degradation model (shown as straight lines in Figure 1.4).

ALTs work best when applied to damage-endurance failures, and in some cases, performance-requirement failures. ALTs are generally not suitable for assessing stress-strength models. This is because ALTs are based on accumulated damage or cumulative performance decline assumptions.

Similar to the damage-endurance model, Equations (3.3) and (3.6) may be used to calculate the probability of failure. In the performance-requirement model, the “damage” distribution at a given time in (3.6) is replaced by the “performance” distribution, the likelihood of exceedance of which from the minimum requirement level (or distribution) determines the probability of failure.

Example 3.1

Consider the stress and strength of a beam in a structure represented by the following normal distributions:

$$\mu_s = 420 \text{ kg/cm}^2 \text{ and } \sigma_s = 32 \text{ kg/cm}^2$$

$$\mu_l = 310 \text{ kg/cm}^2 \text{ and } \sigma_l = 72 \text{ kg/cm}^2$$

What is the reliability of this structure?

Solution

$$SM = \frac{420 - 310}{\sqrt{32^2 + 72^2}} = 1.4$$

with $SM = z = 1.4$ and using a standard normal value,

$$R = \Phi(1.4) = 0.9198$$

Example 3.2

A random variable representing the strength of a nuclear power plant containment building follows a lognormal distribution with a mean strength of 0.905 MPa and standard deviation of 0.144 MPa. Three possible accident scenarios can lead to high-pressure conditions inside the containment building that may exceed its strength. The pressures cannot be calculated precisely, but can be represented as another random variable that follows a lognormal distribution.

- a) For a given accident scenario that causes a mean pressure load inside the containment building of 0.575 MPa with a standard deviation of 0.117 MPa, calculate the probability that the containment fails.
- b) If the four scenarios are equally probable and each leads to high-pressure conditions inside the containment with the following mean and standard deviations, calculate the probability that the containment fails.

μ (MPa)	0.575	0.639	0.706	0.646
σ (MPa)	0.117	0.063	0.122	0.061

- c) If the containment strength distribution is divided into the following failure mode contributors with the mean failure pressure and standard deviation indicated, repeat part a.

Failure Location	Mean Pressure, μ_s (MPa)	Standard Deviation, σ_s (MPa)
Liner tear	0.910	1.586E-3
Basemat	0.986	1.586E-3
Cylinder hoop	1.089	9.653E-4
Wall-basemat	1.131	1.586E-3
Cylinder meridional	1.241	1.034E-3
Dome	1.806	9.653E-4
Personnel air lock door	1.241	1.655E-3

Solution

If s is a lognormally distributed random variable representing strength, and l is a normally distributed random variable representing pressure stress (load), then the random variable, $Y = \ln(s) - \ln(l)$, is also normally distributed. For the lognormal distribution with mean and standard deviation of μ_y and σ_y , the respective mean and logarithm standard deviation of the normal distribution, μ_t and σ_t , can be obtained (see Modarres, et. al, 2017).

$$R = \Phi(SM) = \Phi\left(\frac{\mu_{s_t} - \mu_{l_t}}{\sqrt{\sigma_{s_t}^2 + \sigma_{l_t}^2}}\right) = \Phi\left(\frac{-0.112 - (-0.574)}{\sqrt{0.158^2 + 0.201^2}}\right) = 0.9649$$

where Φ is the cumulative normal distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

- a) The probability of containment failure: $F = 1 - R = 0.0351$
- b) Because the four scenarios are equally probable, the system is equivalent to a series system, such that: $R = R_1 \times R_2 \times R_3 \times R_4$.

$$R_1 = \Phi(SM_1) = \Phi(1.81) = 0.9649$$

$$R_2 = \Phi(SM_2) = \Phi(1.83) = 0.9664$$

$$R_3 = \Phi(SM_3) = \Phi(1.07) = 0.8577$$

$$R_4 = \Phi(SM_4) = \Phi(1.79) = 0.9633$$

The probability of containment failure: $F = 1 - R = 1 - R_1 R_2 R_3 R_4 = 0.2296$

- c) Because each failure mode may cause a system failure, this case can be treated as a series system. Because we know the median of the lognormal distribution instead of the mean, it takes several algebra steps to solve for the respective means and standard deviations.

$$R_a = \Phi(SM_a) = \Phi(2.38) = 0.9913$$

$$R_b = \Phi(SM_b) = \Phi(2.78) = 0.9973$$

$$R_c = \Phi(SM_c) = \Phi(3.27) = 0.9995$$

$$R_d = \Phi(SM_d) = \Phi(3.46) = 0.9997$$

$$R_e = \Phi(SM_e) = \Phi(3.92) \approx 1$$

$$R_f = \Phi(SM_f) = \Phi(5.78) \approx 1$$

$$R_g = \Phi(SM_g) = \Phi(3.92) \approx 1$$

The probability of containment failure: $F = 1 - R = 1 - R_a R_b R_c R_d R_e R_f R_g = 0.0122$.

In order to estimate damage and endurance or performance and requirement as a function of time (or number of cycles), we need to understand the failure mechanisms and underlying degradation processes leading to damage and reduction of endurance limits. Many failure mechanisms can be directly linked to the degradation of parts on an item, and PoF models allow one to extrapolate to an

assumed age or failure time based on the measurements of degradation or performance over time. To reduce testing or inspection time further, tests and inspections can be performed at elevated stresses, and the degradation at these elevated stresses can be measured, resulting in a type of analysis known as accelerated degradation.

In some cases, it is possible to directly measure the degradation over time, as with the wear of a bearing or with the propagation of cracks in structures under random loading (causing fatigue crack growth). In other cases, direct measurement of degradation might not be possible without invasive or destructive measurement techniques that would directly affect the subsequent performance of the item. In such cases, the degradation can be estimated by measuring certain performance characteristics, such as using resistance to infer the degradation of a dielectric material. In both cases, it is necessary to define a level of degradation at which a failure will occur, the endurance limit. With this endurance limit, basic mathematical models are used to extrapolate failure probabilities—and thus, the performance measurements—over time to the point where the failure would occur. This is done at different damage levels (caused by applied stresses) assigning stress levels with time or cycles-to-failure. Once the times-to-failure at the corresponding stress levels have been determined, it is merely a matter of analyzing the extrapolated failure times like conventional accelerated time-to-failure data.

Once the level of failure (or the degradation level that would constitute a failure) is defined, the degradation over time should be measured. The uncertainty in the results is directly related to the number of items and amount of data obtained in the observations at each stress level, as well as in the amount of overstressing with respect to the normal operating conditions.

3.4.2. FORMS OF DEGRADATION AND PERFORMANCE MODELS

The degradation and performance of the items subject to AT are measured over time, either continuously or at predetermined intervals. The measured performance informs the extrapolation to use conditions to estimate the failure time using specific models. Common degradation models have the following forms:

$$\begin{aligned} \text{Linear: } & \theta = at + b \\ \text{Exponential: } & \theta = be^{at} \\ \text{Power: } & \theta = bt^a \\ \text{Logarithmic: } & \theta = a \ln(t) + b \end{aligned} \tag{3.7}$$

where θ represents the performance or cumulative degradation of an item, t represents time or number of cycles, and a and b are model parameters. Once the model parameters a and b are estimated from ALT data, the time or number of cycles to failure can be estimated based on endurance levels or maximum sustainable cumulative damage. Extrapolation involves inherent inaccuracies that can be mitigated by using mechanistic models (where possible) or expanding the range of data to limit the extrapolation length.

Combining degradation models is also useful for some applications. For example, a power-exponential model may be expressed as $\theta = bt^{a_1}e^{a_2t}$.

Example 3.3

Assume that the model describing the median cumulative damage follows a power model $\theta = 3.2 \times 10^{-10}t^{3.5}, 0 \leq \theta \leq 1$, where θ is the median cumulative damage and t is time or number of cycles. If the median cumulative damage θ can be represented as a truncated normal distribution bounded by 0 and 10 with a standard deviation of 0.1, and the endurance limit D can be represented

as a Gaussian distribution with a mean of 1 and standard deviation of 0.1, then determine the distribution for the time-to-failure.

Solution

A simple simulation (such as MATLAB) can be used to obtain the time to failure distribution. The objective of the routine is to randomly select endurance limits and cumulative damage curves and determine their intersection points: that is, the times-to-failure. The endurance limit described by a Gaussian or normal distribution $NOR(1, 0.1)$ is illustrated on the y-axis of Figure 3.5.

The median cumulative curves are obtained by taking the inverse value of the truncated normal using the previously defined power model as the mean. When the times-to-failure are all obtained from a sample size of about 10,000, the following distribution is obtained as illustrated in Figure 3.6.

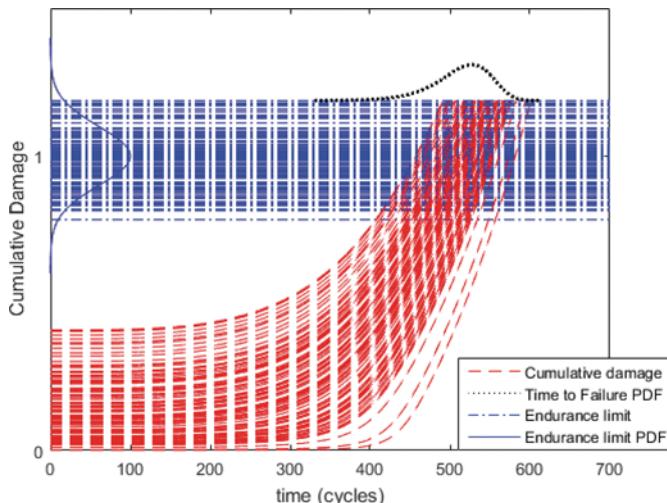


Figure 3.5: Cumulative damage as a truncated normal pdf overlapping the endurance limit described as a Gaussian probability density function

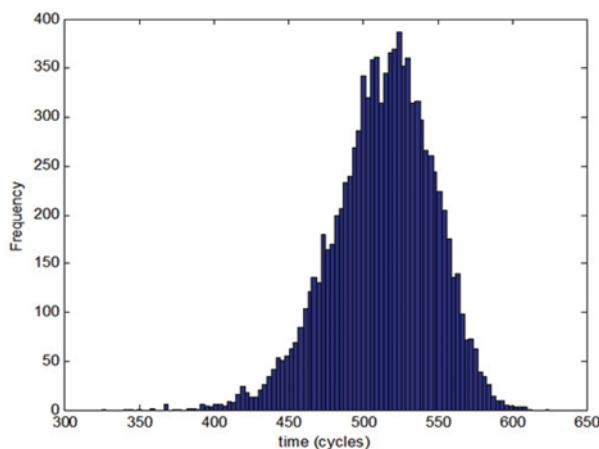


Figure 3.6: Histogram of the time-to-failure distribution

REFERENCES

- Levin, M., & Kalal, T. (2003). *Improving Product Reliability*. Chichester: John Wiley and Sons.
- Modarres, M., Kaminskiy, M., & Krivtsov, V. (2017). *Reliability Engineering and Risk Analysis*, 3rd Ed. Boca Raton: CRC Press.

Chapter 4: Analysis of Accelerated Life Testing Data and Physics-Based Reliability Model Development

4.1. INTRODUCTION

Data from accelerated life tests require analysis to estimate parameters of the life-stress models and associated uncertainties. Such analysis allows life characteristics under normal operating conditions to be predicted. For a variety of reasons, engineers may wish to obtain reliability results for their products more quickly than they can obtain the data under normal operating conditions. In this chapter, formal quantitative methods to select a life-stress model and proper applications of such models will be discussed.

4.2. ACCELERATED LIFE DATA ANALYSIS METHODS

In Chapter 2 and in this chapter the physics of failure concepts based on the study of failure mechanisms are discussed. The main motivation for this is to understand how the physics behind the failure mechanism being observed could be used to model and plan tests for accelerating failures in items. Doing this allows us to construct a model which defines the relationship between the item's life at a given stress level and the item's expected life at other stress levels. Recall that in typical life data analysis, the objective is to establish a life distribution that describes the times to failure of an item, at the use level. This is normally achieved by collecting times to failure or non-failure (censored) data, then fitting such data to an appropriate probability distribution function such as the exponential, lognormal, or Weibull distribution. In accelerated life data analysis, however, the time to failure probability distribution at the use stress level is not explicitly given; rather, one must derive an estimate of the use stress level life distribution from accelerated life test data, which is typically achieved through extrapolation of the life-stress models such as the one depicted in Figure 3.3 of Chapter 3. The remaining sections of this chapter will outline the main ideas of ALT methods, tools and techniques for data analysis of accelerated life test data.

4.3. BASICS OF ALT DATA ANALYSIS

In ALT data analysis, the main challenge is to infer life information at the use stress level from the life tests conducted at some elevated stress levels. More specifically, recall that in quantitative accelerated testing, the aim is not only to determine where and how failures occur to improve the item, but also to:

- Understand how failure is accelerated by a stress agent and fit an acceleration model (life-stress model and associated time to failure distribution model) to the observed accelerated test data;
- Gather sufficient data at multiple high stress levels to accurately extrapolate a life-stress model and assess the life distribution at the use stress level(s).

To appreciate the process, consider a simplified scenario whereby an item/component is subjected to two levels of elevated stresses, and the times to failure of the item/component at these stresses are recorded. At each stress level, a distribution of the time to failure can be determined, based on prior knowledge or through goodness of fit tests. To obtain the distribution of the times to failure at the use stress level, one must be able to extrapolate from the ALT data using an appropriate life-stress model (e.g. linear, exponential), known as the accelerated life model. Figure 4.1 shows this simple case of a linear (or log-linear) model used to extrapolate the ALT data at higher stresses using median life

vs. stress line. In this case, a selected percentile (e.g. median or 10th percentile) of the time to failure distribution at each of the elevated stress levels is extrapolated to the use stress level, thereby providing an estimate of the failure time at use stress conditions associated with the selected percentile. It is obvious from this that the more test levels of elevated stress the item is subjected to, the more accurate the extrapolation will be, assuming that failures occur due to a common failure mechanism, and that the test is conducted appropriately.

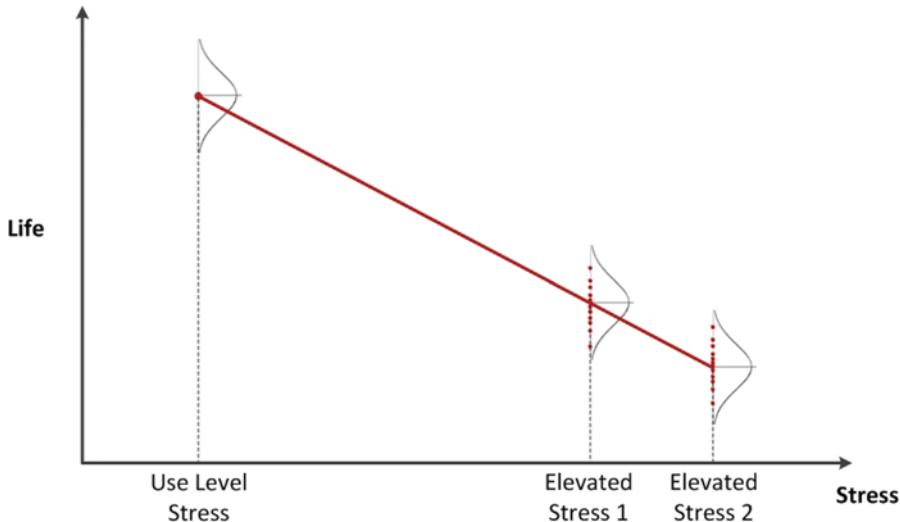


Figure 4.1: Simple linear extrapolation for eliciting use-level information from ALT data

Since ALT data analysis relies on the accuracy and strength of the applied extrapolation, there are a number of ALT planning considerations to take into account when conducting these tests. In addition, a number of common complications exist in the analysis of ALT data that engineers must be aware of. These complications are discussed in more detail in Chapter 6 along with the discussion on ALT test planning considerations.

4.4. TYPES OF COLLECTED ACCELERATED LIFE TEST DATA

The quality and accuracy of estimates produced by an accelerated life test are directly related to the “goodness” of the collected data. There are generally two types of observed data: *complete* and *censored* failure data. Failure data is generally referred to as complete when failure occurrence time (or interval) is known, while censored data refers to several cases: the case with no failure; the case where a failure was not observed or the time of failure is unknown; the case with a failure that is not caused by the same failure mode of interest; or the case where an operating item is removed from the test for some other reason. For example, one either observes that an item failed at a certain time (complete data) or that it operated successfully up to a certain time and our observation ceased thereafter (right censored data).

An example of a visual representation for complete data is shown in Figure 4.2. Here, we see five units tested which all fail at different known times. On the other hand, it is also possible that some of the units tested do not fail by the time the test is stopped, or fail but not due to the same failure mode of interest. In this case, we have right-censored (suspended) data. Figure 4.3 shows an example of right-censored data in which, of the five units that were tested, three failed (complete failures) and two censored (suspended) at different times. In most accelerated life testing, it is common to have some censoring schemes. This type of popular censoring is also known as *multiply censored* data, since the censored data on the right have different running times intermixed with failure times. Multiply

censored data arise when units are tested at different times, resulting in different test start and running times when the data are recorded. If units are started together at a test condition and those data are analyzed before all units fail, the data would be singly censored as depicted in Figure 4.4. Note that in all these examples, right censoring is applicable. Left censored data arise when the exact time or interval of a failed unit is unknown; rather it is only known that failure occurred before a certain time.

Another classification worth noting is that censored data can be either time censored, where data is observed after a predetermined time (Type 1 censoring), or failure censored, where data is observed after a predetermined number of failures occurred (Type 2 censoring).

There are also cases where exact times to failure of the items under test are not known, but an interval of time in which the failure occurred is recorded. In this case, the data is known as interval, or grouped, data. Figure 4.5 illustrates this case, in which a solid line represents the interval in which the unit fails, and a dotted line shows an inspection interval during which it does not fail. Such data can also contain right and left censored observations.

4.5. LIFE-STRESS MODELS

In Chapter 2, we presented an outline of the process used to determine the PoF to understand a physically informed empirical relation between life and stresses. This is a crucial element of ALT, as it illuminates the relationship between characteristics of life and applied stress. Combined with the life distribution, it provides a single model, known as the accelerated stress life distribution model, which the ALT data analysis can be based on. As discussed in Chapter 2, in terms of determining appropriate accelerated life models, the PoF approach is inherently a probabilistic approach, since it combines the knowledge obtained from the underlying mechanistic degradation and failure with ALT data to establish the probability distribution of life of the component under a given stress.

The main message of this section is that in ALT data analysis, one of the most crucial steps is the development of the accelerated stress vs. life (or life vs. stress) model. The available forms of such models are expressed by Equations (3.5-3.8). In each case, parameters a and b should be estimated from the ALT data.

1. Linear:

$$L = aS + b \quad (4.1)$$

where L = Life, S = Stress, and a and b are model parameters. Although only relevant in the simplest empirical cases, many variables of other models can be transformed such that it can be fitted to a linear relationship.

2. Exponential:

$$L = be^{aS}, \text{ or } \ln(L) = \ln(b) + aS \quad (4.2)$$

The exponential life-stress relationship is very common for modeling ALT data, as it conveniently represents the life-stress relationship resulting from many naturally occurring failure phenomena. The *Arrhenius Relationship* is a special case of an exponential life-stress model where the stress agent is represented as the inverse of the absolute temperature, T , given by:

$$L = be^{\frac{Ea}{kT}} = be^{\frac{a}{T}} \quad (4.3)$$

where E_a and K are parameters known as the activation energy and Boltzman's constant. An alternative to the Arrhenius relationship used for temperature acceleration is the *Eyring relationship*:

$$L = \frac{b}{T} e^{\frac{a}{T}} \quad (4.4)$$

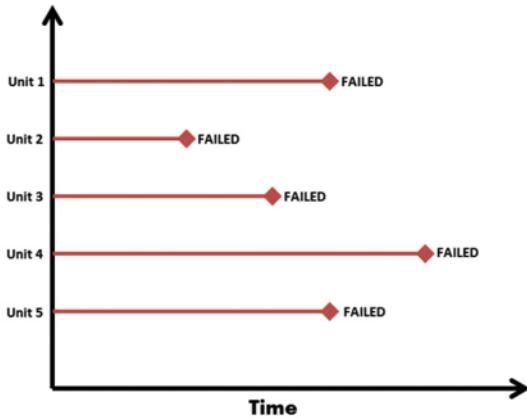


Figure 4.2: Complete data

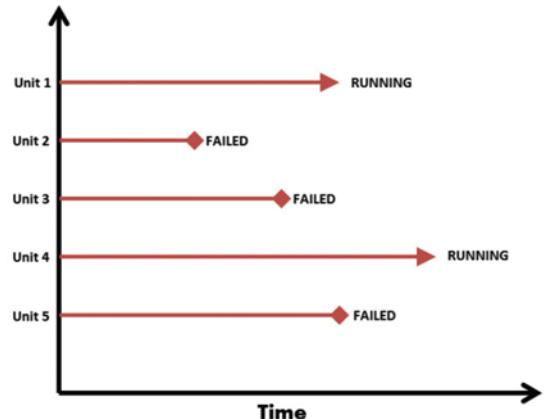


Figure 4.3: Censored data (multiply censored)

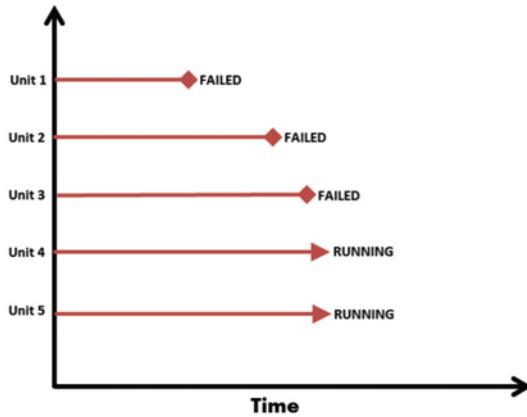


Figure 4.4: Singly censored data

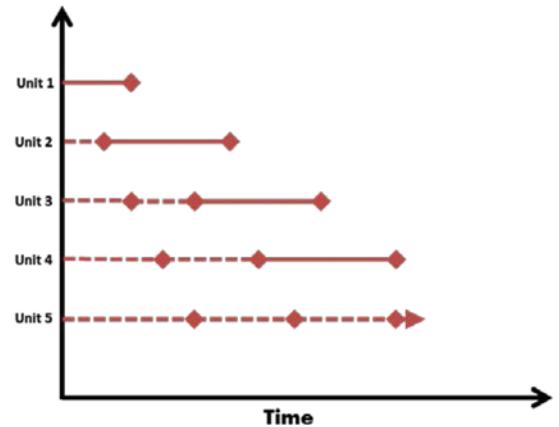


Figure 4.5: Interval or grouped censored data

3. Power:

$$L = b(S)^a, \text{ or} \quad (4.5)$$

$$\ln(L) = \ln(b) + a\ln(S)$$

In this popular life-stress model, life has a power relationship. The so-called S-N relationship describing low-cycle fatigue in metals, as discussed in Chapter 2, is an example of the power relationship. The *Inverse Power Rule* is merely a simple transformation of the power rule, given by:

$$L = \frac{b}{(S)^a} \quad (4.6)$$

4. Logarithmic:

In some cases, a linear extrapolation can be obtained using a logarithmic transformation on the accelerating variable as follows:

$$L = a \ln(S) + b \quad (4.7)$$

5. Multi-Stress Models:

For multivariable accelerated tests, that is, those that have more than one accelerating variable, the multi-stress exponential model is a suitable statistical model that can be used as a basis for the life-stress model. The multi-exponential model is given by

$$\begin{aligned} L &= e^{a_0 + a_1 S_1 + \dots + a_n S_n}, \text{ or} \\ \ln(L) &= a_0 + a_1 S_1 + \dots + a_n S_n \end{aligned} \quad (4.8)$$

A simple example of a model that incorporates two accelerating variables is the dual-exponential model such as the so-called *Temperature-Humidity* model. This relationship is given as follows:

$$\text{Temperature-Humidity: } L = A e^{\left(\frac{a}{T} + \frac{b}{S}\right)} \quad (4.9)$$

Other multi-stress models may be combined versions of the above. For instance, the Power-Exponential model used when temperature and a non-thermal stress are involved is given by:

$$\text{Temperature-Nonthermal: } L = \frac{c}{s^b e^{-\frac{a}{T}}} \quad (4.10)$$

When applied to ALT, the life variable in the life-stress models is represented by a pdf built on the basis of scattered times-to-failure observed at elevated stresses, since at the elevated stress failures occur earlier. The life-stress models are then used to extrapolate certain quantiles of the pdf(s) at elevated stresses to lower use stress levels. The concept is depicted in Figure 4.6 when the stress agent (τ_{max} -- the maximum shear stress) causes wear. Based on two ALT tests at elevated stresses (τ_{max1} and τ_{max2}), the pdfs of life distributions are found, and based on the applicable life-stress, selected percentiles of those pdfs are extrapolated to the use stress level τ_{max} . For example, consider developing the damage distribution in a damage-endurance model due to wear of a bearing in rotating machinery as a function of time. As expected from the PoF analysis, the higher wear values lead to lower life of the bearing. Note that a main assumption is that the mechanism of failure remains the same. Thus, the life distribution shape factor remains the same at all times as the damage progresses.

Figure 4.6 illustrates this concept, using the PoF-based damage-endurance model for the failure mechanism wear (accelerated by the stress τ_{max}). In Figure 4.6 the damage-endurance model in the form of Equation (4.5) is plotted on a log-log scale. Observations at higher stress levels (τ_{max1} and τ_{max2}) shorten the test time and yield better estimates of the probabilistic distribution of the life of the item. Usually, lognormal or Weibull distributions best represent the form of these distributions. Life distribution models at high stress levels are then extrapolated to the use level along the life acceleration model. Data from accelerated tests are then used to estimate the parameters of both the

life-stress and distribution models. The wear agent is directly related to the distribution of damage of the bearing over time.

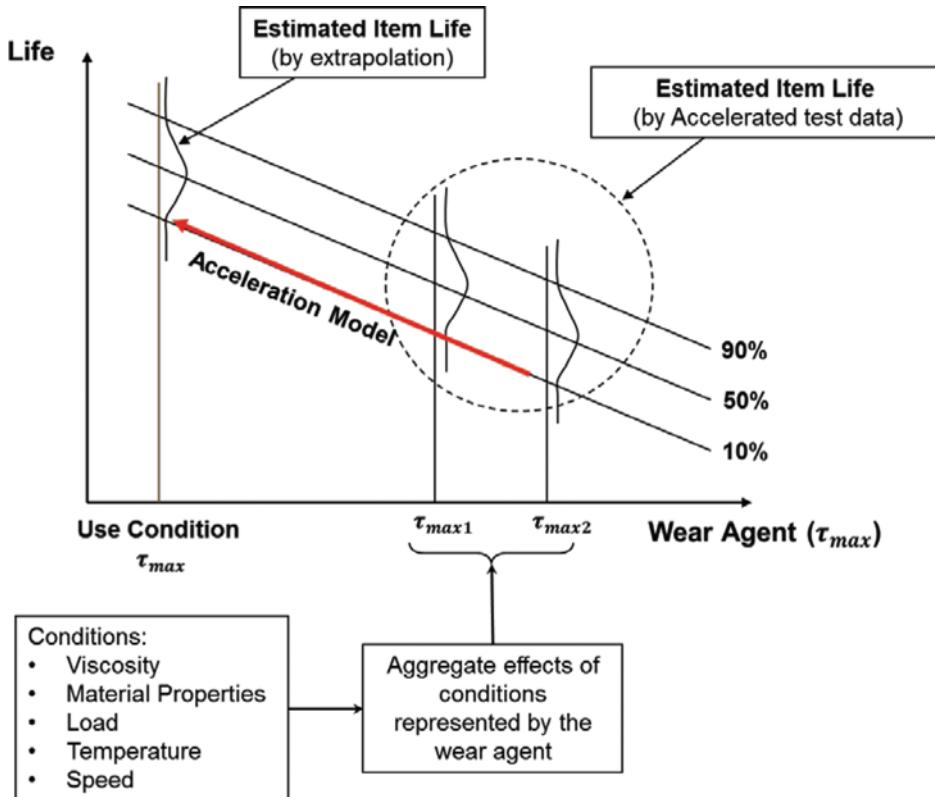


Figure 4.6: Wear (damage) life-stress modeling concept

The maximum shear stress in the vicinity of the surface is an aggregate agent of failure responsible for material removal in the abrasive wear process. The wear rate (and thus the damage) can be expressed as a power relationship similar to Equation (4.5). Having the wear rate, the life of the item can be estimated by considering an endurance limit beyond which the bearing can be considered failed. This end state can be defined based either on the surface roughness (i.e. RMS of journal and bearing roughness) or on a certain amount of material removal from the bearing (e.g., the instance that the protective coating thickness of the bearing is removed). Based on this description, the damage of the bearing can be related to the wear agent as

$$\dot{W} \propto \left[\frac{\tau_{max}}{\tau_{yp}} \right]^n$$

$$\left(\frac{1}{D} \right) \propto L = \frac{C_0}{\dot{W}}$$
(4.11)

Therefore,

$$L = C \left[\frac{\tau_{yp}}{\tau_{max}} \right]^n = \frac{K}{[\tau_{max}]^n} \quad (4.12)$$

where

\dot{W} = Wear rate

C_0 = End state wear

C = Constant

L, D = Life and damage of the bearing in terms of thickness removed (dependent variable)

K, n = Constants to be estimated from the accelerated test results

τ_{yp} = Material shear yield point

τ_{max} = Maximum shear stress in the vicinity of the surface (independent variable)

Note that the presence of τ_{yp} (shear yield point of the bearing material) also enables this model to consider the temperature degradation of the item's material, which can be critical for coated type bearings. Assuming that the life is distributed lognormally, then the joint distribution of life-stress can be obtained as follows

$$f(t|\mu, \sigma) = \frac{1}{t\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(\ln t - \mu)^2}{\sigma^2} \right] \quad (4.13)$$

From Equation (4.11) one concludes that $\mu = -\ln K - n \ln(\tau_{max})$, where μ in Equation (4.13) is the mean (and median) of the natural logarithms of the time of failure.

Therefore,

$$f(t|\mu, \sigma) = \frac{1}{t\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(\ln t + \ln K + n \ln \tau_{max})^2}{\sigma^2} \right] \quad (4.14)$$

Equation (4.14) represents the complete conditional distribution of life (pdf of time to failure) given a specific stress level τ_{max} . The objectives of the next three sections of this chapter are to develop and estimate the parameters of this complete model, including uncertainties associated with the model parameters.

An important metric of interest in ALT is the Acceleration Factor (AF) discussed earlier. AF is the ratio of the life at a use level of stress, L_{use} , to the life at the elevated level of stress, L_{acc} . Therefore,

$$AF = \frac{L_{use}}{L_{acc}} \quad (4.15)$$

Finally, reliability at the use level is perhaps the most important metric of interest when conducting an ALT, as it allows warranty levels on the item to be established, as well as providing insights as to how the item may meet stated reliability goals.

There are three main methods for model parameter estimation in ALT data analysis of the ALT models such as the one shown in Equation (4.14).

1. Plotting method

2. Maximum likelihood estimation method
3. Bayesian estimation method

The subsequent three sections will discuss these methods.

4.6. PROBABILITY PLOTTING METHOD FOR ALT MODEL ESTIMATION

The plotting/graphical method for obtaining parameter estimates from ALT data is the simplest of the three methods listed above. Although less accurate than the other methods, it provides a quick and easy way of interpreting the collected data. It can also provide estimates of an item's life distribution at any stress, and most importantly, allows assessment of data validity and model fit to data.

The plotting method for ALT data analysis involves developing two types of plots. The first step is to determine the most appropriate time-to-failure life distribution for the collected ALT data at elevated stresses. The times-to-failure collected may be represented by any life distribution. One of the most important aspects of ALT analysis is that at any given level of stress, life itself will be governed by a particular probability distribution for an item. That is, the same distribution form having the same shape parameter should be developed as representative of the time to failure models for all stress levels, including the use stress level. If the analysis shows good fits with different distributions or the same distribution but with different shape parameters, then the underlying physical mechanism may have changed (shifted, etc.). In these cases the ALT should be repeated to make sure that the failure mechanism does not change.

For ALT, the most common life distributions used are the exponential, Weibull, and lognormal distributions. One way to determine the best fitting distribution from the collected failure data is to conduct a goodness of fit (GOF) test for each data set against each potential probability distribution (e.g. Chi-square GOF, Kolmogorov-Smirnov GOF, or a simple calculation of the correlation coefficient for each case). Once the most appropriate life distribution is selected, a probability plot of the data at each stress level can be generated, using common plotting techniques such as Kaplan-Meier, or the rank-increment approach.

Once a plot of the life distribution at each accelerated stress level is generated, the second step in the plotting method is to plot the linear version of the life-stress model for one or more nominal life characteristics of the time to failure distributions. Eliciting information about use level conditions, based on the plotted ALT data, helps plotting the life-stress model. To do this, a “nominal” life characteristic must be chosen and plotted. There is no specific definition for nominal life; depending on the goals of the ALT, the analyst usually determines it. Typically, nominal life is a specific characteristic of the life distribution—usually the median, 10-percentile or other relevant distribution percentiles (Nelson W., 2004). The plotting method then determines estimates by solving for the slope and intercept of the plotted life-stress line, and this is achieved through a simple regression technique, as will be discussed in the following section. Note that graphical methods work best for the simple case of one-stress accelerated life test, since the method does not easily accommodate the added complexity that multiple stresses introduce into the analysis.

4.6.1. LIFE-STRESS MODEL BY REGRESSION

The regression technique used in the plotting method of ALT data analysis is fairly straightforward. Take, for example, the inverse power law that from the previous section is given by $\ln(t) = \ln(b) - a \ln(S)$ which provides a linear relationship between life and load. The relationship above can then be used as a basis for performing a simple linear regression.

Consider the set of data provided in Table 4.1 that represents the ALT data of a component in which temperature is the stress-causing agent of failure. Further assume that it has already been shown that

the scatter in times to failure data listed in Table 4.1 can be best modeled by the lognormal distribution, and that the Arrhenius model (exponential) is the most appropriate life-stress relationship for this particular ALT.

Table 4.1: Times to failure observed in an ALT

Temperature (°C)	Number of Units Tested	Recorded Failure Times (hrs)								
150	3	2350	2560	2980						
200	9	220	250	330	370	380	460	460	510	610

Assuming that the use level temperature is 80°C, we want to estimate life at this use level using the probability plotting method. The failure data in Table 4.1 are representative of complete failures.

As a first step, a multiple probability plot of the data at each accelerated temperature is constructed. The plotting points in this case are using the rank adjustment method, $F(t) = \frac{i-0.375}{n+0.25}$, as presented in Table 4.2. The multiple lognormal probability plots are presented as shown in Figure 4.7.

Table 4.2: Plotting points using rank adjustment method

Temperature	150 °C = 423 °K		200 °C = 473 °K	
	i	t _i (hrs)	CDF [F(t)x100%]	t _i (hrs)
1	2350	19.23	220	6.76
2	2560	50.00	250	17.57
3	2980	80.77	330	28.38
4			370	39.19
5			380	50.00
6			460	60.81
7			460	71.62
8			510	82.43
9			610	93.24

The multiple lognormal probability plots are presented in Figure 4.7. The results are:

- For T=423°K, $\mu_t = 7.87$, and $\sigma_t = 0.145$;
- For T=473°K, $\mu_t = 5.94$, and $\sigma_t = 0.357$.

Note that for lognormal distributions, the median life $t_{0.5} = e^{\mu_t}$, and therefore the parameter denoted as μ above (and hereafter for lognormal distributions), is equivalent to $t_{0.5}$ (median time to failure). Therefore, the following are gathered:

$$\mu_{423^\circ K} = e^{7.87} \cong 2617 \text{ hrs}$$

$$\mu_{473^\circ K} = e^{5.94} \cong 381 \text{ hrs}$$

We can see from Figure 4.7 that the shape parameters σ_t in both cases of the elevated stress levels are reasonably close to each other, and hence, the plotting approach can reasonably conclude that the failures experienced at each accelerated temperature are governed by the same failure mechanism. Because the plotting method involves extrapolation using linear regression, there is no need to re-plot the life distributions with a common shape parameter. By determining that the test is governed by a common failure mechanism (i.e., no competing failure mechanisms), then the variability in the shape parameter (which impacts the estimated life values from each plot) is taken into account through the regression process. Other methods involve redrawing the plots using a weighted average of the scale

parameter (i.e., parallel plots), so that the final nominal life value will fall on a straight line, and no regression is required to obtain the slope and intercept of the life vs. stress plot.

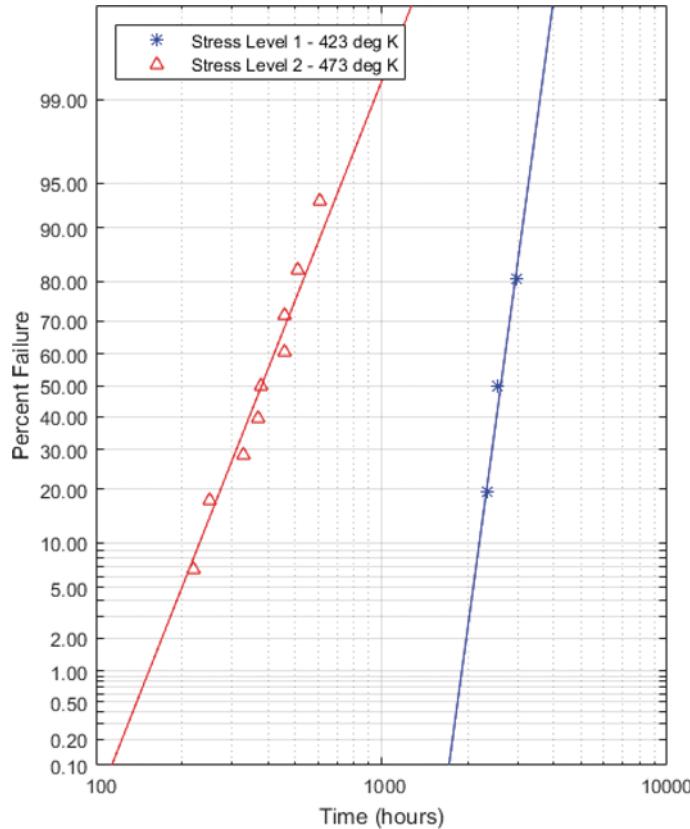


Figure 4.7: Lognormal data plotting at each stress level

Next, from the Arrhenius life-stress relationship, we know that:

$$L = A e^{\frac{E_a}{KT}} \quad (4.16)$$

where A is a constant to be determined, T is the temperature (in Kelvin), K is Boltzmann's constant ($8.617 \times 10^{-5} \text{ eV} \cdot \text{K}^{-1}$) and note that $1/K = 11605$, and E_a is the activation energy (in eV). Taking the log of each side of the Arrhenius relationship yields the following:

$$\ln(L) = \ln(A) + \frac{E_a}{K} \frac{1}{T} = \ln(A) + 11605 E_a \frac{1}{T} \quad (4.17)$$

Here, we have established a linear relationship between life and stress (represented by the inverse of temperature $1/T$). For the lognormal distribution, the value of $\ln(\text{Life})$ that we can use for the above equation is the median value calculated previously (the suitable nominal value of life for analysis purposes). Hence, we can obtain estimates for the parameters of the Arrhenius model through the least squares regression method, which yields the following:

$$\hat{\beta}_0 = \ln(A) = -10.359; \hat{A} = 3.17 \times 10^{-5}$$

$$\hat{\beta}_1 = \frac{E_a}{K} = 7711; \hat{E}_a = 0.665 \text{ eV}$$

Therefore, the life-stress relationship is given by:

$$Life = 3.17 \times 10^{-5} e^{\frac{7711}{T}}$$

Applying the above regression model, we can obtain an extrapolated estimate for median life at the use condition of 80 °C (353°K) as follows:

$$Life(T = 353 K) = 3.17 \times 10^{-5} e^{\frac{7711}{353}} \cong 9.73 \times 10^4 \text{ hrs}$$

We can apply the same procedure for any percentile value of interest; however, note how the plotting method can provide only crude point estimates for the parameters of interest. Moreover, it does not provide any indication of the confidence intervals associated with the uncertainty for each parameter.

Example 4.1

Consider a set of complete failures data provided in Table 4.3 that represents the ALT data of a component in which temperature is the stress-causing agent of failure. Further assume that it has already been shown that the scatter in times to failure data listed in Table 4.3 can be best modeled by the Weibull distribution, and that the Arrhenius model (exponential) is the most appropriate life-stress relationship for this particular ALT. Assuming that the use level temperature is 333 °K, we want to estimate life at this use level using the plotting method.

Table 4.3: Time to failures observed in an ALT

Temperature (°K)	Number of Units Tested	Recorded Failure Times (hrs)							
406	8	248	456	528	731	813	965	972	1528
436	6	164	176	289	319	386	459		
466	6	92	105	155	184	219	235		

Solution

As the first step, a multiple probability plot of the data at each accelerated temperature is constructed. The plotting points in this case are calculated using the rank adjustment method, $F(t) = \frac{i-0.375}{n+0.25}$, as shown in the Table 4.4. Excel is used for Weibull analysis in this example. For the Weibull distribution, the CDF is $F(x) = 1 - e^{-(\frac{x}{\alpha})^\beta}$. With some effort, the Weibull cumulative distribution function can be transformed so that it appears in the familiar form of a straight line: $\ln \left[\ln \left(\frac{1}{1-F(x)} \right) \right] = \beta \ln x - \beta \ln \alpha$. Note that the characteristic life α in the Weibull distribution corresponds to the 63.2% of life.

Table 4.4: CDF by rank adjustment method

Temperature	406 °K		436 °K		466 °K	
	i	t _i (hrs)	CDF [F(t) × 100%]	t _i (hrs)	CDF [F(t) × 100%]	t _i (hrs)
1	248	7.58	164	10.00	92	10.00
2	456	19.70	176	26.00	105	26.00
3	528	31.82	289	42.00	155	42.00
4	731	43.94	319	58.00	184	58.00
5	813	56.06	386	74.00	219	74.00
6	965	68.18	459	90.00	235	90.00
7	972	80.30				
8	1528	92.42				

The multiple Weibull probability plots are presented in Figure 4.8. The Weibull analysis results are:

For T=406 °K, $\hat{\beta}_1 = 2.01$, and $\hat{\alpha}_1 = 895$;

For T=436 °K, $\hat{\beta}_2 = 2.56$, and $\hat{\alpha}_2 = 341$;

For T=466 °K, $\hat{\beta}_3 = 2.79$, and $\hat{\alpha}_3 = 187$.

We can see that the shape parameters in both cases of the elevated stress levels are reasonably close to each other, and hence, the plotting approach can reasonably conclude that the failures experienced at each accelerated temperature are governed by the same failure mechanism. Because the plotting method involves extrapolation using linear regression, there is no need to re-plot the life distributions with a common shape parameter. By determining that the test is governed by a common failure mechanism (i.e., no competing failure mechanisms), then the variability in the shape parameter (which impacts the estimated life values from each plot) is taken into account through the regression process. Other methods involve redrawing the plots using a weighted average of the scale parameter (i.e. parallel plots), so that the final nominal life value will fall on a straight line, and no regression is required to obtain the slope and intercept of the life vs. stress plot.

Next, we can obtain estimates for the parameters of the Arrhenius model through the least squares regression approach as shown in

Figure 4.9. The Arrhenius model is given by:

$$L = A e^{\frac{E_a}{KT}} \quad (4.18)$$

where A is a constant to be determined, T is the temperature (in Kelvin), K is Boltzmann's constant (note that $1/K = 11605$), and E_a is the activation energy. Taking the log of each side of the Arrhenius relationship yields the following

$$\ln(L) = \ln(A) + \frac{E_a}{K} \times \frac{1}{T} \quad (4.19)$$

Here, we have established a linear relationship $Y = mX + b$ between life and stress (represented by inverse of temperature ($1/T$)), where $Y = T$; $m = \frac{E_a}{K}$; $X = \frac{1}{T}$. The relationship above can then be

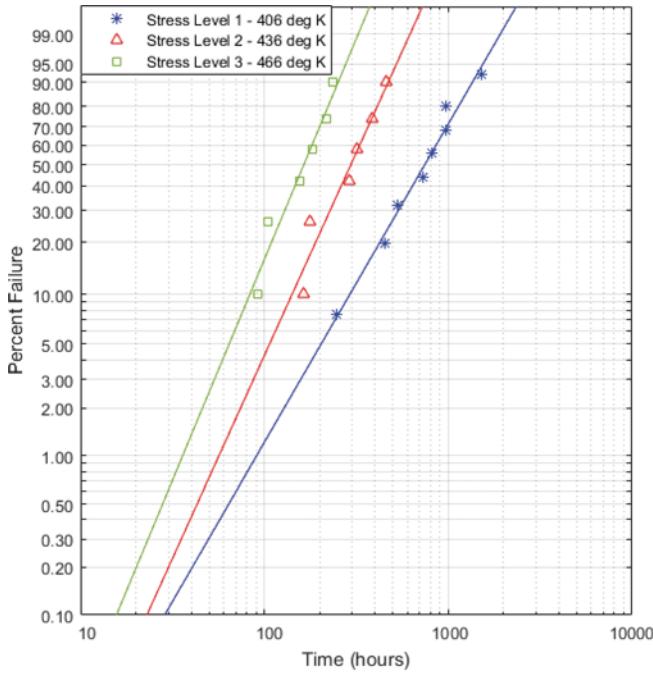


Figure 4.8: Weibull data plotting at each stress level

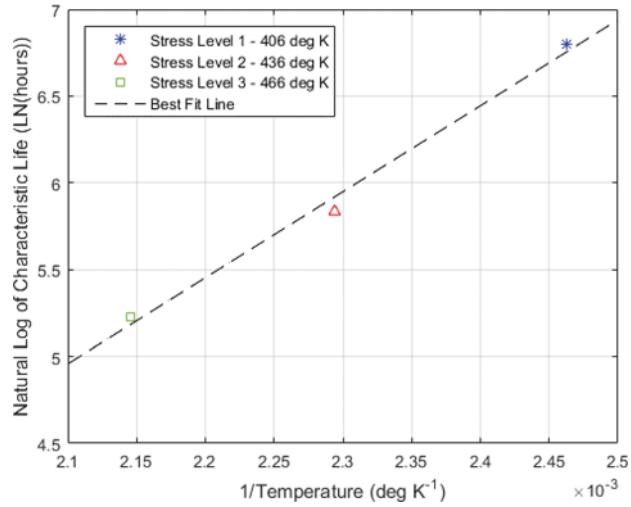


Figure 4.9: Least squares regression for Arrhenius model

used here, we have established a linear relationship $Y = mX + b$ between life and stress (represented by inverse of temperature $1/T$), where $Y = T$; $m = \frac{E_a}{K}$; $X = \frac{1}{T}$. The relationship above can then be used as a basis for performing a simple linear regression.

Therefore, we can obtain the following:

$$\hat{m} = \frac{\widehat{E_a}}{K} = 4935.9 \frac{\text{eV}}{\text{°K}}; \quad \widehat{E_a} = 0.4253 \text{ eV}$$

$$\hat{b} = \ln(\hat{A}) = -5.4091; \quad \hat{A} = 4.48 \times 10^{-3} \text{ hr.}$$

Therefore, the life-stress relationship is given by:

$$\text{Life} = 4.48 \times 10^{-3} e^{\frac{4935.9}{T}}$$

Applying the above regression model, we can obtain an extrapolated estimate for median life at the use condition of 333 °K as follows:

$$\text{Life}_{(T=333\text{°K})} = 4.48 \times 10^{-3} e^{\frac{4935.9}{333}} \cong 1.23 \times 10^4 \text{ hrs}$$

We can apply the same procedure for any percentile value of interest; however, note how the plotting method can provide only crude point estimates for the parameters of interest. Moreover, it does not provide any indication of the confidence intervals associated with the uncertainty for each parameter.

4.6.2. SUMMARY OF PLOTTING METHOD FOR ANALYZING ALT DATA

From the example provided in , the plotting approach for analyzing ALT data is summarized below:

1. Examine a scatter plot of failure time versus the accelerating variable (stress).
2. Fit distributions (using probability-plotting techniques) individually to the ALT data at each elevated level of the accelerating variable. Plot the fitted probability lines to form multiple probability plots and determine parameters of each distribution and corresponding percentiles of life for each stress level tested (e.g., 10% life, median and 90% life).
3. Fit an overall life-stress model. This will depend on either a determined empirical relationship or an already well-established physical relationship (through study of physics of failure) between life and the accelerating variable.
4. Perform a regression analysis based on the ALT life at a given percentile (e.g., 10% life or median life) versus stress to obtain estimates of metrics of interest at applicable use levels.
5. Compare the combined model from Step 3 with the individual analysis in Step 2 to check for evidence of lack of fit for the overall model.

Note that even though the plotting method for ALT data analysis is simple, it presents a number of disadvantages as follows:

- Graphical methods cannot be employed in cases where only censored data exists (i.e. no failures occur) at one or more stress levels. Disregarding the data can lead to serious inaccuracies, since the result would not represent the true failure behavior of the unit under test.
- It is not always possible to linearize the life-stress function in order to obtain estimates for model parameters.
- Confidence intervals on all of the results cannot be estimated through the plotting method. A simple (but highly unreliable!) way to generate confidence intervals through the plotting technique involves placing additional regression lines that would provide the maximum and minimum extrapolation results (see example in Figure 4.10). This is not a recommended way of providing confidence intervals, as it has no formal mathematical basis, but may be useful in communicating rough estimates in some scenarios.

4.7. MAXIMUM LIKELIHOOD ESTIMATION APPROACH TO ALT DATA ANALYSIS

The maximum likelihood estimation (MLE) method is a formal analytical approach to generating estimates for parameters of interest, along with uncertainties associated with those parameters, through the calculation of confidence intervals. The MLE method has a strong statistical basis for calculating estimates. Its main disadvantage is the fact that it does not provide any visualization of fit of the data. In some situations, this limitation may significantly impact the analysis. For instance, in cases where there is an appreciable difference in the shape parameter for each data set corresponding to accelerated stress levels, then simply calculating the MLE estimate of the parameter may hide the fact that other failure mechanisms may exist in the ALT data, in addition to the one that is assumed. In reality, however, it is advisable that the MLE method be carried out in conjunction with an initial graphical

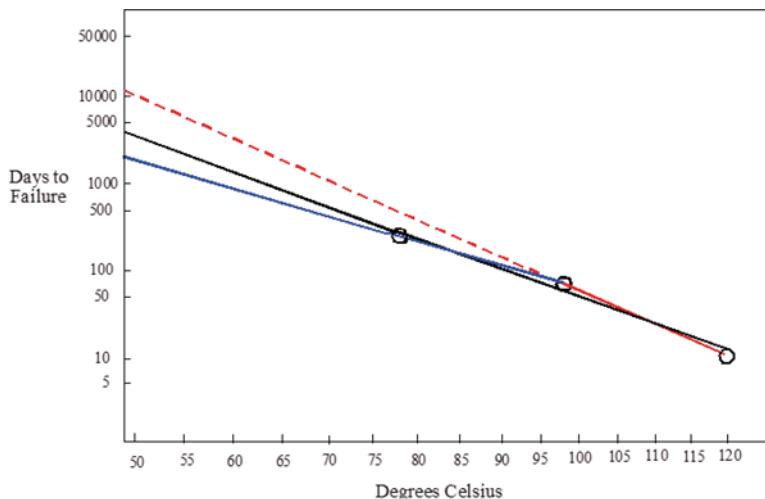


Figure 4.10: Example of rudimentary approach to estimating confidence interval using graphical method for data analysis

plot of the data, so that secondary failure modes or mechanisms can be ruled out prior to getting the precise estimates provided by the MLE approach.

In essence, in the MLE approach the likelihood function that describes the probability of the observed data occurring determines specific values of parameters of interest such that the likelihood function is maximized. Naturally, the first step in this process is to determine the likelihood function itself in light of the data observed. The key question to ask when trying to determine the likelihood function is, “What is the probability of the given failure time or cycle (and censored times) occurring?” Hence, in the case of accelerated life tests, the likelihood function is often the product of the probability distribution that describes the life of the item under test (e.g. exponential, lognormal, Weibull time to failure distribution), as it captures the parameters of interest that we are trying to determine. For instance, for the Arrhenius life-stress model, we may want to find an estimate of the constant A that would maximize the likelihood function, which is the function describing the probability of the data being observed, and which therefore would depend on the censoring type applicable to the data. A general form for the likelihood function l that is applicable for ALT right-censored data analysis is given below, where $f(\cdot)$ is the pdf of the time-to-failure, $F(\cdot)$ is the cdf of the time to failure or time between failures, and $\boldsymbol{\theta}_M$ is a vector of unknown parameters.

For complete failures and right censored data:

$$l = \prod_{i=1}^{N_c} [f(t_i; \boldsymbol{\theta}_M)]^{n_i} \cdot \prod_{j=1}^{N_r} [1 - F(t_j; \boldsymbol{\theta}_M)]^{n_j} \quad (4.19)$$

where n_i is the number of failures in the i^{th} time-to-failure data point; n_j is the number of right censored (suspensions) in the j^{th} censored data point; N_c is number of complete failures; and N_r is the number of right-censored data points.

For complete failures and left censored data,

$$l = \prod_{i=1}^{N_c} [f(t_i; \boldsymbol{\theta}_M)]^{n_i} \cdot \prod_{j=1}^{N_l} [F(t_j; \boldsymbol{\theta}_M)]^{n_j} \quad (4.20)$$

where n_i is the number of failures in the i^{th} time to failure data point; n_j is the number of left censored (suspensions) in the j^{th} censored data point; N_c is number of complete failures; and N_l is the number of left censored data points.

For interval failures and interval censored data in N_I intervals:

$$l = \prod_{i=1}^{N_I} [F(t_i^{UI}; \boldsymbol{\theta}_M) - F(t_i^{LI}; \boldsymbol{\theta}_M)]^{n_i} \cdot \prod_{j=1}^{N_I} [1 - F(t_j^{UI}; \boldsymbol{\theta}_M) + F(t_j^{LI}; \boldsymbol{\theta}_M)]^{n_j} \quad (4.21)$$

where n_i is the number of failures in the i^{th} time interval; n_j is the number of censored (suspensions) in the j^{th} time interval; N_I is the number of failures and censored intervals; t_i^{UI} is the upper limit of i^{th} time interval of failure; t_i^{LI} is the lower limit of the i^{th} time interval of failure; t_j^{UI} is the upper limit of j^{th} time interval of suspensions; and t_j^{LI} is the lower limit of j^{th} time interval of suspensions.

Note that Equations (4.19) and (4.20) assume only non-interval data, and (4.21) assumes only interval data. If there are both interval data and non-interval data, (4.19) and (4.20) may be multiplied by (4.21) to arrive at combinations of complete, right censored and interval data, or compete, left censored and interval data.

Since the main goal of the MLE approach is to maximize the likelihood function, the function needs to be differentiated and equated to zero to determine the value of the parameter of interest that would maximize the likelihood function. Therefore, it is often convenient to take the log-likelihood function for easier computation of the derivative. The log-likelihood function for ALT data containing complete and right censored data is given by:

$$\ln(l) = \sum_{i=1}^{N_c} n_i \cdot \ln[f(t_i; \boldsymbol{\theta}_M)] + \sum_{j=1}^{N_r} n_j \cdot \ln[1 - F(t_j; \boldsymbol{\theta}_M)] \quad (4.22)$$

The maximum likelihood estimate then finds $\boldsymbol{\theta}_M$ such that the probability of data observed is the maximum. This is done by taking the derivative of $\ln(l)$ with respect to $\boldsymbol{\theta}_M$, equating it to 0 and solving for $\boldsymbol{\theta}_M$, to find $\hat{\boldsymbol{\theta}}_M$. And as mentioned previously, this is usually more manageable if the log-likelihood function is used instead. That is,

$$\frac{\partial \ln(l)}{\partial \theta_M} = 0 \quad (4.23)$$

Other expressions in Equations (4.19)-(4.21) can also be similarly converted to the log-likelihood of the form (4.22), the derivatives of which with respect to the parameters in the log-likelihood function will form a system of equations to estimate the parameters.

Example 4.2

Consider an accelerated test of a stainless steel pressure vessel which is subjected to corrosion, and has shown that accelerated corrosion occurring at an elevated temperature ($220\text{ }^{\circ}\text{C}$) yields the following days-to-failure for a sample of 4 vessels: 98, 157, 194, 283 days. The life-stress relationship is shown by $L = be^{\frac{a}{T}}$, where T is the temperature in Kelvin (the stress agent), a is known to be $4228\text{ }^{\circ}\text{K}$, and life, L , is expressed in terms of days-to-failure, as a random variable follows an exponential distribution. Outline the process for obtaining the maximum likelihood estimation for the parameter b .

Solution

Denote that the exponential distribution of the time-to-failure is $f(t) = \lambda e^{-\lambda t}$. Using MTTF (which is $1/\lambda$), in the exponential it corresponds to the 63.2% of life, and by incorporating this into the given life-stress relationship, we can describe the life-stress relationship as follows:

$$MTTF = \frac{1}{\lambda} = L = be^{\frac{a}{T}}, \text{ or } \lambda = \frac{1}{be^{\frac{a}{T}}}$$

Therefore, the pdf of life given temperature (stress) T is $f(t) = \frac{1}{be^{\frac{a}{T}}} e^{-\frac{1}{be^{\frac{a}{T}}}t}$. Using the MLE approach, we can take the above conditional *pdf* (i.e., pdf of time to failure conditioned on a temperature T) and develop the likelihood function for complete failure data to $l = \prod_{i=1}^{N_c} f(t_i)$. The log-likelihood function in this case is given by:

$$\ln(l) = \sum_{i=1}^{N_c} \ln \left\{ \frac{1}{be^{\frac{a}{T}}} e^{-\frac{1}{be^{\frac{a}{T}}}t_i} \right\} = \sum_{i=1}^{N_c} \left[\ln \left\{ \frac{1}{be^{\frac{a}{T}}} \right\} - \frac{t_i}{be^{\frac{a}{T}}} \right]$$

Therefore, according to Equation (4.23),

$$\frac{\partial \ln(l)}{\partial b} = \sum_{i=1}^{N_c} \frac{-1}{b} - \left(\frac{-t_i}{b^2 e^{\frac{a}{T}}} \right) = \left(\frac{1}{b} \right) \sum_{i=1}^{N_c} \left(\frac{t_i}{be^{\frac{a}{T}}} - 1 \right)$$

Substituting the given t_i 's (in days) and solving for b at $\frac{\partial \ln(l)}{\partial b} = 0$ yields:

$$\hat{b} = 0.0345 \text{ days}$$

Having obtained this MLE estimate for the unknown parameter b , we can obtain estimates of life at any stress level. For instance, if the use temperature level is $80\text{ }^{\circ}\text{C}$ ($353\text{ }^{\circ}\text{K}$), then an estimate of the MTTF (mean life) at the use temperature level is given by:

$$MTTF_{use} = be^{\frac{a}{T_{use}}} = 0.0345e^{\frac{4228}{353}} = 5489.2 \text{ days}$$

4.8. CONFIDENCE INTERVALS FOR MLE

In ALT data analysis, parameters of the life-stress model should be estimated and thereby predict other reliability metrics of interest under test or use stress level conditions. Such metrics may include predicted reliability at a given time, mean life, or expected number of failures by the end of warranty of time. There is always some uncertainty associated with these estimates because of the scatter and limitation of data on which the life-stress models were based. One of the major advantages of the MLE approach compared with the plotting method in ALT data analysis is that MLE allows for the calculation of confidence intervals associated with MLE parametric estimates. This allows analysts to describe the amount of uncertainty associated with the calculated estimates of ALT model parameters. In ALT data analysis, we often have to deal with censored data, and one common method of calculating confidence intervals for model parameters and metrics of interest is through the use of the Fisher information matrix.

The Fisher information matrix is used to determine the variance and covariance of estimated parameters. In the simple case of a two-parameter probability distribution, based on the likelihood and log-likelihood function $\Lambda = \ln(l)$, the Fisher information matrix, F for a two-parameter life-stress distribution model is given by:

$$F = \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \theta_1^2} & -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \theta_2} \\ -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \theta_2^2} \end{bmatrix} \quad (4.24)$$

where θ_1 and θ_2 are specific parameters estimated by MLE. By using the MLE values for these parameters, (i.e. $\hat{\theta}_1$ and $\hat{\theta}_2$), and then inverting the matrix, we can obtain an estimate of the covariance matrix as follows:

$$\begin{aligned} F^{-1} &= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \theta_1^2} & -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \theta_2} \\ -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \theta_2^2} \end{bmatrix}_{\hat{\theta}_1, \hat{\theta}_2}^{-1} \\ &= \begin{bmatrix} \text{var}(\hat{\theta}_1) & \text{cov}(\hat{\theta}_1, \hat{\theta}_2) \\ \text{cov}(\hat{\theta}_2, \hat{\theta}_1) & \text{var}(\hat{\theta}_2) \end{bmatrix} \end{aligned} \quad (4.25)$$

Now in general, MLE estimates of the distribution parameters tend to the normal distribution for large sample sizes. Therefore, if $\hat{\theta}$ is the MLE point estimate for θ , based on a large sample of n units, and if $z = \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}}$, then $\Pr(x \leq z) \rightarrow \Phi(z)$, where $\Phi(z)$ is the CDF of the standard normal distribution

at the value of z , and is given by $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$. Given a significance level α (i.e. confidence level of $1 - \alpha$), approximate confidence interval on the parameter θ are defined as

$$\text{Two-sided: } \Pr\left(-Z_{1-\frac{\alpha}{2}} < \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}} < Z_{1-\frac{\alpha}{2}}\right) \cong 1 - \alpha \quad (4.26)$$

$$\text{One-sided: } \Pr\left(-Z_\alpha < \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}}\right) = \Pr\left(Z_\alpha > \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}}\right) \cong 1 - \alpha$$

where Z_α and $Z_{1-\frac{\alpha}{2}}$ are defined as the standard normal inverse value at the confidence level of $1-\alpha$ (for one-sided confidence interval) and $1 - \frac{\alpha}{2}$ (for two-sided confidence interval). That is,

$$\begin{aligned} Z_\alpha &= \Phi^{-1}(1 - \alpha) \text{ (one-sided confidence intervals)} \\ Z_{1-\frac{\alpha}{2}} &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ (two-sided confidence intervals)} \end{aligned} \quad (4.27)$$

Further simplifying (4.27), we can obtain the approximate two-sided confidence bounds on the parameter θ at a significance level α , as $\left(\hat{\theta} - Z_{\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\theta})} < \theta < \hat{\theta} + Z_{\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\theta})}\right)$. For the one-sided confidence interval, the upper confidence limit would be $\theta < \hat{\theta} + Z_\alpha \cdot \sqrt{\text{var}(\hat{\theta})}$, and the lower confidence limit would be $\theta > \hat{\theta} - Z_\alpha \cdot \sqrt{\text{var}(\hat{\theta})}$.

Some model parameters have a constraint that they can only be positive (e.g., the shape parameter β for the Weibull distribution). In this case, we instead treat $\ln(\theta)$ as normally distributed for large sample sizes. The two-sided approximate confidence bounds on θ , at a significance level α , then become:

$$\Pr\left(\hat{\theta} e^{-\frac{z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\theta})}}{\hat{\theta}}} < \theta < \hat{\theta} e^{\frac{z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\theta})}}{\hat{\theta}}}\right) \cong 1 - \alpha \quad (4.28)$$

For one-sided confidence intervals on positive parameters, the upper limit is $\hat{\theta} < \hat{\theta} e^{\frac{z_{1-\alpha} \cdot \sqrt{\text{var}(\hat{\theta})}}{\hat{\theta}}}$ and the lower limit is: $\hat{\theta} > \hat{\theta} e^{-\frac{z_{1-\alpha} \cdot \sqrt{\text{var}(\hat{\theta})}}{\hat{\theta}}}$.

Example 4.3

For the Weibull distribution, determine the covariance matrix and the $(1-\gamma)\%$ two-sided interval.

$$\text{cov}(\hat{\alpha}, \hat{\beta}) = [I(\hat{\alpha}, \hat{\beta})]^{-1} = \begin{bmatrix} \text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\beta}) \\ \text{cov}(\hat{\beta}, \hat{\alpha}) & \text{var}(\hat{\beta}) \end{bmatrix} = \frac{1}{n_F} \begin{bmatrix} 1.1087 \frac{\hat{\alpha}}{\hat{\beta}} & 0.2570 \hat{\alpha} \\ 0.2570 \hat{\alpha} & 0.6079 \hat{\beta}^2 \end{bmatrix}$$

Solution

The two-sided confidence limits on scale parameter

$$\hat{\alpha} \exp\left(-\frac{\frac{z_{1-\gamma}}{2}\sqrt{\text{var}(\hat{\alpha})}}{\hat{\alpha}}\right) \leq \alpha \leq \hat{\alpha} \exp\left(\frac{\frac{z_{1-\gamma}}{2}\sqrt{\text{var}(\hat{\alpha})}}{\hat{\alpha}}\right)$$

The two-sided confidence interval for the shape parameter

$$\hat{\beta} \exp\left(-\frac{\frac{z_{1-\gamma}}{2}\sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}\right) \leq \beta \leq \hat{\beta} \exp\left(\frac{\frac{z_{1-\gamma}}{2}\sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}\right)$$

4.9. MLE APPROACH TO ESTIMATING PARAMETERS OF COMMON DISTRIBUTIONS

This section will describe the MLE parameter estimation of three common life distributions, when the data are composed of complete and right-censored data. These distributions are as follows:

- (a) Exponential distribution
- (b) Weibull distribution
- (c) Lognormal distribution

These distributions are used to describe the life of a unit under test at a given stress level. Since these are only “single” distributions, they are not a function of stress, but rather, only the failure (or suspension) times gathered from the test, assuming a common stress level for each of those observed times.

4.9.1. EXPONENTIAL LIFE DISTRIBUTION

The exponential distribution is a widely-used probability distribution, mainly due to its simplicity. This wide use is often coupled with some concerns due to its assumptions of a constant hazard rate that means no wear out, which, naturally, does not apply in many cases. Nevertheless, in accelerated life testing, the exponential distribution is commonly used to describe the life distribution at each stress level. The MLE approach to estimating the failure rate parameter λ of the single exponential distribution is outlined below.

The likelihood function for a sample of single failure per N_c group of data (i.e., n_i 's are all one) and assuming no censored data points is given by:

$$l = \prod_{i=1}^n \lambda e^{-\lambda t_i} = \lambda^n e^{-\lambda \sum_{i=1}^n t_i} \quad (4.29)$$

where t_i is the time of i^{th} failure; n is the number of failures; λ is the failure rate.

Taking the log of both sides gives:

$$\ln(l) = \Lambda = n \ln(\lambda) - \lambda \sum_{i=1}^n t_i \quad (4.30)$$

Taking the derivative of Λ with respect to λ and equating to zero allows the maximum likelihood point estimate for λ , denoted as $\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i}$.

4.9.2. WEIBULL LIFE DISTRIBUTION

The Weibull distribution is a popular and flexible probability distribution and can model many types of failures in reliability engineering. The 2-parameter Weibull pdf is given by

$$f(t) = \frac{\beta t^{\beta-1}}{\alpha^\beta} e^{-(\frac{t}{\alpha})^\beta}; f(t) \geq 0, t \geq 0, \beta > 0, \alpha > 0, \text{ where } \alpha \text{ is the scale parameter, and } \beta \text{ is the shape parameter (slope of the Weibull plot curve).}$$

Given the above pdf, the log likelihood function for the 2-parameter Weibull distribution for complete failure and right censored data according to (4.22) is:

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{\beta}{\alpha} \left(\frac{t_i}{\alpha} \right)^{\beta-1} e^{-\left(\frac{t_i}{\alpha}\right)^\beta} \right] - \sum_{j=1}^{N_r} n_j \cdot \left(\frac{t_j}{\alpha} \right)^\beta \quad (4.31)$$

The MLE solution is then obtained by solving for the pair of parameters (α and β) such that $\frac{\partial \Lambda}{\partial \alpha} = 0$ and $\frac{\partial \Lambda}{\partial \beta} = 0$.

where

$$\begin{aligned} \frac{\partial \Lambda}{\partial \alpha} &= -\frac{\beta}{\alpha} \sum_{i=1}^{N_c} n_i + \frac{\beta}{\alpha} \sum_{i=1}^{N_c} n_i \cdot \left(\frac{t_i}{\alpha} \right)^\beta + \frac{\beta}{\alpha} \sum_{j=1}^{N_r} n_j \cdot \left(\frac{t_j}{\alpha} \right)^\beta \\ \frac{\partial \Lambda}{\partial \beta} &= \frac{1}{\beta} \sum_{i=1}^{N_c} n_i + \sum_{i=1}^{N_c} n_i \cdot \ln \left(\frac{t_i}{\alpha} \right) - \sum_{i=1}^{N_c} n_i \cdot \left(\frac{t_i}{\alpha} \right)^\beta \ln \left(\frac{t_i}{\alpha} \right) - \sum_{j=1}^{N_r} n_j \cdot \left(\frac{t_j}{\alpha} \right)^\beta \ln \left(\frac{t_j}{\alpha} \right) \end{aligned} \quad (4.32)$$

Example 4.4

Consider times to failure (in hours) as follows: 16, 34, 53, 75, 93, 120; and assuming a 2-parameter Weibull distribution, find the MLE estimates of the parameters. Estimate the parameters of the pdf using MLE.

Solution

Since we have non-grouped data with no censored units, $N_c=6$ and $n_i=1$ for all i 's. Therefore,

$$\begin{aligned} \frac{\partial \Lambda}{\partial \alpha} &= -\frac{6\beta}{\alpha} + \frac{\beta}{\alpha} \sum_{i=1}^6 \left(\frac{t_i}{\alpha} \right)^\beta = 0 \\ \frac{\partial \Lambda}{\partial \beta} &= \frac{6}{\beta} + \sum_{i=1}^6 \ln \left(\frac{t_i}{\alpha} \right) - \sum_{i=1}^6 \left(\frac{t_i}{\alpha} \right)^\beta \ln \left(\frac{t_i}{\alpha} \right) = 0 \end{aligned}$$

Solving the above equations simultaneously yields $\hat{\alpha} = 73.53$, $\hat{\beta} = 1.93$.

4.9.3. LOGNORMAL LIFE DISTRIBUTION

The lognormal distribution is another widely used probability distribution in reliability engineering, often representing cycles to failure in degradation mechanisms. The lognormal distribution applies to a random variable if its logarithm is normally distributed. It is defined as

$$f(t) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln t - \mu_t}{\sigma_t} \right)^2}; f(t) \geq 0, t > 0, -\infty < \mu_t < \infty; \sigma_t > 0 \quad (4.33)$$

where μ_t = mean of the logarithms of the times to failure, and σ_t = standard deviation of the natural logarithms of the times to failure. Given the above pdf, the log-likelihood function for the lognormal distribution, according to Equation (4.22), when only complete data and right-censored data are present is

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{1}{\sigma_t t_i} \phi \left(\frac{\ln(t_i) - \mu_t}{\sigma_t} \right) \right] + \sum_{j=1}^{N_r} n_j \cdot \ln \left[1 - \Phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right) \right] \quad (4.34)$$

where ϕ is the standard normal distribution value, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, and Φ is the cumulative normal distribution value, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$.

Note that the censored part of the log-likelihood equations is added to (instead of subtracted from) the failure part of the equation. This is because for the lognormal distribution, there is no closed form solution to the CDF, $F(t)$. Hence, the logarithm of the entire expression $1 - \Phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right)$ needs to be evaluated as part of the log likelihood equation. The MLE solution is then obtained by solving for the pair of parameters (μ_t and σ_t) such that $\frac{\partial \Lambda}{\partial \alpha} = 0$ and $\frac{\partial \Lambda}{\partial \beta} = 0$, where

$$\begin{aligned} \frac{\partial \Lambda}{\partial \mu_t} &= \frac{1}{\sigma_t^2} \sum_{i=1}^{N_c} n_i \cdot [\ln(t_i) - \mu_t] + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{\phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right)}{1 - \Phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right)} \\ \frac{\partial \Lambda}{\partial \sigma_t} &= \sum_{i=1}^{N_c} n_i \cdot \left\{ \frac{[\ln(t_i) - \mu_t]^2}{\sigma_t^3} - \frac{1}{\sigma_t} \right\} + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{\left[\frac{\ln(t_j) - \mu_t}{\sigma_t} \right] \phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right)}{1 - \Phi \left(\frac{\ln(t_j) - \mu_t}{\sigma_t} \right)} \end{aligned} \quad (4.35)$$

Once the estimates of $\hat{\mu}_t$ and $\hat{\sigma}_t$ are determined, the actual mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of the lognormal distribution can be obtained from the following relationships:

$$\begin{aligned} \hat{\mu} &= e^{\hat{\mu}_t + \frac{1}{2}\hat{\sigma}_t^2} \\ \hat{\sigma} &= \sqrt{\left(e^{2\hat{\mu}_t + \frac{1}{2}\hat{\sigma}_t^2} \right) \left(e^{\hat{\sigma}_t^2} - 1 \right)} \end{aligned} \quad (4.36)$$

Example 4.5

Consider the following times to failure data (in hrs): 144, 385, 747, 1144, 1576, and 2612. Assuming a lognormal distribution, obtain MLE estimates of the parameters.

Solution

Since we have non-grouped data with no censored data, $F=6$ and $N_i=1$ for all i 's. Therefore,

$$\frac{\partial \Lambda}{\partial \mu_t} = \frac{1}{\sigma_t^2} \sum_{i=1}^6 [\ln(t_i) - \mu_t] = 0, \text{ and } \frac{\partial \Lambda}{\partial \sigma_t} = \sum_{i=1}^6 \left\{ \frac{[\ln(t_i) - \mu_t]^2}{\sigma_t^3} - \frac{1}{\sigma_t} \right\} = 0$$

Substituting the given values of T_i and solving the above equations simultaneously yields:

$$\hat{\mu}_t = 6.64$$

$$\hat{\sigma}_t = 1.04$$

The mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of the times-to-failure can be estimated using:

$$\hat{\mu} = e^{\hat{\mu}_t + \frac{1}{2}\hat{\sigma}_t^2} = e^{6.64 + \frac{1}{2}(1.04)^2} \cong 1313 \text{ hrs}$$

$$\hat{\sigma} = \sqrt{\left(e^{2\hat{\mu}_t + \frac{1}{2}\hat{\sigma}_t^2}\right)\left(e^{\hat{\sigma}_t^2} - 1\right)} = \sqrt{\left(e^{2 \times 6.64 + \frac{1}{2}1.04^2}\right)\left(e^{1.0443^2} - 1\right)} \cong 1406 \text{ hrs}$$

4.10. MLE-BASED PARAMETER ESTIMATION FOR DIFFERENT LIFE-STRESS MODELS

Section 4.9 outlined three life distributions and the corresponding MLE estimates for their respective parameters. However, accelerated life models not only deal with determining the life distribution at each tested stress level, but more importantly, life distribution conditional on the level of stress and estimation of parameters of the corresponding life-stress relationship. Based on the distribution of life-stress models, we can extrapolate and predict the life distribution and other metrics of interest at the use stress level. For this reason, a distribution of life conditional on the level of stress is used, and its relevant parameters are estimated based on the ALT results using the MLE approach. This section will outline the MLE approach for determining the parameters of some common ALT models.

4.10.1. THE EXPONENTIAL LIFE-STRESS MODEL

The exponential life-stress model such as the Arrhenius model is one of the most common models used in accelerated life testing of items subject to temperature stresses. The Arrhenius model derived from the well-known Arrhenius reaction rate expression:

$$R(T) = A e^{-\frac{E_a}{kT}} \quad (4.37)$$

where R is the reaction rate, A is thermal constant, E_a is the activation energy, the energy that a molecule must possess in order to participate in the reaction (i.e., it is a measure of the effect that temperature has on the reaction), K is Boltzmann's constant ($8.618 \times 10^{-5} \text{ eV K}^{-1}$), and T is the absolute temperature (in Kelvin).

The general exponential model is formulated by assuming that life is proportional to the inverse reaction rate, R in Equation (4.37), represented by

$$L(T) = b e^{\frac{a}{T}} \quad (4.38)$$

where $L(T)$ is life at temperature T (in Kelvin) and a and b are the model parameters.

The linear form of Equation (4.38) is often used in ALT in form of

$$\underbrace{\ln(L(T))}_{\substack{\text{dependent} \\ \text{variable (life)}}} = \underbrace{\ln(b)}_{\text{intercept}} + \underbrace{a}_{\text{slope}} \cdot \underbrace{\left(\frac{1}{T}\right)}_{\substack{\text{independent} \\ \text{variable (stress)}}}. \quad (4.39)$$

The corresponding acceleration factor for the Arrhenius model using Equation (4.38) is given by

$$AF = \frac{L_{use}}{L_{acc}} = \frac{be^{\frac{a}{T_{use}}}}{be^{\frac{a}{T_{acc}}}} = e^{\left(\frac{a}{T_{use}} - \frac{a}{T_{acc}}\right)} \quad (4.40)$$

4.10.2. EXPONENTIAL LIFE-STRESS MODEL WITH WEIBULL LIFE DISTRIBUTION

In the case where a Weibull distribution describes the random variation in life at each accelerated stress level, and the exponential model describing life-stress relationship, the parameters of the combined model may be estimated from the MLE approach. For this purpose, it is important to first define this conditional combined model. By replacing the Weibull scale parameter α (which corresponds to 63.2% of expended life) with the life dependent variable in Equation (4.38), the joint conditional exponential life-stress Weibull distribution (conditional on temperature T) may be obtained as

$$f(t, T) = \frac{\beta t^{\beta-1}}{\left(be^{\frac{a}{T}}\right)^{\beta}} e^{-\left(\frac{t}{be^{\frac{a}{T}}}\right)^{\beta}} = \frac{\beta}{be^{\frac{a}{T}}} \left(\frac{t}{be^{\frac{a}{T}}}\right)^{\beta-1} e^{-\left(\frac{t}{be^{\frac{a}{T}}}\right)^{\beta}} \quad (4.41)$$

The expressions for mean and reliability metrics from the model in Equation (4.41) are $\mu = be^{\frac{a}{T}} \cdot \Gamma\left(\frac{1}{\beta} + 1\right)$ and $R(t, T) = e^{-\left(\frac{t}{be^{\frac{a}{T}}}\right)^{\beta}}$, respectively. Note that if the parameter, a , is positive, then reliability increases as stress decreases. The corresponding hazard rate function is $\lambda(t, T) = \frac{f(t, T)}{R(t, T)} = \frac{\beta}{be^{\frac{a}{T}}} \left(\frac{t}{be^{\frac{a}{T}}}\right)^{\beta-1}$.

The exponential life-stress with the Weibull life model for complete and right censored ALT data has the general log-likelihood function

$$\begin{aligned} \Lambda = & \sum_{i=1}^{N_c} n_i \cdot \ln[f(t_i, T_i; a, b, \beta)] \\ & + \sum_{j=1}^{N_r} n_j \cdot [1 - \ln[F(t_j, T_j; a, b, \beta)]] \end{aligned} \quad (4.42)$$

Therefore, the corresponding log-likelihood function for Equation (4.41) would be

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{\beta}{be^{\frac{a}{T_i}}} \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^{\beta-1} e^{-\left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^\beta} \right] - \sum_{j=1}^{N_r} n_j \cdot \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right)^\beta \quad (4.43)$$

The MLE solution (parameter estimates $\hat{\beta}, \hat{a}, \hat{b}$) will be obtained by solving Equation (4.43) for β, a , and b such that $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$, and $\frac{\partial \Lambda}{\partial b} = 0$ where

$$\begin{aligned} \frac{\partial \Lambda}{\partial \beta} &= \frac{1}{\beta} \sum_{i=1}^{N_c} n_i + \sum_{i=1}^{N_c} n_i \cdot \ln \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right) - \sum_{i=1}^{N_c} n_i \cdot \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^\beta \ln \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right) \\ &\quad - \sum_{j=1}^{N_r} n_j \cdot \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right)^\beta \cdot \ln \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right) = 0 \\ \frac{\partial \Lambda}{\partial a} &= -\beta \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} + \beta \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^\beta + \beta \sum_{j=1}^{N_r} n_j \cdot \frac{1}{T_j} \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right)^\beta = 0 \\ \frac{\partial \Lambda}{\partial b} &= -\frac{\beta}{b} \sum_{i=1}^{N_c} n_i + \frac{\beta}{b} \sum_{i=1}^{N_c} n_i \cdot \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^\beta + \frac{\beta}{b} \sum_{j=1}^{N_r} n_j \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right)^\beta = 0 \end{aligned} \quad (4.44)$$

From the above, we have three equations to solve for three unknowns. Solutions of such systems usually require computer numerical tools such as MATLAB or MS Excel Solver. Confidence bounds for model parameters were discussed in Section 4.7. For the Arrhenius-Weibull model, the local Fisher information matrix, evaluated using the MLE estimates of model parameters (in this case, $\hat{\beta}, \hat{a}, \hat{b}$) is given by:

$$\begin{aligned} F &= \begin{bmatrix} \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{a}) & \text{cov}(\hat{\beta}, \hat{b}) \\ \text{cov}(\hat{a}, \hat{\beta}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{b}) \\ \text{cov}(\hat{b}, \hat{\beta}) & \text{cov}(\hat{b}, \hat{a}) & \text{var}(\hat{b}) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \beta^2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial a} & -\frac{\partial^2 \Lambda}{\partial \beta \partial b} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \beta} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial b} \\ -\frac{\partial^2 \Lambda}{\partial b \partial \beta} & -\frac{\partial^2 \Lambda}{\partial b \partial a} & -\frac{\partial^2 \Lambda}{\partial b^2} \end{bmatrix}^{-1} \end{aligned} \quad (4.45)$$

As discussed in Section 4.8, for positive parameters, we treat the logarithm of parameters as tending to the normal distribution for large sample sizes. Therefore, for the positive parameters β and b the two-sided confidence interval at a significance level of α is given by

$$\left(\hat{\beta} e^{-\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} < \beta < \hat{\beta} e^{\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right) \text{ and } \\ \left(\hat{b} e^{-\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})}}{\hat{b}}} < b < \hat{b} e^{\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})}}{\hat{b}}} \right) \quad (4.46)$$

Since the parameter α can be both positive and negative, we can treat the parameter itself as tending to the normal distribution for large sample sizes. Therefore, the two-sided confidence interval for parameter α , at a significance level of α , is given by $(\hat{\alpha} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\alpha})} < \alpha < \hat{\alpha} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\alpha})})$.

Example 4.6

Consider the individual ALT censored data shown in the table below. Find estimates of the Arrhenius-Weibull life-stress model and the corresponding confidence intervals assuming 90% confidence bounds.

Solution

Application of MLE on Equation (4.43) produces the point and confidence interval estimates for the Arrhenius-Weibull life model.

Failure Data			Censored Data		
i	t_i (hrs)	T_i ($^{\circ}$ C)	j	t_j (hrs)	T_j ($^{\circ}$ C)
1	14	200	1	80+	100
2	20	200	2	84+	100
3	21	200			
4	23	200			
5	35	200			
6	36	200			
7	39	200			
8	55	150			
9	56	150			
10	60	150			
11	67	150			
12	77	150			
13	70	100			
14	71	100			
15	75	100			
16	78	100			
17	89	100			

$$\hat{\beta} = 6.42; \hat{b} = 2.13; \hat{a} = 1406.08$$

$$5.18 < \hat{\beta} < 7.95$$

$$2.02 < \hat{b} < 2.24$$

$$1405.97 < \hat{a} < 1406.18$$

The corresponding Weibull multi plot covering each stress level is shown in Figure 4.11.

$$\hat{\beta} = 3.55; \hat{b} = 14.27; \hat{a} = 199.81$$

This corresponds to the mean values for $\hat{\beta}$, \hat{b} , and \hat{a} .

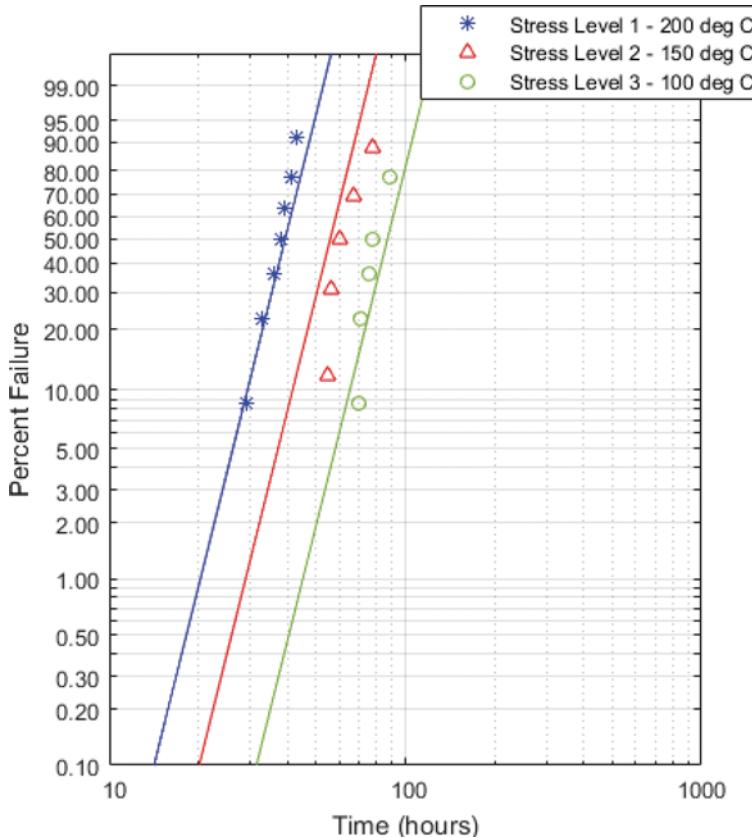


Figure 4.11: Weibull multi plot for Arrhenius-Weibull life-stress model

Example 4.7

Consider an accelerated test where temperature is the accelerating variable and 20 units were tested to failure assuming complete failures (no censoring). Eight units were tested at 406 °K, and six units each at 436 °K and 466 °K, with times to failure tabulated below. Assuming an Arrhenius-Weibull life-stress relationship, find the parameters of the model, using both the plotting method and MLE method and compare the results. Find the expected life at the use level temperature of 353 °K and compare the results.

Temperature	Times to Failure (hrs)								
	406 °K	248	456	528	731	813	965	972	1528
436 °K	164	176	289	319	386	459			
466 °K	92	105	155	184	219	235			

Solution

Using median ranks, the Weibull multi plot covering each stress level is shown in Figure 4.12.

From Figure 4.12, we can see that the Weibull shape parameters for each separate plot are reasonably close to one another (average $\beta = 2.46$), supporting the assumption of the Weibull life distribution for the ALT time to failure data. This plot also shows the estimates for the Weibull parameters α and β at each stress level. Given that the life-stress relationship is described by the Arrhenius relationship,

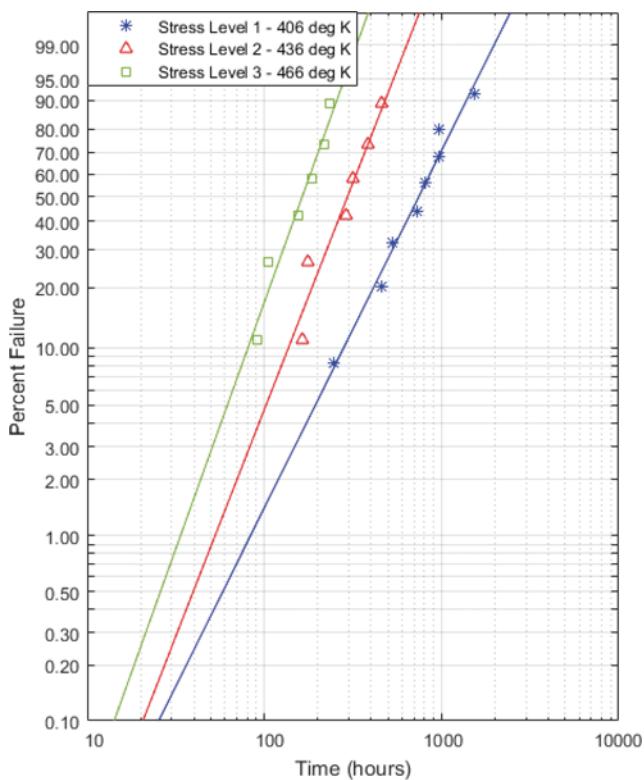


Figure 4.12: Multi plot for exponential life-stress with Weibull life

$L(T) = be^{\frac{a}{T}}$ and its logarithm $\ln[L(T)] = \ln(b) + a\left(\frac{1}{T}\right)$, we can estimate the parameters of the above equation with the estimated life-stress values from the plot. Hence, we can obtain estimates for the parameters of the Arrhenius model through the least squares regression method, which yield the following:

$$\ln(b) = -5.49; \text{ or } \hat{b} = 4.12 \times 10^{-3}$$

$$\hat{a} = 4971$$

From this we can extrapolate to the use temperature of 353 °K to obtain an expected characteristic life value of 5378 hrs. The plot of this extrapolation is shown in Figure 4.13.

The MLE estimates of the model parameters are

$$\hat{\beta} = 2.69; \hat{b} = 2.4 \times 10^{-3}; \hat{a} = 5201$$

To illustrate how the MLE equations are constructed, the expression for $\frac{\partial \Lambda}{\partial a}$ is shown below. As discussed, numerical methods are typically used to determine the solution to these equations.

$$\frac{\partial \Lambda}{\partial a} = -\beta \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} + \beta \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} \left(\frac{t_i}{be^{\frac{a}{T_i}}} \right)^\beta + \beta \sum_{j=1}^{N_r} n_j \cdot \frac{1}{T_j} \left(\frac{t_j}{be^{\frac{a}{T_j}}} \right)^\beta$$

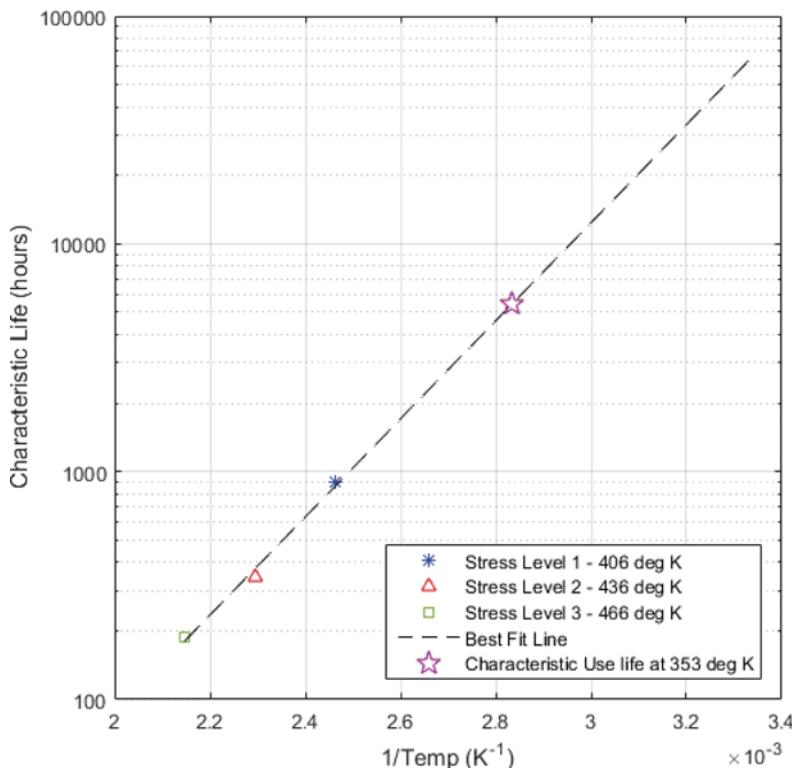


Figure 4.13: Extrapolation for exponential life-stress with Weibull time to failure

Since there were no censored points in this example, only the first 2 summation terms are applicable. Therefore,

$$\begin{aligned} \frac{\partial \Lambda}{\partial a} = -\beta \left\{ \frac{1}{406} + \frac{1}{436} + \frac{1}{436} + \frac{1}{436} + \frac{1}{436} + \frac{1}{436} + \frac{1}{436} + \right. \\ \left. \frac{1}{466} + \frac{1}{466} + \frac{1}{466} + \frac{1}{466} + \frac{1}{466} + \frac{1}{466} \right\} + \beta \left\{ \frac{1}{406} \left(\frac{248}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{456}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{528}{Ce^{406}} \right)^\beta + \right. \\ \left. \frac{1}{406} \left(\frac{731}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{813}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{965}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{972}{Ce^{406}} \right)^\beta + \frac{1}{406} \left(\frac{1528}{Ce^{406}} \right)^\beta + \frac{1}{436} \left(\frac{164}{Ce^{436}} \right)^\beta + \right. \\ \left. \frac{1}{436} \left(\frac{176}{Ce^{436}} \right)^\beta + \frac{1}{436} \left(\frac{289}{Ce^{436}} \right)^\beta + \frac{1}{436} \left(\frac{319}{Ce^{436}} \right)^\beta + \frac{1}{436} \left(\frac{386}{Ce^{436}} \right)^\beta + \frac{1}{436} \left(\frac{459}{Ce^{436}} \right)^\beta + \frac{1}{466} \left(\frac{92}{Ce^{466}} \right)^\beta + \right. \\ \left. \frac{1}{466} \left(\frac{105}{Ce^{466}} \right)^\beta + \frac{1}{466} \left(\frac{155}{Ce^{466}} \right)^\beta + \frac{1}{466} \left(\frac{184}{Ce^{466}} \right)^\beta + \frac{1}{466} \left(\frac{219}{Ce^{466}} \right)^\beta + \frac{1}{466} \left(\frac{235}{Ce^{466}} \right)^\beta \right\} = 0 \end{aligned}$$

For the MLE method, we can also find the confidence intervals associated with the model. For a confidence level of 90% (significance level of $\alpha=0.1$), we obtain:

$$\begin{aligned} \left(\hat{\beta} \cdot e^{-\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} < \beta < \hat{\beta} \cdot e^{\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right) \Rightarrow 2.02 < \beta < 3.59 \\ \left(\hat{b} \cdot e^{-\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})}}{\hat{b}}} < b < \hat{b} \cdot e^{\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})}}{\hat{b}}} \right) \Rightarrow 2.00 \times 10^{-4} < b < 2.36 \times 10^{-2} \\ \left(\hat{a} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} \right) \Rightarrow 4218 < a < 6186 \end{aligned}$$

4.10.3. EXPONENTIAL LIFE-STRESS MODEL WITH LOGNORMAL LIFE DISTRIBUTION

In the case where a lognormal distribution describes the life distribution at each accelerated stress level, to determine the MLE model parameters for an exponential life-stress model, the median of the lognormal distribution, μ , serves as the nominal value for life variable (dependent variable) in the exponential life-stress model. That is,

$$\mu = e^{\mu_t} = L(T) = b e^{\frac{a}{T}} \quad (4.47)$$

The linear version of Equation (4.47) would be

$$\mu_t = \ln(b) + \frac{a}{T} \quad (4.48)$$

We can then substitute the μ_t in Equation (4.48) for the same term in the lognormal pdf to obtain the conditional life-stress exponential life-stress model with lognormal distribution representing the life model as follows:

$$f(t, T) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) - \ln(b) - \frac{a}{T}}{\sigma_t} \right)^2} \quad (4.49)$$

Accordingly, we may estimate mean life and reliability as $\mu = e^{\mu_t + \frac{1}{2}\sigma_t^2} = e^{\ln(C) + \frac{B}{T} + \frac{1}{2}\sigma_t^2}$ and $R(t, T) = \int_t^\infty f(t, T) dt = \int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(C) - \frac{B}{T}}{\sigma_t}\right)^2} dt$, respectively. The lognormal reliability function has no closed form solution, and therefore must be numerically solved. Similarly, the hazard rate may

$$\text{be obtained from } \lambda(t, V) = \frac{f(t, V)}{R(t, V)} = \frac{\frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(C) - \frac{B}{V}}{\sigma_t}\right)^2}}{\int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(C) - \frac{B}{V}}{\sigma_t}\right)^2} dt}.$$

The exponential life-stress with lognormal pdf has a log-likelihood function that can be written using Equation (4.22) in the following form:

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{1}{\sigma_t t_i} \phi \left(\frac{\ln(t_i) - \ln(b) - \frac{a}{T_i}}{\sigma_t} \right) \right] + \sum_{j=1}^{N_r} n_j \cdot \ln \left[1 - \Phi \left(\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right) \right] \quad (4.50)$$

The MLE solution (parameter estimates $\hat{\sigma}_t$, \hat{a} , \hat{b}) will be obtained by solving variables σ_t , a , and b in the three equations $\frac{\partial \Lambda}{\partial \sigma_t} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$, and $\frac{\partial \Lambda}{\partial b} = 0$. These equations are

$$\begin{aligned} \frac{\partial \Lambda}{\partial \sigma_t} &= \sum_{i=1}^{N_c} n_i \left\{ \frac{\left[\ln(t_i) - \ln(b) - \frac{a}{T_i} \right]^2}{\sigma_t^3} - \frac{1}{\sigma_t} \right\} \\ &\quad + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{\left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right] \cdot \phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]} \\ &= 0 \end{aligned} \quad (4.51)$$

$$\begin{aligned} \frac{\partial \Lambda}{\partial a} &= \frac{1}{\sigma_t^2} \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} \left[\ln(t_i) - \ln(b) - \frac{a}{T_i} \right] \\ &\quad + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{1}{T_j} \left\{ \frac{\phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]} \right\} = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial \Lambda}{\partial b} &= \frac{1}{b \cdot \sigma_t^2} \sum_{i=1}^{N_c} n_i \left[\ln(t_i) - \ln(b) - \frac{a}{T_i} \right] \\ &\quad + \frac{1}{b \cdot \sigma_t} \sum_{j=1}^{N_r} n_j \cdot \left\{ \frac{\phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) - \ln(b) - \frac{a}{T_j}}{\sigma_t} \right]} \right\} = 0 \end{aligned}$$

From Equation (4.51), we have three equations to solve for three unknowns. Solutions of such systems usually require numerical tools. Confidence intervals for the model parameters using the local Fisher information matrix are given by:

$$\begin{aligned} F &= \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{b}) \\ \text{cov}(\hat{a}, \hat{\sigma}_t) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{b}) \\ \text{cov}(\hat{b}, \hat{\sigma}_t) & \text{cov}(\hat{b}, \hat{a}) & \text{var}(\hat{b}) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial b} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial b} \\ -\frac{\partial^2 \Lambda}{\partial b \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial b \partial a} & -\frac{\partial^2 \Lambda}{\partial b^2} \end{bmatrix}^{-1} \end{aligned} \quad (4.52)$$

For positive parameters, we treat the logarithm of parameters as tending to the normal distribution for large sample sizes. Therefore, for positive parameters σ_t and b , the two-sided confidence interval at a

significance level of α is given by $\left(\hat{\sigma}_t \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} < \sigma_t < \hat{\sigma}_t \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} \right)$ and
 $\left(\hat{b} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{b})}}{\hat{b}}} < b < \hat{b} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{b})}}{\hat{b}}} \right)$.

Since the parameter a can be both positive and negative, we can treat the parameter itself as tending to the normal distribution for large sample sizes. Therefore, the two-sided confidence interval for parameter a , at a significance level of α is given by $\left(\hat{a} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{a})} \right)$.

4.10.4. THE EYRING LIFE-STRESS MODEL

The Eyring life-stress relationship is a special form of the exponential life-stress model that is commonly used when the acceleration variable is thermal in nature (e.g., temperature or relative humidity). The Eyring relationship is given by:

$$L(T) = \frac{1}{T} e^{-(c - \frac{a}{T})} \quad (4.53)$$

where a and c are the model parameters. Note that the Eyring relationship is a slight variation of the Arrhenius model (by rearranging terms), with the Eyring relationship only differing due to the new $\frac{1}{T}$ term in Equation (4.53). Rearranging (4.53) and making it log-linear, then,

$$\ln(L(T)) = \underbrace{-c}_{\text{intercept}} + \underbrace{\frac{a}{T}}_{\text{slope}} \cdot \underbrace{\left(\frac{1}{T}\right)}_{\substack{\text{independent} \\ \text{variable}}} - \ln(T) \quad (4.54)$$

The acceleration factor for the Eyring model is given by:

$$AF = \frac{L_{use}}{L_{acc}} = \frac{\frac{1}{T_{use}} e^{-(c - \frac{a}{T_{use}})}}{\frac{1}{T_{acc}} e^{-(c - \frac{a}{T_{acc}})}} = \frac{T_{acc}}{T_{use}} e^{a(\frac{1}{T_{use}} - \frac{1}{T_{acc}})} \quad (4.55)$$

4.10.5. THE EYRING-WEIBULL MODEL

In the case where a Weibull distribution describes the life distribution at each accelerated stress level, the method for determining the Eyring life-stress model parameters is outlined below.

First, the Weibull scale parameter α (which corresponds to 63.2% of life expended) as the nominal value for Weibull life model should be replaced by the life dependent variable in the Eyring model. Accordingly, $\alpha = L(T) = \frac{1}{T} e^{-(c - \frac{a}{T})}$. Therefore, the conditional life-stress model is given by

$$f(t, T) = \beta \cdot T \cdot e^{(c - \frac{a}{T})} \left(t \cdot T \cdot e^{(c - \frac{a}{T})} \right)^{\beta-1} e^{-\left(t \cdot T \cdot e^{(c - \frac{a}{T})} \right)^\beta} \quad (4.56)$$

The expressions for mean life, reliability and hazard rate relationships are $\mu = \frac{1}{T} e^{(c - \frac{a}{T})} \Gamma(\frac{1}{\beta} + 1)$, $R(t, T) = e^{-\left(t \cdot T \cdot e^{(c - \frac{a}{T})} \right)^\beta}$ and $\lambda(t, T) = \frac{f(t, T)}{R(t, T)} = \beta \cdot T \cdot e^{(c - \frac{a}{T})} \left(t \cdot T \cdot e^{(c - \frac{a}{T})} \right)^{\beta-1}$, respectively.

The Eyring-Weibull log-likelihood function for complete and right censored data is

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left\{ \beta \cdot T_i \cdot e^{(c - \frac{a}{T_i})} \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^{\beta-1} e^{-\left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta} \right\} - \sum_{j=1}^{N_r} n_j \cdot \left[t_j T_j e^{(c - \frac{a}{T_j})} \right]^\beta \quad (4.57)$$

Note that Equation (4.57) can account for either non-grouped sets of data, or when only a single data point exists at each group (i.e., where both n_i and n_j are unity for all i 's and j 's). The MLE solution will be obtained by solving for β , a , and c such that $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$ and $\frac{\partial \Lambda}{\partial c} = 0$. That is,

$$\begin{aligned}
\frac{\partial \Lambda}{\partial \beta} &= \frac{1}{\beta} \sum_{i=1}^{N_c} n_i + \sum_{i=1}^{N_c} n_i \cdot \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] \\
&\quad - \sum_{i=1}^{N_c} n_i \cdot \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] \\
&\quad - \sum_{j=1}^{N_r} n_j \cdot \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta \cdot \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] = 0, \\
\frac{\partial \Lambda}{\partial a} &= \beta \sum_{i=1}^{N_c} n_i - \beta \sum_{i=1}^{N_c} n_i \cdot \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta - \beta \sum_{j=1}^{N_r} n_j \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta \\
&= 0 \\
\frac{\partial \Lambda}{\partial c} &= \frac{1}{\beta} \sum_{i=1}^{N_c} n_i + \sum_{i=1}^{N_c} n_i \cdot \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] \\
&\quad - \sum_{i=1}^{N_c} n_i \cdot \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] \\
&\quad - \sum_{j=1}^{N_r} n_j \left[t_i T_i e^{(c - \frac{a}{T_i})} \right]^\beta \ln \left[t_i T_i e^{(c - \frac{a}{T_i})} \right] = 0
\end{aligned} \tag{4.58}$$

From Equation (4.58), we have three equations to solve for three unknowns, the solution of which usually requires numerical tools. Confidence intervals for the Eyring-Weibull model using the local Fisher information matrix are given by:

$$\begin{aligned}
F &= \begin{bmatrix} \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{a}) & \text{cov}(\hat{\beta}, \hat{c}) \\ \text{cov}(\hat{a}, \hat{\beta}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{c}) \\ \text{cov}(\hat{c}, \hat{\beta}) & \text{cov}(\hat{c}, \hat{a}) & \text{var}(\hat{c}) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \beta^2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial a} & -\frac{\partial^2 \Lambda}{\partial \beta \partial c} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \beta} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial c} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \beta} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c^2} \end{bmatrix}^{-1}
\end{aligned} \tag{4.59}$$

Therefore, for positive Eyring-Weibull parameter β , the two-sided confidence interval at a significance level of α is given by $\left(\hat{\beta} e^{-\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} < \beta < \hat{\beta} e^{\frac{z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right)$.

Since the parameter a and c can be both positive and negative, we can treat the parameter itself as tending to the normal distribution for large sample sizes. Therefore, the two-sided confidence interval

for parameter a and c , at a significance level of α , is given by $\left(\hat{a} - Z_{1-\frac{\alpha}{2}}\sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1+\frac{\alpha}{2}}\sqrt{\text{var}(\hat{a})}\right)$ and $\left(\hat{c} - Z_{1-\frac{\alpha}{2}}\sqrt{\text{var}(\hat{c})} < c < \hat{c} + Z_{1-\frac{\alpha}{2}}\sqrt{\text{var}(\hat{c})}\right)$.

4.10.6. THE EYRING-LOGNORMAL MODEL

In the case where a lognormal distribution describes the time to failure life distribution at each accelerated stress level, to determine the MLE of the model parameters, assign the median of the lognormal distribution, μ , as the nominal value for life-stress based on the Eyring model,

$$\mu = e^{\mu_t} = \frac{1}{T} e^{-\left(c - \frac{a}{T}\right)} \quad (4.60)$$

Or in log-linear form as

$$\mu_t = -\ln(T) - c + \frac{a}{T} \quad (4.61)$$

We can then substitute Equation (4.61) as the median parameter of the lognormal pdf to obtain the conditional Eyring life-stress with the lognormal model, as the life distribution would be

$$f(t, T) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) + \ln(T) + c - \frac{a}{T}}{\sigma_t}\right)^2} \quad (4.62)$$

The expressions for the mean and reliability functions derived from the joint life-stress pdf model in Equation (4.62) are $\mu = e^{\mu_t + \frac{1}{2}\sigma_t^2} = e^{-\ln(T) - c + \frac{a}{T} + \frac{1}{2}\sigma_t^2}$ and $R(t, T) = \int_t^\infty f(t, T) dt = \int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) + \ln(T) + c - \frac{a}{T}}{\sigma_t}\right)^2} dt$, respectively. The solution for the lognormal reliability function has

no closed form solution, and therefore must be solved using numerical methods. Also, the hazard rate associated with the pdf in Equation (4.62) is

$$\lambda(t, T) = \frac{f(t, T)}{R(t, T)} = \frac{\frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) + \ln(T) + c - \frac{a}{T}}{\sigma_t}\right)^2}}{\int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) + \ln(T) + c - \frac{a}{T}}{\sigma_t}\right)^2} dt}. \quad (4.63)$$

The Eyring-lognormal log-likelihood function is derived from the general log-likelihood function of Equation (4.22) as follows:

$$\ln(L) = \Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{1}{\sigma_t t_i} \phi \left(\frac{\ln(t_i) + \ln(T_i) + c - \frac{a}{T_i}}{\sigma_t} \right) \right] \\ + \sum_{j=1}^{N_r} n_j \cdot \ln \left[1 - \Phi \left(\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right) \right] \quad (4.64)$$

The MLE solution (to estimate parameters $\hat{\sigma}_t$, \hat{a} , \hat{c}) will be obtained by solving for σ_t , a , and c such that $\frac{\partial \Lambda}{\partial \sigma_t} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$ and $\frac{\partial \Lambda}{\partial c} = 0$. These relationships are

$$\frac{\partial \Lambda}{\partial \sigma_t} \\ = \sum_{i=1}^{N_c} n_i \left\{ \frac{\left[\ln(t_i) + \ln(T_i) + c - \frac{a}{T_i} \right]^2}{\sigma_t^3} - \frac{1}{\sigma_t} \right\} \\ + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{\frac{\left[\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j} \right]}{\sigma_t} \cdot \phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]} = 0 \quad (4.65)$$

$$\frac{\partial \Lambda}{\partial c} = -\frac{1}{\sigma_t^2} \sum_{i=1}^{N_c} n_i \cdot \frac{1}{T_i} \left[\ln(t_i) + \ln(T_i) + c - \frac{a}{T_i} \right] \\ - \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{1}{T_j} \left\{ \frac{\phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]} \right\} = 0$$

$$\frac{\partial \Lambda}{\partial a} = \frac{1}{\sigma_t^2} \sum_{i=1}^{N_c} n_i \left[\ln(t_i) + \ln(T_i) + c - \frac{a}{T_i} \right] + \\ \frac{1}{\sigma_t} \sum_{j=1}^{N_r} n_j \cdot \frac{1}{T_j} \left\{ \frac{\phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln(t_j) + \ln(T_j) + c - \frac{a}{T_j}}{\sigma_t} \right]} \right\} = 0$$

From Equation (4.65), we have three equations to solve for three unknowns. Solutions of such systems usually require numerical tools. For the Arrhenius-lognormal model, the local Fisher information matrix, evaluated using the MLE estimates of model parameters (in this case, $\hat{\sigma}_t$, \hat{a} , \hat{c}), is given by:

$$\begin{aligned}
F &= \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t, \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{c}) \\ \text{cov}(\hat{a}, \hat{\sigma}_t) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{c}) \\ \text{cov}(\hat{c}, \hat{\sigma}_t) & \text{cov}(\hat{c}, \hat{a}) & \text{var}(\hat{c}) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial c} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial c} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c^2} \end{bmatrix}^{-1} \quad (4.66)
\end{aligned}$$

For the positive Eyring-lognormal parameters σ_t , the two-sided confidence interval at a significance level of α is given by $\left(\hat{\sigma}_t \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} < \sigma_t < \hat{\sigma}_t \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} \right)$. Since the parameter a and

c can be both positive and negative, we can treat the parameter itself as tending to the normal distribution for large sample sizes. Therefore, the two-sided confidence interval for parameter a and c , at a significance level of α , is given by $\left(\hat{a} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{a})} \right)$ and $\left(\hat{c} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{c})} < c < \hat{c} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{c})} \right)$.

4.10.7. POWER LIFE-STRESS MODEL

The power life-stress model is another popular accelerated life model, commonly used in applications where the applied stresses are non-thermal in nature (e.g. fatigue). The inverse power life-stress model is the most useful form given by:

$$L(S) = \frac{1}{aS^n} \quad (4.67)$$

where $L(S)$ = life at stress S , and a and n are model parameters to be determined from ALT analysis. The log-linear form of Equation (4.67) is

$$\ln(L(S)) = \underbrace{-\ln(a)}_{\text{intercept}} + \underbrace{\frac{n}{S}}_{\text{slope}} \cdot \underbrace{\ln(S)}_{\text{independent variable}} \quad (4.68)$$

The acceleration factor for the IPL model is given by:

$$AF = \frac{L_{use}}{L_{acc}} = \frac{\frac{1}{KS_{use}^n}}{\frac{1}{KS_{acc}^n}} = \left(\frac{S_{acc}}{S_{use}} \right)^n \quad (4.69)$$

4.10.8. POWER LIFE-STRESS WITH WEIBULL LIFE MODEL

When a Weibull distribution describes the life at each accelerated stress level, the Weibull scale parameter α , which corresponds to 63.2% of life, may be chosen as the nominal value for life in the power life-stress with the Weibull life distribution

$$\alpha = L(S) = \frac{1}{aS^n} \quad (4.70)$$

Therefore, the conditional power life-stress with the Weibull distribution representing the time-to-failure is given by:

$$f(t, S) = \frac{\beta t^{\beta-1}}{\left(\frac{1}{aS^n}\right)^\beta} e^{-\left(\frac{t}{aS^n}\right)^\beta} = \beta aS^n (aS^n t)^{\beta-1} e^{-(aS^n t)^\beta} \quad (4.71)$$

The expressions for mean and reliability functions derived from the model in Equation (4.71) are $\mu = \frac{1}{aS^n} \Gamma\left(\frac{1}{\beta} + 1\right)$ and $R(t, S) = e^{-(aS^n t)^\beta}$, respectively. Note from above, that if the parameter β is positive, then reliability increases as stress decreases. Also, the corresponding hazard rate function is $\lambda(t, S) = \frac{f(t, S)}{R(t, S)} = \beta aS^n (aS^n t)^{\beta-1}$.

The Power-Weibull log-likelihood function for complete and right censored data is

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln[f(t_i, S_i; \beta, a, n)] + \sum_{j=1}^{N_r} n_j \cdot \ln[F(t_j, S_j; \beta, a, n)] \quad (4.72)$$

When combined the reliability function $R(t, S)$ with Equation (4.71), the likelihood function in Equation (4.72) can be represented by

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\beta aS_i^n (aS_i^n t_i)^{\beta-1} e^{-(aS_i^n t_i)^\beta} \right] - \sum_{j=1}^{N_r} n_j \cdot (aS_j^n t_j)^\beta \quad (4.73)$$

The MLE solution (for parameter estimates $\hat{\beta}, \hat{a}, \hat{n}$) will be obtained by solving for β, a , and n such that $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$ and $\frac{\partial \Lambda}{\partial n} = 0$, where

$$\begin{aligned} \frac{\partial \Lambda}{\partial \beta} &= \frac{1}{\beta} \sum_{i=1}^{N_c} n_i + \sum_{i=1}^{N_c} n_i \ln(aS_i^n t_i) - \sum_{i=1}^{N_c} n_i (aS_i^n t_i)^\beta \ln(aS_i^n t_i) \\ &\quad - \sum_{j=1}^{N_r} n_j (aS_j^n t_j)^\beta \ln(aS_j^n t_j) = 0 \end{aligned} \quad (4.74)$$

$$\frac{\partial \Lambda}{\partial a} = \frac{\beta}{a} \sum_{i=1}^{N_c} n_i - \frac{\beta}{a} \sum_{i=1}^{N_c} n_i (aS_i^n t_i)^\beta - \frac{\beta}{a} \sum_{i=1}^{N_r} n_j (aS_j^n t_j)^\beta = 0$$

$$\begin{aligned}\frac{\partial \Lambda}{\partial n} &= \beta \sum_{i=1}^{N_c} n_i \ln(S_i) - \beta \sum_{i=1}^{N_c} n_i \cdot \ln(S_i) (a S_i^n t_i)^\beta \\ &\quad - \beta \sum_{j=1}^{N_r} n_j \cdot \ln(S_j) (a S_j^n t_j)^\beta = 0\end{aligned}$$

From Equation (4.74), we have three equations to solve for three unknowns.

For the Power-Weibull model, the local Fisher information matrix, evaluated using the MLE estimates of model parameters (in this case, $\hat{\beta}, \hat{a}, \hat{n}$) is given by:

$$\begin{aligned}F &= \begin{bmatrix} \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{a}) & \text{cov}(\hat{\beta}, \hat{n}) \\ \text{cov}(\hat{a}, \hat{\beta}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{n}) \\ \text{cov}(\hat{n}, \hat{\beta}) & \text{cov}(\hat{n}, \hat{a}) & \text{var}(\hat{n}) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \beta^2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial a} & -\frac{\partial^2 \Lambda}{\partial \beta \partial n} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \beta} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial n} \\ -\frac{\partial^2 \Lambda}{\partial n \partial \beta} & -\frac{\partial^2 \Lambda}{\partial n \partial a} & -\frac{\partial^2 \Lambda}{\partial n^2} \end{bmatrix}^{-1} \quad (4.75)\end{aligned}$$

For the positive Power-Weibull parameters β and a , the two-sided confidence interval at a significance

level of α is given by $\left(\hat{\beta} \cdot e^{\frac{z_{\alpha/2}}{2}} < \beta < \hat{\beta} \cdot e^{-\frac{z_{\alpha/2} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right)$ and $\left(\hat{a} \cdot e^{-\frac{z_{1-\alpha/2} \sqrt{\text{var}(\hat{a})}}{\hat{a}}} < a < \hat{a} \cdot e^{\frac{z_{1-\alpha/2} \sqrt{\text{var}(\hat{a})}}{\hat{a}}} \right)$, respectively. Since the parameter n can be both positive and negative, the two-sided

confidence interval at a significance level of α is given by $\left(\hat{n} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} < n < \hat{n} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} \right)$.

4.10.9. POWER LIFE-STRESS WITH LOGNORMAL MODEL

In the case where a lognormal distribution describes the life distribution at each accelerated stress level, the median of the lognormal distribution, μ , would serve as the nominal value for life in the power life-stress model according to

$$\mu = e^{\mu_t} = L(S) = \frac{1}{a S^n} \quad (4.76)$$

The log-linear form of Equation (4.76) is

$$\mu_t = -\ln(a) - n \ln(S) \quad (4.77)$$

By substituting Equation (4.77) as the median of the lognormal pdf, the conditional power-lognormal given stress is obtained as follows:

$$f(t, S) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) + \ln(a) + n \ln(S)}{\sigma_t} \right)^2} \quad (4.78)$$

The expressions for the mean life and reliability functions based on Equation (4.78) are $\mu = e^{\mu_t + \frac{1}{2}\sigma_t^2} = e^{-\ln(a) - n \ln(S) + \frac{1}{2}\sigma_t^2}$ and $R(t, S) = \int_t^\infty f(t, S) dt = \int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) + \ln(a) + n \ln(S)}{\sigma_t} \right)^2} dt$, respectively.

The power-lognormal reliability function has no closed form solution, and therefore must be solved using numerical methods. The hazard rate of the power-lognormal model in Equation (4.78) is

$$\lambda(t, S) = \frac{f(t, S)}{R(t, S)} = \frac{\frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) + \ln(a) + n \ln(S)}{\sigma_t} \right)^2}}{\int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) + \ln(a) + n \ln(S)}{\sigma_t} \right)^2} dt} \quad (4.79)$$

The power-lognormal log-likelihood function based on Equation (4.78) is derived from the general log-likelihood function in Equation (4.22) for the lognormal distribution as follows:

$$\begin{aligned} \Lambda = & \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{1}{\sigma_t t_i} \phi \left(\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right) \right] \\ & + \sum_{j=1}^{N_r} n_j \cdot \ln \left[1 - \Phi \left(\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right) \right] \end{aligned} \quad (4.80)$$

The MLE solution for parameter estimates $\hat{\sigma}_t$, \hat{a} , \hat{n} will be obtained by solving for σ_t , a , and n such that $\frac{\partial \Lambda}{\partial \sigma_t} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$ and $\frac{\partial \Lambda}{\partial n} = 0$, where

$$\begin{aligned} \frac{\partial \Lambda}{\partial \sigma_t} = & \sum_{i=1}^{N_c} n_i \left\{ \frac{[\ln t_i + \ln a + n \ln S_i]^2}{\sigma_t^3} - \frac{1}{\sigma_t} \right\} \\ & + \frac{1}{\sigma_t} \sum_{j=1}^{N_r} N_j \cdot \frac{\left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right] \cdot \phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]} \\ = & 0 \end{aligned} \quad (4.81)$$

$$\begin{aligned} \frac{\partial \Lambda}{\partial n} = & -\frac{1}{\sigma_t^2} \sum_{i=1}^{N_c} n_i \cdot \ln(S_i) [\ln t_i + \ln a + n \ln S_i] \\ & - \frac{1}{\sigma_t^2} \sum_{j=1}^{N_r} n_j \cdot \ln(S_i) \left\{ \frac{\phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]} \right\} = 0 \end{aligned}$$

$$\begin{aligned}\frac{\partial \Lambda}{\partial a} = & -\frac{1}{a \cdot \sigma_t^2} \sum_{i=1}^{N_c} n_i [\ln t_i + \ln a + n \ln S_i] \\ & - \frac{1}{a \cdot \sigma_t} \sum_{j=1}^{N_r} n_j \cdot \left\{ \frac{\phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]}{1 - \Phi \left[\frac{\ln t_i + \ln a + n \ln S_i}{\sigma_t} \right]} \right\} = 0\end{aligned}$$

From the above, we have three equations to solve for three unknowns. The local Fisher information matrix, evaluated using the MLE estimates of model parameters $\hat{\sigma}_t$, \hat{a} , \hat{n} , is given by

$$\begin{aligned}F = & \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t, \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{n}) \\ \text{cov}(\hat{a}, \hat{\sigma}_t) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{n}) \\ \text{cov}(\hat{n}, \hat{\sigma}_t) & \text{cov}(\hat{n}, \hat{a}) & \text{var}(\hat{n}) \end{bmatrix} \\ = & \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial n} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial n} \\ -\frac{\partial^2 \Lambda}{\partial n \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial n \partial a} & -\frac{\partial^2 \Lambda}{\partial n^2} \end{bmatrix}^{-1} \quad (4.82)\end{aligned}$$

For positive IPL-lognormal parameters σ_t and a , the two-sided confidence interval at a significance level of α is given by $\left(\hat{\sigma}_t \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} < \sigma_t < \hat{\sigma}_t \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} \right)$ and $\left(\hat{a} e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})}}{\hat{a}}} < a < \hat{a} e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})}}{\hat{a}}} \right)$. Since parameter n can be both positive and negative, the two-sided confidence

interval for this parameter at a significance level of α is given by $\left(\hat{n} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} < n < \hat{n} + Z_{1+\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} \right)$.

Example 4.8

A 2024-T3 Aluminum structure was low-cycle fatigue tested to failure at two different stress values, 400 and 450 MPa, and the following data were obtained. Assume lognormal-IPL relationship. Estimate the number of cycles to failure at 300 MPa.

	Stress Range [MPa]	
	400	450
Cycles to failure	16	8
	18	9
	20	9
	21	10
	24	11
	27	13
	29	14

Solution

Using Equation (4.80), the parameters are calculated through the maximum likelihood estimator approach, $\hat{a} = 2.288 \times 10^{-18}$; $\hat{n} = 6.266$; and $\widehat{\sigma_t} = 0.195$. Accordingly, the mean life of the life-stress lognormal model at 300 MPa is calculated from the expected value expression for a lognormal distribution:

$$t = \exp \left[-\ln(a) - n \ln(S) + \frac{1}{2} \sigma_t^2 \right] = \exp[-\ln(2.288 \times 10^{-18}) - 6.266 \ln(300) + \frac{1}{2} (0.195)^2] = 134.2 \text{ cycles}$$

Example 4.9

Time to failure of a journal bearing is believed to follow a lognormal distribution. Accelerated life tests revealed the median life of the bearing as 2000 hours in the first test, where radial force of 750 Newtons was applied on the bearing and it was reduced 1500 hours when the radial force was increased to 800 Newtons in the second test. What is the expected median life for the bearing in normal operation where the radial force is 400 Newtons?

Solution

Radial force on the bearing is the acceleration variable in these tests. Higher radial force results in higher pressure in the lubricant film and ultimately higher normal stress on the surface. The friction has also a linear relationship with the normal load being applied on the bearing. τ_{max} , which is the

maximum shear stress in the vicinity of mating surfaces, is estimated from: $\tau_{max} = k_e \sqrt{\left(\frac{\sigma_n}{2}\right)^2 + \tau_f^2}$ where τ_{max} = maximum shearing stress; k_e = stress concentration factor; σ_n = normal stress on the surface; τ_f = friction generated shear stress; f = friction factor. Since normal stress σ_n is a linear function of bearing radial force, the τ_{max} becomes a linear function of applied radial force using the power life-stress model, so we have:

$$\frac{t_1}{t_2} = \frac{2000}{1500} = \frac{\left(\frac{a}{\tau_{max1}^n}\right)}{\left(\frac{a}{\tau_{max2}^n}\right)} = \left(\frac{800}{750}\right)^n \rightarrow n = 4.46$$

$$\frac{t_{use}}{t_2} = \frac{T_{use}}{1500} = \frac{\left(\frac{a}{\tau_{use}^n}\right)}{\left(\frac{a}{\tau_{max2}^n}\right)} = \left(\frac{800}{400}\right)^{4.46} \rightarrow T_{use} = 33013 \text{ hrs}$$

Example 4.10

A cylindrical tank made of steel alloy that contains chromium and molybdenum is subjected to degradation from the corrosion-fatigue mechanism. Estimate the mean crack growth rate at the stress intensity of 30 ksi-in^{1/2}, in a wet gas environment, according to the following data. Assume this is subject to the lognormal-IPL relationship.

ΔK [ksi-in ^{1/2}]	40	50	60	70
da/dN [in/cycle]	4.0×10^{-3}	8.1×10^{-3}	9.0×10^{-3}	9.8×10^{-3}

Solution

Crack growth behavior is generally modeled using the Paris equation, $\frac{da}{dN} = C(\Delta K)^m$, where a =crack length; N = number of cycles; ΔK =stress intensity factor range; and C , m = material-related parameters.

The resistance of the material to crack growth can be written as the inverse of the crack growth rate, $\frac{1}{da/dN} = \frac{1}{C(\Delta K)^m}$. Then, set $\frac{1}{da/dN} = R(S)$, where S represents the stress intensity factor and $R(S)$ the crack resistance. Since crack resistance is proportionally related to life (i.e., $R(S) \propto L(S)$) where $L(S)$ is life of the tank at stress S , then $L(S) = \frac{1}{C(\Delta K)^m}$. Assuming that life distribution of items under the corrosion-fatigue mechanism follows a lognormal life distribution, and the inverse power law describes the life-stress model, then Equation (4.78) is the applicable model. Also, Equation (4.80) would yield the MLE solutions to the associated parameters. According to Equation (4.80), $C = 1.496 \times 10^{-5}$; $m = 1.554$; and $\sigma_t = 0.819$.

Then, the mean life of the IPL-lognormal model at stress intensity 30 ksi-in^{1/2} is calculated from the expected value expression for a lognormal distribution:

$$\begin{aligned}\bar{a}' &= \exp \left[-\ln(C) - m \ln(\Delta K) + \frac{1}{2} \sigma_{a'}^2 \right] \\ &= \exp \left[-\ln(1.496 \times 10^{-5}) - 1.554 \ln(30) + \frac{1}{2} (0.819)^2 \right] \\ &= 473.4 \text{ cycles/in; or } 2.11 \times 10^{-3} \text{ in/cycles}\end{aligned}$$

4.10.10. DUAL-STRESS EXPONENTIAL LIFE-STRESS MODEL

This model is used for life-stress situations involving two independent stress conditions, each accelerating the degradation and reducing the life. This model can be considered as a variation of the Eyring model used in electronics when temperature and humidity are considered as agents. The Eyring relationship can be used when temperature or humidity are the accelerating variable alone. However, the dual stress exponential model can be used when both of these variables independently affect the life. The dual-stress exponential life-stress model is given by

$$L(T, H) = c e^{\left(\frac{a}{T} + \frac{b}{H}\right)} \quad (4.83)$$

where a , b and c are the model parameters to be determined, H is one stress agent, such as the relative humidity (expressed as a percentage), and T is the second stress, such as temperature (expressed in absolute units, °K). The log-linear version of Equation (4.83) is

$$\ln(L(T, H)) = \ln(c) + \frac{a}{T} + \frac{b}{H} \quad (4.84)$$

Note that since this model involves two different types of accelerating variables, keeping one of the stresses constant while varying the other can only generate a linear life-stress plot. Depending on which stress is kept constant, either the parameter a or b can become the slope of the linear plot. It is also important to note that using the dual-stress exponential model for ALT requires careful consideration of the choice of stress levels to be used during the test. The acceleration factor for the dual-stress life-stress model is given by:

$$AF = \frac{L_{use}}{L_{acc}} = \frac{ce^{\left(\frac{a}{T_{use}} + \frac{b}{H_{use}}\right)}}{ce^{\left(\frac{a}{T_{acc}} + \frac{b}{H_{acc}}\right)}} = e^{a\left(\frac{1}{T_{use}} - \frac{1}{T_{acc}}\right) + b\left(\frac{1}{H_{use}} - \frac{1}{H_{acc}}\right)} \quad (4.85)$$

4.10.11. DUAL-STRESS EXPONENTIAL LIFE-STRESS MODEL WITH WEIBULL LIFE DISTRIBUTION

In the case where a Weibull pdf model describes the life distribution at each accelerated stress level, the Weibull scale parameter α may be used as the nominal value for life in the dual-stress exponential life-stress log-linear model, as described by

$$f(t, T, H) = \frac{\beta}{c} \cdot e^{-\left(\frac{a}{T} + \frac{b}{H}\right)} \left[\frac{t}{c} \cdot e^{-\left(\frac{a}{T} + \frac{b}{H}\right)} \right]^{\beta-1} e^{-\left[\frac{t}{c} e^{-\left(\frac{a}{T} + \frac{b}{H}\right)}\right]^\beta} \quad (4.86)$$

The expressions for the mean life and reliability corresponding to Equation (4.86) are $\mu = ce^{\left(\frac{a}{T} + \frac{b}{H}\right)} \cdot \Gamma\left(\frac{1}{\beta} + 1\right)$ and $R(t; T, H) = e^{-\left[\frac{t}{c} e^{-\left(\frac{a}{T} + \frac{b}{H}\right)}\right]^\beta}$. Also the hazard rate is $\lambda(t; T, H) = \frac{f(t, T, H)}{R(t; T, H)} = \frac{\beta}{c} \cdot e^{-\left(\frac{a}{T} + \frac{b}{H}\right)} \left[\frac{t}{c} \cdot e^{-\left(\frac{a}{T} + \frac{b}{H}\right)} \right]^{\beta-1}$.

The log-likelihood function for complete and right censored data is derived from the general log-likelihood function as

$$\begin{aligned} \ln(L) = \Lambda = & \sum_{i=1}^{N_c} n_i \cdot \ln[f(t_i, T_i, H_i; a, b, c)] \\ & + \sum_{j=1}^{N_r} n_j \cdot \ln[F(t_j, T_j, H_j; a, b, c)] \end{aligned} \quad (4.87)$$

Therefore, according to

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left\{ \frac{\beta}{c} \cdot e^{-\left(\frac{a}{T_i} + \frac{b}{H_i}\right)} \left[\frac{t_i}{c} \cdot e^{-\left(\frac{a}{T_i} + \frac{b}{H_i}\right)} \right]^{\beta-1} e^{-\left[\frac{t_i}{c} \cdot e^{-\left(\frac{a}{T_i} + \frac{b}{H_i}\right)}\right]^\beta} \right\} - \sum_{j=1}^{N_r} n_j \cdot \left[\frac{t_j}{c} \cdot e^{-\left(\frac{a}{T_j} + \frac{b}{H_j}\right)} \right]^\beta \quad (4.88)$$

The MLE solution for the estimates $\hat{\beta}, \hat{a}, \hat{b}, \hat{c}$ will be obtained by solving $\frac{\partial \Lambda}{\partial \beta} = 0, \frac{\partial \Lambda}{\partial a} = 0, \frac{\partial \Lambda}{\partial b} = 0$ and $\frac{\partial \Lambda}{\partial c} = 0$. The local Fisher information matrix, evaluated using the MLE estimates of model parameters (in this case, $\hat{\beta}, \hat{a}, \hat{b}, \hat{c}$) is given by:

$$F = \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t, \hat{c}) & \text{cov}(\hat{\sigma}_t, \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{b}) \\ \text{cov}(\hat{c}, \hat{\sigma}_t) & \text{var}(\hat{c}) & \text{cov}(\hat{c}, \hat{a}) & \text{cov}(\hat{c}, \hat{b}) \\ \text{cov}(a, \hat{\sigma}_t) & \text{cov}(\hat{a}, \hat{c}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{b}) \\ \text{cov}(\hat{b}, \hat{\sigma}_t) & \text{cov}(\hat{b}, \hat{c}) & \text{cov}(\hat{b}, \hat{a}) & \text{var}(\hat{b}) \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial c} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial b} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial c^2} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c \partial b} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a \partial c} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial b} \\ -\frac{\partial^2 \Lambda}{\partial b \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial b \partial c} & -\frac{\partial^2 \Lambda}{\partial b \partial a} & -\frac{\partial^2 \Lambda}{\partial b^2} \end{bmatrix}^{-1} \quad (4.89)$$

For positive model parameters β and c , the two-sided confidence interval at a significance level of α is given by $\left(\hat{\beta} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} < \beta < \hat{\beta} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right)$ and $\left(\hat{c} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{c})}}{\hat{c}}} < c < \hat{c} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{c})}}{\hat{c}}} \right)$.

Since the parameter b and a can be both positive and negative, the two-sided confidence interval at a significance level of α is given by $\left(\hat{b} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})} < b < \hat{b} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{b})} \right)$ and $\left(\hat{a} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} \right)$.

4.10.12. DUAL-STRESS EXPONENTIAL LIFE-STRESS MODEL WITH LOGNORMAL LIFE DISTRIBUTION

In the case where a lognormal distribution describes the life distribution at each accelerated stress level, the median of the lognormal distribution, μ , serves as the nominal value for the dual-stress exponential life-stress model according to

$$\mu = e^{\mu_t} = ce^{\frac{a+b}{T+H}} \quad (4.90)$$

The log-linear version of Equation (4.90) is

$$\mu_t = \ln(c) + \frac{a}{T} + \frac{b}{H} \quad (4.91)$$

Therefore, the lognormal the conditional dual-stress exponential life-stress distribution model is as follows:

$$f(t, T, H) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t)-\ln(c)-\frac{a}{T}-\frac{b}{H}}{\sigma_t}\right)^2} \quad (4.92)$$

The mean life and reliability functions according to Equation (4.92) are $\mu = e^{\mu_t + \frac{1}{2}\sigma_t^2} = e^{\ln(c) + \frac{a}{T} + \frac{b}{H} + \frac{1}{2}\sigma_t^2}$, and $R(t, T, H) = \int_t^\infty f(t, T, H) dt = \int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t)-\ln(c)-\frac{a}{T}-\frac{b}{H}}{\sigma_t}\right)^2} dt$,

respectively. The lognormal reliability function has no closed form solution, and therefore must be solved using numerical methods or through the use of standard normal tables. The corresponding

$$\text{hazard rate function for Equation (4.92) is } \lambda(t, T, H) = \frac{f(t, T, H)}{R(t, T, H)} = \frac{\frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t)-\ln(c)-\frac{a}{T}-\frac{b}{H}}{\sigma_t}\right)^2}}{\int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t)-\ln(c)-\frac{a}{T}-\frac{b}{H}}{\sigma_t}\right)^2} dt}.$$

The log-likelihood function to estimate parameters of the pdf in Equation (4.92) for complete and right-censored data is

$$\begin{aligned} \Lambda &= \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{1}{\sigma_t t_i} \phi \left(\frac{\ln(t_i) - \ln(c) - \frac{a}{T_i} - \frac{b}{H_i}}{\sigma_t} \right) \right] \\ &\quad + \sum_{j=1}^{N_r} n_j \cdot \ln \left[1 - \Phi \left(\frac{\ln(t_j) - \ln(c) - \frac{a}{T_j} - \frac{b}{H_j}}{\sigma_t} \right) \right] \end{aligned} \quad (4.93)$$

The MLE solution to estimate parameters σ_t, a, b and c are obtained from $\frac{\partial \Lambda}{\partial \sigma_t} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$, $\frac{\partial \Lambda}{\partial b}$ and $\frac{\partial \Lambda}{\partial c} = 0$. The corresponding confidence intervals estimated from the local Fisher information matrix, evaluated using the MLE estimates of the model parameters (in this case, $\hat{\sigma}_t, \hat{a}, \hat{b}$, and \hat{c}), are given by:

$$\begin{aligned}
F &= \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t, \hat{c}) & \text{cov}(\hat{\sigma}_t, \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{b}) \\ \text{cov}(\hat{c}, \hat{\sigma}_t) & \text{var}(\hat{c}) & \text{cov}(\hat{c}, \hat{a}) & \text{cov}(\hat{c}, \hat{b}) \\ \text{cov}(a, \hat{\sigma}_t) & \text{cov}(\hat{a}, \hat{c}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{b}) \\ \text{cov}(\hat{b}, \hat{\sigma}_t) & \text{cov}(\hat{b}, \hat{c}) & \text{cov}(\hat{b}, \hat{a}) & \text{var}(\hat{b}) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial c} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial b} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial c^2} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c \partial b} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a \partial c} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial b} \\ -\frac{\partial^2 \Lambda}{\partial b \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial b \partial c} & -\frac{\partial^2 \Lambda}{\partial b \partial a} & -\frac{\partial^2 \Lambda}{\partial b^2} \end{bmatrix}^{-1} \quad (4.94)
\end{aligned}$$

For the positive thermal-humidity lognormal model parameters σ_t and c , the two-sided confidence interval at a significance level of α is given by $\left(\hat{\sigma}_t \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} < \sigma_t < \hat{\sigma}_t \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} \right)$ and

$\left(\hat{c} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{c})}}{\hat{c}}} < c < \hat{c} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{c})}}{\hat{c}}} \right)$. Since the parameters a and b can be both positive and

negative, the two-sided confidence interval for parameters a and b , at a significance level of α , is given by $\left(\hat{a} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{a})} \right)$ and $\left(\hat{b} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{b})} < b < \hat{b} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{b})} \right)$.

4.10.13. POWER-EXPONENTIAL LIFE-STRESS MODEL

The power-exponential life-stress model is used when a combination of temperature (or humidity) acceleration variable along with a second non-thermal stress (e.g. mechanical or voltage) accelerated variable in an ALT is used. The power-exponential model is given by:

$$L(T, S) = cS^{-n}e^{\frac{a}{T}} \quad (4.95)$$

where S is the non-thermal stress, T is the thermal stress T (expressed in absolute units, °K or relative humidity), and a , C and n are the model parameters to be determined. The log-linear version of this model is

$$\ln(L(T, S)) = \ln(c) - n \ln(S) + \frac{a}{T} \quad (4.96)$$

Note that since this model involves two different but independent types of stresses (accelerants), a linear life-stress plot can only be generated by keeping one of the stresses constant while varying the other. Depending on which stress is kept constant, either the parameter a or n can become the slope of the linear plot. By keeping non-thermal stress constant, the relationship becomes $\ln(L(T, S)) =$

$\underbrace{\{\ln(c) - n \ln(S)\}}_{\text{intercept}} + \underbrace{\frac{a}{S}}_{\text{slope}} \cdot \underbrace{\left(\frac{1}{T}\right)}_{\text{indep. variable}}$, and by keeping the thermal stress constant, the relationship

becomes $\ln(L(T, S)) = \underbrace{\{\ln(c) + \frac{a}{T}\}}_{\text{intercept}} - \underbrace{\frac{n}{S}}_{\text{slope}} \cdot \underbrace{\ln(S)}_{\text{variable}}$.

According to Equation (4.95), the acceleration factor for the power-exponential life-stress model is given by

$$AF = \left(\frac{S_{acc}}{S_{use}} \right)^n e^{a(\frac{1}{T_{use}} - \frac{1}{T_{acc}})} \quad (4.97)$$

4.10.14. POWER-EXPONENTIAL LIFE-STRESS MODEL WEIBULL LIFE DISTRIBUTION

In the case where a Weibull distribution describes the life distribution for the power-exponential model, the Weibull scale parameter α will serve as the nominal value. That is,

$$\alpha = L(T, S) = c S^{-n} e^{\frac{a}{T}} \quad (4.98)$$

Therefore, the joint life-stress Weibull distribution model is given by

$$f(t, T, S) = \frac{\beta S^n e^{-\frac{a}{T}}}{c} \cdot \left(\frac{t S^n e^{-\frac{a}{T}}}{c} \right)^{\beta-1} e^{-\left(\frac{t S^n e^{-\frac{a}{T}}}{c} \right)^\beta} \quad (4.99)$$

The corresponding expressions for the mean life and reliability functions of the pdf in Equation (4.99)

are $\mu = \frac{c}{S^n e^{-\frac{a}{T}}} \cdot \Gamma\left(\frac{1}{\beta} + 1\right)$ and $R(t, T, S) = e^{-\left(\frac{t S^n e^{-\frac{a}{T}}}{c}\right)^\beta}$, respectively. The corresponding hazard rate for (4.99) is given by $\lambda(t, T, S) = \frac{f(t, T, S)}{R(t, T, S)} = \frac{\beta S^n e^{-\frac{a}{T}}}{c} \left(\frac{t S^n e^{-\frac{a}{T}}}{c} \right)^{\beta-1}$. The power-exponential with Weibull life distribution has the log-likelihood function

$$\Lambda = \sum_{i=1}^{N_c} n_i \cdot \ln \left[\frac{\beta S_i^n e^{-\frac{a}{T_i}}}{a} \cdot \left(\frac{S_i^n e^{-\frac{a}{T_i}}}{a} t_i \right)^{\beta-1} e^{-\left(\frac{S_i^n e^{-\frac{a}{T_i}}}{c} t_i \right)^\beta} \right] - \sum_{j=1}^{N_r} N_j \cdot \left(\frac{S_j^n e^{-\frac{a}{T_j}}}{c} t_j \right) \quad (4.100)$$

The MLE solution will be performed to obtain estimates $\hat{\beta}, \hat{a}, \hat{c}$ and \hat{n} such that $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$, $\frac{\partial \Lambda}{\partial c} = 0$, and $\frac{\partial \Lambda}{\partial n} = 0$. The confidence intervals for the model parameters will be obtained from the local Fisher information matrix, evaluated using the MLE estimates of model parameters (in this case, $\hat{\beta}, \hat{a}, \hat{c}, \hat{n}$) from

$$\begin{aligned}
F &= \begin{bmatrix} \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{a}) & \text{cov}(\hat{\beta}, \hat{c}) & \text{cov}(\hat{\beta}, \hat{n}) \\ \text{cov}(\hat{a}, \hat{\beta}) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{c}) & \text{cov}(\hat{a}, \hat{n}) \\ \text{cov}(\hat{c}, \hat{\beta}) & \text{cov}(\hat{c}, \hat{a}) & \text{var}(\hat{c}) & \text{cov}(\hat{c}, \hat{n}) \\ \text{cov}(\hat{n}, \hat{\beta}) & \text{cov}(\hat{n}, \hat{a}) & \text{cov}(\hat{n}, \hat{c}) & \text{var}(\hat{n}) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \beta^2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial a} & -\frac{\partial^2 \Lambda}{\partial \beta \partial c} & -\frac{\partial^2 \Lambda}{\partial \beta \partial n} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \beta} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial c} & -\frac{\partial^2 \Lambda}{\partial a \partial n} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \beta} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c^2} & -\frac{\partial^2 \Lambda}{\partial c \partial n} \\ -\frac{\partial^2 \Lambda}{\partial n \partial \beta} & -\frac{\partial^2 \Lambda}{\partial n \partial a} & -\frac{\partial^2 \Lambda}{\partial n \partial c} & -\frac{\partial^2 \Lambda}{\partial n^2} \end{bmatrix}^{-1} \quad (4.101)
\end{aligned}$$

For the positive T-NT-Weibull parameters β and c , the two-sided confidence interval at a significance

level of α are $\left(\hat{\beta} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} < \beta < \hat{\beta} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}} \right)$ and $\left(\hat{c} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{c})}}{\hat{c}}} < c < \hat{c} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{c})}}{\hat{c}}} \right)$.

Since the parameter a and n can be both positive and negative, the two-sided confidence interval for parameters a and n , at a significance level of α , are $\left(\hat{a} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{a})} \right)$ and $\left(\hat{n} - Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} < n < \hat{n} + Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{n})} \right)$.

4.10.15. POWER-EXPONENTIAL LIFE-STRESS MODEL LOGNORMAL LIFE DISTRIBUTION

In the case where a lognormal distribution describes the life distribution of the power-exponential life-stress model by assigning the median of the lognormal distribution, μ as the nominal value the following expression is found:

$$\mu = e^{\mu_t} = cS^{-n}e^{\frac{a}{T}} \quad (4.102)$$

The equivalent log-linear of Equation (4.102) is

$$\mu_t = \ln(c) - n \cdot \ln(S) + \frac{a}{T} \quad (4.103)$$

We can then substitute Equation (4.103) as the scale parameter μ_t into the lognormal pdf to obtain the power-exponential lognormal life distribution as follows:

$$f(t, T, S) = \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t) - \ln(c) + n \cdot \ln(S) - \frac{a}{T}}{\sigma_t} \right)^2} \quad (4.104)$$

The expressions for associated mean life and reliability functions are $\mu = e^{\mu_t + \frac{1}{2}\sigma_t^2} = e^{\ln(c) - n.\ln(S) + \frac{a}{T} + \frac{1}{2}\sigma_t^2}$, and $R(t, T, S) = \int_t^\infty f(t, T, S) dt = \int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(c) + n.\ln(S) - \frac{a}{T}}{\sigma_t}\right)^2} dt$,

respectively. The solution for the lognormal reliability function has no closed form solution, and therefore must be solved using numerical methods. The corresponding hazard rate is $\lambda(t, T, S) =$

$$\frac{f(t, T, S)}{R(t, T, S)} = \frac{\frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(c) + n.\ln(S) - \frac{a}{T}}{\sigma_t}\right)^2}}{\int_t^\infty \frac{1}{\sigma_t t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t) - \ln(c) + n.\ln(S) - \frac{a}{T}}{\sigma_t}\right)^2} dt}.$$

The power-exponential lognormal life distribution has the log-likelihood function in Equation (4.105), assuming complete and right-censored data.

$$\begin{aligned} \Lambda = & \sum_{i=1}^{N_c} n_i \cdot \ln \left\{ \frac{1}{\sigma_t t_i} \phi \left[\frac{\ln(t_i) - \ln(c) + n_i \cdot \ln(S_i) - \frac{a}{T_i}}{\sigma_t} \right] \right\} \\ & + \sum_{j=1}^{N_r} n_j \cdot \ln \left\{ 1 - \Phi \left[\frac{\ln(t_j) - \ln(c) + n_j \cdot \ln(S_j) - \frac{a}{T_j}}{\sigma_t} \right] \right\} \end{aligned} \quad (4.105)$$

The MLE solution (parameter estimates for $\hat{\sigma}_t, \hat{a}, \hat{c}, \hat{n}$) will be obtained by solving for σ_t, a, c and n such that $\frac{\partial \Lambda}{\partial \sigma_t} = 0$, $\frac{\partial \Lambda}{\partial a} = 0$, $\frac{\partial \Lambda}{\partial c} = 0$, and $\frac{\partial \Lambda}{\partial n} = 0$. The corresponding Fisher information matrix for the power-exponential life-stress with lognormal life model, evaluated using the MLE estimates of model parameters (in this case, $\hat{\sigma}_t, \hat{a}, \hat{c}, \hat{n}$), is given by

$$\begin{aligned}
F &= \begin{bmatrix} \text{var}(\hat{\sigma}_t) & \text{cov}(\hat{\sigma}_t, \hat{a}) & \text{cov}(\hat{\sigma}_t, \hat{c}) & \text{cov}(\hat{\sigma}_t, \hat{n}) \\ \text{cov}(\hat{a}, \hat{\sigma}_t) & \text{var}(\hat{a}) & \text{cov}(\hat{a}, \hat{c}) & \text{cov}(\hat{a}, \hat{n}) \\ \text{cov}(\hat{c}, \hat{\sigma}_t) & \text{cov}(\hat{c}, \hat{a}) & \text{var}(\hat{c}) & \text{cov}(\hat{c}, \hat{n}) \\ \text{cov}(\hat{n}, \hat{\sigma}_t) & \text{cov}(\hat{n}, \hat{a}) & \text{cov}(\hat{n}, \hat{c}) & \text{var}(\hat{n}) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \sigma_t^2} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial a} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial c} & -\frac{\partial^2 \Lambda}{\partial \sigma_t \partial n} \\ -\frac{\partial^2 \Lambda}{\partial a \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial a^2} & -\frac{\partial^2 \Lambda}{\partial a \partial c} & -\frac{\partial^2 \Lambda}{\partial a \partial n} \\ -\frac{\partial^2 \Lambda}{\partial c \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial c \partial a} & -\frac{\partial^2 \Lambda}{\partial c^2} & -\frac{\partial^2 \Lambda}{\partial c \partial n} \\ -\frac{\partial^2 \Lambda}{\partial n \partial \sigma_t} & -\frac{\partial^2 \Lambda}{\partial n \partial a} & -\frac{\partial^2 \Lambda}{\partial n \partial c} & -\frac{\partial^2 \Lambda}{\partial n^2} \end{bmatrix}^{-1} \quad (4.106)
\end{aligned}$$

For positive parameters σ_t and c , the two-sided confidence interval at a significance level of α is given

by $\left(\hat{\sigma}_t \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} < \sigma_t < \hat{\sigma}_t \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{\sigma}_t)}}{\hat{\sigma}_t}} \right)$ and $\left(\hat{c} \cdot e^{-\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{c})}}{\hat{c}}} < c < \hat{c} \cdot e^{\frac{Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{c})}}{\hat{c}}} \right)$. Since

the parameters a and n can be both positive and negative, the two-sided confidence interval at a significance level of α is given by $(\hat{a} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{a})} < a < \hat{a} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{a})})$ and $(\hat{n} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{n})} < n < \hat{n} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\text{var}(\hat{n})})$.

Example 4.11

Consider 12 electronic devices that are subjected to a continuous accelerated life test, whose results (in terms of complete failure data) are shown in the table below. Assume that the physics-of-failure identifies temperature and voltage as the acceleration variables and supports application of the power-exponential model with a lognormal life distribution. Determine the parameters of this model and associated 90% two-sided confidence intervals. ALT data for the T-NT example are given below.

Time (hrs)	Temp (°K)	Voltage (V)
620	348	3
632	348	3
658	348	3
822	348	3
380	348	5
416	348	5
460	348	5
596	348	5
216	378	3
246	378	3
332	378	3
400	378	3

Solution

Using the MLE equations provided in the previous section, we obtain the following parameter estimates, $\hat{\sigma}_t = 0.183$, $\hat{c} = 0.035$, $\hat{a} = 3729.65$, $\hat{n} = 0.777$, and the following 90% confidence bounds:

$$0.140 < \sigma_t < 0.237$$

$$0.0055 < c < 0.2270$$

$$3004 < a < 4455$$

$$0.453 < n < 1.101$$

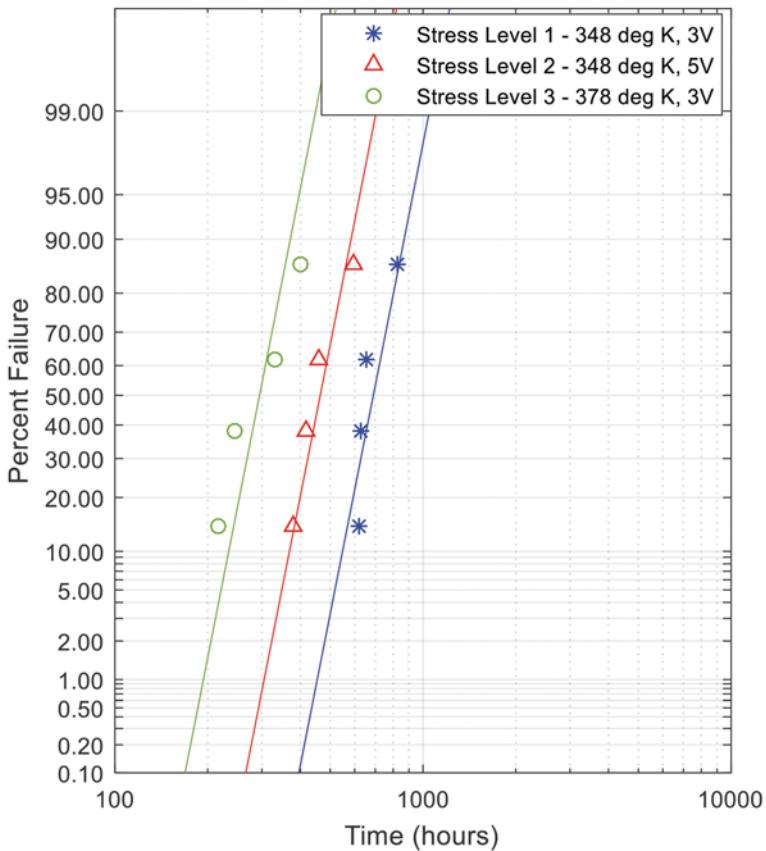


Figure 4.14: Lognormal multi plot for T-NT life-stress model

4.11. PROPORTIONAL HAZARDS (PH) MODEL

The life-stress relationships presented in the previous section are either single or dual stress relationships and some of them are members of the family of Proportional Hazards (PH) Models. This

family is often of great utility in terms of modeling material, component, or system life in an accelerated testing situation. A number of commonly used and well-understood life stress models belong to this class, including the Arrhenius model, the inverse power law (IPL) model, temperature-humidity model, the generalized Eyring model, and so forth.

If we let S be the covariate that represents stress of a drug, and if $h(t)$ is the hazard rate, then Equation (4.107), which is a measure of the effect that the stress has on the expected life can be expressed as:

$$h(t; \gamma | S; \theta) = h_0(t; \gamma)g(S; \theta) \quad (4.107)$$

Where θ is a parameter and $g(S; \theta)$ is a modifier function. This model is readily expandable to more than one stress and parameter as

$$h(t; \gamma | S; \theta) = h_0(t; \gamma)g(S; \theta) \quad (4.108)$$

The modifier function can be a combination of different independent models such as exponential, inverse power law, etc. Equation (4.95) is an example of this kind of combined models. Also dependent stresses can be modeled by PH model which is explained in Section 4.13.

A commonly used linear form of the modifier function includes product of the row vector of stress (S) and the column vector of coefficients (θ), or

$$g(S; \theta) = e^{\sum_{i=1}^n \theta_i S_i} \quad (4.109)$$

Therefore, the full PH model can be written as

$$h(t; \gamma | S; \theta) = h_0(t; \gamma)e^{\sum_{i=1}^n \theta_i S_i} \quad (4.110)$$

One important point to be made is that the values of the covariates (S_i) can be the raw data themselves, or some useful transformation (logarithms, reciprocals, etc.,) of them. Equation (4.83) is an example of dual stress exponential life stress model in reciprocal form.

4.11.1. THE PARAMETRIC PH MODEL WITH AN EXAMPLE

Suppose we know (or assume) that the failure times for a particular component operating under a constant (but arbitrary) covariate vector S are Weibull distributed with shape parameter β and characteristic life α .

We know that the hazard function for the Weibull distribution is

$$h_{weibull} = \frac{\beta}{\alpha^\beta} t^{\beta-1} \quad (4.111)$$

This can be used as the baseline hazard rate,

$$h(t; \alpha_0, \beta | S; \theta) = \frac{\beta}{\alpha_0^\beta} t^{\beta-1} e^{\sum_{j=1}^m \theta_j S_j} \quad (4.112)$$

It worth noting that this model is not same as dual-stress exponential life-stress model with weibull life distribution that is discussed in 4.10.11.

Assuming $S_0 = 1$, and $\theta_0 = -\beta \ln \alpha_0$:

$$h(t|S; \theta) = \beta t^{\beta-1} e^{(-\beta \ln \alpha_0)(1)} e^{\sum_{j=1}^m \theta_j S_j} = \beta t^{\beta-1} e^{\sum_{j=0}^m \theta_j S_j} \quad (4.113)$$

The reliability function of the PH-Weibull model is

$$R(t|S; \theta) = \exp \left[-t^\beta e^{\sum_{j=0}^m \theta_j S_j} \right] \quad (4.114)$$

And the PDF for the PH-Weibull distribution will be

$$f(t|S; \theta) = \beta t^{\beta-1} \exp \left[\sum_{j=0}^m \theta_j S_j - t^\beta e^{\sum_{j=0}^m \theta_j S_j} \right] \quad (4.115)$$

Mettas (2000, p. 2) concludes by developing the log-likelihood function, when failure and suspension times are known exactly, as

$$\Lambda = \sum_{i=1}^{n_f} \ln \left\{ \beta t_i^{\beta-1} \exp \left[\sum_{j=0}^m \theta_j S_{i,j} - t_i^\beta e^{\sum_{j=0}^m \theta_j S_{i,j}} \right] \right\} - \sum_{i=1}^{n_s} t_i^\beta e^{\sum_{j=0}^m \theta_j S_{i,j}} \quad (4.116)$$

Where n_f and n_s are the number of failures and suspensions, and t_i is the time of the i^{th} failure or suspension. As was done in the non-parametric case, taking the partial derivative of Equation (4.116) with respect to each parameter ($\beta, \alpha_0, \dots, \alpha_m$), setting equal to 0, and then simultaneously solving the system of equations, provides the parameter maximum likelihood estimates.

Example 4.12

Parametric PH-Weibull model with two continuous covariates

Consider a component whose reliability may be affected by both the ambient temperature and relative humidity under which it operates. Five samples at each of three (T, H) combinations were run until failure. The stress levels and failure times are given in Table 4.5.

Table 4.5: Stress levels and failure times

Temp (°K)		Relative Humidity (%)		Time to failures (hr)
T	1/T	RH	1/RH	
393	2.54×10^{-3}	0.60	1.68	102, 115, 151, 196, 210
353	2.83×10^{-3}	0.85	1.18	184, 193, 244, 265, 281
353	2.83×10^{-3}	0.60	1.67	298, 311, 337, 347, 405

Analyzing the data using the reciprocals of temperature and humidity as the covariates x_1 and x_2 provides the following MLE parameter estimates.

$$\hat{\beta} = 6.17, \hat{\theta}_0 = 11.05, \hat{\theta}_1 = -14474, \hat{\theta}_2 = -4.38$$

Suppose that the actual anticipated usage conditions for this component were 20°C (293 K, or $1/T = 3.41 \times 10^{-3}$) and 40% relative humidity ($RH = 0.40$, or $1/RH = 2.5$). Then the Weibull parameters for the usage conditions would be $\beta = 6.17$ and

$$\begin{aligned}\eta &= (\exp \left[\sum_{j=0}^m \theta_j S_j \right])^{-1/\beta} \\ &= [e^{(11.05 - 14474 \times 3.41 \times 10^{-3} - 4.38 \times 2.5)}]^{-1/6.17} \\ &= 2952 \text{ hrs}\end{aligned}$$

If the mission time for the component was 1,500 hrs then the estimated mission reliability would be

$$R(1500 \text{ hrs}) = e^{-\left(\frac{1500}{2952}\right)^{6.17}} = 0.98$$

4.12. BAYESIAN ESTIMATION APPROACH TO ALT MODEL PARAMETER ESTIMATION

Bayesian estimation methods for ALT data analysis are closely related to MLE methods. Bayesian methods, however, allow ALT data to be combined with “prior” information in the form of probability distributions about the ALT parameters, to produce a posterior distribution of the parameters. This posterior distribution is then used to quantify uncertainty characteristics (mean, probability bounds) about the parameters and functions of the parameters. Hence, the idea behind this approach is the use of Bayes’ theorem to generate a posterior distribution for a parameter of interest, based on the likelihood of the observed data, given those parameters and prior distribution of those parameters. Figure 4.15 is a schematic of the approach (Modarres et. al, 2017):

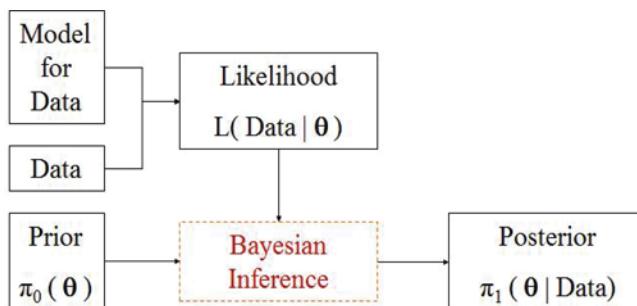


Figure 4.15: Schematic of Bayesian approach to estimation of ALT model parameters

From Bayes’ theorem, the mathematical expression for the posterior distribution for a vector of parameters of interest θ is given by:

$$\frac{\pi_1(\theta | Data)}{posterior} = \frac{\overbrace{l(Data|\theta)}^{likelihood} \overbrace{\pi_0(\theta)}^{prior}}{\underbrace{\int l(Data|\theta) \pi_0(\theta) d\theta}_{normalizing factor}} \quad (4.117)$$

When the posterior distribution $\pi_1(\theta | \text{Data})$ is found, one can develop an estimate of some reliability metric, M , which depends on the parameter θ . This metric might be reliability, the time to failure pdf, or the hazard rate, and is found from the *posterior predictive distribution* given by

$$M(x|\text{Data}) = \int_{\forall\theta} M(x|\theta) \pi_1(\theta|\text{Data}) d\theta \quad (4.118)$$

Note that each value of θ from the posterior distribution $\pi_1(\theta|\text{Data})$ corresponds to a specific member of the model representing the metric of interest $M(x|\theta)$, and the probability of that specific member being the correct value is governed by the probability that θ is the correct value of the model parameter (which is obtained from $\pi_1(\theta|\text{Data})$). In the inference of $M(x|\text{Data})$, we therefore need to consider the entire distribution $\pi_1(\theta|\text{Data})$.

Estimation methods such as the MLE yield point and interval estimates of the parameter and not probability distributions in the way Bayesian estimation methods achieve. Point estimates can be easily incorporated into the Bayesian estimation method by substituting them into the function representing the metric of interest $M(x|\theta)$; however, Bayesian estimation yields a range of possible parameter values with assigned pdf's. For instance, in the case of the Weibull distribution, where the applicable vector of parameters $\theta = \{\alpha, \beta\}$, the MLE's $\hat{\alpha}$ and $\hat{\beta}$ can be used in the Weibull function to generate the estimated Weibull distribution.

Example 4.13

Estimate the mean value and 90% two-sided probability interval of the parameter θ of a binomial distribution (for example, representing items tested to reveal defective ones). Assume a uniform prior distribution between 0 and 0.1 describes parameter θ , and the binomial likelihood function is $l(\text{Data}|\theta) = C_5^{227}\theta^5(1-\theta)^{222}$, where C is the combination term. The likelihood function represents observed data, for example involving 227 inspections involving 5 defective units.

Solution

Prior distribution can be written as $\pi_0(\theta) \sim \text{Unif}(0,0.1) = \frac{1}{0.1-0} = 10$

Therefore, the normalizing factor of Equation (4.117) denominator is given by $\int_0^{0.1} C_5^{227}\theta^5(1-\theta)^{222} \cdot 10 d\theta = 10C_5^{227} \cdot \{9.128 \times 10^{-13}\}$, and from Equation (4.118) the posterior of θ is

$$f(\theta|\text{Data}) = \frac{10C_5^{227} \cdot \theta^5(1-\theta)^{222}}{10C_5^{227} \cdot \{9.128 \times 10^{-13}\}} = \frac{\theta^5(1-\theta)^{222}}{9.128 \times 10^{-13}}$$

The mean value of θ is obtained from $\hat{\theta} = \int_0^{0.1} \theta \cdot f(\theta|\text{Data}) d\theta = \int_0^{0.1} \frac{\theta^6(1-\theta)^{222}}{9.128 \times 10^{-13}} d\theta = 0.0262$.

Similarly the 90% probability bounds of θ may be obtained from $\int_0^{\theta_{0.05}} f(\theta|\text{Data}) d\theta = 0.05$, and $\int_0^{\theta_{0.95}} f(\theta|\text{Data}) d\theta = 0.95$. Therefore, the 90% two-sided probability bound would be $0.0115 < \theta < 0.0456$.

4.12.1. PRIOR INFORMATION FOR BAYESIAN ESTIMATION

The prior distribution used in Bayesian estimation effectively captures all of the information known about the parameter or parameters of interest before we collect the data. In general, there are two possible sources of prior information. They can either come from subjective opinion (e.g. expert

opinion), or from relevant historical data. Prior knowledge of model parameters may be available in one of the following formats:

- *Degenerate priors* – refers to parameters that are absolutely known.
- *Non-informative priors* – parameters that are completely or mostly unknown. The distribution of these types of priors is typically expressed as approximately uniform over a range of values. They are applied when there is relatively little or very limited information available to describe the parameter of interest.
- *Informative, non-degenerate priors* – lie in between above categories. Informative priors apply when we have some information about the parameters of the model before we collect the data. For instance, they may come from historical data relating to similar components to the one under test, previous test results, or computational analysis.

For ALT data analysis, the most important motivation for using prior information is to combine it with the collected ALT data to arrive at a more accurate estimation of the life-stress and time-to-failure models. Some examples of proper prior distributions for scalar parameters include normal prior distribution with mean a and standard deviation b , such that $f(\beta) = \frac{1}{b} \phi_{nor}\left(\frac{\beta-a}{b}\right)$; $-\infty < \beta < \infty$ and uniform prior distribution between a and b such that $f(\beta) = \frac{1}{b-a}$; $a < \beta < b$. This prior distribution does not express preference over specific values of β in the given range and therefore is largely an informative prior. Other prior distribution is beta distribution for parameters that can only vary between zero and one and isosceles triangle prior distribution with base (range) between a and b . For a positive parameter β , it is often more convenient to specify the prior pdf in terms of $\ln(\beta)$.

4.12.2. A BAYESIAN ESTIMATION ALT DATA ANALYSIS EXAMPLE

Consider a component that has been tested at elevated stresses for fatigue as its main failure mechanism. A standard fatigue test was performed to explore fatigue properties of the component material. There are also a few data available from an accelerated life test of the component itself working under high stress conditions. These data are presented in Table 4.6.

Table 4.6: Data from fatigue test of material (prior data) and component ALT data

Fatigue Test in Material Lab		Accelerated life Test of Component	
Stress Amplitude (Ksi)	Cycles to Failure	Stress Amplitude (Ksi)	Cycles to Failure
78.90	45,000	59.6	2,900,000
74.02	240,000	58.7	1,400,000
68.16	800,000	56.2	9,000,000
63.27	1,500,000	57.2	10,000,000+
62.05	2,700,000	55.3	10,000,000+
59.61	7,800,000	45.1	10,000,000+
59.61	10,000,000		
58.63	26,000,000+		
57.65	12,000,000+		
57.41	22,000,000+		

(+) Censored data

Assuming an inverse power life-stress model is applicable along with the lognormal for life distribution, the Bayesian inference approach to express model parameters requires a definition of a proper likelihood using the actual ALT data and an appropriate prior distribution from available knowledge and information about the parameters of the ALT models. Assuming the power life-stress model with

lognormal life model described by Equation (4.78), the likelihood function assuming non-group data is given by:

$$\begin{aligned} l(data|K, n, \sigma_t) &= \prod_{i=1}^{N_c} \left\{ \frac{1}{\sigma_t t_i} \phi \left(\frac{\ln(t_i) + \ln(K) + n \ln(S_i)}{\sigma_t} \right) \right\} \\ &\times \prod_{j=1}^{N_r} \left\{ 1 - \Phi \left(\frac{\ln(t_j) + \ln(K) + n \ln(S_j)}{\sigma_t} \right) \right\} \end{aligned} \quad (4.119)$$

Next, we establish the prior distribution. From the given material laboratory data, the following MLE estimates are obtained for this life-stress model

Parameter	5.00%	Mean	95.00%
K	3.39×10^{-48}	1.61×10^{-42}	7.59×10^{-37}
n	16.45	19.58	22.71
σ_t	0.28	0.49	0.84

From the above, if we are to assume a uniform distribution for $\ln(a)$ and lognormal distribution for n and σ so their percentiles match their estimated upper and lower MLE limits, and further assuming a 90% confidence that these are upper and lower bounds of the model parameters, then suitable priors for the parameters are given by:

$$\text{For } K: \ln(K) \sim \text{Unif}[\ln(3.39 \times 10^{-48}), \ln(7.59 \times 10^{-37})] = \text{Unif}(-109.3, -83.16)$$

$$a_n = \log[\sqrt{n_{(5\%)} \times n_{(95\%)}}] = \frac{\log(22.71) + \log(16.45)}{2} = 2.96$$

$$\text{For } n: \quad b_n = \frac{\log[\sqrt{n_{(5\%)} / n_{(95\%)}}]}{Z_{(95\%)}} = \frac{(\log(22.71) - \log(16.45))}{1.645} = 0.098 \quad \left. \right\} \rightarrow n \sim \text{lnorm}(2.96, 0.098)$$

$$a_{\sigma_t} = \log[\sqrt{\sigma_{(5\%)} \times \sigma_{(95\%)}}] = \frac{\log(0.84) + \log(0.28)}{2} = -0.72$$

$$\text{For } \sigma: \quad b_{\sigma_t} = \frac{\log[\sqrt{\sigma_{(5\%)} / \sigma_{(95\%)}}]}{Z_{(95\%)}} = \frac{(\log(0.84) - \log(0.24))}{1.645} = 0.38 \quad \left. \right\} \rightarrow \sigma_t \sim \text{lnorm}(-0.72, 0.38)$$

Given the above prior distributions for model parameters, and assuming independence of the parameters, we can build the prior joint distribution of parameters as follows:

$$\pi_0(K, n, \sigma_t) = \frac{g(K) \cdot h(n) \cdot i(\sigma_t)}{\int \int \int g(K) \cdot h(n) \cdot i(\sigma_t) dK dn d\sigma_t} \quad (4.120)$$

Finally, applying Bayes' theorem, the joint posterior distribution of parameters would be:

$$\begin{aligned} \pi_1(K, n, \sigma_t | Data) &= \frac{l(data|(K, n, \sigma_t) \cdot \pi_0(K, n, \sigma_t)}{\int \int \int l(data|(K, n, \sigma_t) \cdot \pi_0(K, n, \sigma_t) dK dn d\sigma_t} \end{aligned} \quad (4.121)$$

The joint posterior distributions such as the one above are usually computationally intensive, especially in the case where conjugate distributions do not apply. These posteriors are usually calculated using

advanced numerical simulations. For example, by using the Markov Chain Monte Carlo simulation developed under WinBUGS, we obtain the results shown in Figure 4.16.

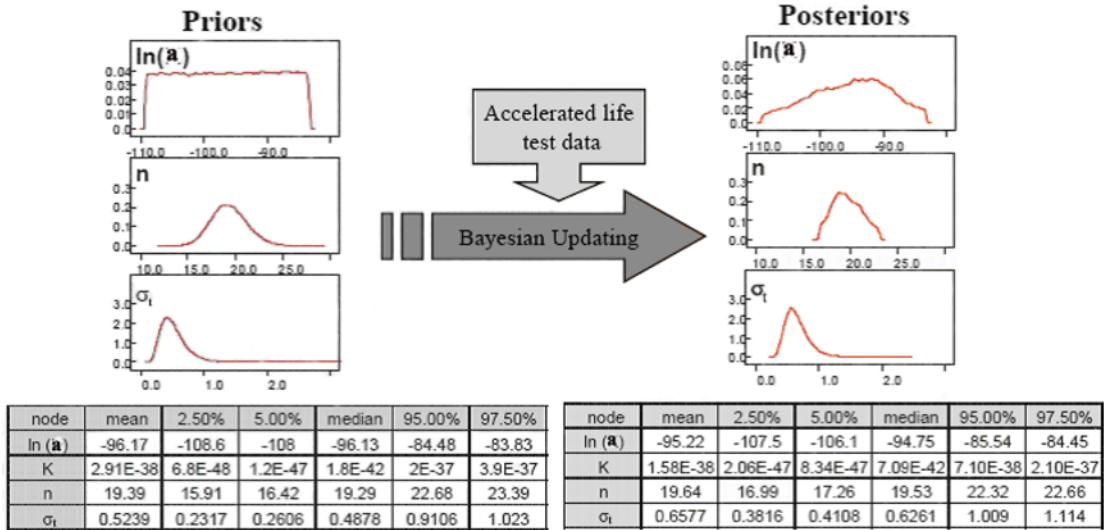


Figure 4.16: Marginal posterior distributions of the parameters

From the discussions regarding the posterior predictive distribution, we can obtain the posterior predictive distribution for the joint pdf representing the accelerated life function ($f(t, S)$ for instance) independent of model parameters as follows:

$$f(t, S | Data) = \int_K \int_n \int_{\sigma_t} f(t, S | K, n, \sigma_t) f(K, n, \sigma_t | Data) dK dn d\sigma_t \quad (4.122)$$

As noted in this example, the Bayesian estimation approach is an excellent method for incorporating an existing degree of belief about a certain parameter of interest that is being analyzed from the observed evidence. Note that by providing the posterior probability distributions of the unknowns of interest, we can obtain confidence bounds and also visualize and analyze the whole distribution – which is an advantage over the MLE approach. With the Bayesian approach, any available prior knowledge of parameters can be integrated in the solution so long as they are represented by distributions. Furthermore, with the MLE approach, one can only include prior knowledge if it has the same data format (i.e. complete failure or censored data), while all censored data can be considered to improve the state of knowledge when using Bayesian estimation.

4.13. DETERMINING STRESS DEPENDENCIES

An area where the PH model is particularly useful is in the investigation of possible interactions between two or more stresses. An interaction between two stresses occurs if the accelerating effect of one stress is dependent on the particular level of a second stress. In Example 4.12, for instance, it may be the case that the accelerating effect of relative humidity depends on the ambient temperature at which the component is operating. Including the interaction in the PH-Weibull model, requires adding an additional covariate term to $g(S, \theta)$ (see Example 4.14). The values of this new covariate are determined by multiplying together the covariate values of the two terms involved in the interaction. Thus, for a two stress model the modifier function would be

$$g(S, \theta) = e^{\theta_0 + \theta_1 S_1 + \theta_2 S_2 + \theta_{12} S_1 S_2} \quad (4.123)$$

Note the cross product term ($S_1 S_2$) in the exponent, and the new coefficient θ_{12} that will need to be estimated. The process for estimating the parameters is the same as outlined before by using MLE or Bayesian estimation approach. However, in order to estimate the interaction coefficient, time-to-failure data would be needed from the four possible combinations of the two stressors at two levels.

Example 4.14

Consider the PH-Weibull model with two continuous covariates and interaction. Table 4.7 provides the same data as Table 4.5 augmented with an additional set of observations at the fourth “corner” of the two-stress, two-level square. Also, the product of $(1/T) \times (1 / RH)$ has been added to the table. These are the values that will be used as a third stress in the model: the interaction stress $S_1 S_2$.

Table 4.7: Time to failure

“Corner”	Temp (°K)		Relative Humidity (%)		$1/(T \times RH)$	Time to failures (hr)
	T	$1/T$	RH	$1/RH$		
T_2, H_1	393	2.54×10^{-3}	0.60	1.67	4.24×10^{-3}	102,115,151,196,210
T_1, H_2	353	2.83×10^{-3}	0.85	1.18	3.33×10^{-3}	184,193,244,265,281
T_1, H_1	353	2.83×10^{-3}	0.60	1.67	4.72×10^{-3}	298,311,337,347,405
T_2, H_2	393	2.54×10^{-3}	0.85	1.18	2.99×10^{-3}	52,68,74,97,120

Analyzing the data provides the following MLE parameter estimates.

$$\hat{\beta} = 5.22, \hat{\theta}_0 = 55.21, \hat{\theta}_1 = -28570, \hat{\theta}_2 = -31.13, \hat{\theta}_{12} = 9669.23$$

Under the same usage conditions from Example 4.12 ($1/T = 3.41 \times 10^{-3}$, $1/RH = 2.5$, and $1 / (T \times RH) = 8.53 \times 10^{-3}$), the Weibull parameters are $\beta = 5.22$ and

$$\begin{aligned} &= [e^{(55.21 - 28570 \times 3.41 \times 10^{-3} - 31.13 \times 2.5 + 9669.23 \times 8.53 \times 10^{-3})}]^{-1/5.22} \\ &= 1333 \text{ hrs} \end{aligned}$$

The two normal use Weibull lines are plotted in Figure 4.17. Although the slopes of the two lines are very nearly parallel, the effect of the interaction term is to shift the Weibull line to the left, which is in the direction of poorer reliability. Therefore, it is important to determine whether or not the interaction term is statistically significant. If it turns out not to be, then one can ignore the term in the model and predict the higher levels of reliability from the right-most line. But if the interaction term is significant, then one is forced to use the left-most line which predicts considerably lower reliability performance. Worse still would be if the interaction term was significant, but still omitted from the model. In this case reliability predictions (from the right-most line) would be highly overstated.

4.13.1. CONFIDENCE BOUNDS

Confidence bounds can be found in the usual manner by beginning first with the Fisher information matrix.

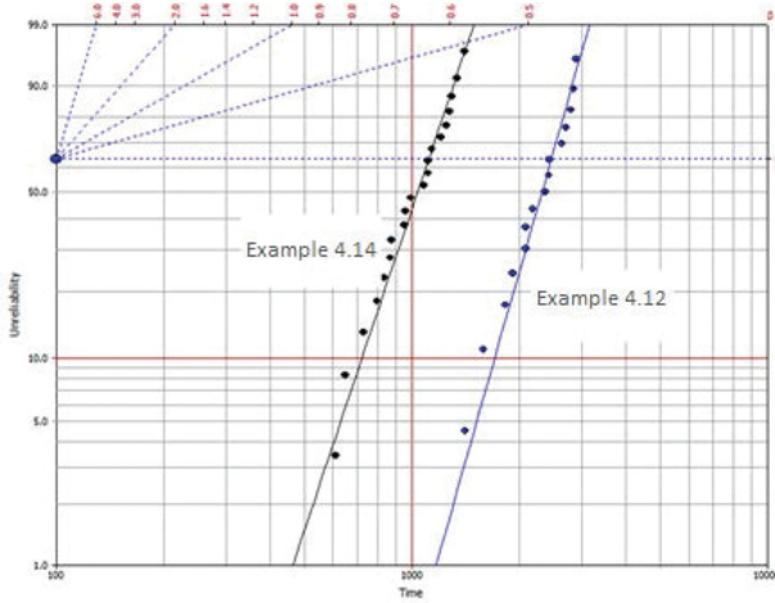


Figure 4.17: Normal Use Weibull Plots from Example 4.12 & Example 4.14, showing the effect of the interaction term.

$$F = \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial \beta^2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial \theta_0} & -\frac{\partial^2 \Lambda}{\partial \beta \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \beta \partial \theta_2} & -\frac{\partial^2 \Lambda}{\partial \beta \partial \theta_{12}} \\ -\frac{\partial^2 \Lambda}{\partial \theta_0 \partial \beta} & -\frac{\partial^2 \Lambda}{\partial \theta_0^2} & -\frac{\partial^2 \Lambda}{\partial \theta_0 \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \theta_0 \partial \theta_2} & -\frac{\partial^2 \Lambda}{\partial \theta_0 \partial \theta_{12}} \\ -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \beta} & -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \theta_0} & -\frac{\partial^2 \Lambda}{\partial \theta_1^2} & -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \theta_2} & -\frac{\partial^2 \Lambda}{\partial \theta_1 \partial \theta_{12}} \\ -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \beta} & -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \theta_0} & -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \theta_2^2} & -\frac{\partial^2 \Lambda}{\partial \theta_2 \partial \theta_{12}} \\ -\frac{\partial^2 \Lambda}{\partial \theta_{12} \partial \beta} & -\frac{\partial^2 \Lambda}{\partial \theta_{12} \partial \theta_0} & -\frac{\partial^2 \Lambda}{\partial \theta_{12} \partial \theta_1} & -\frac{\partial^2 \Lambda}{\partial \theta_{12} \partial \theta_2} & -\frac{\partial^2 \Lambda}{\partial \theta_{12}^2} \end{bmatrix} \quad (4.124)$$

Taking the inverse results in

$$F^{-1} = \begin{bmatrix} Var(\beta) & Cov(\beta, \theta_0) & Cov(\beta, \theta_1) & Cov(\beta, \theta_2) & Cov(\beta, \theta_{12}) \\ Cov(\theta_0, \beta) & Var(\theta_0) & Cov(\theta_0, \theta_1) & Cov(\theta_0, \theta_2) & Cov(\theta_0, \theta_{12}) \\ Cov(\theta_1, \beta) & Cov(\theta_1, \theta_0) & Var(\theta_1) & Cov(\theta_1, \theta_2) & Cov(\theta_1, \theta_{12}) \\ Cov(\theta_2, \beta) & Cov(\theta_2, \theta_0) & Cov(\theta_2, \theta_1) & Var(\theta_2) & Cov(\theta_2, \theta_{12}) \\ Cov(\theta_{12}, \beta) & Cov(\theta_{12}, \theta_0) & Cov(\theta_{12}, \theta_1) & Cov(\theta_{12}, \theta_2) & Var(\theta_{12}) \end{bmatrix}^{-1} \quad (4.125)$$

$\hat{\beta}, \hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_{12}$

The $(1 - \alpha)100\%$ confidence bounds for the various parameters are found as

$$\text{For } \beta: \hat{\beta} e^{\pm \Phi_{\alpha}^{-1} \times \frac{\sigma_{\hat{\beta}}}{\hat{\beta}}} \quad (4.126)$$

$$\text{For } \theta_i: \hat{\theta}_i \pm \Phi_{\frac{\alpha}{2}}^{-1} \times \sigma_{\hat{\theta}_i}$$

Example 4.15

Find 90% 2-sided confidence bounds on Example 4.14 parameters.

Solution

Referring back to the data and the parameter estimates found in Example 4.14, and recalling that the log-likelihood function (Λ) is given in equation (4.116); then after evaluating the various second partial derivatives in the Fisher matrix and taking the inverse, we have

$$F^{-1} = \begin{bmatrix} 0.90 & 8.66 & -4564 & -5.11 & 15877 \\ 8.66 & 687.84 & -268135 & -462.04 & 168393 \\ -4564 & -268135 & 106550375 & 179042 & -65002953 \\ -5.11 & -462.04 & 179042 & 319.40 & -116707 \\ 1587 & 168392 & -65002953 & -116707 & 42854524 \end{bmatrix}$$

For the 90% 2-sided confidence interval, $\alpha = 0.10$, so the standard normal variate is $\Phi_{\alpha/2}^{-1} = 1.64$. Using this, the parameter estimates from Example 4.14, and the variances for those estimates from the diagonal of F^{-1} , the confidence intervals are easily found.

Parameter	Point Estimate	90% 2-sided Confidence Bound
β	5.22	(3.87, 7.04)
θ_0	55.21	(12.07, 98.35)
θ_1	-28570	(-45549, -11591)
θ_2	-31.13	(-60.52, -1.73)
θ_{12}	9669	(-1099, 20437)

What is of particular interest is the confidence interval for the interaction parameter θ_{12} . The fact that the interval brackets zero (0) provides evidence that the interaction term may not be statistically significant, and would best be left out of the model. Since all of the other θ_i confidence intervals avoid capturing zero within the interval, it suggests that these other parameters are statistically significant.

4.14. SUMMARY OF THE ALT STEPS AND COMMON PROBLEMS IN PRACTICE

The previous sections covered the overall process and analytical methods for analyzing ALT data. In this section the steps involved are summarized, then common problems to avoid during the ALT process are discussed. The steps of an ALT process are summarized below.

1. A physics of failure analysis is performed, from which the likely dominant failure mechanisms are determined. According to the literature, the PoF analysis should identify possible failure mechanisms and physical models that quantitatively describe the degradation process and failure caused by such mechanisms. The analysis should use sound engineering judgments if such literature is lacking. The analysis would then identify the most dominant failure mechanisms, the stress agents that accelerate the dominant failure mechanism, and other important variables that influence initiation and progression of the failure mechanism.
2. The dominant failure mechanisms to accelerate during the test are selected. Usually all are selected, but occasionally due to time or resource constraints a subset may be selected

- for testing. Justification and implications of not testing a dominant failure mechanism on the final results should be discussed and documented.
3. The various stress agents and their limits in accelerating each failure mechanism should be carefully studied. Stresses will affect multiple failure mechanisms. Therefore they must be selected to ensure the failure mechanism is accelerated without inducing any extraneous failure mechanisms. From the PoF analysis and for each failure mechanism appropriate life-stress models should be selected or developed. Prior values for the life-stress model parameters are selected from past experiences and, if needed, from engineering judgments.
 4. A careful test plan is developed. For each of the dominant failure mechanisms to be tested, appropriate stress agents to be accelerated and the magnitude of the stress are determined using the life-stress model of Step 3. The type of accelerated test is determined, whether constant load acceleration, step stress acceleration, or another appropriate test. Based on the PoF analysis, the stress magnitude for each test or step of test (in step-stress accelerated) must be selected so that failure mechanisms dominant under usage stress will remain dominant at the elevated stresses selected for testing. The type of data collection method is determined (i.e., complete, censored, or interval data). Finally, the number of units to be tested and the length of the test at each elevated stress level are determined.
 5. The test is performed. During the test, root-causes of failures should be performed to determine which failure mechanism caused the failure.
 6. The test data are interpreted and analyzed, which includes extrapolating the accelerated test results to normal operating conditions using the life-stress model and its parameters estimated from the ALT data. In some instances, and especially in the case of the step-stress test, this step and Step 5 may be performed while the test is being conducted. Based on the results of the analysis, the untested stress levels may be adjusted if needed.
 7. If more than one failure mechanism is considered and they can be independently accelerated, Steps 4-6 are repeated for the remaining failure mechanisms. For tests involving simultaneous acceleration of multiple failure mechanisms, the combined mechanism is considered as one composite mechanism. In the former case, competing risk models can be used to combine the time-to-failure distributions of the failure mechanisms to form a composite life distribution. In the latter case, the end result of the data analysis would be the composite life distribution model. In the case of composite failure mechanism, one should ensure that one failure mechanism will not mask another one. This issue will be further addressed later in this section.

4.15. TIME VARYING STRESS TESTS

For the purpose of reducing test units and test duration, time-varying stress methods have been applied. The most common type of time-varying stress test is the step-stress test, which is mainly used in the following situations:

- Only a small number of systems are available,
- Extremely long or specialized test equipment is required,
- Limited environmental chamber capability and/or test fixtures are involved,
- Very expensive support equipment is a test limit.

In these types of tests, the units under test are subjected to a predetermined low level of stress, for a predetermined period (or number of failures). The stress is then stepped up to an increased level, and the procedure is repeated for a number of additional stress levels. Since stresses are increased at each level, a cumulative damage effect occurs, and thus the test follows a time-varying stress profile (i.e. at higher stress levels, failures occur more quickly). The test is terminated when either a predefined

number of failures have occurred (all units or subset of units), or when the predetermined test duration has been reached.

As mentioned above, time-varying stress methods are based on cumulative exposure models (i.e. the cumulative damage theory), since at each new level of stress the unit under test has already been subjected to periods of some lower-level stresses. Some of the challenges and good practices applicable to step-stress testing and the interpretation of test data from such tests include the following:

- Each step in the step-stress continuum needs to be far enough apart to get meaningful results. Many small, very close steps usually make the analysis more difficult and results will be ambiguous. As many as ten steps have been recommended in practice; however, three to six steps make for more manageable and meaningful test results.
- The accuracy of the test is inversely proportional to the length of each step. In other words, the damage accumulated in each step should not take much of the total life such that there is enough life to exhibit in next step.
- It is desirable that each successively higher step provides a more significant contribution to percent of the population failing than the step prior. This limits the number of steps for data analysis.
- The stresses should be considered cumulative in most situations, since many failures involve a “wear out” phenomenon such as oxidation, diffusion, or fatigue crack growth. This cumulative damage is very common for materials, electronics and/or mechanically dominated systems exposed to applicable stresses.
- The steps should be approximately equal in duration; however, this is neither a requirement nor an impediment. The key is not just the step length but also more importantly monitoring the failure accumulated in each step.
- The failure mechanisms should remain relatively constant over a stress range. That is, the relationship between log stress and log life is linear over the range of stresses involved. This is desirable but not strictly required.
- It is also desirable that when multiple stresses are present, one must ensure that they do not interact, or at the very least, have a very weak interaction.

Recall the cumulative damage theory discussed in Chapter 2. For step-stress testing, a unit will be subjected to a certain stress level, S_1 for a time period t_1 . Suppose t_2 denotes the time to failure of the unit at stress S_1 . According to the linear damage theory, the percentage of life consumed by the unit under test, at time t_1 , after being subjected to stress S_1 is given by $\%life\ at\ t_1 = \frac{t_1}{t_2}$.

The cumulative exposure model holds the assumption that the remaining life of items depends only on the current cumulative damage and current stress, regardless of how the damage is accumulated. Moreover, if kept at the current stress, survivors will fail according to the life distribution for that stress level but starting at the previously accumulated damage.

In Figure 4.18(a), there is a stress pattern with four steps (S_1, S_2, S_3, S_4). For each of the constant stresses S_i , the corresponding CDF of time to failure $F_i(t)$ is indicated by a blue curve in Figure 4.18(b), $i = 1,2,3,4$. The CDF for life under the step-stress pattern is $F_0(t)$ shown in Figure 4.18(c).

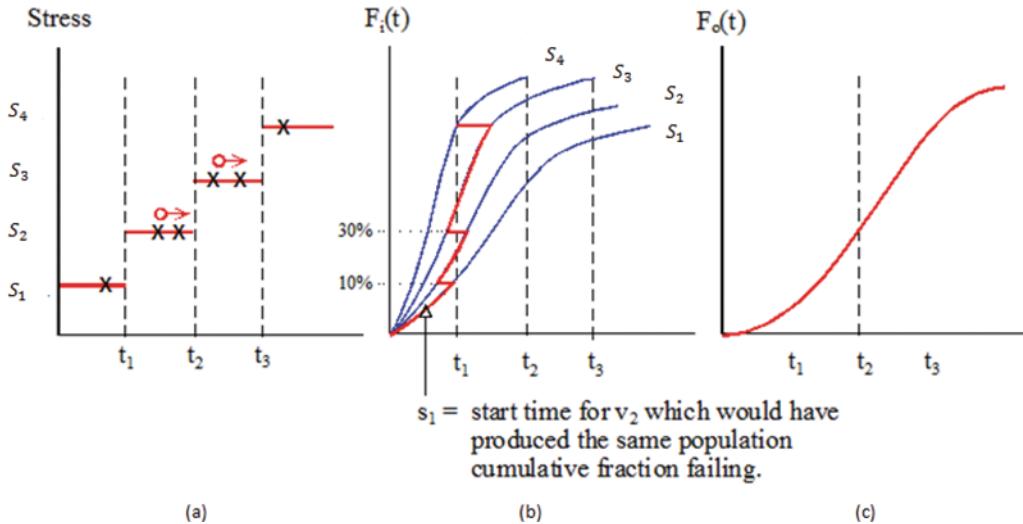


Figure 4.18: Step-stress profile and corresponding life distributions

In Figure 4.18(b), the arrows show that the items first follow the CDF for S_1 up to the time of the first step t_1 . When the stress increases from S_1 to S_2 , the survived item continues along the CDF for S_2 , starting at the accumulated damage. Similarly, when the stress increases from S_2 to S_3 , and from S_3 to S_4 , etc., the survived items continue along the next CDF, starting at the accumulated damage. This model is mathematically expressed as follows to obtain the $F_0(t)$:

The damage of item accumulated in step 1 is $F_0(t) = F_1(t), 0 \leq t \leq t_1$.

Then step 2 has an equivalent start time s_1 , which would have produced the same population cumulative damage; that is, s_1 , is the solution of: $F_2(s_1) = F_1(t_1)$.

The cumulative damage of items in step 2 by total time t is $F_0(t) = F_2[(t - t_1) + s_1], t_1 \leq t \leq t_2$.

Similarly, step 3 has the equivalent start time s_2 that is the solution of $F_3(s_2) = F_2[(t_2 - t_1) + s_1]$.

In general, step i has the equivalent start time s_{i-1} , which is the solution of

$$F_i(s_{i-1}) = F_{i-1}[(t_{i-1} - t_{i-2}) + s_{i-2}]$$

and

$$F_0(t) = F_i[(t - t_{i-1}) + s_{i-1}], t_{i-1} \leq t \leq t_i$$

Thus, $F_0(t)$ for the step-stress pattern is made up of segments of the $F_i(t)$ as shown in Figure 4.18(c).

A typical stress profile for a step stress test is shown below in Figure 4.19. This example has four steps and shows step 1 as a stress level just below the expected maximum operating condition for the unit under test. Beyond this the overstress region exists, where failures are expected to occur more frequently. Note that there exists a region beyond the overstress region that should be avoided in step stress tests. This region is known as the “sudden death” region since the units under test are not expected to be able to tolerate such stress levels, and failures other than those caused by the cumulative damage assumption will take place. This region should be avoided, as it will not offer any useful life information for ALT, but rather is more suitable for qualitative tests such as HALT.

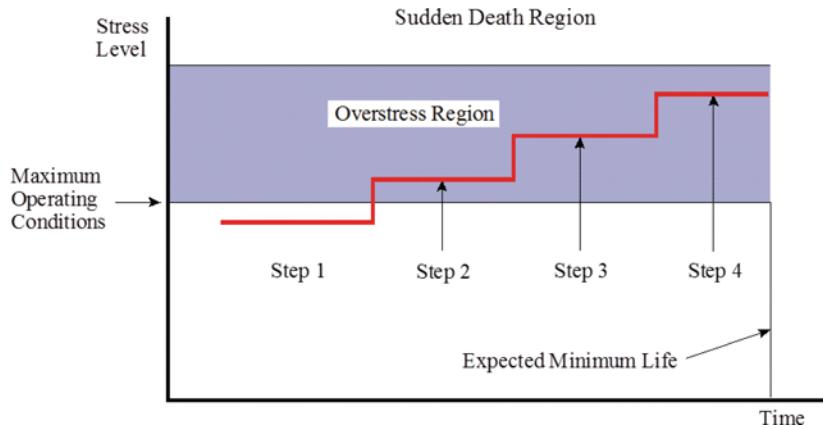


Figure 4.19: Example stress profile for a step-stress test

Some basic assumptions for step-stress testing are as follows:

- Constant stress is used at each step in the test
- For any constant stress S , the life distribution shape parameter must remain constant (to indicate that the same failure mode and mechanism applies at each stress level)
- Life and stress have a known relationship (i.e., the life-stress model) which remains applicable at all steps
- The remaining life of units under test depends only on the current cumulative fraction failed and current stress – regardless of how the damage has accumulated. Moreover, if held at the current stress, survivors will fail according to the cdf for that stress, but starting at the previously accumulated fraction failed (Nelson W. , 2004).

4.16. STEP-STRESS ANALYSIS AND MODEL DEVELOPMENT

As discussed in previous sections, ALT models are based on an underlying life distribution and a relationship between life and stress. The same applies to ALT models involving time varying stresses; however, the equations are more complicated, as they need to account for the cumulative damage effect at each step. As with regular ALTs, the objective of data analysis is to obtain the parameters of the model for failure distribution and life-stress relationship. Two ways that will be discussed for step-stress ALT analysis are the plotting method and the MLE method. Each of these methods will be discussed in the next two subsections.

4.16.1. PLOTTING METHOD FOR STEP-STRESS DATA ANALYSIS

One of the easiest and quickest methods for step-stress data analysis is the plotting method, which uses acceleration factors to obtain model parameter estimates. The steps involved in this method are outlined below.

Step 1: Given a step-stress scenario, create a table that shows failure distribution at each step. This table will be used to develop a life distribution plot based on the data under the assumption of no damage accumulation.

Step 2: Plot life distribution curves for each step (assuming no accumulated damage, which also means the units entering each step are undamaged). Obtain estimate of the parameters of the life-stress and life distribution models from plot, using the usual procedure for ALT plotting (as discussed in Sections 4.6 and 4.8). Assume time to failure at each step is the total time at that step (i.e., disregard the accumulated time of the previous steps).

Step 3: From the results of Step 2 determine the acceleration factor between all steps. That is, use

$$AF_{i \rightarrow j} = \frac{t_i}{t_j} \quad (4.127)$$

where $AF_{i \rightarrow j}$ is the acceleration factor between step j and a previous step i , and t_i and t_j are the actual lengths of test times for steps i and j . Based on these acceleration factors, find the equivalent time in a given step that produces the same damage the unit received while running in all the lower steps and accounts for the difference between the test time at each step. The adjusted times are the equivalent times at previous steps plus the time taken at that step. Realistically, one can now construct another plot of the life distribution by repeating step 2, to obtain updated parameter estimates. However, a quicker and simpler (but less accurate method) is to obtain new estimates of the parameters via the adjusted acceleration factors approach (see step 4).

Step 4: Determine new estimates of parameters affected by adjusted AFs. For calculating the new acceleration factor from step i to j , the following formula is used:

$$AF_{i \rightarrow j}^{Adj} = \frac{\%fail_j}{\%fail_i} \times AF_{i \rightarrow j} \quad (4.128)$$

where $AF_{i \rightarrow j}^{Adj}$ is the adjusted acceleration factor between steps i and j , and $\%fail_i$ and $\%fail_j$ are the percent of the total number of units tested that failed in steps i and j , respectively. Note that care must be taken when using the plotting method, as it can give inaccurate results and thus should only be used in less complicated cases involving few steps.

Nevertheless, to gain an appreciation of this simple method, consider the case where a unit is subjected to a voltage stress, starting at 2V and increasing to 7V. Further, eleven units are tested using the same stress profile. Units that fail are removed from the test and their total time on test is recorded. Table 4.8 summarizes the details of the step-stress test and the results. Further, assume that the life at each step is characterized by the Weibull distribution, and that the inverse power law best represents the relationship between life and stress.

Table 4.8: Details of the step-stress example and failures observed

Step	Stress (volts)	Test Time (hrs)	Test Result (time of failure occurrence)	Cumulative Damage
1	2	250	0 fails in 11	
2	3	100	3 fails in 11 (30; 60; 80)	27.27%
3	4	20	3 fails in 8 (2; 10; 16)	54.55%
4	5	10	3 fails in 5 (1; 4; 8)	81.82%
5	6	10	2 fails in 2 (1; 5)	100.00%
6	7	10	0 fails in 0	

Step 2 of the plotting method requires the life distribution curves at each step. Note that no accumulated damage is considered: in step 2, for example, the probability plot is drawn based on the stress level 3V, where three items fail at 30 hrs, 60 hrs, and 80 hrs, and 8 items censored at 100 hrs. The plot with life distribution curves at each step is shown below in Figure 4.20. The corresponding estimates of the parameters of the Weibull distribution are shown in Table 4.9.

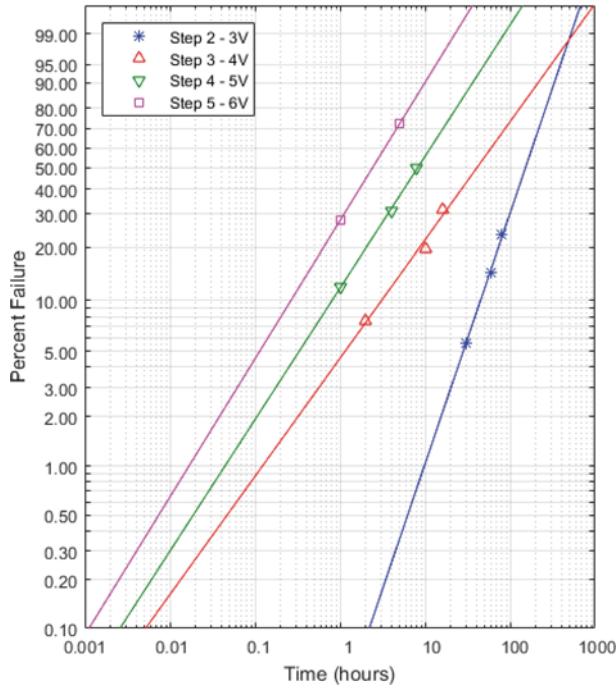


Figure 4.20: Weibull multi plot for step 2 of example

Table 4.9: Estimate of the parameters of Weibull distribution

Stress (Volts)	β	α
3	1.545	193.17
4	0.729	67.94
5	0.811	12.98
6	0.851	3.74

The initial estimate for β is obtained from the average β from the plot:

$$\hat{\beta} = \frac{1.545 + 0.729 + 0.811 + 0.851}{4} = 0.984.$$

Assuming that PoF analysis showed that the life-stress model is best described by the inverse power model, initial estimates of the life-stress model parameters K and n are obtained from regression of the log-linear version of the inverse power law model, $\ln(t) = -\ln(K) - n \ln(S)$. The value for $\ln(t)$ may be associated with a given quantile of the plotted distributions (e.g., 10% or the scale parameter 63.2%) for each plot. Applying these to the above equation yields the following set of simultaneous equations:

$$\begin{cases} \ln(193.17) = -\ln(K) - n \cdot \ln(3) \\ \ln(67.94) = -\ln(K) - n \cdot \ln(4) \\ \ln(12.98) = -\ln(K) - n \cdot \ln(5) \\ \ln(3.74) = -\ln(K) - n \cdot \ln(6) \end{cases}$$

Next, we need to find estimates of \hat{K} and \hat{n} such that the errors $\sum_i \{\ln(t_i) - [-\ln(K) - n \cdot \ln(S)]\}$ are minimized. This is basically a simple optimization problem, but it can be solved by trial and error too, which results in: $\hat{K} = 7.24 \times 10^{-6}$, and $\hat{n} = 5.77$.

Assuming that Weibull is the most adequate life distribution, using the above parameter estimates, step 3 is performed to determine the acceleration factor between all steps using Equation (4.127). The results are shown in Table 4.10. Based on these acceleration factors, find the equivalent time in a given step that produces the same damage the unit received while running in all the lower steps and accounting for the difference between the test time at each step.

Taking cumulative damage into account, step 4 should be performed to compute the equivalent time for each step, and then adjust a and n as follows. For the inverse power life-stress model, the acceleration factor is given as $AF = \frac{L_{use}}{L_{acc}} = \left(\frac{S_{acc}}{S_{use}}\right)^n$. By calculating the acceleration factor between each applicable step, we can determine the equivalent times (to produce the same damage in the current step) due to the actual times the item spent in the previous steps:

$$t_j^e = t_j + \sum_{i < j} \frac{t_i}{AF_{i \rightarrow j}} \quad (4.129)$$

For the case of inverse power life-stress relationship, Equation (4.133) can be written as

$$t_j^e = t_j + \sum_{i < j} \frac{t_i}{\left(\frac{S_j}{S_i}\right)^n} = t_j + \sum_{i < j} \left(t_i \left(\frac{S_i}{S_j}\right)^n \right) \quad (4.130)$$

Table 4.10: Acceleration factor between all steps

From	To	AF	Equivalent Time (hours)
Step 1	Step 2	10.39	250 @ step 1 = 24.06 @ step 2
Step 1	Step 3	54.69	250 @ step 1 = 4.57 @ step 3
Step 1	Step 4	198.33	250 @ step 1 = 1.26 @ step 4
Step 1	Step 5	568.21	250 @ step 1 = 0.44 @ step 5
Step 1	Step 6	1383.59	250 @ step 1 = 0.18 @ step 6
Step 2	Step 3	5.26	100 @ step 2 = 19.00 @ step 3
Step 2	Step 4	19.09	100 @ step 2 = 5.24 @ step 4
Step 2	Step 5	54.69	100 @ step 2 = 1.83 @ step 5
Step 2	Step 6	133.17	100 @ step 2 = 0.75 @ step 6
Step 3	Step 4	3.63	20 @ step 3 = 5.52 @ step 4
Step 3	Step 5	10.39	20 @ step 3 = 1.92 @ step 5
Step 3	Step 6	25.30	20 @ step 3 = 0.79 @ step 6
Step 4	Step 5	2.87	10 @ step 4 = 3.49 @ step 5
Step 4	Step 6	6.98	10 @ step 4 = 1.43 @ step 6
Step 5	Step 6	2.43	10 @ step 5 = 4.11 @ step 6

From Equation (4.130), we get the equivalent times of failure t_j^e , in each step in Table 4.11.

Next, we determine the new estimates of parameters affected by adjusted AFs. Adjusted acceleration factors are obtained from Equation (4.128) and listed in Table 4.12.

Table 4.11: Computation of equivalent time that produces the same amount of damage

Step	Step Actual Time	Step Equivalent Time	Cumulative Damage
1	250	250	
2	100	124.06	27.27%
3	20	43.57	54.55%
4	10	22.01	81.82%
5	10	17.68	100.00%

Table 4.12: Adjusted acceleration factor

Steps (from i to j)	Adjusted AFs	Adjusted n
2 to 3	5.69	6.05
2 to 4	16.91	5.54
2 to 5	25.72	4.69
3 to 4	2.97	4.88
3 to 5	4.52	3.72
4 to 5	1.52	2.30
Average n		4.53

After repeating the process (i.e., the calculations that lead to Tables 4.10, 4.11, and 4.12) over a few iterations, the result converges to values listed in Table 4.13.

Based on the results in Table 4.13, the final adjustment in acceleration factors will result in final estimates of the parameters as shown in Table 4.14.

Table 4.13: Final results of equivalent time computation after a few iterations

Step	Step Actual Time	Step Equivalent Time	Cumulative Damage
1	250	250	
2	100	144.56	27.27%
3	20	62.53	54.55%
4	10	34.20	81.82%
5	10	25.75	100.00%

Table 4.14: Final acceleration factors and estimation of the parameters of life-stress model

Steps (from i to j)	Adjusted AFs	Adjusted n
2 to 3	4.62	5.32
2 to 4	12.68	4.97
2 to 5	20.58	4.36
3 to 4	2.74	4.52
3 to 5	4.45	3.68
4 to 5	1.62	2.66
Average n		4.25
Optimized n		4.36

Note that the optimized n above was determined by minimizing the error associated with the difference between the calculated acceleration factors at each step above and the acceleration factors calculated from estimated parameters. Since the plotting method only provides rough estimates, using the average value of n for these calculations is more suitable. From the above, we can update the Weibull plot based on total time experienced at each step as shown in Figure 4.21 and summarized in Table 4.15.

Table 4.15: Step stress equivalent times for Weibull plot

STEP	STRESS	DURATION	EFFECTIVE FAILURE TIMES	# FAILED	# AT START	CUMULATIVE FAILURE %
2	3	144.56	74.56,104.56,124.56	3	11	27.27%
3	4	62.53	44.53,52.53,58.53	3	8	54.55%
4	5	34.20	25.20,28.20,32.20	3	5	81.82%
5	6	25.75	16.75,20.75	2	2	100.00%

From the results of the shape parameters in Figure 4.21, the updated estimate for β is

$$\hat{\beta} = \frac{2.99 + 5.82 + 6.87 + 6.40}{4} = 5.52$$

Updated estimates of K and n are obtained using log linear inverse power regression as follows

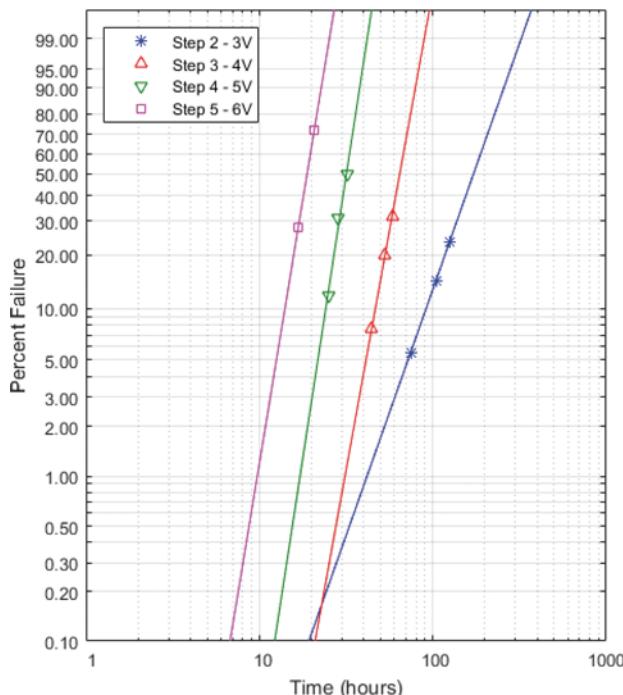


Figure 4.21: Final Weibull plots for the corrected data of step-stress test example

$$\begin{cases} \ln(194.25) = -\ln(K) - n \cdot \ln(3) \\ \ln(68.70) = -\ln(K) - n \cdot \ln(4) \\ \ln(33.53) = -\ln(K) - n \cdot \ln(5) \\ \ln(19.96) = -\ln(K) - n \cdot \ln(6) \end{cases}$$

Next, we obtain estimates for the parameters of the model through the least squares regression approach to obtain the final estimates as:

$$\hat{K} = 1.44 \times 10^{-4} \text{ and } \hat{n} = 3.29.$$

It is useful to compare the original estimates assuming no damage accumulation versus new estimates after correcting for damage accumulation. Table 4.16 summarizes the comparison.

Table 4.16: Comparisons of original estimates and final estimates after correcting for cumulative damage

Parameter	Original	Adjusted
$\hat{\beta}$	0.984	5.52
\hat{K}	7.24×10^{-6}	1.44×10^{-4}
\hat{n}	5.77	3.29

MTTF of device at the use stress level may also be calculated as

$$MTTF = \frac{1}{KS^n} \Gamma\left(\frac{1+\beta}{\beta}\right) = \frac{1}{1.44 \times 10^{-4} \times 2^{3.29}} \Gamma\left(\frac{1+5.52}{5.52}\right) = 655 \text{ hrs}$$

Example 4.16

Consider the case where a unit is subjected to both temperature and humidity. The range of 65 °C to 95 °C is selected for temperature, while 70% to 95% is selected for humidity. It appears that these temperature and humidity ranges are sufficient to conduct a step-stress test and do not represent conditions causing immediate catastrophic failures. The steps that are at least 5 °C and 5% RH are meaningful. This is because of limited environmental chamber technology and control in many situations. Fifteen samples are placed into the step-stress test and tested using the same tentative test plan as shown in the Table 4.17. Units that fail are removed from the test and their total time on test is recorded. Further, assume that the life at each step is characterized by the Weibull distribution, and that the thermal-humidity relationship best represents the relationship between life and stress.

Table 4.17: Tentative step-stress test plan

Step	Temperature (°C)	Humidity (%)	Test Time (Days)
1	65	70	7
2	75	80	12
3	85	85	8
4	95	95	7

Three of the most likely test outcomes are presented in Table 4.18. Discuss the validity of each of them, and choose the best result for the plotting method.

Table 4.18: Hypothetical results summary

Step	Test Result 1	Test Result 2 (days of failure occurrence)	Test Result 3
1	0 failure in 15	0 fails in 15	3 fails in 15
2	0 failure in 15	2 fails in 15 (8; 11)	8 fails in 12
3	1 failure in 15	3 fails in 13 (3; 5; 7)	3 fails in 4
4	2 fails in 14	3 fails in 10 (1; 3; 6)	1 fails in 1

Solution

Test Result 1: Shows too few failures in general through Step 4. This makes conclusions difficult to determine. The test might continue.

Test Result 2: Represents a good compromise of failures at each step. It is rare that results are as nice as in test 2. This is the prime reason to try to adjust the fallout in each of the steps and spread the failures more evenly as suggested by the ground rules. Flexibility is useful in selecting the duration of each step. Sufficient failures from each step would be a minimum of 3 in this sample of 15, but in all cases, no more than 1/3 of the remaining total.

Test Result 3: Shows too many failures in Step 2, Steps 3 and Step 4, and then has too few samples left to provide meaningful results.

Therefore, the test result 2 is the most feasible. If testing results like 1 or 3 are obtained, the step-stress test may have to be revised or even redesigned. Therefore, the detail of testing result 2 is shown as Table 4.19, which would be focused on in the following analysis.

Table 4.19: Detail of testing result 2

Step	DGS	Number in Group	Failed (F) Suspended	State End Time	Humidity RH	Temperature °K	Cumulative Damage
1	1	15	S	7	70%	525	
	2	1	F	8	80%	535	
2	3	1	F	11	80%	535	13.33%
	4	13	S	12	80%	535	
	5	1	F	3	85%	545	
3	6	1	F	5	85%	545	33.33%
	7	1	F	7	85%	545	
	8	10	S	8	85%	545	
	9	1	F	1	95%	555	
4	10	1	F	3	95%	555	53.33%
	11	1	F	6	95%	555	
	12	7	S	7	95%	555	

The life distribution curves at each step are shown below in Figure 4.22, and the corresponding estimate of the parameters of the Weibull distribution is shown in Table 4.20.

Table 4.20: Estimate of the parameters of Weibull distribution

Step	Temperature (°C)	Humidity (%)	β	α
2	75	80	2.89	23.21
3	85	85	1.71	16.59
4	95	95	0.82	25.58

The initial estimate of the parameters of the temperature-humidity Weibull is obtained from the arithmetic average of β from the plot and yields a point estimate of

$$\hat{\beta} = \frac{2.89 + 1.71 + 0.82}{4} = 1.81$$

Assuming that PoF analysis showed that the life-stress model is best described by the inverse power model, initial estimates of the life-stress model parameters a and n are obtained from regression of the log-linear version of the thermal-humidity relationship: $\ln(t) = \ln(A) + \frac{\phi}{V} + \frac{b}{U}$. The value for $\ln(t)$ may be associated with a given percentile of the plotted distributions (e.g., 10% or the scale parameter 63.2%) for each plot. Applying these to the above equation yields the following set of simultaneous equations.

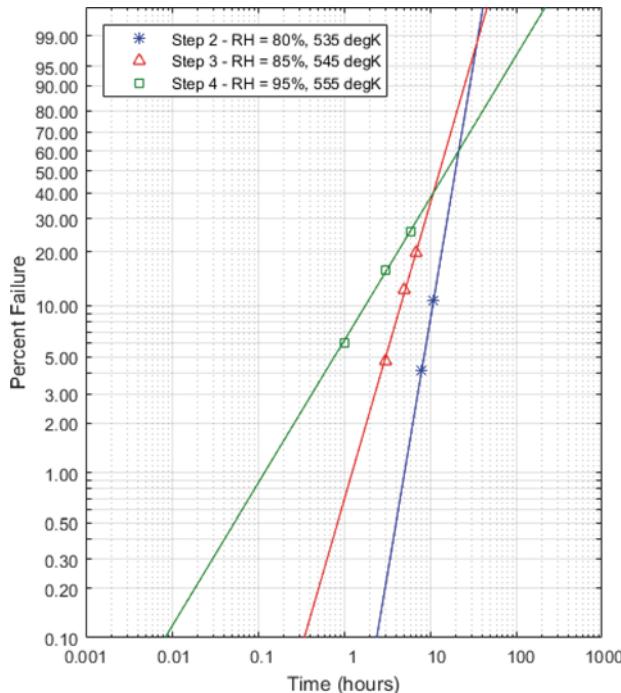


Figure 4.22: Weibull multi plot for step 2 of example

$$\begin{cases} \ln(23.21) = \ln(A) + \frac{\phi}{348} + \frac{b}{0.8} \\ \ln(16.59) = \ln(A) + \frac{\phi}{358} + \frac{b}{0.85} \\ \ln(25.58) = \ln(A) + \frac{\phi}{368} + \frac{b}{0.95} \end{cases}$$

Solving the above set of equations yields $\hat{A} = 4.55 \times 10^{-14}$, $\hat{b} = 18494$, and $\hat{\phi} = 17.28$

The corresponding acceleration factors between all steps are determined using the above parameter estimates. For the temperature-humidity life-stress model, the acceleration factor is given as

$$AF = \frac{L_{use}}{L_{acc}} = \exp \left[\phi \left(\frac{1}{v_U} - \frac{1}{v_A} \right) + b \left(\frac{1}{u_U} - \frac{1}{u_A} \right) \right]. \quad (4.131)$$

Based on these acceleration factors, find the equivalent time in a given step that produces the same damage the unit received while running in all the lower steps and accounts for the difference between the test time at each step. The results are shown in Table 4.21.

Table 4.21: Acceleration factor between all steps

From	To	AF	Equivalent Time
Step 1	Step 2	1.8731	7 days @ step 1 = 3.74 days @ step 2
Step 1	Step 3	3.4874	7 days @ step 1 = 2.00 days @ step 3
Step 1	Step 4	6.2913	7 days @ step 1 = 1.11 days @ step 4
Step 2	Step 3	1.8618	12 days @ step 2 = 6.45 days @ step 3
Step 2	Step 4	3.3488	12 days @ step 2 = 3.57 days @ step 4
Step 3	Step 4	1.804	8 days @ step 3 = 4.43 days @ step 4

By calculating the acceleration factor between each applicable step, we can determine the equivalent times (to produce the same damage in the current step) due to the actual times the item spent in the previous steps. For the case of temperature-humidity life-stress relationship, Equation (4.129) can be written as

$$t_j^e = t_j + \sum_{i < j} \frac{t_i}{\exp \left[\phi \left(\frac{1}{v_U} - \frac{1}{v_A} \right) + b \left(\frac{1}{u_U} - \frac{1}{u_A} \right) \right]} \quad (4.132)$$

From Equation (4.135), we get the results shown in Table 4.22.

Table 4.22: Computation of equivalent time that produces the same amount of damage

Step	Step Actual Time (days)	Step Equivalent Time (days)	Cumulative Damage
1	7	7	
2	12	3.74+12 = 15.74	13.33%
3	8	2.00+6.45+8 = 16.45	33.33%
4	7	1.11+3.57+4.43+7 = 16.11	53.33%

Determine the new estimates of parameters affected by the adjusted AFs. Adjusted acceleration factors are obtained from Equation (4.128) and are listed in Table 4.23.

Table 4.23: Adjusted acceleration factor

Steps (from i to j)	Adjusted AFs
2 to 3	$= \frac{33.33\%}{13.33\%} \times \frac{15.74}{16.45} = 2.39$
2 to 4	$= \frac{53.33\%}{13.33\%} \times \frac{15.74}{16.45} = 3.91$
3 to 4	$= \frac{53.33\%}{33.33\%} \times \frac{15.74}{16.45} = 1.63$

With this new estimate of the acceleration factors, one can calculate better estimates of β , A , b , and ϕ . With the new estimate, we then calculate improved acceleration factors. After several more passes through this process, the best estimates of β , A , b , and ϕ with the data set can be obtained. This is the model that should be extrapolated to the use level.

$$\begin{cases} 2.39 = \exp \left[\phi \left(\frac{1}{535} - \frac{1}{545} \right) + b \left(\frac{1}{80} - \frac{1}{85} \right) \right] \\ 3.91 = \exp \left[\phi \left(\frac{1}{535} - \frac{1}{555} \right) + b \left(\frac{1}{80} - \frac{1}{95} \right) \right] \\ 1.63 = \exp \left[\phi \left(\frac{1}{545} - \frac{1}{555} \right) + b \left(\frac{1}{85} - \frac{1}{95} \right) \right] \end{cases}$$

Next, we need to find estimates of ϕ and b such that the errors $\Delta|AF_{3,2} - 2.39|$, $\Delta|AF_{4,2} - 3.91|$, $\Delta|AF_{4,3} - 1.63|$ are minimized (with AF's in the error values calculated from the estimated parameters). This is basically a simple optimization problem which can be solved by trial and error too.

4.16.2. MAXIMUM LIKELIHOOD ESTIMATION METHOD FOR STEP-STRESS DATA ANALYSIS

Section 4.9 discussed the MLE approach to data analysis for constant time stress application. For time-varying stress tests such as step-stress tests, the MLE approach is similar. For step-stress testing, the likelihood and log-likelihood functions, Equations (4.19), (4.20), (4.21) and (4.22), still apply. However, the added complication is with the mathematical formulation of the pdf $f(\cdot)$ and CDF $F(\cdot)$ functions, since each of these needs to account for the damage accumulated at each step. To develop the mathematical formulation for these, consider the step-stress test shown below in Figure 4.23, with stresses S_1 , S_2 , and S_3 . Further, assume that the underlying life distribution is the Weibull distribution and that an inverse power relationship between life and stress exists.

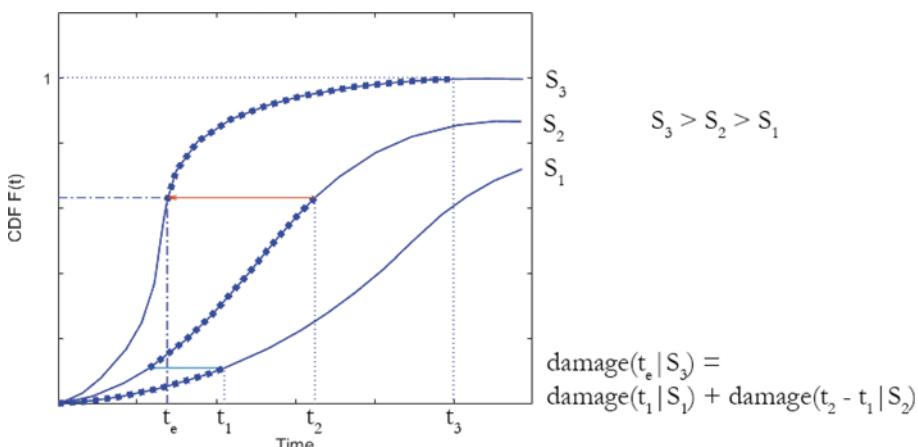


Figure 4.23: Step stress profile and corresponding life distributions

From the power life-stress with the Weibull life model Equation (4.71), the cdf model is $F(t, S) = 1 - e^{-(KS^n t)^\beta}$.

Therefore, the conditional cdf given the stress used in each step shown in Figure 4.23 would be

$$F_1(t, S_1) = 1 - e^{-(KS_1^n t)^\beta} \quad (4.133)$$

$$F_2(t, S_2) = 1 - e^{-(KS_2^n t)^{\beta}}$$

$$F_3(t, S_3) = 1 - e^{-(KS_3^n t)^{\beta}}$$

Using the cdf functions above, we should incorporate the cumulative exposure model to adjust these expressions for the damage experienced by the item prior to entering each step. The cumulative exposure model assumes that the remaining life of the test units depends only on the damage accumulated by the units up to that point and that there is no “memory” of exactly how the damage was accumulated. Figure 4.23 shows $F_0(t)$, which illustrates the effective cumulative distribution, as contributed by each portion of $F_i(t)$ as shown in Figure 4.23. Now, since the units are held at a constant stress at each step, the surviving units will fail in accordance with the distribution at the current step, but with a starting time corresponding to the total time accumulated up to the beginning of the current step.

Units failing at the first step have not experienced any other stresses and will fail in accordance with the cumulative probability of failure $F(t, S_1)$. On the other hand, units that make it to the second step will fail in accordance with $F(t, S_2)$, but will have accumulated some equivalent time, t_1^e , that accounts for the time t_1 spent by the item in a previous step under stress S_2 . Hence, in general the equivalent cdf at the i^{th} step is more accurately presented as:

$$F_i(t, S_i) = 1 - e^{-(KS_i^n [(t - t_{i-1}) + t_{i-1}^e])^{\beta}} \quad (4.134)$$

where t_{i-1}^e is the equivalent time (same cumulative probability of failure and thus the same damage) of step (i-1) if the item operated under the stress of step (i). For example, the probability that the units will fail at a time t (somewhere between t_1 and t_2) while experiencing stress S_2 is equivalent to the probability that the units will fail after accumulating $(t - t_1)$ plus an additional time of τ_1 , to account for the damage accumulated at stress S_1 . The equivalent time t_1^e is obtained from

$$\begin{aligned} F(t_1, S_1) &= F(t_1^e, S_2) \\ 1 - e^{-(KS_1^n t_1)^{\beta}} &= 1 - e^{-(KS_2^n t_1^e)^{\beta}} \\ S_1^n t_1 &= S_2^n t_1^e \\ t_1^e &= t_1 \left(\frac{S_1}{S_2} \right)^n, \end{aligned} \quad (4.135)$$

which is the same as (4.137). Or in general the equivalent time is obtained from

$$t_{i-1}^e = (t_{i-1} - t_{i-2} + t_{i-2}^e) \left(\frac{S_{i-1}}{S_i} \right)^n \quad (4.136)$$

Once the CDF for each step has been determined as per above, the pdf can also be determined from the following:

$$f_i(t, S_i) = -\frac{d}{dt} [F_i(t, S_i)] \quad (4.137)$$

The log-likelihood function for this case when complete and right censored data exist would be

$$\Lambda = \sum_{i=1}^{N_c} n_i \ln \left[\beta K S_i^n (K S_i^n t_i^e)^{\beta-1} e^{-(K S_i^n t_i^e)^{\beta}} \right] - \sum_{j=1}^{N_r} n_j \cdot (K S_j^n t_j^e)^{\beta} \quad (4.138)$$

The MLE solution (parameter estimates $\hat{\beta}, \hat{K}, \hat{n}$) will be obtained by solving for β, K , and n such that $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial K} = 0$ and $\frac{\partial \Lambda}{\partial n} = 0$. The derivatives are similar to Equation (4.74) with the times to failure and right-censored time t_i and t_j changed to t_i^e and t_j^e respectively.

Example 4.17

For the data given in Table 4.8, stresses are

$$\begin{aligned} S_1 &= 2, & 0 < t \leq 250 \\ S_2 &= 3, & 250 < t \leq 350 \\ S_3 &= 4, & 350 < t \leq 370 \\ S_4 &= 5, & 370 < t \leq 380 \\ S_5 &= 6, & 380 < t \leq 390 \\ S_6 &= 7, & 390 < t \leq \infty \end{aligned}$$

Using the log-likelihood function defined in Equation (4.138) estimate the parameters $\hat{\beta}, \hat{K}, \hat{n}$.

Solution

The parameter estimates for β, K , and n are obtained by solving for $\frac{\partial \Lambda}{\partial \beta} = 0$, $\frac{\partial \Lambda}{\partial K} = 0$ and $\frac{\partial \Lambda}{\partial n} = 0$ for the log-likelihood Equation (4.138). The equivalent times to failure may be computed by Equations (4.135) – (4.136) which are also equivalent to Equation (4.130). Below for example are the equivalent failure times at stress level 2 (3V) and stress level 3 (4V),

$$\begin{aligned} t_i^{e@3V} &= t_i + (250hr) \left(\frac{2V}{3V} \right)^n \\ t_i^{e@4V} &= t_i + \left[20hr + (250hr) \left(\frac{2V}{3V} \right)^n \right] \left(\frac{3V}{4V} \right)^n = t_i + (250hr) \left(\frac{2V}{4V} \right)^n + (20hr) \left(\frac{3V}{4V} \right)^n \end{aligned}$$

The equivalent times to survival t_j^e for each stress level are calculated similarly. The parameter MLE values below are found using a MATLAB script,

$$\begin{aligned} \hat{\beta} &= 2.30 \\ \hat{K} &= 1.09 \times 10^{-5} \\ \hat{n} &= 5.07 \end{aligned}$$

These results are more accurate than the approximate approach used in Section 4.16.1.

Example 4.18

Consider the following example where the maximum shearing stress model may express the wear life of a bearing as follows:

$$N = C \left[\frac{\tau_{yp}}{\tau_{max}} \right]^n$$

where

N = wear life cycles

C, n = constants to be determined from the test results

τ_{yp} = material yield point

τ_{max} = maximum shear stress in the vicinity of the surface.

A step-stress test (i.e. for stress τ_{max} measured in psi) is performed where τ_{max} is increased in a stepwise fashion to measure the wear life in cycles. Clearly, wear occurs in all steps and is cumulative as one goes through the steps. The following information is recorded:

Time (cycles)	Stress Level (psi)
0-2500	250
2500-3500	750
3500-3700	1500

The objective of this test is to determine the parameters of the model and the life at normal operating life of 200 psi. Seven units were tested, and wear failures occurred at the following cycles: 2800, 3100, 3300, 3520, 3600, and 3660. Assuming the Weibull life model, determine the parameters of the distribution and constants C and n (assume $\tau_{yp} = 1400$ psi).

Solution

Rearranging the given life-stress relationship, we get the IPL relationship:

$$N = C \left[\frac{\tau_{yp}}{\tau_{max}} \right]^n = \frac{1}{a \cdot \tau_{max}^n}$$

where $a = \frac{1}{C \cdot \tau_{yp}^n}$; therefore, $C = \frac{1}{a \cdot \tau_{yp}^n}$

This problem then becomes a matter of estimating a and n from the given step-stress data. The given stress profile can be expressed in terms of equivalent times as

Actual time of failure/censoring	Stress τ_{max} when failure occurred	Equivalent time of failure/censoring
2800	750	$300 + 2500 \left(\frac{250}{750} \right)^n$
3100	750	$600 + 2500 \left(\frac{250}{750} \right)^n$
3300	750	$800 + 2500 \left(\frac{250}{750} \right)^n$
3520	1500	$20 + 2500 \left(\frac{250}{1500} \right)^n + 1000 \left(\frac{750}{1500} \right)^n$
3600	1500	$100 + 2500 \left(\frac{250}{1500} \right)^n + 1000 \left(\frac{750}{1500} \right)^n$
3660	1500	$160 + 2500 \left(\frac{250}{1500} \right)^n + 1000 \left(\frac{750}{1500} \right)^n$
3700+	1500	$200 + 2500 \left(\frac{250}{1500} \right)^n + 1000 \left(\frac{750}{1500} \right)^n$

From equivalent failure times (including one censored point), the log-likelihood function Equation (4.138) should be used to estimate $\hat{\beta}$, \hat{a} , and \hat{n} as follows: $\hat{\beta} = 2.23$, $\hat{a} = 19633$, and $\hat{n} = 2.26$, which equates to $C = 390.74$.

4.16.3. BAYESIAN INFERENCE METHOD FOR STEP-STRESS DATA ANALYSIS

As we discussed previously, the accuracy of the commonly used maximum likelihood estimation (MLE) approach mainly depends on whether the model parameters are known precisely. In reality, the information we can get on model parameters always has much uncertainty due to limited samples. Therefore, the Bayesian inference method is a good approach for conducting the inferences from the step-stress accelerated life test.

To demonstrate the application of Bayesian inference, consider the step stress test conceptually shown in Figure 4.24 involving n units, initially tested at a low stress level, S_1 . The stress level remains at S_1 until the stress changing point t_1 at which the stress is increased to a higher stress level S_2 , with the j -th failures observed shown as t_j^* and censored units as t_j^{*+} . The test continues until a predetermined censoring time t_n . The stress levels are represented as S_i , $i = 0, 1, 2 \dots n$. The normal stress level is denoted as S_0 . We are interested in estimating the t_F , which is the F -th percentile of the lifetime distribution for stress at the normal stress condition S_0 .

Assume that the item's failure time is subject to the Weibull distribution at each stress level, and the life-stress is subject to the inverse power relationship. Therefore, we can assume that the Weibull scale parameter α_i is a function of the stress S_i , $i = 0, 1, 2 \dots n$.

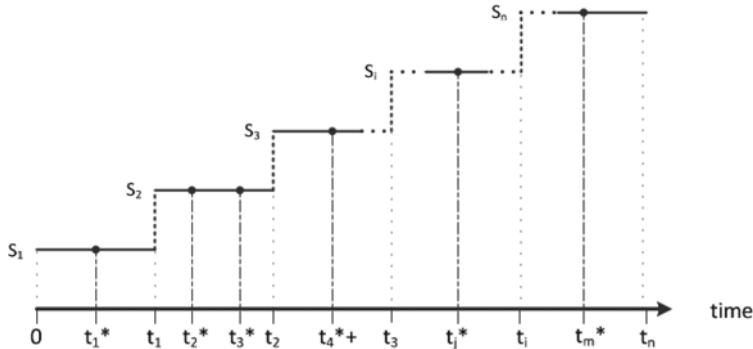


Figure 4.24: A conceptual step-stress test

$$\alpha_i = \frac{1}{AS_i^p} \quad (4.139)$$

where A and p are unknown parameters.

The corresponding cdf and pdf for the failure data points at step S_i are as below:

$$\text{CDF: } F_i(t_c, S_i) = 1 - \exp \left[-(t_c AS_i^p)^\beta \right] \quad (4.140)$$

$$\text{PDF: } f_i(t_c, S_i) = \beta AS_i^p (t_c AS_i^p)^{\beta-1} \exp \left[-(t_c AS_i^p)^\beta \right]$$

where β is the shape parameter, α_i is the scale parameter, and t is the time of failure data point at step S_i .

The probability of obtaining a censored observation is:

$$R_i(t_r, S_i) = \exp \left[-(t_c A S_i^p)^{\beta} \right] \quad (4.141)$$

where β is the shape parameter, α_i is the scale parameter, and t_r is the time of right censored data point at step S_i .

Taking the cumulative exposure model into account, we can rewrite the above equations as following:

$$\begin{aligned} \text{CDF: } F_i(t_{c_{ij}}^*, S_i) &= 1 - \exp \left\{ - \left[(t_{c_{ij}}^* - t_{i-1} + \tau_{i-1}) A S_i^p \right]^{\beta} \right\} \\ \text{PDF: } f_i(t_{c_{ij}}^*, S_i) &= \beta A S_i^p \left[(t_{c_{ij}}^* - t_{i-1} + \tau_{i-1}) A S_i^p \right]^{\beta-1} \exp \left\{ - \left[(t_{c_{ij}}^* - t_{i-1} + \tau_{i-1}) A S_i^p \right]^{\beta} \right\} \end{aligned} \quad (4.142)$$

where β is the shape parameter, α_i is the scale parameter, A and p are unknown parameters, and $t_{c_{ij}}^*$ is the time of the j^{th} failure data point at step S_i , $\tau_{i-1} = (t_{i-1} - t_{i-2} + \tau_{i-2}) \left(\frac{S_{i-1}}{S_i} \right)^p$

Therefore, there are three parameters A , p , and β in the above model, and t_F is a function of A , p , and β :

$$t_F(S_i) = \exp \left\{ - \ln A - p \ln S_i + \frac{1}{\beta} \ln[-\ln(1 - F)] \right\} \quad (4.143)$$

The probability of obtaining a censored observation is:

$$R_i(t_{r_{ik}}, S_i) = \exp \left[-(t_{r_{ik}} A S_i^p)^{\beta} \right] \quad (4.144)$$

where β is the shape parameter, α_i is the scale parameter, and $t_{r_{ik}}$ is the time of the k^{th} right censored data point at step S_i .

There are two ways to construct the likelihood functions:

- 1) Construct the likelihood function by each step:

$$l = \prod_{i=1}^n \prod_{j=1}^{c_{ij}} f_i(t_{c_{ij}}^*, t_{c_{i2}}^* \dots t_{c_{ij}}^*) \prod_{k=1}^{r_{ik}} [1 - F_i(t_{r_{i1}}^*, t_{r_{i2}}^* \dots t_{r_{ik}}^*)] \quad (4.145)$$

where n = number of steps; c_{ij} = number of failure data points at step i ; $t_{c_{ij}}^*$ = time of j^{th} failure data point at step i ; r_{ik} = number of right censored data points at step i ; and $t_{r_{ik}}^*$ = time of k^{th} right censored data point at step i .

- 2) Construct the likelihood function by converting all the previous steps equivalently to the final step n :

$$l = \prod_{j=1}^{j'} f_n(t_{c_{n1}}^*, t_{c_{n2}}^* \dots t_{c_{nj'}}^*) \prod_{k=1}^{k'} [1 - F_n(t_{r_{n1}}^*, t_{r_{n2}}^* \dots t_{r_{nk'}}^*)] \quad (4.146)$$

where $j' = \sum_{i=1}^n c_{ij}$ = total number of failure data points; $k' = \sum_{i=1}^n r_{ik}$ = total number of censored data points; $t_{c_{nj'}}^*$ = time of the j'^{th} failure data point; and $t_{r_{nk'}}^*$ = time of the k'^{th} right censored data point. $t_{c_{nj'}}^* = (t_{c_{ij}}^* - t_i) \left(\frac{s_n}{s_i}\right)^p$; $t_{r_{nk'}}^* = (t_{r_{ik}}^* - t_i) \left(\frac{s_n}{s_i}\right)^p$

The likelihood function generated by the second way is used for the following analysis. Denote the joint prior distribution of the three model parameters by $f(A, p, \beta)$. Assume that we had the experimental data, which consist of n independently and identically distributed (iid) observations $\vec{t} = (t_{c_{n1}}^*, t_{c_{n2}}^* \dots t_{c_{nj'}}^*; t_{r_{n1}}^*, t_{r_{n2}}^* \dots t_{r_{nk'}}^*)$. According to the Bayes theorem, the joint posterior distribution of the model parameters given the data t is:

$$\pi_1(A, n, \beta | t) = \frac{l(t|A, p, \beta) \pi_0(A, p, \beta)}{\iiint l(t|A, p, \beta) \pi_0(A, p, \beta) dAdpd\beta} \quad (4.147)$$

To solve the above posterior, multiple levels of integration are necessary, which are often analytically intractable, and sometimes even a numerical integration cannot be directly obtained. In this case, Markov Chain Monte Carlo (MCMC) simulation is the easiest way to get reliable results without evaluating integrals. WinBUGS is used to implement MCMC simulation.

One MCMC alternating conditional sampling called Gibbs sampling, which is particularly useful in high dimensional problems, is used in WinBUGS. Each Gibbs sampling iteration cycles through the unknown parameters, drawing a sample of one parameter conditional on the latest values of all the other parameters. With a number of iterations, the samples drawn on one parameter can be regarded as simulated observations from its posterior distribution. Then the functions of the model parameters, such as the F^{th} percentile of the lifetime distribution at the normal stress condition, denoted by $t_F(S_0)$, can also be conveniently sampled.

Example 4.19

Consider a step-stress test of cable insulation. The test was run to estimate life at a design stress of 400 volts/mm. The data in Table 4.24 illustrate the test data. Each specimen was first held for 10 minutes each at 5 kV, 10 kV, 15 kV, and 20 kV before it went into step 1 at 26 kV. In steps 1 through 10, a specimen has the same hold time at each voltage (15 minutes, 1 hour, 4 hours, or 16 hours). Table 4.24 shows the step number and the total time on test when a specimen broke down, and its insulation thickness (used to calculate its stress as the voltage divided by the thickness). Assume that this tested cable 1 is comparable to another cable 2 tested before, so we can use the MLE result of the test of cable 2 as the prior information. Assume that cable 2 was tested and the data shown in Table 4.25 were obtained. Update the models parameters using the Bayesian estimation.

Solution

Equations (4.139)-(4.147) and the Bayesian inference expression in Equation (4.147) are used to solve this problem. A MATLAB routine that uses Metropolis-Hastings MCMC was developed to compute the values of the prior and posterior parameters. Assuming non-informative uniform priors for the parameters $UNIF(0,100)$ (i.e., uniformly distributed). We obtain the posterior for the parameters using the data from Table 4.24. The results for the posterior for the parameters are shown in Table 4.26.

Table 4.24: Step-stress test data on cable 1 (* Before step 1, each specimen was held 10 minutes each at 5, 10, 15, and 20 kV; + denotes a right censored)

Voltage Pattern		Specimen Data			
Step*	Kilovolts	Hold (min)	Final Step	Total Time to Failure (min)	Thickness (mm)
1	26	15	5	102	27
		15	5	113	27
2	28.5	15	5	113	27
3	31	60	6	370+	29.5
		60	6	345	29.5
4	33.4	60	6	345+	28
5	36	240	6	1249	29
		240	6	1333	29
6	38.5	240	6	1333+	29
		240	5	1096.9	29
7	41	240	6	1250.8	30
		240	5	1097.9	29
8	43.5	960	3	2460.9	30
		960	3	2460.9+	30
9	46	960	3	2700.4	30
10	48.5	960	4	2923.9	30
		960	2	1160	30
		960	3	1962.9	30
		960	1	363.9+	30
		960	1	898.4+	30
		960	5	4142.1	30

Table 4.25: Prior distribution results of cable 2

Parameter	Lower Bound	Upper Bound
A	4.38×10^{-10}	5.69×10^{-5}
p	1.37×10^{-4}	2.78
β	1.74×10^{-4}	12.10

Table 4.26: Posterior distribution results from data of cable 1 and priors of cable 2

Parameter	Lower Bound	Upper Bound
A	3.70×10^{-9}	5.68×10^{-5}
p	1.12×10^{-3}	1.82
β	4.68×10^{-3}	12.10

REFERENCES

- Chatterjee, K., & Modarres, M. (2012). A probabilistic physics-of-failure approach to prediction of steam generator tube rupture frequency. *Nuclear Science and Engineering*, 170(2), 136-150.
- Collins, J. (1993). Failure of Materials in Mechanical Design (2nd ed.). New York: John Wiley & Sons.
- Hall, P., & Strutt, J. (2003). Probabilistic physics-of-failure models for component reliabilities using Monte Carlo simulation and Weibull analysis: a parametric study. *Reliability Engineering & System Safety*, 80(3), 233-242.
- Mettas, A. (2000). *Modeling & Analysis for Multiple Stress-Type Accelerated Life Data*. Retrieved November 27, 2015, from www.weibull.com:
http://www.reliasoft.com/pubs/2000rm_088.pdf?_ga=1.65677038.130520085.1448665843
- Modarres, M., Kaminskiy, M., & Krivtsov, V. (2017). *Reliability engineering and risk analysis: A practical guide 3rd Ed.* Boca Raton: CRC Press.
- Nelson, W. (2004). Accelerated Testing - Statistical Models, Test Plans, and Data Analysis. New York: John Wiley & Sons.
- Paris, P., Gomez, M., & Anderson, W. (1961). A rational analytic theory of fatigue. *The Trend in Engineering*, 13, 9-14.
- Weibull, W. (1959). *Statistical evaluation of data from fatigue and creep rupture tests, Part I: fundamental concepts and general methods*. Sweden: Wright Air Development Center.
- Weibull, W. (1961). *Fatigue testing an analysis of results*. London: Pergamon Press.

Chapter 5: Analysis of Accelerated Degradation Data and Reliability Model Development

5.1. INTRODUCTION

For Prognosis and Health Management (PHM) applications and when dealing with highly reliable components, accelerated life testing alone may not be sufficient to obtain reliability and prognosis results in the available time frame. Although there may be an option to infer reliability characteristics based on censored test samples, this requires assumptions about the distribution and often leads to high uncertainties in the results. In these situations, the use of degradation modeling and analysis may be more suitable. Degradation analysis involves a probabilistic modeling of a failure mechanism degradation path and comparison of a projected distribution of the cumulative damage to a predefined failure threshold or endurance limit (Wang & Coit, 2007). Figure 5.1 shows an example of a monotonically increasing cumulative degradation model due to a hypothetical failure mechanism. The degradation model shown in Figure 5.1 is probabilistic, since the value of the cumulative degradation on the y-axis is represented by a pdf whose random variable represents the amount of cumulative degradation. That is, at any point along the degradation path, there is a unique distribution of the amount of degradation (e.g. amount of wear, light intensity, and crack length) for a population of similarly degrading components. It is often observed that as time increases, a larger variance in the degradation measure's distribution exists, because the amount of the mean cumulative degradation increases (although it is usually the case that the coefficient of variation remains constant). At any point in time, reliability can be estimated as the probability that the degradation measure remains below the endurance limit that describes the failure threshold value. In the case where the path is decreasing (as is the case for the performance degradation), reliability would be calculated as the probability that the degradation measure is greater than the requirement limit.

To reduce testing time even further, we can perform the tests at elevated levels of stresses and measure the changes that occur with respect to the degradation measure. What results is a series of degradation paths, each corresponding to a particular stress level, which can then be used as a basis for modeling and extrapolation for estimating the reliability metrics or mean remaining life at the use stress level. An example of this is shown in Figure 5.2, where three degradation paths are plotted for three different stress levels. This type of analysis is known as accelerated degradation analysis. In some cases, it is possible to directly measure the degradation over time (e.g. propagation of crack). However there are also some cases where obtaining degradation data may be difficult, and may require destructive or invasive means that in turn may have an effect on the component's performance (e.g. taking apart a motor to measure wear). In all cases it is necessary to establish a predefined level of degradation (endurance or requirement) beyond which failure is considered to occur.

It must be noted that although considering the degradation of components is a more accurate way of predicting life characteristics, this approach also has challenges associated with gathering and evaluating degradation data. For example, it involves complexities and uncertainties to measure the degradation, pinpoint the location of degradation, assess and incorporate measurement errors, and evaluate the detection of degradation or damage areas and assess probability of detection itself. These types of tests are more expensive to perform and more complex to analyze than the traditional case of accelerated life testing. Some of the practical advantages that degradation data can offer to reliability analysis include (Meeker & Escobar, 1998):

- They can provide considerably more reliability information (especially in cases where few or no failures exist) than traditional censored failure-time data.

- Observation of degradation may allow direct modeling of the failure mechanism, thereby providing more credible and precise reliability estimates and a stronger basis for extrapolation and prognosis estimations.

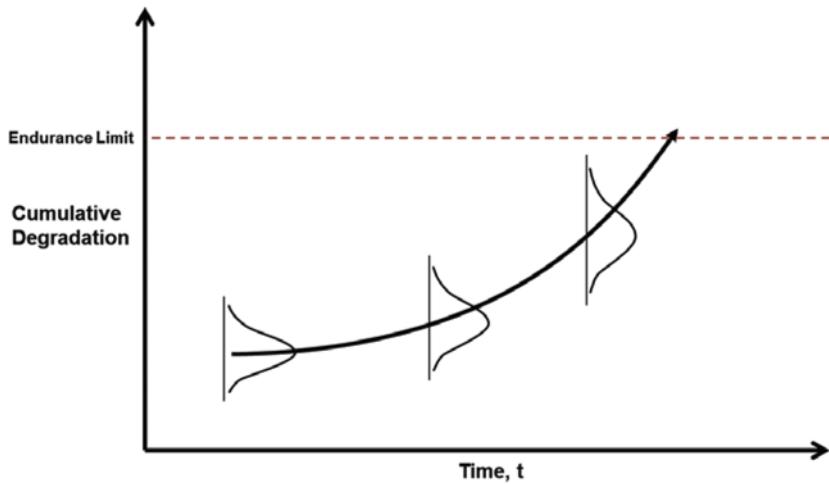


Figure 5.1: Example of a degradation path

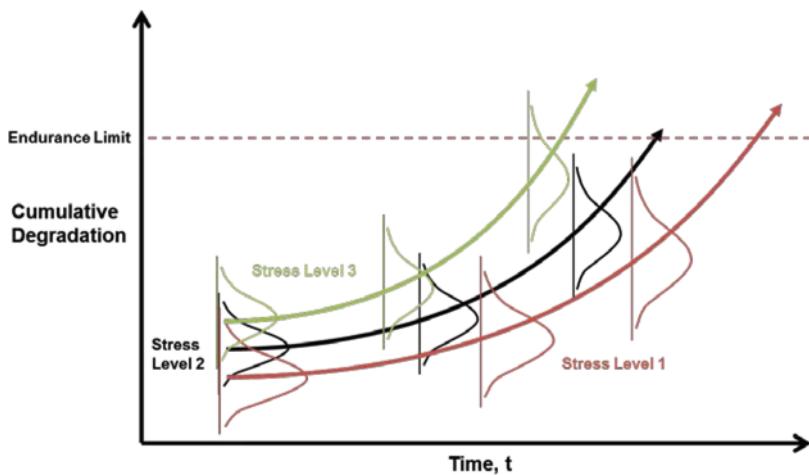


Figure 5.2: Example of degradation paths from an accelerated degradation test for three stress levels

Similar to normal accelerated life testing, accelerated degradation testing and analysis requires more complex machinery and mathematical models to extrapolate the performance measurements over time to the point of the set failure criteria. This is performed at different stress levels, and therefore, each time to failure has a corresponding stress level. The amount of confidence in the results obtained in accelerated degradation tests is directly related to the number of units being tested, the number of units at each stress level, and the acceleration factor (Reliasoft Corporation, 1996). The degradation measures are taken over time, either continuously or at intervals. Once this information has been collected, the next task is to extrapolate the performance measurements to the defined failure criteria, so that degradation at use conditions can be predicted. This section will present some common models, methods of analysis and methods of inference for accelerated degradation tests.

5.2. DEGRADATION MODELS

For accelerated degradation testing, there are two stages of modeling. First, we need to model the relationship between the degradation measure and the time (or cycles) taken to accumulate the amount of degradation. In the second stage, we can make predictions about when the unit under test will reach the predefined threshold level (endurance or requirement limit) to satisfy the failure criteria. These time values can then be used as “pseudo” failure times for input into an ALT model that relates times to failure with stress level. This second modeling activity is exactly the same as the regular ALT modeling and data analysis techniques described in chapter 4. Therefore, this section will focus mainly on the modeling of the relationship between time and a particular cumulative degradation measure. The entire process will then be illustrated via examples.

As mentioned above, failures can generally be traced to an underlying degradation process. Figure 5.3 illustrates conceptual examples of three general shapes for cumulative degradation curves: linear, convex, and concave. These degradation paths can be modeled to a large extent by the same mathematical forms discussed for ALT’s life-stress models conceptually represented by Equation (3.7). The task for reliability engineers and scientists is to find the most appropriate models from these forms or some combination or modified versions that apply to particular degradation applications.

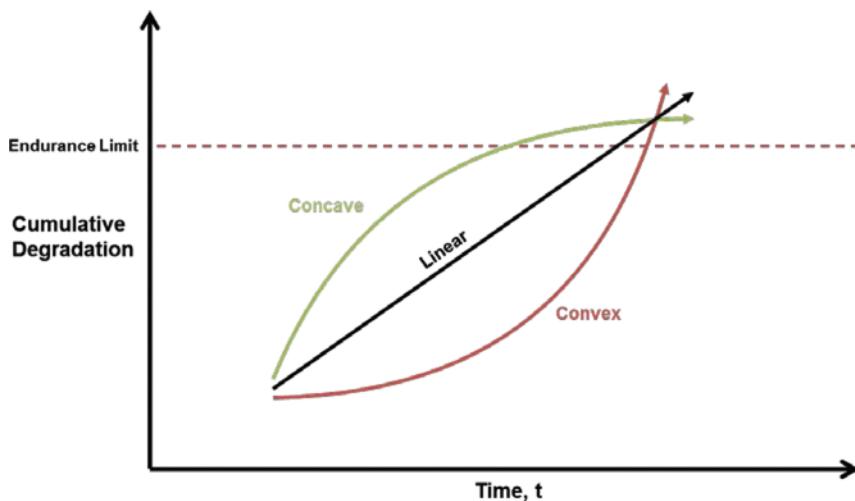


Figure 5.3: Possible shapes for univariate cumulative degradation vs. time

5.2.1. SIMPLE DEGRADATION MODEL WITHOUT VARIATION

Linear Degradation

Linear degradation occurs in several simple processes, such as in wear (e.g. wear of brake pads, wear of tires). For example, if $D(t)$ is the amount of tread wear at time t in an automobile tire, and the wear rate is $\frac{dD(t)}{dt} = C$, then:

$$D(t) = D(0) + mt \quad (5.1)$$

where $D(0)$ is the initial amount of tread wear and m is a parameter describing the rate of degradation. This simple model is graphically presented in Figure 5.4.

Convex Degradation

Convex degradation applies to those failure mechanisms whose rate of degradation is increasing with the level of degradation. Such models are used, for example, in modeling the growth of fatigue cracks. If we let $a(t)$ denote the size of a crack at cycle N , then Paris' model described in Chapter 2 leads to a convex degradation curve similar to that in Figure 2.27 representing fatigue crack size.

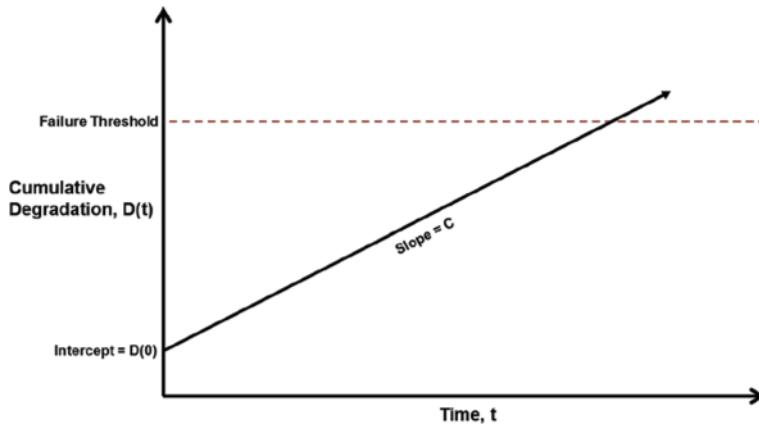


Figure 5.4: A linear degradation model

Concave Degradation

An example of concave degradation is presented in (Meeker & Escobar, 1998), involving a model for growth of failure-causing conducting filaments of chlorine-copper compounds in printed circuit boards. Failure occurs when the filaments reach from one copper-plated filament through-hole to another. The model denotes $A_1(t)$ as the amount of chlorine available for reaction and $A_2(t)$ as proportional to the amount of failure-causing chlorine-copper compounds at time t . Under appropriate conditions, copper combines with chlorine A_1 to form the chlorine-copper compound A_2 with the rate constant, k_1 . The chemical reaction is thus given diagrammatically by $A_1 \xrightarrow{k_1} A_2$. The rate equations for this process are $\frac{dA_1}{dt} = -k_1 A_1$; and $\frac{dA_2}{dt} = k_1 A_2$. The solution to this system of differential equations gives $A_1(t) = A_1(0) e^{-k_1 t}$ and $A_2(t) = A_2(0) + A_1(0)[1 - e^{-k_1 t}]$, where $A_1(0)$ and $A_2(0)$ are the initial values. Letting $A_2(\infty) = A_1(0) + A_2(0)$, and if $A_2(0) = 0$, then the solution for $A_2(t)$ can be expressed as $A_2(t) = A_2(\infty)[1 - e^{-k_1 t}]$, which represents a concave curve.

5.2.2. CONSIDERATION OF THE VARIATION IN DEGRADATION MODEL AND FAILURE TIME

The simple models discussed in 5.2.1 do not take into consideration variability that may influence test results. In reality, there are factors that do cause variability in the degradation curves, as well as the extrapolated failure times. The two main sources of variability are unit-to-unit variability and variability due to the operating and environmental conditions.

Some examples of unit-to-unit variability include:

- **Initial conditions.** Individual units will inevitably vary with respect to the amount of initial degradation and flaws that exist (e.g. wear of brake pad, tire, embedded flaws and initial cracks and dents). This variability in turn results in the variability of time to reach the predetermined failure threshold.

- **Material properties.** Constants relating to material properties will vary from unit to unit, and this again will result in the variability of the estimated time to reach the predetermined failure threshold.
- **Component geometry and dimensions.** The variability between units in terms of exact dimension, weight and geometry will result in the variability in degradation rates.
- **Within-unit variability.** It is common for spatial variability (e.g. varied defect densities and sizes) in different locations to exist within a unit.

Variability due to operating and environmental conditions can also result in variation in the degradation rate between units under test. For example, any variations in temperature and stress will result in variability in the failure causing process, in fatigue applications.

5.2.3. GENERAL DEGRADATION PATH MODEL

A general degradation path model (Meeker & Escobar, 1998) was proposed for the analysis of degradation data (assuming no measurement errors and detection probability) at a fixed level of stress, to determine an estimate of the time-to-failure distribution. In this model, $D(t)$ denotes the true degradation path of a particular unit. Values of $D(t)$ can be monitored continuously, but in practical applications, they are often sampled at discrete points in time. The observed sample degradation path for some unit i at sample time t_{ij} is a unit's actual degradation path $D(t)$ plus a model error, as given by

$$y_{ij} = D_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, n; \quad j = 1, \dots, m_i \quad (5.2)$$

where $D_{ij} = D(t_{ij} | \Theta)$ is the degradation path given the vector of parameters $\Theta = \{\theta_1, \dots, \theta_k\}$ and t is the actual time (operation time), cycles, expended life or an index of age. The actual degradation path of unit i at sampling number j is t_{ij} . $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon)$ and is a residual deviation (model error) for unit i at t_{ij} (corresponding to the sampling number j).

The total number of inspections to measure the cumulative damage on unit i is denoted by m_i . Note that time t could be represented as real time, operating time, or some other appropriate quantitative measure such as miles for automobile tires or the number of loading cycles for fatigue applications. Figure 5.5 illustrates the concept described by Equation (5.2). Note that this diagram only illustrates the case for degradation path of unit i . Multiple curves will exist to represent each unit under test. Figure 5.5 shows data points (cumulative degradation or damage) for unit i , at each measurement time t_{ij} . In reality, there will only be one measurement of the degradation variable at each time, even though for each unit, the measurement time need not be the same. For instance, the first measurement time for unit 1, t_{11} , does not necessarily need to be the same as the first measurement time for unit 2, t_{21} .

The scales of y and t can be chosen to simplify the form of $D(t_{ij} | \theta_{1i}, \dots, \theta_{ki})$. The choice of a degradation model requires not only determination of the mathematical form of $D(t_{ij} | \theta_{1i}, \dots, \theta_{ki})$, but also estimation of the parameters in $\theta_{1i}, \dots, \theta_{ki}$. Especially since the elements of the vector Θ may be correlated, the parameters' covariance or the joint distribution of the parameters of this vector (in a Bayesian estimation) should be estimated. Meeker and Escobar, in Statistical Methods for Reliability Data (1998), describe the use of a general family of transformations to a multivariate normal distribution with mean vector μ_θ and covariance matrix Σ_θ .

It is generally reasonable to assume that the parameters $\theta_{1i}, \dots, \theta_{ki}$ of the cumulative degradation vs. time model parameters are random and independent of the measurement error ε_{ij} . It is also possible to assume that ε_{ij} are independently and identically distributed (iid). Since each degradation observation y_{ij} is taken sequentially, there is potential for autocorrelation between the ε_{ij} 's, especially for closely spaced readings. However, in many practical applications involving the modeling of degradation of units from a population or process, provided that the model fit is adequate and measurement processes are in control, this autocorrelation is typically weak. Moreover, variability is dominated by unit-to-unit variability in the θ values, and point estimates of regression models are not seriously affected by autocorrelation. Although in some cases ignoring autocorrelation can result in standard errors that are seriously biased, this is not as much of a problem when confidence intervals are employed.

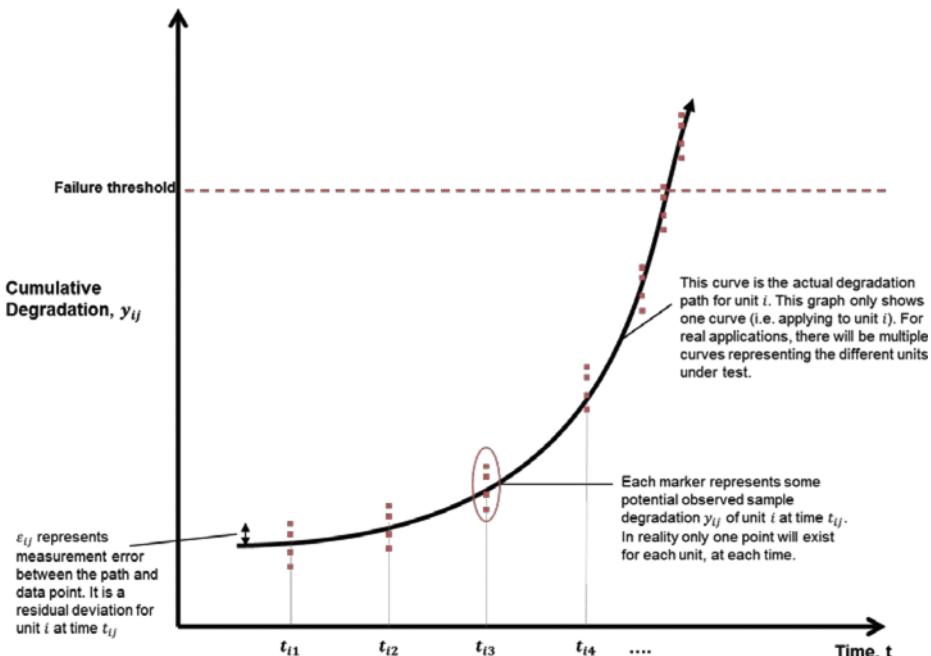


Figure 5.5: General degradation path model for unit i

5.2.4. APPROXIMATE ACCELERATED LIFE DEGRADATION ANALYSIS

A simple (but only approximately correct) method of analyzing degradation data consists of a two-step process. The first step involves a separate analysis for each unit under test to predict the time at which the unit will reach the predetermined critical degradation level (endurance) that corresponds to a failure. These times can be referred to as “pseudo” failure times. The second step involves analyzing these pseudo failure times as a complete sample of failure times to estimate the cdf $F(t)$ representing the time to failure. The method can be formally summarized as follows:

1. For each unit i use the path model $D(t, \Theta_i)$ to establish $y_{ij} = D_{ij} + \varepsilon_{ij}$, where the observed sample path data $(t_{i1}, y_{i1}), \dots, (t_{im}, y_{im})$ should be used to find the maximum likelihood estimate of the vector of unknown degradation model parameters Θ_i . This can be done using least squares regression.
2. Define the degradation limit (endurance limit) D_f and solve equation $D(t, \Theta_i) = D_f$ for t and

denote this solution as \hat{t}_i . Note that D_f itself may be treated as a random variable and the corresponding limit represented by a pdf instead of a fixed value.

3. Repeat the procedure for each sample path to obtain the pseudo failure times $\hat{t}_1, \dots, \hat{t}_n$.
4. Based on the scatter of $\hat{t}_1, \dots, \hat{t}_n$, establish a time to failure distribution relating to the failure times, and a life-stress relationship to relate the pseudo failure times with the stress that is the cause of degradation.

It must be noted that this approximate method can give adequate analysis when the degradation paths are relatively simple (e.g. clearly linear in nature). The fitted form of the path model should properly represent the accumulation of damage. The degradation assessment relies on a large amount of data to estimate the model parameters because several data points at multiple locations along the degradation would be needed to establish the assumed model with reasonable confidence. When established, the resulting amount of measurement error should be small. The model should not be used to extrapolate far beyond the observed degradation data to predict the pseudo failure times.

Clearly, this approximate approach ignores the prediction error in \hat{t}_i and does not take into account the model error ε_{ij} in the observed sample paths. Also, the distributions fitted to the pseudo failure times will not necessarily correspond to the distribution induced by the degradation model. Finally, in some applications, there may be sample paths where not enough information exists to estimate all of the path parameters (for instance, if the path model has an asymptote but the data have not yet reached this asymptote). This might require fitting different models for different sample paths in order to predict the \hat{t}_i 's.

5.2.5. MAXIMUM LIKELIHOOD APPROACH TO ESTIMATING ACCELERATION DEGRADATION MODEL PARAMETERS

For now we assume ε_{ij} autocorrelation is weak and can be ignored. Further we assume that an additive error model described by the normal distribution can represent the error. In this case the likelihood of individual data points y_{ij} and t_{ij} can be estimated using this normal distribution and Equation (5.2). If a correlation exists then we could, for example, rely on time series approaches to model ε_{ij} . Assume that the vector Θ may be represented by the mean values μ_Θ , and the covariance matrix Σ_Θ . Then the likelihood of the data, given the parameters of the model, can be written as

$$\begin{aligned} l(y_{ij}, t_{ij} | \mu_\Theta, \Sigma_\Theta, \sigma_\epsilon) = \\ \prod_{i=1}^n \int_{\Theta} \prod_{j=1}^{m_i} \frac{\phi(z_{ij})}{\sigma_\epsilon} f(\Theta | \mu_\Theta, \Sigma_\Theta) d\Theta \end{aligned} \quad (5.3)$$

where $z_{ij} = \frac{y_{ij} - D(t_{ij}, \Theta)}{\sigma_\epsilon}$, and m_i = number of sampling (i.e., degradation inspections) associated with unit i, and $f(\Theta | \mu_\Theta, \Sigma_\Theta)$ = joint pdf of the vector of the parameters $\Theta = \{\theta_1, \dots, \theta_k\}$ having the mean μ_Θ and the covariance matrix Σ_Θ . Subsequently, by maximizing $l(y_{ij}, t_{ij} | \mu_\Theta, \Sigma_\Theta, \sigma_\epsilon)$ with respect to $\mu_\Theta, \Sigma_\Theta, \sigma_\epsilon$ (i.e., use $\frac{\partial l}{\partial \mu_\Theta} = 0, \frac{\partial l}{\partial \Sigma_\Theta} = 0, \frac{\partial l}{\partial \sigma_\epsilon} = 0$) the MLE estimates of the ADT parameters are found. This will be very difficult, unless $D(\cdot)$ is a linear or log-linear function.

A simpler version (5.3) can be written without the pdf model $f(\Theta | \mu_\Theta, \Sigma_\Theta)$ as

$$l(y_{ij}, t_{ij} | \mu_\Theta, \Sigma_\Theta, \sigma_\epsilon) = \prod_{i=1}^n \int_{\Theta} \prod_{j=1}^{m_i} \frac{\phi(z_{ij})}{\sigma_\epsilon} d\Theta \quad (5.4)$$

where $z_{ij} = \frac{y_{ij} - D(t_{ij}, \Theta)}{\sigma_\epsilon}$ and, for example, for the special case of a power degradation model, $D(t_{ij}) = \theta_1 t_{ij}^{\theta_2}$, $z_{ij} = \frac{(y_{ij} - \theta_1 t_{ij}^{\theta_2})}{\sigma_\epsilon}$, and $\Sigma_\Theta = \begin{bmatrix} \text{Var}(\theta_1) & \text{Cov}(\theta_1, \theta_2) \\ \text{Cov}(\theta_2, \theta_1) & \text{Var}(\theta_2) \end{bmatrix}$. As such, $\frac{\partial l}{\partial \Theta} = \frac{\partial l}{\partial \Sigma_\Theta} = \frac{\partial l}{\partial \sigma_\epsilon} = 0$ would yield $\hat{\Theta}$, $\hat{\Sigma}_\Theta$, and $\hat{\sigma}_\epsilon$.

If there are experimental and/or detection errors involved, then this should be formally accounted for. The MLE approach allows for this by incorporating such errors into characterizing the degradation data by using the linear additive error models represented by the functions $\epsilon_{ij}^e = y_{ij} - y_{ij}^e$ and $\epsilon_{ij}^m = y_{ij} - y_{ij}^m$, where ϵ^e is the measurement error (at the time associated with time t_{ij}). Note that y_{ij} is the true degradation and y_{ij}^e is the measured degradation, ϵ^m is the degradation model error, and y_{ij}^m is the degradation model values associated with time t_{ij} . Accordingly,

$$\epsilon = \epsilon^m - \epsilon^e = y_{ij}^e - y_{ij}^m = y_{ij}^e - D(t_{ij}, \Theta) \quad (5.5)$$

Formal accounting of measurement, model and detection uncertainties in the context of degradation analysis will be further discussed in Chapter 7.

If the normal pdf error model $\epsilon_{ij} \sim N(0, \sigma_{\epsilon_{ij}})$ is no longer valid, then lognormal may also be possible, in which case the experimental and model error would change from the additive form to multiplicative form. For example, $\epsilon_{ij}^e = y_{ij}/y_{ij}^e$, and $\epsilon_{ij}^m = y_{ij}/y_{ij}^m$ and will be $\epsilon_t = y_{ij}^e/D(t_{ij}, \Theta)$. Accordingly, (5.3) can be written as

$$\begin{aligned} l(y_{ij}, t_{ij} | \mu_\Theta, \Sigma_\Theta, \sigma_{\epsilon_t}) \\ = \prod_{i=1}^n \int_{\Theta} \prod_{j=1}^{m_i} \frac{\phi(z_{ij,t})}{t_{ij,t} \sigma_{\epsilon_t}} f(\Theta | \mu_\Theta, \Sigma_\Theta) d\Theta \end{aligned} \quad (5.6)$$

where $z_{ij,t} = \frac{\ln(y_{ij}) - \ln(D(t_{ij}, \Theta))}{\sigma_{\epsilon_t}}$.

$$\begin{aligned} l(y_{e,ij}, t_{ij}, b_e, \sigma_e | \sigma_m, b_m, y_m) \\ = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \left(\frac{y_{e,ij}}{y_{m,ij}} \right) \sqrt{\sigma_m^2 + \sigma_e^2}} e^{-\frac{1}{2} \times \frac{\left[\ln \left(\frac{y_{e,ij}}{y_{m,ij}} \right) - (b_m - b_e) \right]^2}{\sigma_m^2 + \sigma_e^2}} \end{aligned} \quad (5.7)$$

Table 5.1: Times-to-failure corresponding to the recorded weight losses

Weight Lost (mg)	Depth at 240 hours (cm)	Rate of loss (cm/hour)	TTF (hours)
11.1	3.40×10^{-4}	1.42×10^{-6}	70659
10.4	3.18×10^{-4}	1.33×10^{-6}	75415
12.1	3.70×10^{-4}	1.54×10^{-6}	64820
11.4	3.49×10^{-4}	1.45×10^{-6}	68800
9.8	3.00×10^{-4}	1.25×10^{-6}	80033

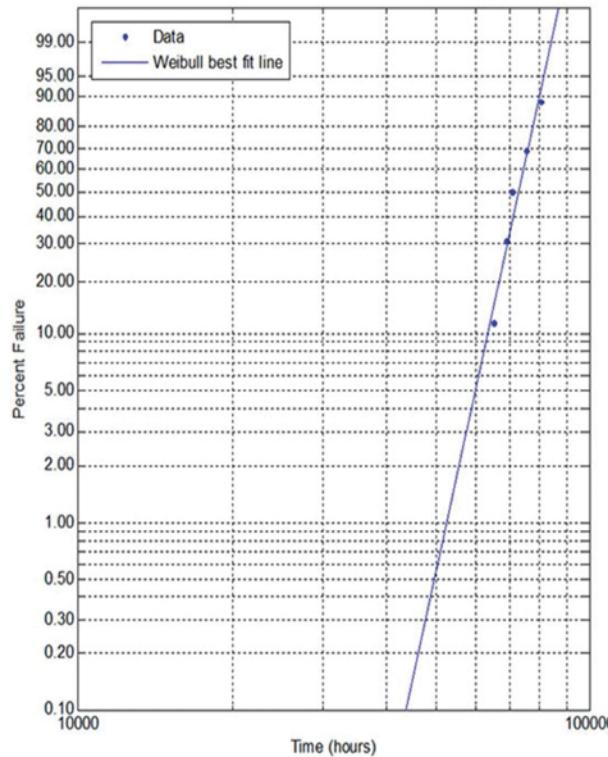


Figure 5.6: Weibull probability plot for calculated failure times

Example 5.1

Five specimens of a new corrosion-resistant material are tested for 240 hours in a highly corrosive environment. The density of the material is 7.6 g/cm³, and the exposed surface area of each specimen is 4.3 cm². At the end of the test period, the measured weight losses in mg were 11.1, 10.4, 12.1, 11.4, and 9.8 (assume no measurement error). If a degradation of 1 mm or more results in a structural failure, predict the failure times for the five specimens. Based on these failure times, determine the probability distribution that best represents the life of the material.

Solution

From the given data, we need to determine the time it takes for 1 mm (depth) of material loss to occur. Manipulating units, we can obtain from the recorded weight loss the volume of material that was lost as follows:

$$\text{Volume lost} = \frac{\text{Mass lost}}{\text{Density}}$$

$$\text{Depth of material loss} = \frac{\text{Volume lost}}{\text{Surface Area}}$$

The time it takes to achieve a depth of 1 mm of degradation is then determined through the rate of material loss given by:

$$\text{Rate of loss (in depth)} = \frac{\text{Depth}}{\text{total time of test}} = \frac{\text{Depth}}{240 \text{ hrs}}$$

From the above, the failure time is determined from:

$$\text{Failure time} = \frac{\text{Failure depth (in cm)}}{\text{Rate of loss}} = \frac{0.1 \text{ cm}}{\text{Rate of loss (cm/hr)}}$$

Accordingly, we obtain the following table outlining the times-to-failure corresponding to the recorded weight losses. Using the data from

Table 5.1, we can fit the times to either the Weibull or lognormal distribution to see which pdf provides the better fit. A MATLAB evaluation produced the plots based on the distribution parameters for the Weibull fit ($\alpha = 74577$ and $\beta = 13.01$) and the lognormal fit ($\mu_t = 11.18$ and $\sigma_t = 0.093$). It was determined that the Weibull distribution provides a better fit than the lognormal distribution. The plot is shown in Figure 5.6 and Figure 5.7.

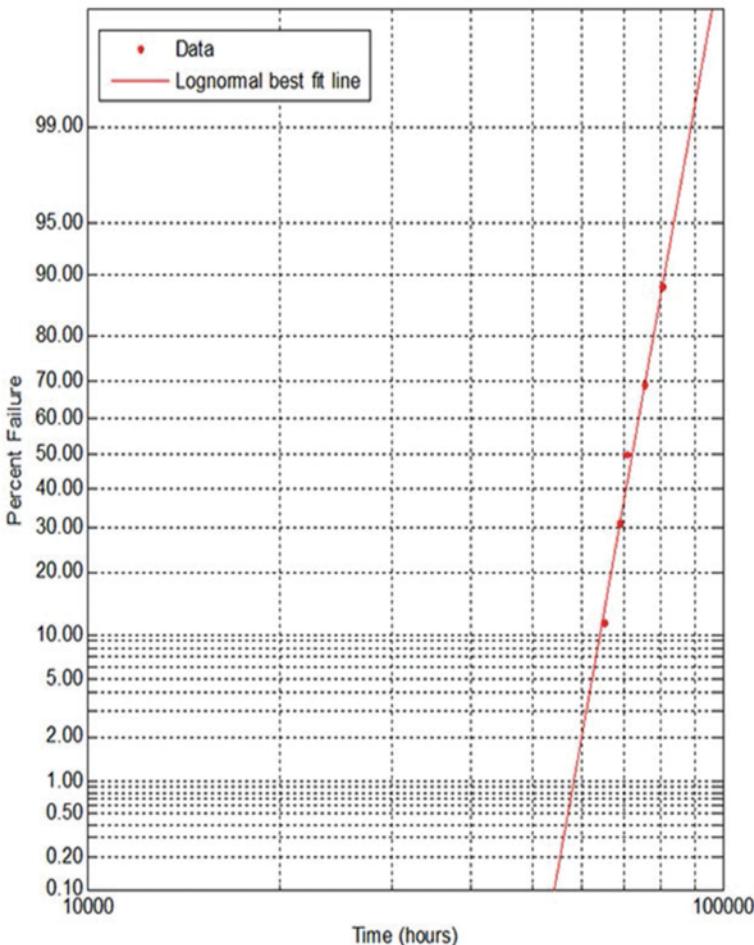


Figure 5.7: Lognormal probability plot for calculated failure times

Example 5.2

Consider the following example involving a test of a particular metal alloy. The sliding test was conducted over a range of different applied weights in order to study the effect of weight and to gain a better understanding of the wear mechanism. The accelerated degradation test data is presented in Table 5.2 and plotted in log-log axes (Figure 5.8). Estimate the parameters of power degradation model using the MLE method, assuming no measurement errors. Repeat the problem assuming 10% multiplicative measurement error.

Table 5.2: Accelerated degradation test data

Weight (grams)	Unit	Cycles (hundreds)							
		2	5	10	20	50	100	200	500
10	1	3.2	4.1	4.5	4.7	5.8	6.8	7.7	9.6
	2	2.7	3.4	3.8	3.9	5.4	5.7	6.3	8.4
	3	2.1	2.7	3.1	3.3	4.0	4.6	5.7	6.6
	4	2.6	3.5	4.0	4.0	5.2	6.1	6.7	8.5
50	5	7.5	7.8	8.2	10.6	12.6	13.3	12.9	14.8
	6	7.5	8.1	9.8	10.9	14.8	16.1	17.3	20.2
	7	7.0	8.9	9.4	11.1	12.4	13.5	16.7	17.3
	8	7.8	8.9	10.0	11.5	13.7	16.2	16.2	21.0
100	9	12.5	15.4	17.2	20.5	24.1	27.0	29.4	37.9
	10	11.0	13.9	16.1	18.6	22.2	27.8	31.0	36.6
	11	13.0	15.1	18.6	20.2	23.9	29.7	31.5	39.6
	12	11.7	13.7	16.7	17.5	22.3	25.3	32.0	38.2

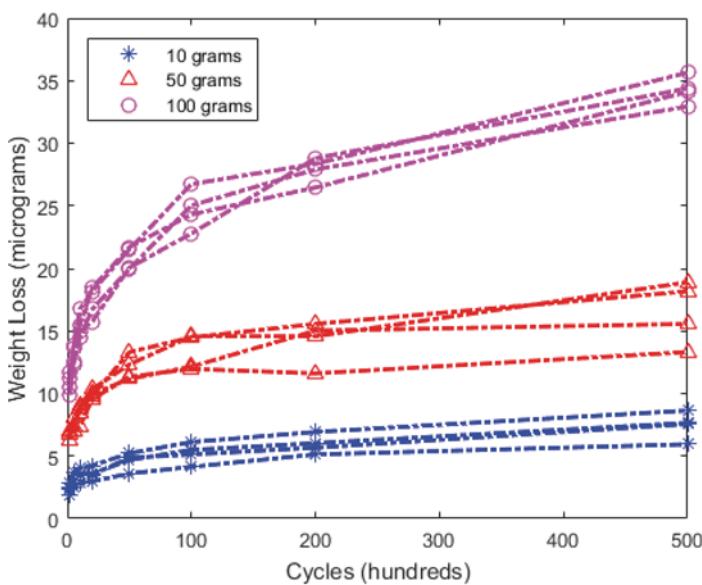


Figure 5.8: Plot of weight loss in micrograms vs. hundreds of cycles

Solution

If we are to assume a simple linear power relationship between the degradation measure (width of scar in microns) to the number of cycles required to accumulate the scar width, then $D(t_{ij}) = \theta_1 t_{ij}^{\theta_2}$ or the log-linear version $\log[D(t_{ij})] = \theta_1 + \theta_2 \cdot \log(t_{ij})$. Now, by using the MLE in Equation (5.3) for the data shown and the degradation model log plot and extrapolating to the time at which the scar width would reach 50 microns, we can obtain estimates of the mean model parameters $\hat{\mu}_{\theta_1}$, $\hat{\mu}_2$, $\hat{\Sigma}_{\theta_1 \theta_2} = \begin{bmatrix} \text{Var}(\theta_1) & \text{Cov}(\theta_1, \theta_2) \\ \text{Cov}(\theta_2, \theta_1) & \text{Var}(\theta_2) \end{bmatrix}$ or $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\Sigma}_{\theta_1 \theta_2}$.

Table 5.3: Pseudo failure times corresponding to the recorded weight

Weight (grams)	Pseudo Failure Times			
10	3634068	6259904	11322469	3604319
50	2563366	44330	237427	75403
100	2388	1815	1691	1930

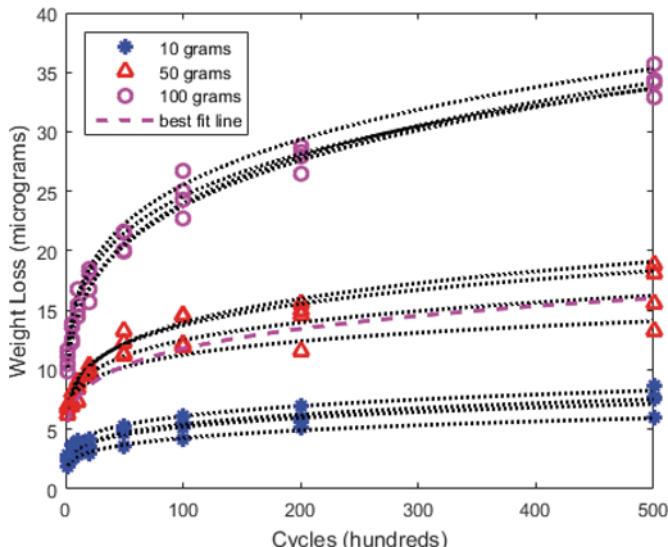


Figure 5.9: Reconstruction of trendline data and best fit line using MATLAB and MLE

Accordingly, using Equation (5.3) or (5.4) and using a MATLAB routine the parameters are estimated as

$$\hat{\mu} = \begin{bmatrix} 1.69 \\ 0.19 \end{bmatrix} \text{ and } \hat{\Sigma}_{\theta_1 \theta_2} = \begin{bmatrix} 0.43 & 2.32 \times 10^{-4} \\ 2.32 \times 10^{-4} & 5.08 \times 10^{-4} \end{bmatrix}.$$

Applying Equations (5.5)-(5.7) and using the same MATLAB routine, the estimated parameters are

$$\hat{\mu} = \begin{bmatrix} 1.58 \\ 0.19 \end{bmatrix} \text{ and } \hat{\Sigma}_{\theta_1 \theta_2} = \begin{bmatrix} 0.43 & 2.83 \times 10^{-4} \\ 2.83 \times 10^{-4} & 5.52 \times 10^{-4} \end{bmatrix}.$$

Example 5.3

Consider the following example of an accelerated degradation test (performance degradation) to model the increasing resistance over time of carbon-film resistors tested at three different levels of temperature (data provided in Table 5.4) as reported by Meeker and Escobar (1998). According to the requirements of the project, failure is considered to have occurred when the resistance has increased by 5%. Use the approximate method of analysis described previously to analyze these data.

Table 5.4: Resistance over time of carbon-film resistors tested at three different levels of temperature

Unit Number	Temperature (°C)	Initial Resistance	Hours			
			452	1030	4341	8084
1	83	217.97	0.28	0.32	0.38	0.62
2	83	217.88	0.22	0.24	0.26	0.38
3	83	224.67	0.41	0.46	0.54	0.81
4	83	215.92	0.25	0.29	0.32	0.48
5	83	219.88	0.25	0.26	0.42	0.57
6	83	219.63	0.32	0.36	0.45	0.58
7	83	218.27	0.36	0.41	0.52	0.7
8	83	217.27	0.24	0.28	0.34	0.55
9	83	219.98	0.33	0.4	0.44	0.85
10	133	218.05	0.4	0.47	0.72	1.05
11	133	219.38	0.88	1.19	2.06	3.15
12	133	218.35	0.53	0.64	0.99	1.6
13	133	217.78	0.47	0.62	1	1.5
14	133	218.28	0.57	0.75	1.26	2.03
15	133	216.38	0.55	0.67	1.09	1.79
16	133	217.65	0.78	0.96	1.48	2.27
17	133	221.91	0.83	1.12	1.96	3.29
18	133	218.47	0.64	0.8	1.23	1.84

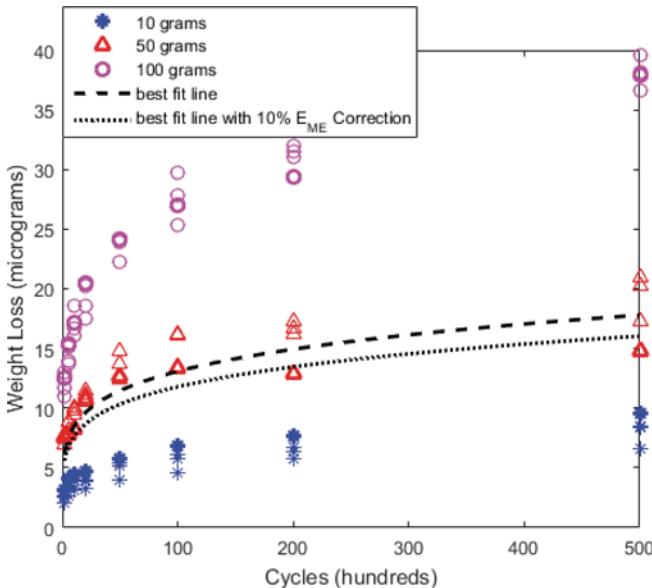


Figure 5.10: The original data along with the best fit lines without and with assumption of 10% measurement error correction

Solution

The plots of the degradation versus time for each of the sample paths using linear or semi-log models do not lead to acceptable models. However, if we use an appropriate transformation, which results in a square root-linear plot (i.e. $\sqrt{y} = mx + c$), it results in an approximately linear plot, a more appropriate model for extrapolation purposes. This square root to linear plot is given in the plot in Figure 5.11.

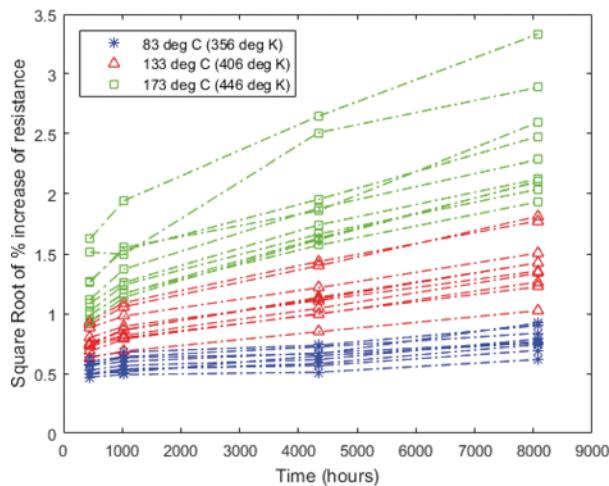


Figure 5.11: Square root to linear plot

As can be seen from the plot in Figure 5.11, there is a clear statistical bias of an upward trend between the first and second data points for each degradation path. After this, the plot is approximately linear, except for one path at the 173 °C group, which clearly has a poor linear fit. We can reconstruct the

above plot with trend lines for each degradation path (using MS Excel) in Figure 5.12 using only the model for times greater than 1000 hours.

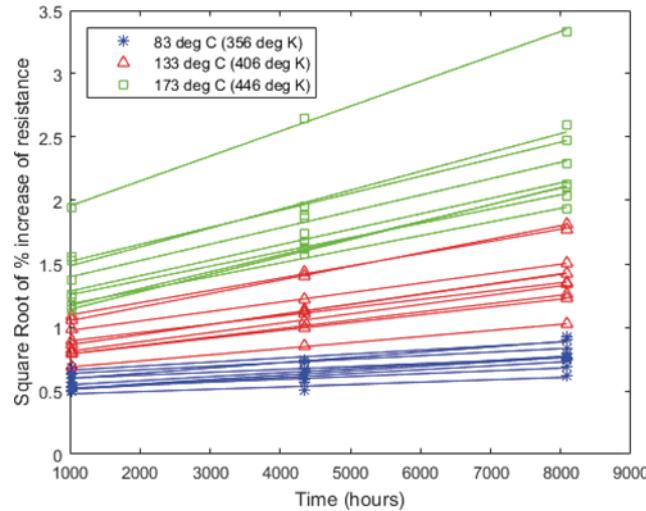


Figure 5.12: Reconstruction of Figure 5.11 with trend lines for each degradation path (using Matlab)

Table 5.5: Pseudo failure times corresponding to the recorded weight

Temperature (°C)	Pseudo Failure Times		
	83	133	173
	40425	12061	2422
	50669	12772	6043
	50682	17886	6334
	54252	18187	7496
	57738	18345	8796
	58368	20343	8960
	72443	21521	9033
	78439	22899	9689
	97914	24370	10758
	33268		

The trend lines can be assessed using the least squares regression technique, from which we can derive pseudo times to failure for further data analysis. Note that this set of pseudo failure times excludes the degradation path that shows a poor linear fit at the 173 °C level. The pseudo times for each path are listed in Table 5.5

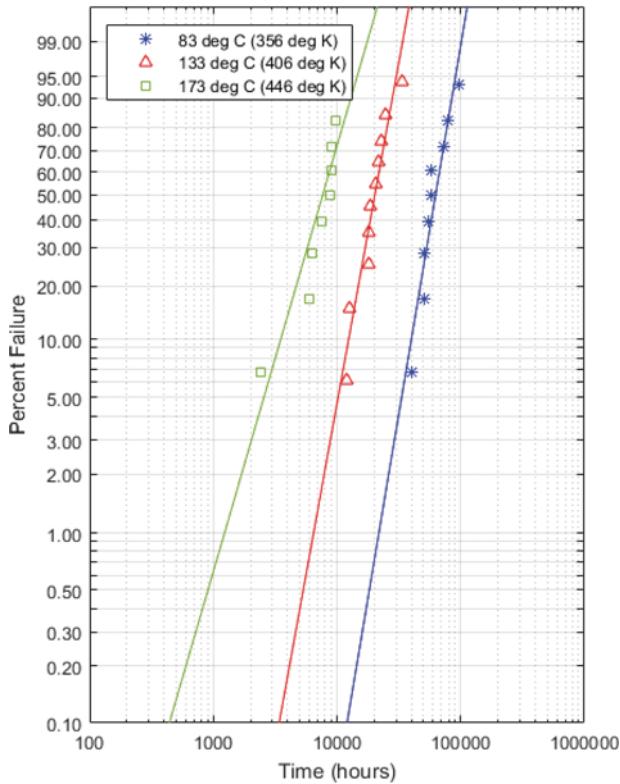
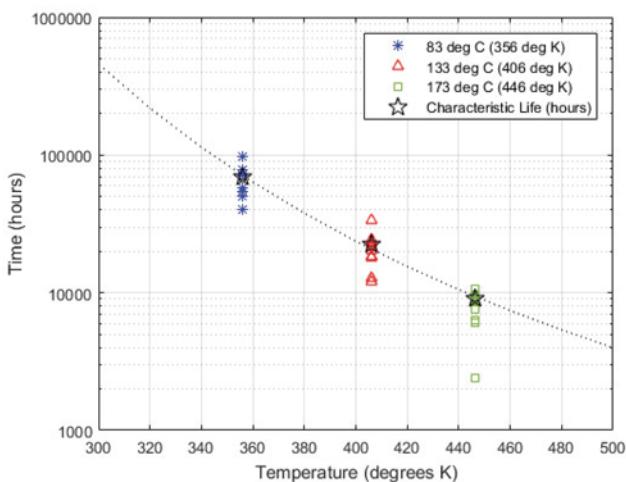


Figure 5.13: Weibull distribution fit for the pseudotimes given in Table 5.5

From this we can use the Arrhenius relationship to determine reliability characteristics at the use level temperature of 50 °C (323 °K). In this case, a Weibull distribution provides a reasonable life distribution fit for the data. The Weibull multi plot is shown in Figure 5.13.



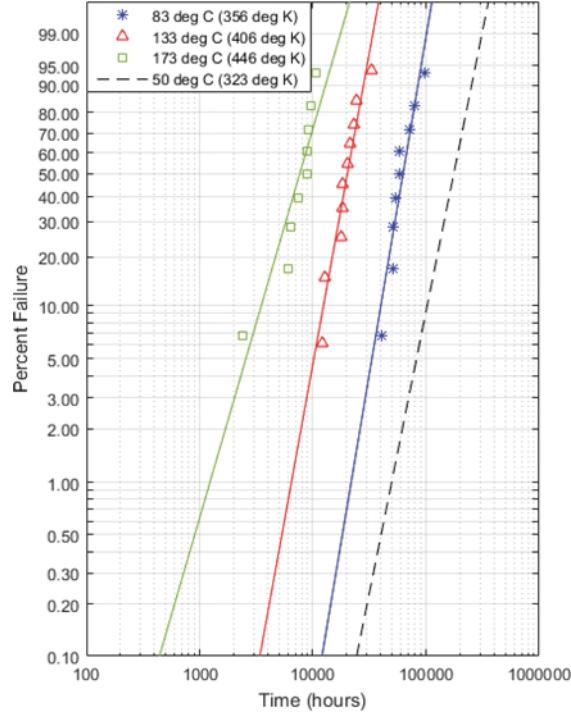


Figure 5.14: Life-stress plot (top), and estimated life distribution at the use temperature of 50 °C (323 °K) (bottom)

Further, a life-stress plot (Figure 5.14, top), and estimated life distribution at the use temperature of 50 °C (323 °K) (Figure 5.14, bottom) can be developed. From this a wide range of other reliability characteristics can be gathered. Furthermore, confidence intervals can be determined to allow for a more flexible estimation of metrics relating to the reliability of these carbon film resistors.

5.2.6. BAYESIAN ESTIMATION OF ADT MODEL PARAMETERS

Estimation of the ADT parameters through a Bayesian approach parallels that of the MLE approach using the normal and lognormal likelihood function described by Equations (5.4) and (5.6). The only difference is that prior values of the parameters also enter the estimation process. The following expression represents the Bayesian inference for the ADT model parameters assuming the normal distribution for the likelihood function:

$$\pi(\boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta}, \sigma_{\varepsilon} | y_{ij}, t_{ij}) = \frac{\left[\prod_{i=1}^n \int_{\boldsymbol{\Theta}_{\beta}} \prod_{j=1}^{m_i} \frac{\Phi_{NOR}(z_{ij})}{\sigma_{\varepsilon}} f(\boldsymbol{\Theta}_{\beta}) d\boldsymbol{\Theta}_{\beta} \right] \pi_0(\boldsymbol{\mu}_{\theta_{\beta}}, \boldsymbol{\Sigma}_{\theta_{\beta}}, \sigma_{\varepsilon})}{\iiint_{\boldsymbol{\mu}_{\theta_{\beta}}, \boldsymbol{\Sigma}_{\theta_{\beta}}, \sigma_{\varepsilon}} \left[\prod_{i=1}^n \int_{\boldsymbol{\Theta}_{\beta}} \prod_{j=1}^{m_i} \frac{\Phi_{NOR}(z_{ij})}{\sigma_{\varepsilon}} f(\boldsymbol{\Theta}_{\beta}) d\boldsymbol{\Theta}_{\beta} \right] \pi_0(\boldsymbol{\mu}_{\theta_{\beta}}, \boldsymbol{\Sigma}_{\theta_{\beta}}, \sigma_{\varepsilon}) d\boldsymbol{\mu}_{\theta_{\beta}} d\boldsymbol{\Sigma}_{\theta_{\beta}} d\sigma_{\varepsilon}} \quad (5.8)$$

Example 5.4

Consider a performance degradation test of LED lights, where temperature and forward current are the two major factors causing output degradation as reported in Pan and Crispin (2010). In the ADT, assume the case of controlled ambient temperature and forward current, and where light output is measured at several time points. The normal use condition of these LEDs is at 30 °C and 400 mA. The test conditions are shown in Table 5.6. The minimum requirement for LED lights is at 70% of its initial value (i.e., the light may be considered failed when falling below this level).

Table 5.6: Three condition values as well as their corresponding temperature, forward current, and number of LEDs

Condition	T (°C)	I (mA)	No. of LEDs
1	77	700	8
2	77	500	14
3	45	700	14

The following nonlinear model for true performance degradation of luminosity of LEDs is proposed in Mitsuom F. (1991) as

$$D_{ij} = \left(1 + \gamma_0(T, I)t_{ij}^{\gamma_1}\right)^{-1}, i = 1, \dots, 36 \text{ and } j = 1, \dots, 4 \quad (5.9)$$

where i = number of test units, j = number of measurement, γ_0 is the degradation rate as a function of testing temperature and forward current, and D_{ij} is the performance level that degrades from a normalized level from 1 to 0 over time.

Perform the MLE of the parameters of the performance degradation model. Also, estimate the 10%, 50% and 90% life estimate at the use condition using this model assuming the degradation test results summarized in Table 5.7.

Solution

The plot of performance degradation of the data gathered and listed in Table 5.7 is plotted in Figure 5.15. Taking the log of both sides of Equation (5.9) produces log-linear trendlines for the data, but this

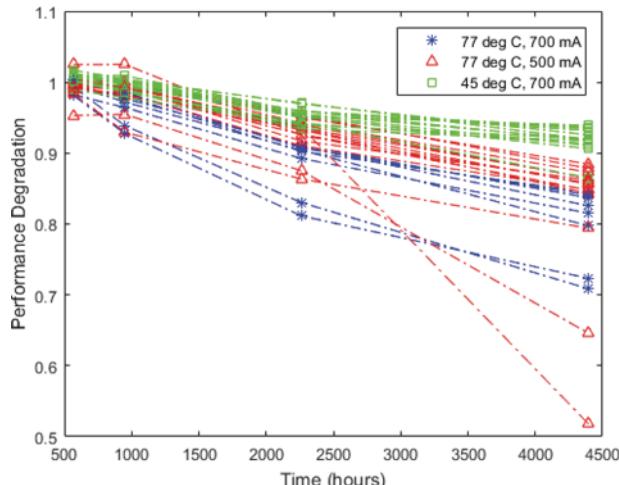


Figure 5.15: Performance degradation data from Table 5.7.

leaves the values of the parameters γ_0 and γ_1 undefined. A combination of MATLAB and MS Excel routines were used to find the values of these parameters that produced the following trendlines in Figure 5.16.

Using the parameter estimates for each set, the failure times are obtained by rearranging Equation (5.9) to solve for the failure time TTF ,

$$TTF_i = \left(\frac{D^{-1} - 1}{\gamma_{0,i}} \right)^{\frac{1}{\gamma_{1,i}}}$$

and by setting the D value to 0.7, which is when failure is supposed to take place. The resulting failure times are listed in Table 5.8

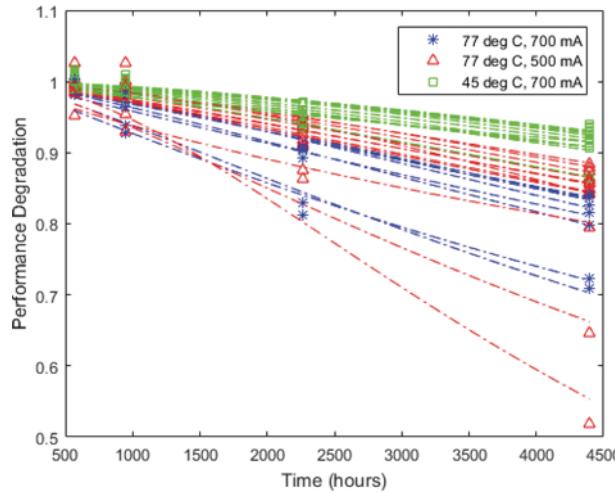


Figure 5.16: Trendlines of performance degradation obtained through MATLAB and Excel.

Table 5.7: Temperature and current at four given time periods for 36 units

Unit	T (°C)	I (mA)	569 (hrs)	950 (hrs)	2261 (hrs)	4397 (hrs)
1	77	700	1.003	0.9848	0.9073	0.8408
2	77	700	0.9943	0.9808	0.9025	0.8442
3	77	700	0.9811	0.9277	0.8115	0.7233
4	77	700	0.9950	0.9722	0.9043	0.7975
5	77	700	0.9930	0.9388	0.8304	0.7085
6	77	700	0.9822	0.9644	0.8925	0.8152
7	77	700	1.0038	0.9860	0.9089	0.8257
8	77	700	0.9971	0.9770	0.9108	0.8368
9	77	500	0.9916	0.9924	0.9175	0.8619
10	77	500	0.9945	0.9929	0.9402	0.8645
11	77	500	0.9859	0.9871	0.9336	0.856
12	77	500	0.9839	0.9297	0.8636	0.7941
13	77	500	0.9917	0.9839	0.9105	0.8484
14	77	500	0.9892	0.9807	0.933	0.8428
15	77	500	1.0014	0.9953	0.9369	0.5177
16	77	500	0.9948	0.9886	0.9264	0.8577
17	77	500	1.0249	1.0252	0.9255	0.8479
18	77	500	0.9966	0.9930	0.9357	0.8654
19	77	500	1.0023	1.0008	0.9427	0.8730
20	77	500	1.0035	0.9996	0.9535	0.8785
21	77	500	0.9982	0.9945	0.9503	0.8845
22	77	500	0.9530	0.9548	0.8753	0.6458
23	45	700	1.001	0.9859	0.9421	0.9130
24	45	700	0.9893	0.9786	0.9398	0.9204
25	45	700	1.0126	0.9998	0.9520	0.9392
26	45	700	1.0065	0.9940	0.9509	0.9063
27	45	700	1.0172	1.0042	0.9697	0.9258
28	45	700	1.0081	0.9999	0.9585	0.9319
29	45	700	1.0077	0.9987	0.9595	0.9177
30	45	700	1.0109	1.0091	0.9584	0.9366
31	45	700	1.0031	0.9872	0.9432	0.8651
32	45	700	1.0083	1.0027	0.9554	0.9192
33	45	700	1.0029	0.9899	0.9344	0.9173
34	45	700	1.0120	1.0029	0.9595	0.9346
35	45	700	1.0099	1.0026	0.9708	0.9194
36	45	700	1.0113	0.9964	0.9532	0.9109

Table 5.8: Failure times corresponding to the recorded temperature and current and the failure conditions of 70% degradation

Temperature (°C)	Time to Failure (hours)		
	77	77	45
	700	500	700
8048	9513	15263	
8539	9586	22339	
4815	9570	15796	
6635	8148	12436	
4442	8837	11936	
7574	8755	15069	
7364	3088	13103	
8174	9235	13860	
	7221	9438	
	9544	12882	
	9499	15373	
	9818	14251	
	10867	12326	
	3865	12211	

The next step is to find the mean life at the given use conditions of 30 °C and 400 mA. This requires a dual input life model, and the most fitting one is the temperature-non-thermal life model given in Chapter 4.

$$L(I, T) = \frac{C}{I^n \exp\left(-\frac{B}{T}\right)}$$

Since the data will be fit to a Weibull plot, this life model will be evaluated as a temperature-non-thermal Weibull life model, where the parameters desired are C , n , B , and β . All failure times listed in Table 5.8 as well as the temperatures and currents listed in Table 5.8 are used to perform the MLE on the following log-likelihood:

$$\Lambda = \sum_{i=1}^{36} \ln \left[\frac{\beta I_i^n \exp\left(-\frac{B}{T_i}\right)}{C} \left[\frac{I_i^n \exp\left(-\frac{B}{T_i}\right)}{C} t_i \right]^{\beta-1} \exp \left\{ - \left[\frac{I_i^n \exp\left(-\frac{B}{T_i}\right)}{C} t_i \right]^\beta \right\} \right]$$

A MATLAB routine for MLE is used to calculate the parameters as

$$\hat{\mu}_{C,n,B,\beta} = \begin{bmatrix} 258 \\ 0.56 \\ 2444 \\ 0.97 \end{bmatrix}; \hat{\Sigma}_{C,n,B,\beta} = \begin{bmatrix} 2961103 & 1942 & 329424 & 4.80 \times 10^{-9} \\ 1941 & 1.79 & 1334 & 3.12 \times 10^{-12} \\ 329424 & 1335 & 2458062 & 4.92 \times 10^{-10} \\ 4.80 \times 10^{-9} & 3.12 \times 10^{-12} & 4.92 \times 10^{-10} & 0.026 \end{bmatrix}$$

Thus, the use life estimate compared to the data looks like the following plot:

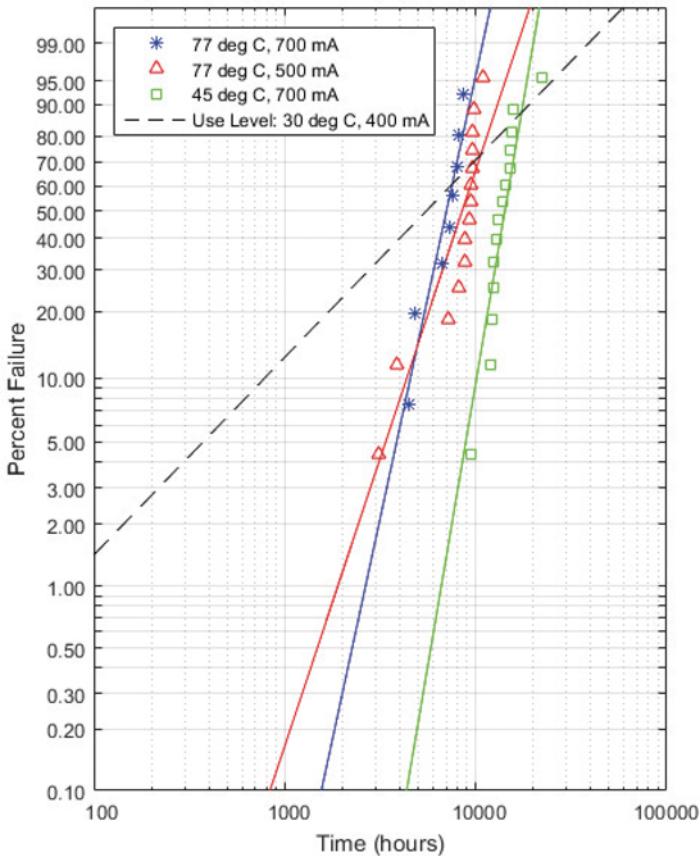


Figure 5.17: Weibull distribution fit with characteristics at the given conditions and the use conditions of 30 °C and 400 mA

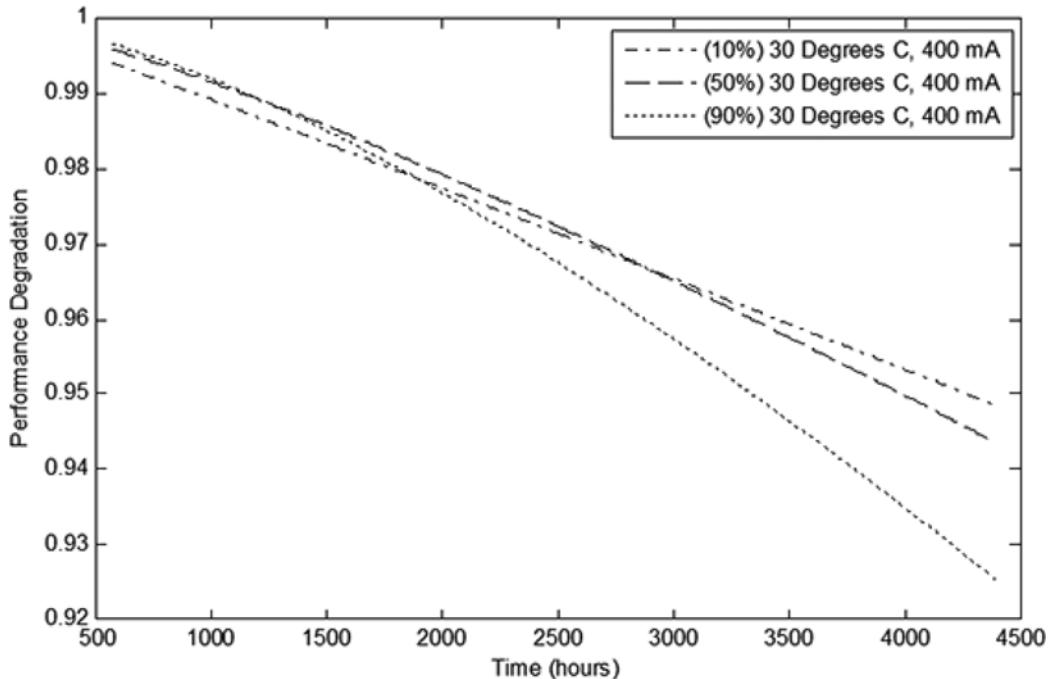
The last step is to use a MATLAB script to calculate the performance degradation at use level. This is done by calculating the value of $\gamma_0(T, I)$ at each confidence bound as

$$\gamma_0(T, I) = \frac{D_{ij}^{-1} - 1}{\left[\frac{c}{I^n \exp(-\frac{B}{T})} \right]^{\gamma_1}}$$

The confidence intervals for the initial parameters end up being as follows in Table 5.9:

Table 5.9: Confidence intervals for degradation parameters

	10%	50%	90%
$-\gamma_0$	6.05×10^{-6}	9.32×10^{-7}	1.75×10^{-7}
γ_1	1.09	1.32	1.55



And this results in the following plot for performance degradation:

Figure 5.18: Mean and 80% confidence bounds of the performance degradation at the use conditions of 30 °C and 400 mA

Example 5.5: Bayesian

Consider an accelerated degradation test of the relative luminosity (proportion of initial luminosity) for LEDs (Hamada et al., 2008). The standard operating temperature for LEDs is 20 °C. The LED failure is defined as occurring when the LED relative luminosity drops to 0.5, i.e., 50% of initial luminosity. The accelerated degradation test that involves testing 25 units each at 25 °C, 65 °C, and 105 °C. The data is completely documented in the original reference, but sample of the data from Unit 1 at all three temperatures is shown in the Table 5.10.

The LED degradation data model follows the form in which an expression for the true degradation of luminosity at time t and temperature T (in degrees Celsius) is:

$$Y_{ijk} = \left[1 + \beta_1 \left\{ t_{ijk} \exp \left[\beta_3 11605 \left(\frac{1}{T_u + 273.15} - \frac{1}{T_i + 273.15} \right) \right] \right\}^{\beta_2} \right]^{-1} + \varepsilon_{ijk}$$

where t_{ijk} is the k^{th} time of the j^{th} unit at the i^{th} temperature. ε_{ijk} is normally distributed $(0, \sigma_\varepsilon^2)$. Note that β_1 and β_2 are positive. T_u is the normal use temperature of 20 °C. That is, the true degradation follows an Arrhenius relationship. Regarding the prior distributions for the parameters β_1 , β_2 , β_3 , and σ_ε we use the uniform distribution UNIF(0,100) for all of them. Estimate the posterior mean and upper and lower values of the degradation model parameters, including the MTTF estimate of the LED at standard operating temperature.

Table 5.10: LED luminosity data of Unit 1 at 25 °C, 65 °C, and 105 °C

Time (hours)	Luminosity drop for Unit 1 at		
	25 °C	65 °C	105 °C
336	0.9704	0.9538	0.9104
672	0.9439	0.8857	0.8549
1008	0.9614	0.8879	0.8196
1344	0.9008	0.8635	0.7986
1680	0.9273	0.835	0.7731
2016	0.8753	0.8165	0.7795
2352	0.8793	0.8034	0.731
2688	0.9106	0.7824	0.7287
3024	0.8572	0.7959	0.6932
3360	0.8572	0.7761	0.7082
3696	0.8698	0.7572	0.6763
4032	0.8369	0.7451	0.6632
4368	0.839	0.7203	0.6797
4704	0.7949	0.7343	0.6347
5040	0.8113	0.7416	0.6336
5376	0.7658	0.7232	0.614
5712	0.8094	0.7175	0.62
6048	0.761	0.7079	0.6346
6384	0.8047	0.6831	0.6151
6720	0.7731	0.6873	0.6021
7056	0.7853	0.6841	0.5802
7392	0.7681	0.6402	0.5924
7728	0.7555	0.632	0.5891
8064	0.7531	0.651	0.5722
8400	0.7574	0.6302	0.5564
8736	0.7496	0.6468	0.5657
9072	0.7396	0.6245	0.556
9408	0.7212	0.6312	0.524
9744	0.7262	0.6366	0.5334

Solution

Begin again by plotting the data in its entirety

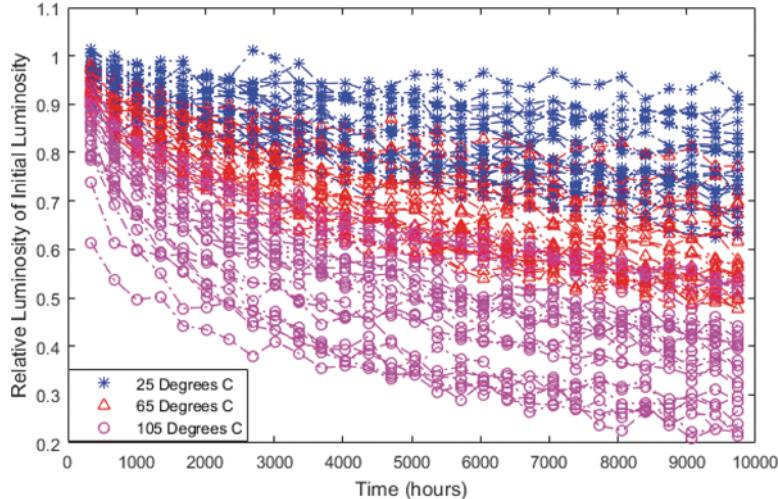


Figure 5.19: Relative luminosity of initial luminosity data

This time the MATLAB Bayesian estimation routine is used to obtain the parameter estimates β_1 , β_2 , β_3 , and σ_ϵ for each unit. The approximate trendlines are displayed on

Figure 5.21.

From the parameter estimates the posterior parameter distributions for the entire set of data are obtained and given as uniform distributions defined in Table 5.11.

Table 5.11: Posterior degradation parameters

	Lower Bound	Upper Bound
β_1	9.69×10^{-6}	0.06
β_2	0.29	1.00
β_3	4.7×10^{-3}	1.33
σ_ϵ	7.1×10^{-3}	0.20

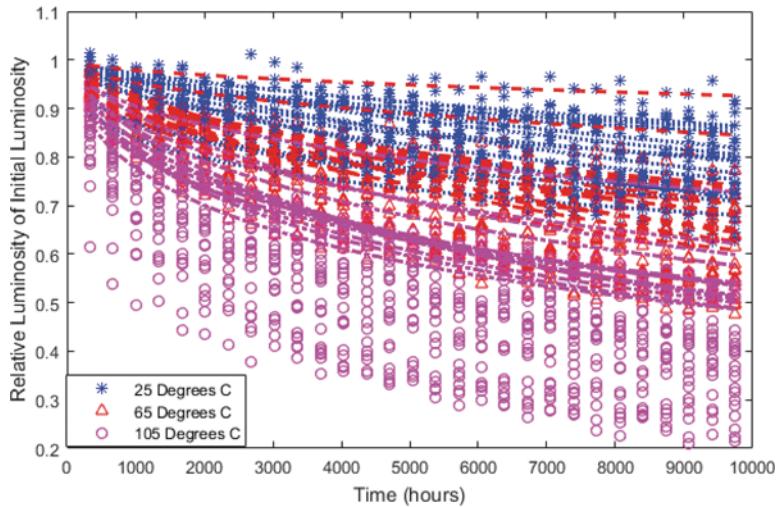


Figure 5.20: Trendlines of relative luminosity based on parameter estimates obtained from the Matlab Bayesian estimation routine

The MTTF estimate at operating temperature is obtained from the trendlines, where the following times to failure can be obtained given that failure occurs at 50% initial luminosity. A temperature vs. time relation plot, shown in Figure 5.20, is made to show the relation and thus obtain the MTTF, which is 121,975 hours. The Weibull distribution fit is shown in Figure 5.22.

Table 5.12: Failure times corresponding to the recorded temperature and the failure conditions of 50% degradation of luminosity

Temperature (°C)	Time to Failure (hours)		
	25	65	105
25	18022	8680	785
25	18090	8805	1794
25	23892	9125	1979
25	28640	9180	2100
25	32558	9707	2140
25	35332	9822	2151
25	37260	10121	2949
25	47982	10387	2972
25	48426	10664	3154
25	49318	12256	3705
25	50158	12320	4310
25	57456	12678	4481
25	58229	12836	4735
25	61522	13276	4957

Table 5.12: Failure times corresponding to the recorded temperature and the failure conditions of 50% degradation of luminosity(Continued)

Temperature (°C)	Time to Failure (hours)		
	25	65	105
65905	13503	5595	
81058	14130	5600	
81708	16135	5882	
83211	19903	6006	
108429	20078	6449	
108960	23178	6909	
121577	25449	9634	
178064	28318	10868	
219197	34039	11182	
241354	56603	11220	
708508	69457	12708	

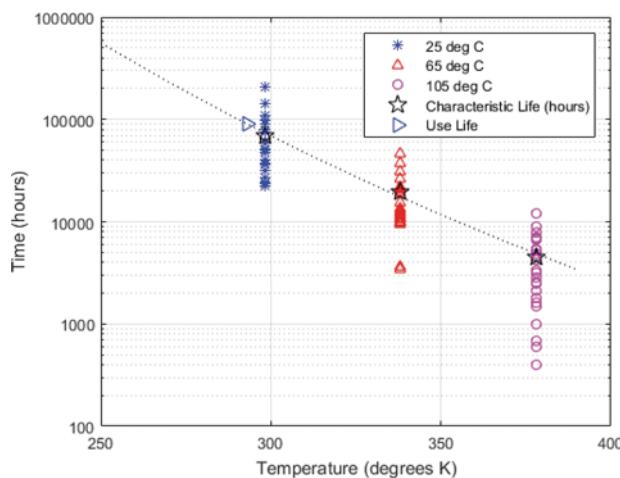


Figure 5.21: Temperature vs. time to failure relationship plot

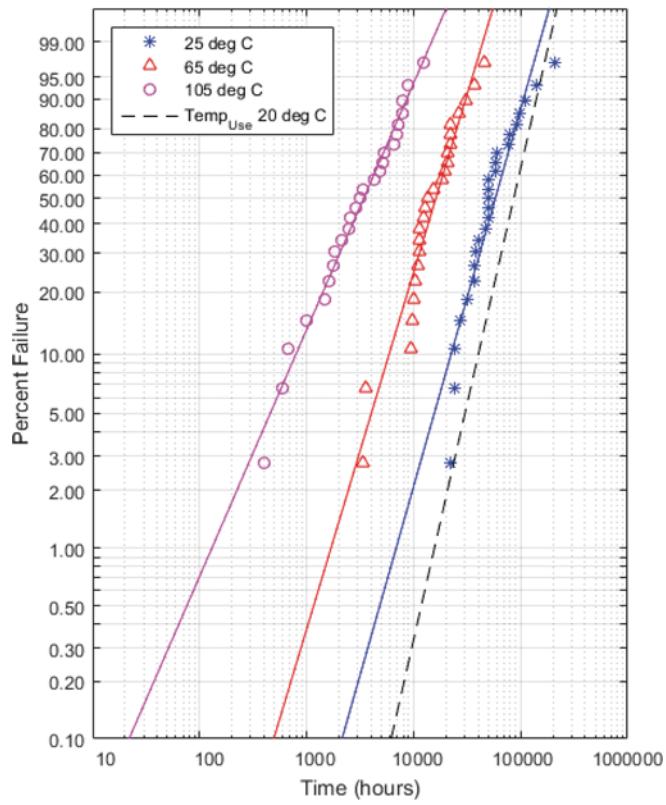


Figure 5.22: Weibull distribution fit with characteristics at the given conditions and the use conditions of 20°C

REFERENCES

- Hamada, Michael S., Wilson, Alyson, Shane Reese, C. Martz, Harry (2008). Chapter 8: Using Degradation Data to Assess Reliability. *Bayesian Reliability*, Springer.
- Meeker, W., & Escobar, L. (1998). *Statistical Methods for Reliability Data*. New Jersey: John Wiley and Sons.
- Meeker, W., & Hahn, G. (1985). *How to Plan an Accelerated Life Test*. Milwaukee: ASQC Quality Press.
- Reliasoft Corporation. (1996). *Accelerated Life Testing Reference*. Tucson: Reliasoft.
- Wang, P., & Coit, D. (2007). Reliability Assessment Based on Degradation Modeling with Random or Uncertain Failure Threshold. Proceedings of the 2007 Reliability & Maintainability Symposium (RAMS). Orlando.

Chapter 6: Accelerated Test Planning

6.1. INTRODUCTION

Planning for ALT and ADT is a critical early step in accelerated testing. Before any life or degradation test is begun, two critical questions must be addressed. These questions are: 1) What are the stress levels of constant or step stress that should be used? 2) How many units should be tested at each stress level? The answers to these two questions are closely interrelated. The literature in test planning is limited and the methods used are often either based on experiential techniques and heavy reliance on anecdotal techniques, or they are based on formal statistical notions containing very limiting assumptions that often conflict with the needs of most practical accelerated testing and requirements. Much of the research on the statistical planning methods has been conducted to search for an optimized combination of stress levels and number of test units to be used at each stress level. For example, Nelson (1990), Meeker and Escobar (1998) and Nelson (2005) provide a substantial review of the state-of-the-art statistical methods. Conversely, in many ALT and ADT test cases, the test time and the total number of test units are selected experientially based on major constraints such as availability and number of test facilities, budget, available time for the test, design and development length, and experience of the test staff. In all the cases ALTs are often in search of a given quantile of the life at the use stress level (e.g., the 10% of life corresponding to that at which 10% of the units are expected to fail). The accuracy of the test results is measured in terms of the size of the confidence interval with a reasonably high confidence level. In this chapter informal and mostly conceptual methods of test planning for single or multiple ALT and ADT, including step stress test, will be discussed. First, some good engineering considerations will be discussed, followed by ALT and ADT testing concepts.

6.2. ISSUES TO CONSIDER PRIOR TO ACCELERATED TESTING

The ALT and ADT methods summarized in Chapter 4 and Chapter 5 offer many benefits in failure and reliability analysis. These tests can be extremely useful for obtaining metrics of interest relating to materials and products. However, there are also some potential problems to be aware of when conducting ALT and ADT. A number of potential problems experienced by practitioners are largely known and can lead to incorrect results. To prevent important mistakes when planning and performing accelerated tests, Meeker and Escobar (1998) describe the following potential errors:

- **Introduction of New Failure Modes**

High levels of accelerating variables (e.g., temperature or voltage) can induce failure modes that would not be observed at normal operating conditions. For instance, instead of simply accelerating a failure-causing chemical process, increased temperature may essentially change certain material properties. In less extreme cases, high levels of accelerating variable (agents of failure) will change the relationship between life and the accelerating variable. For example, when temperature exceeds the recrystallization temperature of the material and is maintained for a sufficient amount of time, grain growth will affect yield strength and other mechanical properties. Similarly, elevating temperature results in significant time dependent strain due to diffusion, super-plasticity and dislocation climb even at low mechanical stress because creep rate is a function of temperature, and a new failure mechanism of creep rupture may dominate.

- **Characterization of Uncertainties**

There are always uncertainties in the life-stress model and the degradation model (model

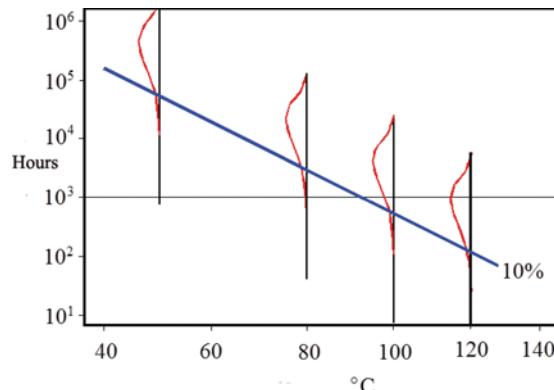


Figure 6.1: Possible results for a typical temperature-accelerated failure mode on an IC device (Meeker, 1998)

uncertainty), and estimates of the model parameters (parameter uncertainties). Basing reliability assessments and decisions on limited information and highly uncertain models, or on point estimates alone, without assessment of confidence in such estimates will often lead to serious errors. It is also important to note that confidence intervals and uncertainty bounds on model parameters do not account for all model uncertainties. Such uncertainties can be tremendously amplified by extrapolation in accelerated testing. In general, performing sensitivity analysis is an important step in any quantitative life analysis involving uncertainty, and is a particularly useful first step for assessing the effects of model uncertainty. For example, one can rerun analyses under different assumed forms of the model to observe the effects that different model assumptions have on important reliability assessments and conclusions. In all steps of the ALT discussed earlier in Chapter 4 and in ADT discussed in Chapter 5, uncertainties should be both qualitatively and quantitatively characterized (when possible) and communicated among the stakeholders and particularly among the accelerated testing team members.

- **Time Scale Differences Associated with Multiple Accelerated Stresses**

Standard acceleration methods generally will not accelerate all time scales in the same manner. A serious drawback of accelerated testing is to assume a simple relationship between life or damage (degradation) and the accelerating stresses when the actual relationship is really more complicated and synergistic. Consider the interaction between relative humidity and vibration stress in a plastic encapsulated device in a humid environment. The main failure mechanism in this case is corrosion of the metallization under operational condition, and the acceleration stress is humidity. However, the reduction in strength of the partially corroded material causes the device to fail, by mechanisms such as a wire break, under mechanical impact before an opening is caused by corrosion alone. Thus, the mechanical stress, which would otherwise be insignificant, becomes the dominant stress enhanced by the presence of corrosion. If life is estimated by considering corrosion only, the result would be excessively optimistic.

- **Recognition of Masked Failure Modes**

Figure 6.1 shows a life-stress relationship of an ALT in a single failure mode case involving temperature acceleration (assuming exponential life-stress model). It is possible that while we focus on one known failure mode in a test such as this, another failure mode that is more prevalent at high stresses may potentially mask the failure mode of interest believed to be the leading failure mode at the use stress level. This is shown in Figure 6.2, in which failure mode 1 masks failure mode 2, which is the prevalent mode of failure at the use stress level and thus the most frequent in the field. A careful PoF analysis prior to planning can avoid this problem.

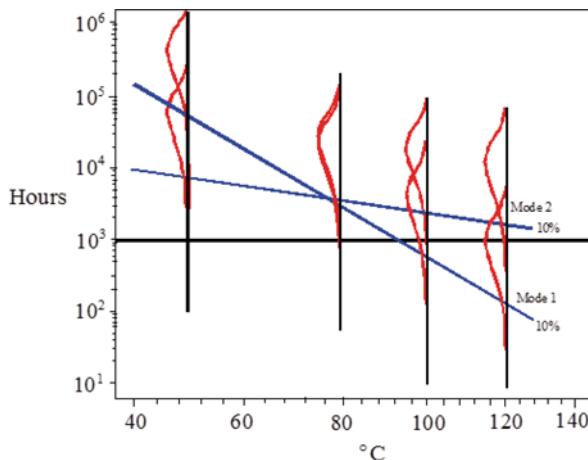


Figure 6.2: Failure mode 2 with lower activation energy, masked at high temperature and unmasked at lower temperature (Meeker, 1998)

- **Incorrect Comparison of Items Tested**

It is sometimes claimed that accelerated testing is not really useful for predicting reliability but is useful for comparing alternatives (e.g., alternative designs, vendors). The thought behind this claim is that accelerated tests generally cannot be expected to adequately replicate actual use conditions and thus properly estimate the life characteristics, but that they are good for comparison of life characteristics between two items. Therefore, it is believed that if Item 1 is better than Item 2 in an accelerated test, then the same would be true in field use, as illustrated in Figure 6.3. However, this belief may not always be true. For example, similar to the masked failure mode, if the comparison involves similar and well-behaved life-stress models, the results can be similar to Figure 6.3, as long as the slope of the life-stress models of the two items are very similar. For example according to Figure 6.3, the test results show that Item 1 has better life characteristics than Item 2 during the use condition, if the same is observed at the high stress accelerated testing.

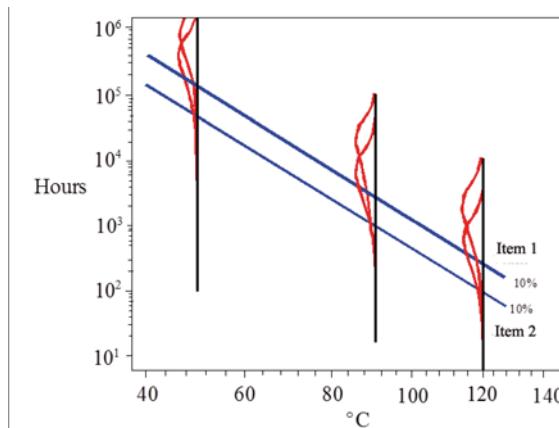


Figure 6.3: Well-behaved comparison of two items (Meeker, 1998)

However, these types of comparisons only apply in these ideal and well-behaved situations. When the life-stress model or failure mechanisms are appreciably different for the same stress

agents, a situation similar to the one shown in Figure 6.2 may arise where the life-stress models of the two items are no longer of the same characteristics (i.e., they have very different slopes) as indicated by the non-parallel lines. Therefore, under these conditions, judging the performance of the item at the use stress level based on the failures observed at high stress levels may be flawed. Again, a careful PoF study should be able to reveal the possibility of such faulty comparisons.

- **Acceleration Stresses of the Failure Mechanism Causes Deceleration of Another**

In some cases, it is possible that increasing what is thought to be an accelerating variable will actually cause deceleration. In one example, tests at increased temperature to accelerate circuit failures predicted few failures at the use level. However, failures in the field pointed at far higher likelihood of failure than predicted because the increased temperature used in the accelerated testing resulted in lower humidity, since humidity was not controlled during the test. The primary failure mechanism observed at the use level was caused by corrosion (due to higher humidity levels) that did not occur at high temperature and low humidity, which were the conditions of testing. For this reason, in most accelerated tests, especially for electronic equipment, both temperature and humidity need to be controlled.

- **Design and Manufacturing Changes After the Accelerated Test**

Changes in the design or manufacturing, even those that appear minor on the surface, could introduce drastic changes that induce new failure mechanisms or failure modes. For example, a new electromechanical device was to be used in a system designed for 20 years of service in a protected environment. An accelerated test of the device was conducted, which estimated a 20-year life (during which no more than 10% of the units were to fail) under normal operating conditions (assuming an average use-rate). After the accelerated test, and as the product was going to production, a material change was made. The change introduced a new material-degradation mechanism and failure mode that caused (or would have caused) all in-service units to fail within 10 years. As such, the assumed or estimated accelerated life or degradation models were no longer valid. These changes must be balanced with other considerations relating to the configuration of the design when it was tested, versus the configuration of the design in service.

- **Accelerated Tests based on Specially Built Prototype Test Units**

This is slightly related to the previous problem, and refers to the risk that seriously incorrect conclusions can result from an accelerated test if units tested are different from the actual production units in material or workmanship. For example, mass production conditions are often harsher than those in a laboratory condition. Cleanliness, precision, material preparation and care in building the prototype versus mass manufactured and shipped units may vary substantially. Materials and parts used in the prototype units might differ from those that will be used in mass production. Highly trained specialists may build the prototype units, leading to higher quality units being tested than those being manufactured in mass production. An analysis should be performed to assess the applicability of the test results from prototype and specialty unit to the ones manufactured at the end.

- **Failure Mechanism Shifting**

A failure mechanism may shift at higher stress levels for the same stress agent. For example, temperature is known to vary the rates of many physical and chemical reactions. Many of the failure mechanisms that cause degradation in electronic devices are fundamentally physical or chemical processes in which temperature is the accelerating agent. However, high temperature

values may also accelerate other failure mechanisms like secondary slow trapping, corrosion in the presence of moisture, creep, electro-migration, intermetallic growth and dielectric breakdown, which may not necessarily be considered as a prevalent failure mechanism at use-level temperature. Another potential failure mechanism shift is that of crack initiation mechanism shifting to crack propagation in fatigue failures. This latter example shows that the life-stress model will not be valid for the entire length of the degradation process in an accelerated fatigue test. The failure mechanism shifting should be addressed as part of the PoF process prior to test planning and performance and analysis of the tests results.

6.3. PLANNING FOR ACCELERATED LIFE TESTS

So far in this book, we have discussed the main uses of ALTs, which are to provide information about the life distribution of materials and products of interest under normal use conditions. However, to be successful, these tests need to be appropriately planned so that the desired information can be extracted out of the test. A poorly planned ALT can result in significant loss of time, effort, and money. It is therefore crucial that prior to conducting the test, a careful plan be developed to aid in accurately estimating reliability characteristics at use conditions, whilst minimizing test time and cost.

The main objectives of ALT planning are to

- Determine the accelerated stress levels and the corresponding sample sizes;
- Estimate duration of the test;
- Define criteria for acceptance testing;
- Determine the number of failures allowed.

The ALT planning steps to realize these objectives are further elaborated in the next section

6.3.1. STEPS FOR ACCELERATED LIFE TESTS

From examining the literature, the following are steps for accelerated life testing using the PoF approach to ALT.

1. A PoF analysis is performed where the likely failure mechanisms expected during the normal use-level stress conditions are chosen for accelerated testing. In this process the most appropriate failure mechanism, empirical mathematical forms of the life-stress models, and life distribution models are chosen using good engineering judgment and by relying on the past experiences. For the dominant failure mechanisms, the stress(es) that accelerate damage and failure are selected and other agents that accelerate them are identified. For step-stress tests, the dominant failure mechanisms should produce cumulative damage as stresses go from one step to another.
2. Life-stress models are examined, and the generic values of their associated parameters are selected from the past experiences and data, engineering judgment, or generic values obtained from the literature.
3. The various accelerating stresses, stress agents that accelerate them, and their failure limits (e.g., endurance limits) are studied in the context of each failure mechanism. The stresses for testing must be selected such that they affect the failure mechanisms of interest and do not introduce extraneous failure mechanisms.
4. Stress agents to be accelerated are examined to assure their reasonableness, and the magnitude of the stresses is determined. This is done mostly based on the PoF analysis, and the planners

should carefully avoid issues that were summarized in Section 6.2 such as the failure mode and mechanism masking.

5. The type of data collection approach (complete, type I, type II, right censored, left censored, and interval data) and duration of the test is also established. For the choices of the life pdf models in the PoF analysis (Step 1), generic values of the shape parameter of those pdf models are chosen from past experiences and engineering judgment.
6. The type of accelerated life test is determined (i.e., constant load acceleration or step stress accelerating), and the sample size at each stress level is selected by apportioning the proper number of units such that 30-70 percent of the units will be expected to fail and the number of units placed in test should be such that about the same number of failures are generated at each stress level tested. For this reason, one needs to allocate more test units at low stress levels than the high stress levels. Generic values of the parameters of the life-stress model and pdf models selected in steps 2 and 5 are used to make a rough estimate of the expected number of failures at each stress level. Generally stresses (or stress agents) are selected so as to minimize the amount of extrapolation necessary. To do this typical ranges of acceleration factors as high as 500-600 for electronics and 40-50 for mechanical units are appropriate for the highest possible stress level. The lowest stress level should be associated with an acceleration factor of 10-20 for electronics and 2-4 for mechanical units. Stress levels in between these should be selected.
7. Step 6 is the most important step and should be revised accordingly if it violates the required constraints of cost, available test facilities, length of the test and number of test units available. As such, this step may need multiple iterations until the most desirable stress levels and the number of unit tests to be tested have been established and optimized.
8. The test is performed and data are gathered. During the test it is often very useful to establish the root causes of the observed failures to determine which failure mechanism caused the failure and which failure mode was responsible for the ultimate manner of failure. For step-stress tests, the stress levels are revised for the next step, if needed. As such, in step stress tests this step allows more flexibility in conducting the test.
9. The test data are interpreted and evaluated according to methods discussed in Chapter 4 including assessment of uncertainties in the life-stress model and uncertainties about the parameters of the corresponding pdf model.
10. The life-stress model estimated from the ALT data in step 9 is extrapolated to normal operating conditions and life characteristics assessed to assure reasonableness of the results.
11. Revise the tests if the data analysis shows shifting failure mechanisms or the results were inappropriate.
12. If more than one failure mechanism is considered, the competing risk model should be used to combine the time-to-failure pdf of corresponding to other failure mechanisms to form a composite distribution.

The above ALT planning guidelines are relatively basic and rely on the experiential approach and assumptions. Many other complicated factors can come into play when planning ALT (e.g. multiple accelerating variables or formal planning methods via MLE). There are no hard and fast rules when it comes to planning accelerated life tests, but careful consideration of the issues highlighted in this section will provide a good foundation for developing an effective test plan. Ultimately, however, it is the engineer's experience and the various constraints and requirements specific to the component under test that will influence the test plan.

6.3.2. OPTIMAL DESIGN OF ACCELERATED LIFE TEST

The accelerated life test plan has been continually improved and optimized to adapt to the current extremely reliable and complex modern products. The improvements can be categorized with respect to whether they are testing related and/or extrapolation related. For example, extending constant-stress testing to time-varying stress testing (Miller & Nelson (1983), Nilsson (1985), Bai, et al (1989), (1991), (1992), (1993)); designing a test with multiple accelerating variables (Escobar & Meeker (1999)); dealing with different sizes of specimens through testing (D.S. Bai (1989)); integrating Bayesian methods to statistical plans for accelerated life tests (DeGroot & Goel (1979), Chaloner & Larntz (1990), Zhang & Meeker (2006)); establishing the accelerated test models considering the uncertainty of parameters (Meeter & Meeker (1994)); and minimizing the asymptotic variance of the maximum likelihood estimator (Escobar & Meeker (1995)).

There is little to no actual data through the stage of test planning, so it is necessary to integrate the historical data, engineering judgment, and expert elicitation. The Bayesian approach is one of the best choices to address the uncertainty of planning values like testing stress, testing time, and number of testing units. A simulation-based approach is generally proposed to address such test planning given certain optimal criteria.

The objective of most testing design is to extrapolate a particular percentile of the life distribution at normal stress conditions given some constraints like limited budget and testing duration. Specifically, the simulation data would be randomly generated based on the identified life-stress relationship and lifetime distribution. Then the likelihood function can be constructed, and the corresponding prior distribution can be specified based on the historical information. Given the optimal criteria, a utility function is proposed, and the ultimate goal is to maximize or minimize the utility function. In this section, two examples are introduced to illustrate how to use the historical information and expert elicitation.

Example 6.1

Consider a simple step-stress accelerated life test plan with two accelerated stress levels. The Bayesian approach is proposed to illustrate the methodology to determine the optimal stress changing time to maximize the estimation precision of the 100pth percentile of the time to failure distribution at the normal stress condition, t_p (Yuan, Liu, & Kuo, 2012). Assume the normal stress level is S_0 , and the accelerated testing stresses are S_1 and S_2 . In particular, n items are tested at a lower stress level S_1 , and then the stress level is changed to the higher stress level S_2 at the stress changing point τ . The test continues until a pre-specified censoring time t_c .

Solution

Step 1: Modeling

Assume the lifetime distribution of testing item is subject to Weibull distribution, so the cdf of time to failure through the test is:

$$F(t) = \begin{cases} 1 - e^{-(\frac{t}{\alpha_1})^\beta} & 0 \leq t \leq \tau \\ 1 - e^{-(\frac{\tau}{\alpha_1} + \frac{t-\tau}{\alpha_2})^\beta} & t \geq \tau \end{cases} \quad (6.1)$$

where β is the shape parameter and α_i is the scale parameter.

Then we can get the probability of obtaining a censored observation:

$$P(t > t_c) = R(t_c) = e^{-(\frac{\tau}{\alpha_1} + \frac{t_c - \tau}{\alpha_2})^\beta} \quad (6.2)$$

Assume the life-stress relationship is an Arrhenius relationship, so the scale parameter θ_i can be a log-linear function of the inverse of stress:

$$\ln \alpha_i = a + \frac{b}{S_i} \quad i = 0, 1, 2 \quad (6.3)$$

where a and b are unknown parameters.

Step 2: Bayesian Update

The joint prior distribution of the three model parameters is denoted by $f(a, b, \beta)$, so then we would get the posterior distribution as:

$$f(a, b, \beta | t) = \frac{L(t|a, b, \beta)f(a, b, \beta)}{\iiint L(t|a, b, \beta)f(a, b, \beta)dadb\beta} \quad (6.4)$$

where the likelihood function $L(t|a, b, \beta)$ is

$$L(t|a, b, \beta) = \prod_{i=1}^{n_1} \beta \cdot e^{-\beta(a + \frac{b}{S_1})} \cdot t_i^{\beta-1} \cdot e^{-[t_i \cdot e^{-(a + \frac{b}{S_1})}]^\beta} \\ \cdot \prod_{j=1}^{n_2} \beta \cdot e^{-\beta(a + \frac{b}{S_2})} \cdot [e^{b(\frac{1}{S_2} - \frac{1}{S_1})}\tau + t_i - \tau]^{\beta-1} \\ \cdot e^{-[\tau e^{-(a + \frac{b}{S_1})} + (t_i - \tau)e^{-(a + \frac{b}{S_2})}]^\beta} \\ \cdot \left\{ e^{-[\tau e^{-(a + \frac{b}{S_1})} + (t_c - \tau)e^{-(a + \frac{b}{S_2})}]^\beta} \right\}^{n_3}$$

$$\begin{aligned} \Lambda &= \ln(L(t|a, b, \beta)) \quad (6.5) \\ &= \sum_{i=1}^{n_1} \left\{ \ln \beta - \beta \ln \left(a + \frac{b}{S_1} \right) + (\beta - 1) \ln t_i \right. \\ &\quad \left. - \left[t_i \cdot e^{-(a + \frac{b}{S_1})} \right]^\beta \right\} \\ &\quad + \sum_{j=1}^{n_2} \left\{ \ln \beta - \beta \ln \left(a + \frac{b}{S_2} \right) + (\beta - 1) \ln t_i \right. \\ &\quad \left. - \left[\tau e^{-(a + \frac{b}{S_1})} + (t_i - \tau)e^{-(a + \frac{b}{S_2})} \right]^\beta \right\} \\ &\quad - n_3 \left[\tau e^{-(a + \frac{b}{S_1})} + (t_c - \tau)e^{-(a + \frac{b}{S_2})} \right]^\beta \end{aligned}$$

where n_1 is the number of failures at stress level S_1 ; n_2 is the number of failures at stress level S_2 ; and n_3 is the number of censored items at time t_c .

Step 3: Prior Specification

There are three unknown parameters: a , b , and β , the uncertainty of which are represented with the joint prior distribution $f(a, b, \beta)$ based on historical data, engineering judgement, and expert opinions. To specify this distribution, it is generally assumed that all the parameters are independent, so the joint prior distribution can be expressed as:

$$f(a, b, \beta) = f(a)f(b)f(\beta) \quad (6.6)$$

where $f(a)$, $f(b)$, and $f(\beta)$ are, respectively, the prior distributions of a , b , and β .

The approach above is analytically convenient, but all the parameters may not actually be independent from each other. An alternative approach is recommended to relax the independency assumptions to construct the joint prior distribution for the shape and scale parameters of the Weibull distribution (Singpurwalla (2008)).

Step 4: Optimization Criteria

The 100pth percentile of the life distribution at the normal stress condition t_p can be expressed as a function of a , b , and β :

$$t_p = e^{\left(a + \frac{b}{\beta}\right)}[-\ln(1 - p)]^{1/\beta} \quad (6.7)$$

Multivariate transformation of random variables can be applied to derive the joint posterior distribution of t_p , a , and b , which is denoted by $f(t_p, a, b | \vec{t})$. From this joint posterior distribution, we can obtain the marginal posterior distribution of t_p as

$$f(t_p | t) = \int \int f(t_p, a, b | t) da db \quad (6.8)$$

The posterior mean and posterior variance of t_p are given with the data t :

$$E(t_p | t) = \int_0^\infty t_p \cdot f(t_p | t) dt_p \quad (6.9)$$

and

$$V(t_p | t) = \int_0^\infty [t_p - E(t_p | t)]^2 \cdot f(t_p | t) dt_p \quad (6.10)$$

Since the posterior variance depends on the unobserved data, the expectation of posterior variance of t_p is used as a marginal expectation of the posterior variance over all possible data as the optimization criterion. The objective is to minimize the expectation of posterior variance of t_p as below:

$$E[V(t_p|t)] = \int V(t_p|t) \cdot f(t) dt \quad (6.11)$$

Step 5: Optimization

Generally, there is no closed form for the complicated integral, so numerical computations like MCMC would be used to address such high dimensional integration to evaluate the expectation of posterior variance of t_p . In particular, data vectors are randomly generated, and then each data vector would be simulated by the following steps:

- Generate randomly a set of values for the three model parameters a , b , and β from the joint prior distribution $f(a, b, \beta)$.
- Use the standard inverse transformation method to generate n i.i.d. failure times from the distribution function given by Equation (6.1). The right-censored observations are the one with simulated failure times greater than the censoring time t_c .
- Evaluate the posterior variance $V(t_p|t)$ for each simulated data vector where numerical multidimensional integration may be used, such as Gibbs sampling, Metropolis-Hastings algorithm, and the Gaussian quadrature method. For example, use Gibbs sampling to draw a random sample of t_p from the marginal posterior distribution $f(t_p|t)$. Estimate $V(t_p|t)$ using the sample variance, and then get the expectation of posterior variance of t_p by averaging $V(t_p|t)$ over all the simulated data vectors.
- Evaluate the expectation of posterior variance of t_p at discrete time points, and then plot the expectation of posterior variance of t_p vs. the stress changing point. The optimal stress changing time would be easily identified from the plot.

For illustrative purposes, a simple step stress accelerated life testing is designed to explore the thermal property of a mechanical component. Assume an Arrhenius relationship for life-stress joint distribution and that the lifetime is subject to Weibull distribution. The normal stress amplitude is $S_0=100$ °K, the accelerated stresses are $S_1=300$ °K and $S_2=400$ °K. Assume there are 30 components available for testing, and the test would be censored at $t_c=100$ hours, and $p=0.001$. According to the historical information, the unknown parameters are assigned with non-informative prior distribution as follows:

Parameter	Prior Distribution
a	Uniform [2.1, 3.9]
b	Uniform [310, 890]
β	Uniform [1.4, 2.9]

Given the initial values of parameters as $a=3.157$, $b=621.9$, and $\beta=2.318$, run the simulations 80 times with 60000 iterations each run with the simulation-based programming. Figure 6.4, it's easy to shows the trending of the expectation of posterior variance and stress changing time.

Example 6.2

Consider the case that only one item is available for testing, and we need to determine the optimal testing stress level S_i to effectively assure the field performance of the item (Martz & Waterman (1978)). Assume only right censored or complete failures are to be observed. Develop a Bayesian model to estimate the required level of test stress that maximizes the expected probability of survival

at the selected design stress level. Engineering experience from similar past tests on similar items may be used to fit the model.

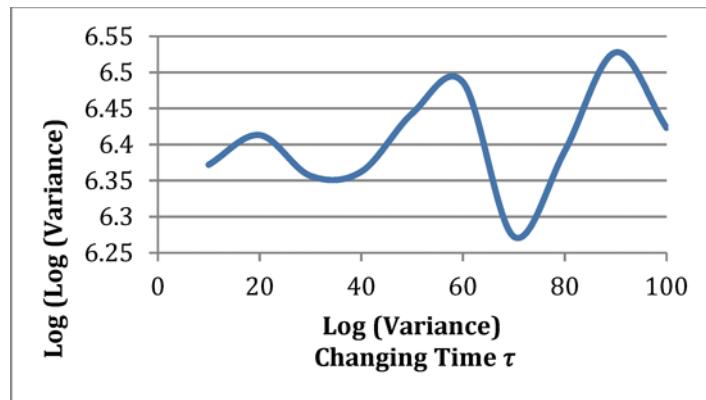


Figure 6.4: Log-variance vs. stress changing time

Solution

Step 1: Modeling

The testing stress is denoted as “k” in items of design stress S_1 . S_k represents the event that the testing item survives with probability $P_k = \Pr(S_k)$ under testing stress k. F_k represents the event that the testing item fails with probability $\Pr(F_k) = 1 - P_k$ under testing stress k. Assume P_k is a random variable with a modified negative-log gamma prior distribution:

$$f(p_k; \alpha, \beta, \delta) = \frac{p_k^{(1/\beta k^\delta)-1} (-\ln p_k)^{\alpha-1}}{\beta^\alpha k^{\alpha\delta} \Gamma(\alpha)} \quad (6.12)$$

where α is positive as the shape parameter, β is positive as the scale parameter, and k is the testing stress, expressed in items of design stress S_1 . Note that there should be an upper limit of test stress beyond the range of practical interest. δ is positive used to rescale k.

The unconditional mean and variance for the above negative-log gamma prior distribution are

$$E(P_k; \alpha, \beta, \delta) = (1 + \beta k^\delta)^{-\alpha} \quad (6.13)$$

and

$$V(P_k; \alpha, \beta, \delta) = (1 + 2\beta k^\delta)^{-\alpha} - (1 + \beta k^\delta)^{-2\alpha} \quad (6.14)$$

Conditional on whether the testing item survives or fails, we can get the two posterior pdfs, the corresponding cdfs and posterior means with the Bayesian approach. $\chi_{2\alpha}^2$ denotes a χ^2 random variable with 2α degrees of freedom.

- Conditional on the testing item survives:

$$f(p_k|S_k; \alpha, \beta, \delta) = \frac{p_k^{(1/\beta k^\delta)} (-\ln p_k)^{\alpha-1}}{\beta^\alpha k^{\alpha\delta} \Gamma(\alpha)(1+\beta k^\delta)^{-\alpha}} \quad (6.15)$$

$$F(p|S_k; \alpha, \beta, \delta) = \text{Prob}\left(\chi_{2\alpha}^2 > \frac{-2(1+\beta k^\delta) \ln p}{\beta k^\delta}\right) \quad (6.16)$$

$$E(P_k|S_k; \alpha, \beta, \delta) = \left(\frac{1+2\beta k^\delta}{1+\beta k^\delta}\right)^{-\alpha} \quad (6.17)$$

- Conditional on the testing item fails:

$$f(p_k|F_k; \alpha, \beta, \delta) = \frac{(1-p_k)p_k^{(1/\beta k^\delta)-1}(-\ln p_k)^{\alpha-1}}{\beta^\alpha k^{\alpha\delta} \Gamma(\alpha)[1-(1+\beta k^\delta)^{-\alpha}]} \quad (6.18)$$

$$\begin{aligned} F(p|F_k; \alpha, \beta, \delta) &= \text{Prob}(P \leq p|F_k; \alpha, \beta, \delta) \\ &= \text{Prob}\left(\chi_{2\alpha}^2 > \frac{-2 \ln p}{\beta k^\delta}\right) \\ &\quad - (1+\beta k^\delta)^{-\alpha} \text{Prob}\left(\chi_{2\alpha}^2 > \frac{-2(1+\beta k^\delta) \ln p}{\beta k^\delta}\right) + (1-(1+\beta k^\delta)^{-\alpha}) \end{aligned} \quad (6.19)$$

$$E(P_k|F_k; \alpha, \beta, \delta) = \frac{(1+\beta k^\delta)^{-\alpha} - (1+2\beta k^\delta)^{-\alpha}}{1-(1+\beta k^\delta)^{-\alpha}} \quad (6.20)$$

Step 2: Optimal Criteria

With such limited information of a single tested item, the expected model is developed to help effectively trade between the two alternatives: success and failure.

- Suppose that the tested item is able to survive a testing stress higher than the design stress. It is reasonable to be more confident that the item is more likely to survive through the design stress. The statement can be defined mathematically as

$$E(P_1|S_k) = g_1(k) \cdot E(P_1|S_1), k \geq 1 \quad (6.21)$$

where a suitable $g_1(k)$ is with the following properties: (a) $g_1(1) = 1$; (b) $\lim_{k \rightarrow \infty}(g_1(k)) = \frac{1}{E(P_1|S_1)}$; and (c) $g_1'(k) > 0$, at all points of continuity of $g_1(k)$. For example, $g_1(k) = k^c$.

- On the other hand, if the tested item fails under a testing stress higher than the design stress, the probability of surviving at the design stress is at least the same as the item which was tested and failed at the design stress. The statement can be defined mathematically as

$$E(P_1|F_k) = g_1(k)E(P_1|F_1), k > 1 \quad (6.22)$$

where a suitable $g_2(k)$ is with the following properties: (a) $g_1(1) = 1$; (b) $\lim_{k \rightarrow \infty} (g_2(k)) = \frac{E(P_1)}{E(P_1|F_1)}$; and (c) $g'_2(k) > 0$, at all points of continuity of $g_1(k)$. For example, $g_2(k) = 1$.

The optimal stress level S_k maximizes the expected probability $E(P_1|\alpha, \beta, \delta)$ that the only testing item survives.

$$\begin{aligned} E(P_1|\alpha, \beta, \delta) &= E(P_1|S_k; \alpha, \beta, \delta) \text{Prob}(S_k) \\ &\quad + E(P_1|F_k; \alpha, \beta, \delta) \text{Prob}(F_k) \end{aligned} \quad (6.23)$$

Substituting the above equation into this expression, we can get the objective function as below:

$$E(P_1|\alpha, \beta, \delta) = [\gamma_1 g_1(k) - \gamma_2 g_2(k)](1 + \beta k^\delta)^{-\alpha} + \gamma_2 g_2(k) \quad (6.24)$$

$$\text{where } \gamma_1 = \left(\frac{1+\beta}{1+2\beta}\right)^\alpha, \gamma_2 = \frac{1-\left(\frac{1+\beta}{1+2\beta}\right)^\alpha}{(1+\beta)^\alpha - 1} = \frac{1-\gamma_1}{(1+\beta)^\alpha - 1}$$

It is useful to solve $\frac{\partial E(P_1|\alpha, \beta, \delta)}{\partial k} = 0$, in order to maximize $E(P_1|\alpha, \beta, \delta)$. The solution yields the desired optimal test stress k_0 .

Step 3: Parameter Estimation based on Expert Elicitation

Since there is only a single test item available, it is reasonable and necessary to integrate the subjective information like expert elicitation. Some questions are proposed prior to the test to solicit answers to complete fitting the model as follows:

- Question 1: at what stress level, k_1 , will the test item have approximately a φ_1 expected chance of survival?
- Question 2: at what stress level, k_2 , will the test item have approximately a φ_2 expected chance of survival?

With both question 1 and question 2, it is easy to obtain

$$\begin{cases} (1 + \beta k_1^\delta)^{-\alpha} = \varphi_1 \\ (1 + \beta k_2^\delta)^{-\alpha} = \varphi_2 \end{cases} \quad (6.25)$$

Therefore, both β and δ can be expressed in terms of α , so there is only one parameter for the mode.

$$\beta = \left(\varphi_i^{-\frac{1}{\alpha}} - 1 \right) k_i^{-\delta}, i = 1, 2 \quad (6.26)$$

$$\delta = \frac{\ln[(\varphi_1^{-\frac{1}{\alpha}} - 1)/(\varphi_2^{-\frac{1}{\alpha}} - 1)]}{\ln[k_1/k_2]} \quad (6.27)$$

since the prior variance of the survival probability is given by

$$V(p; \alpha, \beta, \delta) = V(\alpha) = (1 + 2\beta k^\delta)^{-\alpha} - (1 + \beta k^\delta)^{-2\alpha} \quad (6.28)$$

The parameter α may be chosen to coincide with the experimenter's prior estimate of the variation at some stress k . Recall the conditional expectations for the posterior, for which two other functions, $g_1(k)$ and $g_2(k)$, must be specified. Then some more historical information and expert information have to be used to estimate the remaining unknown parameters.

6.4. PLANNING FOR ACCELERATED DEGRADATION TESTS

Accelerated Degradation Test (ADT) is applied to address the case with rare failure to provide information about the performance degradation over time. The degradation information is able to provide a more insightful view into the product service to support the product design and evaluation. Ideally, the degradation information can be obtained without damaging the testing units. For example, by changing the head and/or flow rate we can directly measure the degradation of a pump. However, for some applications the properties of the testing units may be changed or damaged by collecting the degradation data. For example, measuring the tensile strength of the polymer material subject to alkaline environment and elevated temperature will damage the test unit. Such degradation tests are called Accelerated Destructive Degradation Tests (ADDT), which can only generate a single measurement for each testing unit. Thus, careful test planning is crucial to extract as much information as possible. The main objectives of ADT test planning are to:

- Determine the appropriate indicator of product performance;
- Determine the accelerated stress levels and the corresponding sample sizes;
- Define the failure in terms of product performance;
- Estimate the duration of the test;
- Determine a suitable model for extrapolation of performance degradation.

The steps of ADT planning are similar to the planning of ALT, although ADT planning is more concerned with performance degradation data and the degradation model to extrapolate the future performance. For applications where performance degrades gradually as usage time increases, failure time T is defined as the time when the degradation level reaches a specified critical level Y . With all the corresponding failure times, the following analysis would be similar to the general ALT methods. Note that the underlying failure mechanisms have to be better understood through the planning of ADT, as it is necessary to track not just failures, but also the changing performance.

Example 6.3

Consider a plan for an accelerated degradation test, where the units tested are subject to decreasing degradation over time. There are n units randomly selected for conducting a degradation experiment at each stress level (Yu, 2006). The testing unit is considered failed once the degradation level reaches the specified transformed threshold y_f . Given the testing unit i under stress level j , the measurements are made for every f units of time until time $t_k = f \cdot k \cdot t_u$, where t_u is a unit of time and k is a positive integer representing a number of measurement. Therefore, the specific test plan can be described in terms of the decision variables: $\varphi = (n, f_j, k_j, t_u)$

The objective is to design an efficient degradation experiment such that the 100p percentile t_p under normal stress condition can be estimated precisely. Suppose that the fitted degradation model is non-linear, as

$$g[y_{ij}(t_{ijk})] = -a_{ij}t_{ijk}^b + \varepsilon_{ij}$$

where $g(\cdot)$ is a suitable function that can be determined by engineering experience, similar product, or pilot studies; y_{ij} is the degradation path of unit i under stress level j ; random t_{ijk} is the k^{th} measurement for unit i under stress level S_j ; $b > 0$ is a fixed and known constant; a_{ij} is a coefficient following a reciprocal Weibull distribution, which means a_{ij}^{-1} is subject to a Weibull distribution with scale parameter α_j and shape parameter β ; and ε_{ij} is the error subject to normal distribution $N(0, \sigma)$.

As the critical threshold is y_f , we can get the product lifetime t :

$$t = \left[-\frac{g(y_f)}{a_{ij}} \right]^{\frac{1}{b}}$$

As $a_{ij}^{-1} \sim \text{Weibull } (\alpha, \beta)$, then $-\ln a_{ij}$ follows the extreme value distribution with location parameter $\ln \alpha$ and scale parameter $\frac{1}{\beta}$. Then it can be seen that t follows a Weibull distribution with scale parameter $\theta_j = \{\alpha_j \cdot [-g(y_f)]\}^{\frac{1}{b}}$ and shape parameter. Thus, the product's lifetime distribution can be expressed as follows:

$$f(t) = b\beta \cdot (\theta_j)^{-\beta} \cdot t^{b\beta-1} \cdot e^{-t^{b\beta} \cdot (\theta_j)^{-\beta}}$$

Thus, the 100p percentile t_p under stress level S_j of the product's lifetime distribution can be expressed as follows:

$$t_{p,j} = \exp \left\{ \ln \theta_j + \frac{1}{\beta} \cdot \ln[-\ln(1-p)] \right\} = \exp \left\{ \frac{1}{b} \ln[-g(y_f)] + \frac{1}{b} \ln \alpha_j + \frac{1}{\beta} \cdot \ln[-\ln(1-p)] \right\}$$

The life-stress relationship is assumed to be Arrhenius, so the scale parameter α_j is a log-linear function of the stress

$$\ln \alpha_j = \gamma_0 + \frac{\gamma_1}{S_j}$$

where γ_0 and γ_1 are unknown parameters to be estimated from the degradation data; and S_j is the stress level.

Therefore, the 100p percentile t_p under normal stress condition S_0 is:

$$t_{p,0} = \exp \left\{ \frac{1}{b} \ln[-g(y_f)] + \frac{\gamma_0}{b} + \frac{\gamma_1}{bS_0} + \frac{1}{\beta} \cdot \ln[-\ln(1-p)] \right\}$$

The joint prior distribution of the three model parameters is denoted by $f(a, b, \beta)$; then we would get the posterior distribution as:

$$f(\beta, \gamma_0, \gamma_1 | t) = \frac{L(t|\beta, \gamma_0, \gamma_1) f(\beta, \gamma_0, \gamma_1)}{\int \int \int L(t|\beta, \gamma_0, \gamma_1) f(\beta, \gamma_0, \gamma_1) d\beta d\gamma_0 d\gamma_1}$$

The likelihood function:

$$l(t) = \prod_i \prod_j \prod_k \left(b\beta \cdot \{\alpha_j \cdot [-g(y_f)]\}^{-\frac{\beta}{b}} \cdot t_{ijk}^{b\beta-1} \cdot e^{-t_{ijk}^{b\beta} \cdot \{\alpha_j \cdot [-g(y_f)]\}^{-\frac{\beta}{b}}} \right)$$

$$\begin{aligned}\Lambda &= \log(L(t|\beta, \gamma_0, \gamma_1)) \\ &= \sum_i \sum_j \sum_k \left(\ln b + \ln \beta - \frac{\beta}{b} (\gamma_0 + \frac{\gamma_1}{S_j}) - \frac{\beta}{b} \ln[-g(y_f)] + (b\beta - 1) \ln t_{ijk} \right. \\ &\quad \left. - t_{ijk}^{b\beta} \cdot \left\{ \exp\left(\gamma_0 + \frac{\gamma_1}{S_j}\right) \cdot [-g(y_f)] \right\}^{-\frac{\beta}{b}} \right)\end{aligned}$$

For illustrative purposes, a simple degradation experiment is designed to explore the performance of an AC power supply. The output current is monitored to indicate its performance. The normal working condition is $S_0=200\text{ }^{\circ}\text{K}$, and the accelerated stresses are $S_1=300\text{ }^{\circ}\text{K}$ and $S_2=500\text{ }^{\circ}\text{K}$. Assume the testing unit fails once the output current drops below 80% of the nominal output current; that is, $y_f=0.8$. The percentile of interest p is 0.01. The unit of inspection time is $t_u=1$ hour.

Based on the engineering experience, the function $g[y_{ij}(t_{ijk})] = 1 - \frac{1}{y_{ij}(t_{ijk})}$, and the constant $b = 0.55$; thus the observed degradation paths can be described by:

$$g[y_{ij}(t_{ijk})] = 1 - \frac{1}{y_{ij}(t_{ijk})} = -a_{ij} t_{ijk}^b + \varepsilon_{ij}$$

For illustrative purposes, let's compare the results with three different test plans:

- Test plan 1: 30 units would be tested at each stress level; inspection frequency for each stress level is 1, 3, and 4 respectively; total number of inspections for each stress level is 300, 150, and 100 respectively.
- Test plan 2: 40 units would be tested at each stress level; inspection frequency for each stress level is 2, 5, and 9 respectively; total number of inspections for each stress level is 250, 130, and 80 respectively.
- Test plan 3: 50 units would be tested at each stress level; inspection frequency for each stress level is 3, 7, and 10 respectively; total number of inspections for each stress level is 200, 100, and 60 respectively.

Assume that the prior information is as follows:

Parameter	Initial Value	Prior Distribution
γ_0	0.65	$0.45 \sim 0.98$
γ_1	1820	$1300 \sim 2200$
β	8.53	$6.42 \sim 9.78$

The likelihood function is

$$\begin{aligned}\Lambda &= \log(L(t|\beta, \gamma_0, \gamma_1)) \\ &= \sum_i \sum_j \sum_l \left\{ \ln b + \ln \beta - \frac{\beta}{b} (\gamma_0 + \frac{\gamma_1}{S_j}) - \frac{\beta}{b} \ln\left(\frac{1}{y_f} - 1\right) + (b\beta - 1) \ln t_{ijk} \right. \\ &\quad \left. - t_{ijk}^{b\beta} \cdot \left[\exp\left(\gamma_0 + \frac{\gamma_1}{S_j}\right) \cdot \left(\frac{1}{y_f} - 1\right) \right]^{-\frac{\beta}{b}} \right\}\end{aligned}$$

The 100p percentile lifetime under normal stress condition is

$$t_{p,0} = \exp \left\{ \frac{1}{b} \ln \left(\frac{1}{y_f} - 1 \right) + \frac{\gamma_0}{b} + \frac{\gamma_1}{bS_0} + \frac{1}{\beta} \cdot \ln[-\ln(1-p)] \right\}$$

We now run the simulations for 200 runs with 60,000 iterations. One can observe the difference between the expectation of posterior variance of each plan from Table 6.1.

Table 6.1: Expected variance of each plan

Test Plan No.	Expectation of Variance of $t_{p,0}$ (hour)
1	6.87×10^{235}
2	6.29×10^{288}
3	4.90×10^{267}

REFERENCES

- Martz, H.F. & Waterman, M.S. (1978). A Bayesian model for determining the optimal stress for a single test unit. *Technometrics*, 20(2), 179-185.
- Meeker, W. Q., & Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. New York: John Wiley & Sons.
- Meeker, W., & Hahn, G. (1985). *How to Plan an Accelerated Life Test*. Milwaukee: ASQC Quality Press.
- Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. New York: John Wiley & Sons.
- Nelson, W. (2005). A Bibliography of Accelerated Test Plans. *IEEE Transaction on Reliability*, 54(2), 194-197.
- Nelson, W. B. (1990). *Accelerated Testing - Statistical Models, Test Plans, and Data Analysis*. Hoboken, NJ: John Wiley and Sons.
- Paris, P., Gomez, M., & Anderson, W. (1961). A rational analytic theory of fatigue. *The Trend in Engineering*, 13, 9-14.
- Singpurwalla, N.D. (2008). *Reliability and Risk: A Bayesian Prospective*. England: John Wiley & Sons.
- Weibull, W. (1959). Statistical evaluation of data from fatigue and creep rupture tests, Part I: fundamental concepts and general methods. Sweden: Wright Air Development Center.
- Weibull, W. (1961). *Fatigue testing an analysis of results*. London: Pergamon Press.
- Yu, H.F. (2006). Designing an accelerated degradation experiment with a reciprocal Weibull degradation rate. *Journal of Statistical Planning and Inference*, 136, 282-297.
- Yuan, T., Liu, X., & Kuo, W. (2012). Planning simple step-stress accelerated life tests using Bayesian methods, *IEEE Transactions on Reliability*, 61(1), 254-263.

Chapter 7: Accounting for Uncertainties and Model Validation

7.1. INTRODUCTION

At the heart of Bayesian analysis are two key tenets that allow information to be inferred upon unknowns of interest: the *evidence* the engineer gathers from a process or test, and the *model* the engineer selects to represent the process or test data. The presence of uncertainties in data and the model do not preclude their use – rather these uncertainties simply need to be properly accounted for using probabilistic relationships. It is here that the skill of the reliability engineer in recognizing and describing sources of uncertainty becomes paramount: analyzing data and models with inherent uncertainty requires the nature of the uncertainty to be understood.

The *unknowns of interest* (such as the parameters that describe time to failure of a unit or amount of damage at a given point in time) dictate subsequent evidence, E , through an appropriate inference model. If this evidence (such as the amount of cumulative damage like crack depth at a given time) is not directly or accurately observable, the physical process of generating it can introduce uncertainty. In such a case, the *observable evidence*, \hat{E} , has a probabilistic relationship that is conditional on the true or underlying *evidence*, E . The subsequent observation process can then introduce additional uncertainty if the process is imperfect. In this case, the observable evidence, \hat{E} , has a probabilistic relationship that is conditional on the *observation*, O .

It is crucial to understand these probabilistic relationships, and the “plane of observation” that is defined by the location of \hat{E} between the observer and the item that separates each source of uncertainty. This plane separates uncertainty into two broad categories in the context of evidence gathering. The first category is that *uncertainty in the physical process (such as the failure mechanism)* involving what is called “*classical error*”: the probability of producing the *observable evidence* conditional on the true underlying evidence— $Pr(\hat{E} | E)$ (Carroll & Stefanski (1990), Fuller (1987), Mallick, Hoffman, & Carroll (2002)). The second category is uncertainty in the *observation process* involving what is called “*Berkson error*”: the probability of the *observable evidence* given what is observed – $Pr(\hat{E} | O)$ (Carroll & Stefanski (1990), Mallick, Hoffman, & Carroll (2002), Darby & Feam (1999)).² These two sources of uncertainty need to be incorporated differently for Bayesian inference to be generated.

7.2. UNCERTAINTIES IN EVIDENCE

The reliability engineer must understand which parts of the uncertainty in his or her observations stem from classical and Berkson errors. Each error must be incorporated differently in the ways outlined below.

7.2.1. CLASSICAL ERROR: UNCERTAINTY IN THE PHYSICAL PROCESS

The incorporation of classical error represents a minor conceptual addition to a Bayesian construct that involves no uncertainty. Recalling that Bayes’ theorem is:

² It is not uncommon for literature to write Berkson error in isolation without considering classical error. Subsequently, Berkson error is often written as $Pr(E | \hat{E})$ where E becomes synonymous with the true underlying evidence and \hat{E} represents what is observed. The taxonomy used in this textbook differs from this representation to preserve consistency, but represents the same fundamental principle.

$$\pi_1(\theta|E) = \frac{L(E|\theta)\pi_0(\theta)}{\int_{\theta} L(E|\theta')\pi_0(\theta')d\theta'} \quad (7.1)$$

where θ is the set of unknowns of interest, E is the evidence, $\pi_0(\theta)$ is the joint probability distribution of the unknowns of interest (that describes a prior state of knowledge), $\pi_1(\theta|E)$ is the joint probability distribution that describes an updated state of knowledge after Bayesian inference from the evidence, and $L(E|\theta)$ is the likelihood function describing the probability of observing the evidence given a particular set of unknowns of interest.

If there is uncertainty in the physical process, then classical error needs to be understood and defined. Consider a corrosion ADT involving a chemical test kit that involves inserting a specially coated paper into an unknown oil solution containing corroded contaminants. The objective of the test kit is to identify the presence of a certain contaminant. If the contaminant is present, the paper is designed to turn red. Conversely, if there is no contaminant the paper is designed to turn green. However, the test kit is not perfect. When the paper is supposed to turn red (i.e. the contaminant is present), there is a 10 percent chance that it will incorrectly turn green. Similarly, when the paper is supposed to turn green (i.e. there is no contaminant), there is a 5 percent chance that the paper will incorrectly turn red.

In this scenario, the true or underlying evidence of corrosion, E , refers to the presence (or otherwise) of the contaminant. The observable evidence, \hat{E} , is the color the test kit paper turns. Therefore, the following classical error can be defined:

$$\Pr(\hat{E}|E) = \begin{cases} & \left| \begin{array}{ll} E = \text{'contaminant present'} & E = \text{'contaminant not present'} \\ \hat{E} = \text{'red'} & 0.9 & 0.05 \\ \hat{E} = \text{'green'} & 0.1 & 0.95 \end{array} \right. \end{cases} \quad (7.2)$$

Classical error is incorporated into Bayesian analysis using a methodology referred to as “likelihood in terms of observation” (Tan & Xi, 2003). The likelihood function in Equation (7.1) is modified to incorporate the observable evidence as demonstrated in the case where there is a finite set of possible, underlying evidence:

$$L(\hat{E}|\theta) = \sum_E \Pr(\hat{E}|E)L(E|\theta) \quad (7.3)$$

where $L(\hat{E}|\theta)$ represents the probability of the observable evidence being generated given a particular set of unknowns of interest. Equation (7.3) can be generalized as follows:

$$L(\hat{E}|\theta) = \int_E f(\hat{E}|E')L(E'|\theta)dE' \quad (7.4)$$

where $f(\hat{E}|E')$ is the PDF that is the continuous generalization of the classical error. Equations (7.3) and (7.4) may be substituted into Equation (7.1) to generate a modified representation of Bayes’ Theorem in Equation (7.5).

$$\pi_1(\theta|\hat{E}) = \begin{cases} \text{finite set of possible evidence:} \\ \frac{\pi_0(\theta) \sum_E \Pr(\hat{E}|E') L(E'|\theta)}{\sum_{\theta'} \pi_0(\theta') \sum_E \Pr(\hat{E}|E') L(E'|\theta')} \\ \text{infinite set of possible evidence:} \\ \frac{\pi_0(\theta) \int_E f(\hat{E}|E') L(E'|\theta) dE'}{\int_{\theta'} \pi_0(\theta') [\int_E f(\hat{E}|E') L(E'|\theta') dE'] d\theta'} \end{cases} \quad (7.5)$$

Equation (7.5) is a “natural” extension of Bayes’ theorem as represented in Equation (7.1), primarily because it simply follows the laws of probability (a quality that is not shared when Berkson error is involved, as outlined later on in this textbook).

Example 7.1

A test engineer is conducting an accelerated test to establish the performance level of a bearing in a system. The unknown of interest is the performance of the bearing, x : 0 if the bearing performs above the minimum required level, 1 if it is close to violating the requirement and 2 if it has fallen below the performance requirement. As the system is functional at the time of testing, the test engineer assigns the following prior information to the unknown of interest, x : that is,

$$\pi_0(\theta) = \pi_0(x) = \begin{cases} 0.75 & x = 0 \\ 0.2 & x = 1 \\ 0.05 & x = 2 \end{cases} \quad (7.6)$$

If the bearing is close to violating the performance requirement, there is an 80 percent probability that it will release its lubricant into a liquid coolant. If it has already violated the requirement, then this chance increases to 100 percent. The lubricant is the test kit contaminant of Equation (7.2).

$$L(E|x) = \begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline E = \text{'contaminant present'} & 0 & 0.8 & 1 \\ E = \text{'contaminant not present'} & 1 & 0.2 & 0 \end{array} \quad (7.7)$$

The test engineer needs to find posterior distribution values of x .

Solution

The evidence, E , is the existence (or otherwise) of the contaminant (lubricant) in the solution (coolant fluid). Upon extracting a sample of coolant, the test engineer places the test paper into it, and it turns red. That is,

$$\hat{E} = \text{'red'} \quad (7.8)$$

Substituting Equations (7.6) and (7.8) into Equation (7.5) yields:

$$\pi_1(\theta|\hat{E}) = \pi_1(x|\hat{E}) = \frac{\pi_0(x) \sum_E \Pr(\hat{E}|E) L(E|x)}{\sum_x \pi_0(x') \sum_E \Pr(\hat{E}|E') L(E'|x')} \quad (7.9)$$

Considering the summation in Equation (7.9), and substituting in Equations (7.7) and (7.2) yields

$$\sum_E \Pr(\hat{E}|E) L(E|x) \quad (7.10)$$

$$\begin{aligned}
&= \sum_{\forall E} \left[\begin{array}{c|cc} & E = \text{'contaminant present'} & E = \text{'contaminant not present'} \\ \hline \hat{E} = \text{'red'} & 0.9 & 0.05 \\ \hat{E} = \text{'green'} & 0.1 & 0.95 \end{array} \right] \\
&\quad \times \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline E = \text{'contaminant present'} & 0 & 0.8 & 1 \\ E = \text{'contaminant not present'} & 1 & 0.2 & 0 \end{array} \right] \\
&= \left[\begin{array}{c|cc} & E = \text{'contaminant present'} \\ \hline \hat{E} = \text{'red'} & 0.9 \\ \hat{E} = \text{'green'} & 0.1 \end{array} \right] \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline E = \text{'contaminant present'} & 0 & 0.8 & 1 \end{array} \right] \\
&+ \left[\begin{array}{c|cc} & E = \text{'contaminant not present'} \\ \hline \hat{E} = \text{'red'} & 0.05 \\ \hat{E} = \text{'green'} & 0.95 \end{array} \right] \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline E = \text{'contaminant not present'} & 1 & 0.2 & 0 \end{array} \right] \\
&= \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0 & 0.72 & 0.9 \\ \hat{E} = \text{'green'} & 0 & 0.08 & 0.1 \end{array} \right] \\
&+ \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.05 & 0.01 & 0 \\ \hat{E} = \text{'green'} & 0.95 & 0.19 & 0 \end{array} \right] = \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.05 & 0.73 & 0.9 \\ \hat{E} = \text{'green'} & 0.95 & 0.27 & 0.1 \end{array} \right]
\end{aligned}$$

Multiplying Equation (7.10) by Equation (7.6) yields:

$$\begin{aligned}
&\pi_0(x) \sum_E \Pr(\hat{E}|E) L(E|x) \\
&= \begin{cases} 0.75 & x=0 \\ 0.20 & x=1 \\ 0.05 & x=2 \end{cases} \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.05 & 0.73 & 0.9 \\ \hat{E} = \text{'green'} & 0.95 & 0.27 & 0.1 \end{array} \right] \\
&= \begin{cases} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.0375 & 0.146 & 0.045 \\ \hat{E} = \text{'green'} & 0.7125 & 0.054 & 0.005 \end{cases}
\end{aligned} \tag{7.11}$$

Summing Equation (7.11) for all possible x provides the denominator in Equation (7.9):

$$\begin{aligned}
&\Sigma_x \pi_0(x') \sum_E \Pr(\hat{E}|E') L(E'|x') dx' = \sum_{\forall x} \left[\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.0375 & 0.146 & 0.045 \\ \hat{E} = \text{'green'} & 0.7125 & 0.054 & 0.005 \end{array} \right] \\
&= \begin{cases} & x=0 \\ \hline \hat{E} = \text{'red'} & 0.0375 \\ \hat{E} = \text{'green'} & 0.7125 \end{cases} + \begin{cases} & x=1 \\ \hline \hat{E} = \text{'red'} & 0.0146 \\ \hat{E} = \text{'green'} & 0.054 \end{cases} + \begin{cases} & x=2 \\ \hline \hat{E} = \text{'red'} & 0.045 \\ \hat{E} = \text{'green'} & 0.005 \end{cases}
\end{aligned} \tag{7.12}$$

$$= \begin{cases} \hat{E} = \text{'red'} & | 0.2285 \\ \hat{E} = \text{'green'} & | 0.7715 \end{cases}$$

Substituting Equations (7.12) and (7.11) into Equation (7.9) yields:

$$\begin{aligned} \pi_1(x|\hat{E}) &= \frac{\pi_0(x) \sum_E \Pr(\hat{E}|E) L(E|x)}{\sum_{x'} \pi_0(x') \sum_E \Pr(\hat{E}|E') L(E'|x')} \\ &= \frac{\begin{array}{c|ccc} & x=0 & x=1 & x=2 \\ \hline \hat{E} = \text{'red'} & 0.0375 & 0.146 & 0.045 \\ \hat{E} = \text{'green'} & 0.7125 & 0.054 & 0.005 \end{array}}{\begin{array}{c|c} \hat{E} = \text{'red'} & 0.2285 \\ \hat{E} = \text{'green'} & 0.7715 \end{array}} \\ &\approx \begin{array}{c|cc} & \hat{E} = \text{'red'} & \hat{E} = \text{'green'} \\ \hline x=0 & 0.1641 & 0.9235 \\ x=1 & 0.6390 & 0.0700 \\ x=2 & 0.1969 & 0.0065 \end{array} \end{aligned} \quad (7.13)$$

The fact that the test engineer observed red test paper means that the posterior distribution of the unknowns of interest is given by the “ $\hat{E} = \text{'red'}$ ” column in Equation (7.13). This means that there is a 16.41 percent chance that the bearing is above performance requirement, a 63.90 percent chance that the bearing is close to violating the requirement, and a 19.69 percent chance that the bearing has already gone below the performance requirement. Even though it is not possible for a performing bearing to leak lubricant into the system (i.e. when $x = 0$), there is still a residual probability that in this instance of observing a red paper that the bearing is indeed functional (which is 16.41 percent as calculated above). This is due to the inherent uncertainty associated with the test process—or the classical error.

The incorporation of *classical error* is as much a question of perception as it is an incorporation of uncertainty. In the case of the test kit paper outlined in the examples above, it is entirely proper to also consider the color the paper turns as the “true evidence” and incorporate the uncertainty (through the probabilistic relationship on the presence of the contaminant or not) in the construction of the likelihood function. Equations (7.3) and (7.4) represent how such a likelihood function would conceptually be constructed.

The additional benefit of how classical error adheres to the laws of probability is that it incorporates multiple levels of forms of evidence between the true or underlying evidence. If there are n such levels with the n^{th} and final level being the observable evidence, then classical error can be written as shown in Equation (7.14):

$$L(\hat{E}_i|\hat{E}_{i-1}) = \begin{cases} \text{finite set of possible observable } (i-1)^{\text{th}} \text{ evidence:} \\ \sum_{\hat{E}_{i-1}} \Pr(\hat{E}_i|\hat{E}_{i-1}) L\left(\hat{E}_{i-1} \middle| \begin{array}{l} E \text{ if } i = 2 \\ \hat{E}_{i-2} \text{ if } i \geq 3 \end{array}\right) \\ \text{infinite set of possible observable } (i-1)^{\text{th}} \text{ evidence: ... for all } i \geq 2 \\ \int_{\hat{E}_{i-1}} f(\hat{E}_i|\hat{E}'_{i-1}) L\left(\hat{E}_{i-1} \middle| \begin{array}{l} E \text{ if } i = 2 \\ \hat{E}_{i-2} \text{ if } i \geq 3 \end{array}\right) d\hat{E}'_{i-1} \end{cases} \quad (7.14)$$

7.2.2. BERKSON ERROR: UNCERTAINTY IN THE OBSERVATION PROCESS

As opposed to uncertainty in the physical process and its associated classical error, Berkson error stems from uncertainty in the observation process itself. In essence, uncertainty in the observation process means that one can only guess what the actual observable evidence was based on the observation of the inspector, sensor or other measuring technique. That is, the probabilistic relationship between the two is expressed as a conditional probability of observable evidence (\hat{E}) occurring, conditional on the observation, $O - \Pr(\hat{E}|O)$.

There are two ways of dealing with Berkson error in a Bayesian construct. The first involves constructing the posterior distribution, $\pi_1(\theta|E)$, as a weighted average of all possible observable evidence posterior distributions, $\pi_1(\theta|\hat{E})$, where weighting is described by Berkson error, $\Pr(\hat{E}|O)$. This method is Jeffrey's rule of probability kinematics (Jeffrey, 1965), as described in Equation (7.15):

$$\pi_1(\theta|O) = \begin{cases} \text{finite set of possible observable evidence:} \\ \sum_{\hat{E}} \pi_1(\theta|\hat{E}) \Pr(\hat{E}|O) \\ \text{infinite set of possible observable evidence:} \\ \int_E \pi_1(\theta|\hat{E}') f(\hat{E}'|O) d\hat{E}' \end{cases} \quad (7.15)$$

where $f(\hat{E} | O)$ is the pdf that is the continuous generalization of Berkson error.

The second way of dealing with Berkson error differs from Jeffrey's rule of probability kinematics in that instead of constructing the posterior distribution as a weighted average, the likelihood function is weighted. This methodology is referred to as Cheeseman's rule (Cheeseman, 1986) and is represented in Equation (7.16):

$$L(O|\theta) = \begin{cases} \text{finite set of possible observable evidence:} \\ \sum_{\hat{E}} L(\hat{E}|\theta) \Pr(\hat{E}|O) \\ \text{infinite set of possible observable evidence:} \\ \int_E L(\hat{E}'|\theta) f(\hat{E}'|O) d\hat{E}' \end{cases} \quad (7.16)$$

Example 7.2

Consider the test engineer from Example 7.1. He has a mild form of color-blindness, which means that some shades of green and red appear grey to him. As he sometimes cannot discern between colors on the test paper, he writes down when the paper appears grey to him. In this case, his observation (O) is an element of the set {red; green; grey}. However, it is known through multiple tests that 40 per cent of the time he observes grey when the actual color is green, and the remaining 60 per cent of the time the actual color is red. It is unknown under what conditions the colors red or green only become apparent to the test engineer as grey, meaning this uncertainty cannot be expressed as a probability that is conditional on the actual color, ruling out the use of classical error. The only applicable error becomes Berkson error:

$$\Pr(\hat{E}|O) = \begin{array}{c|ccc} & O = \text{'green'} & O = \text{'grey'} & O = \text{'red'} \\ \hline \hat{E} = \text{'green'} & 1 & 0.4 & 0 \\ \hat{E} = \text{'red'} & 0 & 0.6 & 1 \end{array} \quad (7.17)$$

Consider the scenario in the previous example, except in this case the test engineer observes the color grey on the test paper. He decides to incorporate this uncertainty by utilizing Jeffrey's rule. Substituting Equations (7.17) and (7.13) into Equation (7.15) yields:

$$\begin{aligned} \pi_1(\theta|O) &= \pi_1(x|O) = \sum_{\hat{E}} \pi_1(x|\hat{E}') \Pr(\hat{E}'|O) \\ &= \sum_{\forall \hat{E}} \left\{ \begin{array}{c|cc} & \hat{E} = \text{'red'} & \hat{E} = \text{'green'} \\ \hline x = 0 & 0.1641 & 0.9235 \\ x = 1 & 0.6390 & 0.0700 \\ x = 2 & 0.1969 & 0.0065 \end{array} \right\} \left\{ \begin{array}{c|ccc} & O = & O = & O = \\ \hline \hat{E} = \text{'green'} & \text{'green'} & \text{'grey'} & \text{'red'} \\ \hat{E} = \text{'red'} & 1 & 0.4 & 0 \\ & 0 & 0.6 & 1 \end{array} \right\} \\ &= \left\{ \begin{array}{c|c} & \hat{E} = \text{'red'} \\ \hline x = 0 & 0.1641 \\ x = 1 & 0.6390 \\ x = 2 & 0.1969 \end{array} \right\} \left\{ \begin{array}{c|ccc} & O = \text{'green'} & O = \text{'grey'} & O = \text{'red'} \\ \hline \hat{E} = \text{'red'} & 0 & 0.6 & 1 \end{array} \right\} \\ &+ \left\{ \begin{array}{c|c} & \hat{E} = \text{'green'} \\ \hline x = 0 & 0.9235 \\ x = 1 & 0.0700 \\ x = 2 & 0.0065 \end{array} \right\} \left\{ \begin{array}{c|ccc} & O = \text{'green'} & O = \text{'grey'} & O = \text{'red'} \\ \hline \hat{E} = \text{'green'} & 1 & 0.4 & 0 \end{array} \right\} \quad (7.18) \\ &= \left\{ \begin{array}{c|ccc} & O = & O = & O = \\ \hline & \text{'green'} & \text{'grey'} & \text{'red'} \\ \hline x = 0 & 0 & 0.0985 & 0.1641 \\ x = 1 & 0 & 0.3834 & 0.6389 \\ x = 2 & 0 & 0.1182 & 0.1969 \end{array} \right\} + \left\{ \begin{array}{c|ccc} & O = & O = & O = \\ \hline & \text{'green'} & \text{'grey'} & \text{'red'} \\ \hline x = 0 & 0.9235 & 0.3694 & 0 \\ x = 1 & 0.0700 & 0.0280 & 0 \\ x = 2 & 0.0065 & 0.0026 & 0 \end{array} \right\} \\ &= \left\{ \begin{array}{c|ccc} & O = \text{'green'} & O = \text{'grey'} & O = \text{'red'} \\ \hline x = 0 & 0.9235 & 0.4679 & 0.1641 \\ x = 1 & 0.0700 & 0.4114 & 0.6389 \\ x = 2 & 0.0065 & 0.1208 & 0.1969 \end{array} \right\} \end{aligned}$$

can be seen that using Jeffrey's rule to incorporate the uncertainty introduced by the test engineer's color blindness and the uncertainties in the physical process examined within the previous examples, if he observes a grey test paper there is a 46.79 percent chance that the bearing is functional, a 41.14 percent chance that the bearing is close to failure, and a 12.08 percent chance that it has already failed.

The key to incorporating uncertainty in the observation process is correctly describing Berkson error. This requires the reliability engineer to understand the probabilistic influences at play. In so doing, there are two specific sources of uncertainties that the reliability engineer needs to be aware of: systematic and stochastic uncertainties. Systematic uncertainties can be thought of as inherent biases in the observation process. From the test kit example, Equation (7.17) demonstrates that the test

engineer has a slight bias for observing red. Stochastic uncertainties can be thought of as inherent inaccuracies in the observation process. Similarly, the fact that Equation (7.17) represents a probabilistic relationship at all demonstrates this fact.

7.2.2.1. Systematic Uncertainties

Systematic uncertainty is a particularly useful concept when considering the challenge of incorporating expert opinion, non-calibrated sensors, or inaccurate measuring tools. The aspirational “perfect” expert is one whose Berkson error involves no bias or uncertainty. That is,

$$\Pr(\hat{E}|O) = \begin{cases} 1 & \text{if } \hat{E} = O \\ 0 & \text{if } \hat{E} \neq O \end{cases} \quad (7.19)$$

Introducing a bias or systematic uncertainty then creates a precise but imperfect expert. That is,

$$\Pr(\hat{E}|O) = \begin{cases} 1 & \text{if } \hat{E} = O + \bar{X} \\ 0 & \text{if } \hat{E} \neq O + \bar{X} \end{cases} \quad (7.20)$$

where \bar{X} is a constant error term (possibly in matrix form) that represents the expert’s bias.

7.2.2.2. Stochastic Uncertainties

Sensor measurements, human-based inspections and expert opinion not only involve bias, they also involve random variability and inaccuracy. A useful way of characterizing stochastic uncertainty is by using a Gaussian distribution if the uncertainty stems from an additive error process (as described by the central limit theorem) or lognormal distribution when the error process is multiplicative. Depending on the number of parameters or variables involved in \hat{E} , a multivariate Gaussian distribution can be incorporated. Should the variable in question be a scalar value, then the expert opinion’s Berkson error can be expressed as:

$$\Pr(\hat{E}|O) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{O-(\hat{E}+\bar{X})}{\sigma}\right)^2} \quad (7.21)$$

where σ is the standard deviation of the Gaussian distribution that is used to characterize the inaccuracy and randomness in the observation.

If the observation randomness and inaccuracy stems from a multiplicative error process, then the lognormal distribution should be used to characterize the expert opinion’s Berkson error. That is,

$$\Pr(\hat{E}|O) = \frac{1}{O\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln O - (\ln \hat{E} + \ln \bar{X})}{\sigma}\right)^2} \quad (7.22)$$

7.2.2.3. Relationship between Berkson and Classical Errors

In spite of a natural desire to classify uncertainty in the observation process as classical error to prevent violation of the laws of probability, this should not be done if the stochastic nature does not permit it. If uncertainty in the observation process can be described by classical error, then what in fact is occurring is an extension in the physical process: what was the observation, O , is in fact the observable evidence, \hat{E} , and the methodology outlined in the previous section should be applied. Fundamentally knowing the probability of \hat{E} conditional on O (Berkson error) does not of itself mean that the engineer can ascertain the probability of O conditional on \hat{E} (which can be rewritten as classical error).

For example, consider the color-blind engineer operating the test kit discussed in Example 7.2. In attempting to establish the probability of his observation, O , given the observable evidence, \hat{E} , the discrete form of Bayes' theorem must be employed:

$$\begin{aligned}\Pr(O = o | \hat{E}) &= \frac{\Pr(\hat{E} | O = o) \Pr(O = o)}{\Pr(\hat{E})} \\ &= \frac{\Pr(\hat{E} | O = o) \Pr(O = o)}{\sum_{o' \in \{\text{'green'}, \text{'red'}\}} \Pr(\hat{E} | O = o') \Pr(O = o')}\end{aligned}\quad (7.23)$$

To solve Equation (7.23), the engineer needs to know $\Pr(O = o)$ or the probability that he will observe either red or green conditional on nothing else. This probabilistic relationship simply sits outside the framework of understanding in such a scenario: you cannot ascertain the probability of any particular observation of a test unless it is itself conditional on the test. More simply, there is no generalized rule that permits Berkson error to be correctly represented by classical error.

This is not to say that there are no scenarios where Berkson error can be used to infer classical error, but such a transformation can only occur if the underlying probabilistic influence is understood. In the case of the example that looks at the bullet-detecting machine on the rifle range and its associated calibration experiment, it is equally valid to generate a probabilistic construction of the observation based on the observable evidence, as the relevant metric is displacement. When considering displacement, one can trivially see that A is as far away from B as B is from A. This then allows the perspective of the engineer to change (if so desired) to treat the estimate of the machine to become the observable evidence, and strictly incorporate classical error. Such reciprocal relationships do not always exist, however, as the above example illustrates.

7.3. PPOF MODEL UNCERTAINTIES, ERRORS, AND VALIDATION

It is not unusual for PPoF or mechanistic model selection to itself contain an inherent level of uncertainty, such as models for crack propagation through structural metals. Models are continually developed and updated as materials engineers improve their understanding of microscopic material behavior. Sometimes, these models compete in terms of the underlying physics, and the reliability engineer may not be certain which one is the most appropriate to use.

In such a scenario, the range of models used by the reliability engineer, and any prior information pertaining to their relevance, can be viewed as part of his or her state of knowledge. That is,

$$UOI = \{M_j, \theta_j\}_{j=1,2,3,\dots,m} \quad (7.24)$$

where M_j is the j^{th} model that the reliability engineer is considering, and θ_j is the set of parameters associated with the model M_j .

The prior knowledge is then represented by the following pdfs:

$$\Pr_0(M_j) \dots j = 1, 2, 3, \dots, m \quad (7.25)$$

$$\pi_0(\theta_j | M_j) \dots j = 1, 2, 3, \dots, m \quad (7.26)$$

If there is some characteristic, measure or function of the unknowns of interest (in this case denoted as α) then the model uncertainty can be inherent in establishing its pdf. That is,

$$f(x \mid \{\Pr(M_j), \pi(\theta_j \mid M_j)\}_{j=1,2,3,\dots,m}) = \sum_{j=1}^m \left[\Pr(M_j) \int_{\theta_j} \Pr(x \mid M_j, \theta'_j) d\theta'_j \right] \quad (7.27)$$

There are several ways in which Bayesian analysis can update the reliability engineer's understanding of model uncertainty. They typically exist within the same generic framework. Consider the evidence set E , which consists of a number of independent observations.

$$E = \{e_1, e_2, e_3, \dots, e_n\} \quad (7.28)$$

We define two subsets of E : E_1 and E_2 . The selection of these subsets defines the methodology that is adopted, and will be examined in detail below. Consider the following pdf:

$$f(E_2 \mid E_1, M_j) = \int_{\theta_j} L(E_2 \mid M_j, \theta'_j) \pi_1(\theta'_j \mid M_j, E_1) d\theta'_j \quad (7.29)$$

This pdf is known as a *predictive density*. The term $\pi_1(\theta'_j \mid M_j, E_1)$ is the updated or posterior distribution of the j^{th} model's parameters based on the information contained in the evidence subset E_1 . It is formally defined as:

$$\pi_1(\theta_j \mid M_j, E_1) = \frac{L(E_1 \mid M_j, \theta_j) \pi_0(\theta_j \mid M_j)}{\int_{\theta_j} L(E_1 \mid M_j, \theta'_j) \pi_0(\theta'_j \mid M_j) d\theta'_j} \quad (7.30)$$

Substituting Equation (7.30) into Equation (7.29) yields:

$$f(E_2 \mid E_1, M_j) = \frac{\int_{\theta_j} L(E_2 \mid M_j, \theta'_j) L(E_1 \mid M_j, \theta'_j) \pi_0(\theta'_j \mid M_j) d\theta'_j}{\int_{\theta_j} L(E_1 \mid M_j, \theta'_j) \pi_0(\theta'_j \mid M_j) d\theta'_j} \quad (7.31)$$

The distribution in (7.31) is used to compare the relative merits of each model.

Bayes Factors (BF)

If E_1 is set to a null set, \emptyset , which implies that E_2 is the entire set E , then Equation (7.31) becomes simplified. The ratio of the resultant predictive densities for two models defines the Bayes factor (Gelfand & Dey, 1993):

$$BF_{1,2} = \frac{f(E \mid M_1)}{f(E \mid M_2)} = \frac{\int_{\theta_1} L(E \mid M_1, \theta'_1) \pi_0(\theta'_1 \mid M_1) d\theta'_1}{\int_{\theta_2} L(E \mid M_2, \theta'_2) \pi_0(\theta'_2 \mid M_2) d\theta'_2} \quad (7.32)$$

In essence, the Bayes factor is the ratio of the probabilities of observing the evidence, E , for two competing models based on prior distributions of each model's parameters. In other words, a Bayes factor is the number of times more probable that the evidence is observed for the first model when compared to the second model.

A conceptual shortcoming of Bayes factors is that the predictive distributions are only based on the prior distribution of each model's parameters. A model with poorly selected prior distributions for its parameters is naturally disadvantaged when subjected to this comparison. Table 7.1 provides Jeffreys' proposed guide on how the Bayes factor can be interpreted for model comparison of models M_1 and M_2 (Jeffreys, 1961).

Table 7.1: Jeffreys' Guideline for Bayes factor of Model M₁ vs. Model M₂

Bayes Factor	Strength of Evidence in Favor of Model M ₁
<1	Negative (support model M ₂)
1-3	Insignificant
3-10	Substantial
10-30	Strong
30-100	Very Strong
>100	Decisive

Pseudo Bayes Factors (PsBF)

The problem with strictly using the posterior understanding of model parameters can be overcome by using the pseudo Bayes factors (Geisser & Eddy, 1979). In this case, the predictive distribution considers one evidence component at a time, allowing all other evidence components to inform the predictive distribution:

$$PsBF_{1,2} = \frac{\prod_{i=1}^n f(E_i | \bar{E}_i, M_1)}{\prod_{i=1}^n f(E_i | \bar{E}_i, M_2)} \quad (7.33)$$

where $E_i = \{e_i\}$ and $\bar{E}_i = \{e_1, e_2, \dots, e_{i-1}, e_{i+1}\}$.

Posterior Bayes Factors (PoBF)

Posterior Bayes factors are based on the predictive distributions that have been updated by the entire evidence (Aitkin, 1991). That is,

$$PoBF_{1,2} = \frac{f(E | E, M_1)}{f(E | E, M_2)} \quad (7.34)$$

P-value

Another method for assessing model validity is using the p-value. The p-value is the probability that the result from a probabilistic process is at least as extreme as what is observed. For example, if the p-value is 0.05, then there is a 5 percent chance that a more extreme outcome will result when compared to what was observed. The p-value is utilized often in probability and statistics (Hubbard & Lindsay, 2008). In Bayesian analysis, the reliability engineer must select a statistic whose probability can be predicted by a model and estimated directly from evidence (Box, 1980). Examples of such statistics include mean, percentiles and other characteristics. If this statistic, denoted by X , is expressed by the evidence-based estimate denoted by \hat{x} , the following equations can be written:

$$\hat{x} = g(E) \quad (7.35)$$

where g is some function that defines the statistic.

$$X_j = f_j(x | M_j, \theta_j) \quad (7.36)$$

which defines the predictive distribution of the statistic for the j^{th} model.

Example 7.3

Consider a simple scenario often used to demonstrate basic Bayesian model comparison and validation. The example generated below is similar to others, such as the one outlined in Geweke (2005). Consider an inspection process during an ADT to detect existence of an initial crack using an acoustic signal

amplitude when exceeding a preset level over constant intervals of time. The result would be one of two states: {0,1}. Consider the evidence gathered from 20 inspection intervals showing the following results:

$$E = \{0,0,0,1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0\} \quad (7.37)$$

The test engineer is considering two models. The first is based on a binomial process where the probability that a particular state is observed is set. The second model is based on the transition process between states where there is a probability that the next interval will be different from the previous one:

$$M_1 \rightarrow \Pr(s) = \begin{cases} p_1 & \text{if } s = 0 \\ 1 - p_1 & \text{if } s = 1 \end{cases} \quad (7.38)$$

where s is the existence of an initial crack at the given inspection interval.

$$M_2 \rightarrow \Pr(s_i|s_{i-1}) = \begin{cases} p_2 & \text{if } s_i = s_{i-1} \\ 1 - p_2 & \text{if } s_i \neq s_{i-1} \end{cases} \quad (7.39)$$

Compare the two models using the Bayes factor method and estimate the parameters of the models.

Solution

For the first model, the likelihood function is based on the binomial distribution:

$$L(E|M_1, p_1) = \binom{20}{10} p_1^{10} (1 - p_1)^{10} \quad (7.40)$$

The likelihood function for the second model is also based on the binomial distribution, but instead of the probability of being based on a given state, it is based on transitions from 0 to 1. For the evidence set E , there are 19 cases that exercise the ability for state transition (it is unknown what state existed before the first evidence element). Therefore, there are two observed state transitions on subsequent demands, and 17 observed demands when the state did not transition. The likelihood function then becomes:

$$L(E|M_2, p_2) = \binom{19}{17} p_2^{17} (1 - p_2)^2 \quad (7.41)$$

Assuming uniform prior distributions for p_1 and p_2 over [0,1],

$$\pi_0(p_1) = \begin{cases} 1 & \text{if } p_1 \in [0,1] \\ 0 & \text{otherwise} \end{cases} \quad (7.42)$$

$$\pi_0(p_2) = \begin{cases} 1 & \text{if } p_2 \in [0,1] \\ 0 & \text{otherwise} \end{cases} \quad (7.43)$$

Substituting Equations (7.40) and (7.41) into Equation (7.32) yields:

$$BF_{1,2} = \frac{\int_{p_1} L(E|M_1, p'_1) \pi_0(p'_1|M_1) dp'_1}{\int_{p_2} L(E|M_2, p'_2) \pi_0(p'_2|M_2) dp'_2} = \frac{0.0476}{0.0500} = 0.9524 \quad (7.44)$$

This suggests that there is a slight preference for the second model, which makes conceptual sense, as it appears as if there are long “runs” of consecutive states in the evidence set. However, this assessment is based on uniform prior distributions of unknowns of interest p_1 and p_2 .

To investigate the effect of updating the models based on the information in the evidence set, Bayesian analysis is useful. As the likelihood function is based on the binomial distribution, the beta distribution

provides a “conjugate prior.” The posterior distribution in such a case is also a beta distribution, meaning that only the parameters need to be updated (sparing complex calculation):

$$X \sim Beta(\alpha, \beta) \rightarrow f_{\beta}(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad (7.45)$$

where β is the beta function.

Further, a uniform distribution (which was previously selected as the prior) is a special case of the beta distribution where both parameters are equal to 1. Exploiting these facts, we get:

$$\pi_1(p_1|E) = \frac{1}{B(11,11)} p_1^{10} (1-p_1)^{10} \quad (7.46)$$

$$\pi_1(p_2|E) = \frac{1}{B(18,3)} p_2^{17} (1-p_2)^2 \quad (7.47)$$

It can be seen from Equation (7.47) that the model M_2 suggests that there is a high probability of the system “dwelling” in the same state at consecutive observations.

Substituting into Equations (7.31) and (7.34) yields:

$$PoBF_{1,2} = \frac{f(E|E, M_1)}{f(E|E, M_2)} = \frac{1.447 \times 10^{-5}}{2.339 \times 10^{-5}} = 0.6188 \quad (7.48)$$

It can be seen that as the models are updated, the second model is becoming more supported. Should a single model not be selected, one can incorporate each updated model, assigning a probability to each and utilizing Equation (7.27). The probability of each is proportional to $f(E|E, M)$ that was calculated in Equation (7.48). Completing Bayesian analysis yields:

$$\Pr(M_1) = \frac{f(E|E, M_1)}{\sum_i f(E|E, M_i)} = \frac{1.447 \times 10^{-5}}{1.447 \times 10^{-5} + 2.339 \times 10^{-5}} = 0.3823 \quad (7.49)$$

for $\pi_1(p_1|E)$

and

$$\Pr(M_2) = \frac{f(E|E, M_2)}{\sum_i f(E|E, M_i)} = \frac{2.339 \times 10^{-5}}{1.447 \times 10^{-5} + 2.339 \times 10^{-5}} = 0.6177 \quad (7.50)$$

for $\pi_1(p_2|E)$

7.4. APPLICATIONS OF MODEL VALIDATION IN ADT

This section discusses the validation process for the models developed to address the physical degradation of systems and/or components. The main activity of model validation is to compare the model predictions to the experimental results, where the difference between them is the model error.

The accuracy of each model can be qualitatively indicated with a scatter plot like Figure 7.1. It is assumed that there are n sample units, and the corresponding degradation is measured once at each time j . The horizontal axis D_e represents the degradation measurement of unit i , and D_m is the vertical axis to denote model prediction of degradation for unit i . Both $D_{m,ij}$ and $D_{e,ij}$ are the specified degradation for unit i at time j . $Y_{t,ij}$ is the true degradation of interest. The deviation between $D_{m,ij}$ and $Y_{t,ij}$ is known as the *error of model prediction*, and the deviation between $D_{e,ij}$ and $Y_{t,ij}$ represents *error of experiment results*. Only when there is no difference between the model prediction and experiment result can the point fall on the dashed diagonal line. Through the model validation process, both the

uncertainty in model predictions and experiment results are taken into account to estimate the uncertainty of the true damage of interest. The model validation process varies based on the type of model errors chosen (such as additive or multiplicative error models).

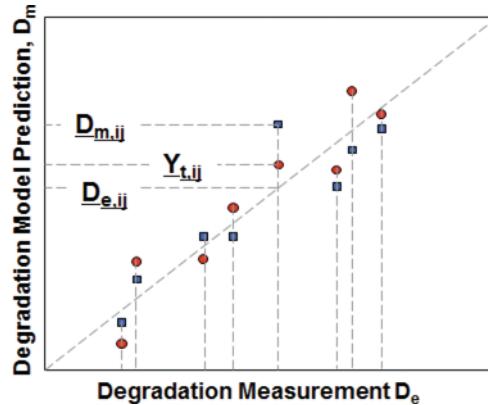


Figure 7.1: Degradation model prediction D_m vs. degradation measurement D_e

Additive Error Model

The additive error model is the most commonly used, which assumes that the errors are subject to addition or aggregation. The introduced error can be negative, zero, or positive, and is described by a Gaussian distribution to represent the difference between the underlying degradation and observable degradation.

$$\begin{aligned} Y_{t,ij} &= D_{e,ij} + E_{e,ij} \quad ; \quad E_e \sim N(b_e, s_e) \\ Y_{t,ij} &= D_{m,ij} + E_{m,ij} \quad ; \quad E_m \sim N(b_m, s_m) \end{aligned} \quad (7.51)$$

where E_e is the error of experiment results, which is normally distributed with the mean b_e and standard deviation s_e ; E_m is the error of model prediction, which is normally distributed with the mean b_m and standard deviation s_m .

When comparing the model error and experimental error, it is easy to obtain the expression for $E_{me,i}$ as the additive error between model prediction and experiment result.

$$E_{me} = (E_m - E_e) \sim N(b_m - b_e, \sqrt{s_m^2 + s_e^2}) \quad (7.52)$$

As the experimental uncertainty can be retrieved from an independent source, a Bayesian updating process can be used to update the prior state of knowledge for a joint distribution of b_m and s_m :

$$\begin{aligned} \pi_1(b_m, s_m | D_{e,ij}, D_{m,ij}, b_e, s_e) \\ = \frac{\pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m)}{\iint_{s_m, b_m} \pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m) db_m ds_m} \end{aligned} \quad (7.53)$$

where $\pi_0(b_m, s_m)$ is the prior joint distribution of the parameters; $\pi_1(b_m, s_m | D_{e,ij}, D_{m,ij}, b_e, s_e)$ is the posterior joint distribution of the parameters; and $L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m)$ is the likelihood function, as

$$\begin{aligned} L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m) \\ = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \sqrt{s_m^2 + s_e^2}} e^{-\frac{1}{2} \left[\frac{D_{m,ij} - D_{e,ij} - (b_m - b_e)}{s_m^2 + s_e^2} \right]^2} \end{aligned} \quad (7.54)$$

The normal distribution assumption is good if the variable of interest is much larger than zero and lightly scattered. Widely scattered data will generate negative lower bounds without any realistic physical interpretation. Moreover, the uncertainty would be reduced when the data is mainly located at lower ranges or higher ranges.

Multiplicative Error Model

In the multiplicative error model, the errors are assigned with lognormal distribution to address the shortcomings of the normal distribution assumption. This model assumes that the model prediction, experiment results, and true damage are all positive or negative. The errors can be represented mathematically as follows:

$$\begin{aligned} \frac{Y_{t,ij}}{D_{e,ij}} &= F_{e,ij}; F_e \sim LN(b_e, s_e) \\ \frac{Y_{t,ij}}{D_{m,ij}} &= F_{m,ij}; F_m \sim LN(b_m, s_m) \end{aligned} \quad (7.55)$$

where F_e is the multiplicative error of experiment results, which is lognormally distributed with the logarithmic mean b_e and logarithmic standard deviation s_e ; F_m is the error of model prediction, which is lognormally distributed with logarithmic mean b_m and logarithmic standard deviation s_m .

With the two equations above, we have $\frac{D_{e,ij}}{D_{m,ij}} = \frac{F_{m,ij}}{F_{e,ij}} = F_{t,ij}$, which is the factor representing the ratio of the multiplicative error of experiment with respect to the model prediction. Given the independency of F_m and F_e , $F_{t,ij}$ is also subject to lognormal distribution F_t with mean $b_m - b_e$, and standard deviation $\sqrt{s_m^2 + s_e^2}$.

The posterior joint distribution of parameters b_m and s_m can be expressed by:

$$\begin{aligned} \pi_1(b_m, s_m | D_{e,ij}, D_{m,ij}, b_e, s_e) \\ = \frac{\pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m)}{\iint_{s_m, b_m} \pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m) db_m ds_m} \end{aligned} \quad (7.56)$$

where $\pi_0(b_m, s_m)$ is the prior joint distribution of the parameters; $\pi_1(b_m, s_m | D_{e,ij}, D_{m,ij}, b_e, s_e)$ is the posterior joint distribution of the parameters; and $L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m)$ is the likelihood function, as:

$$\begin{aligned} L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m) \\ = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \left(\frac{D_{e,ij}}{D_{m,ij}} \right) \sqrt{s_m^2 + s_e^2}} e^{-\frac{1}{2} \left[\frac{\ln \left(\frac{D_{e,ij}}{D_{m,ij}} \right) - (b_m - b_e)}{s_m^2 + s_e^2} \right]^2} \end{aligned} \quad (7.57)$$

Note that the distribution of model prediction error can also be used to infer the distribution of the true damage of interest: $Y_{t,ij} = F_{m,ij} \cdot D_{m,ij}$. Therefore, the actual damage $Y_t \sim LN(\ln(D_m) + b_m, s_m)$.

Example 7.4

Consider an experiment performed to study the crack initiation in a standardized specimen using acoustic emission (AE) and optical microscopy measurement (Keshtgar and Modarres, 2013), for which a probabilistic model for small fatigue cracks was developed:

$$a(N) = \alpha \cdot I(N) + \beta \quad (7.58)$$

where $a(N)$ indicates the true length of the small crack after N loading cycles, $I(N)$ is the calculated AE intensity at cycle, and N , α and β are the unknown model parameters.

A qualitative validation of the developed model was performed using graphical comparisons between model predictions and experimental data, as shown in Figure 7.2.

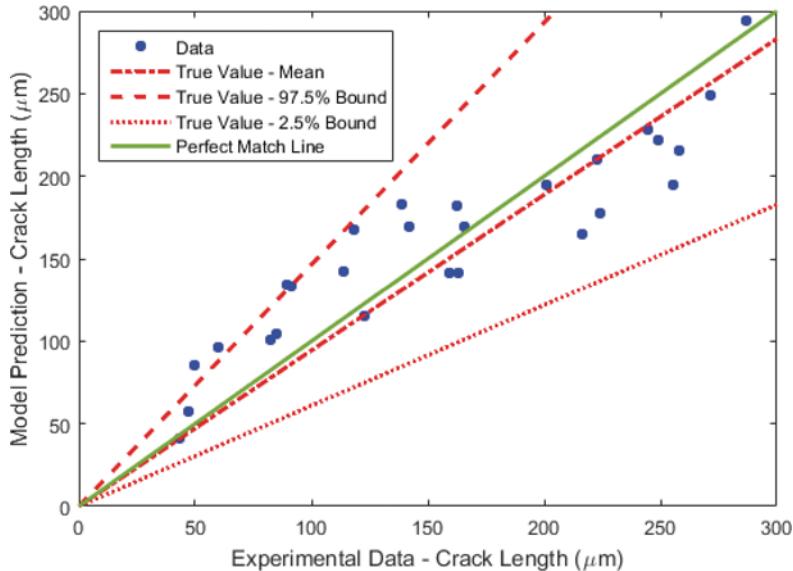


Figure 7.2: Comparison of AE model prediction and experimental results

An advanced simulation-based methodology can be employed to validate the model based on a Bayesian estimation approach. The developed model is validated using a new set of data that was not used in model development process.

With multiplicative error of the model prediction, the likelihood of the observed data considering the probability of detection (POD), measurement errors, measured and model-estimated crack is

$$\begin{aligned} L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m, m, s) \\ = \prod_{i=1}^n \left(POD(D_{e,ij} | m, s) \right) \left(\frac{1}{\sqrt{2\pi} \left(\frac{D_{e,ij}}{D_{m,ij}} \right) \sqrt{s_m^2 + s_e^2}} \right) e^{-\frac{1}{2} \frac{\left[\ln \left(\frac{D_{e,ij}}{D_{m,ij}} \right) - (b_m - b_e) \right]^2}{s_m^2 + s_e^2}} \end{aligned} \quad (7.59)$$

where $POD(D_{e,ij}|m, s) = \frac{e^{\frac{\pi}{\sqrt{3}}(\frac{\log D_{e,ij}-m}{s})}}{1+e^{\frac{\pi}{\sqrt{3}}(\frac{\log D_{e,ij}-s}{s})}}$ with $m \cong 27$ and $s \cong 9$. The mean of experimental error b_e is 0.027, and the corresponding standard deviation s_e is 0.0818.

The posterior joint distribution of b_m and s_m is expressed as

$$\begin{aligned} \pi_1(b_m, s_m | D_{e,ij}, D_{m,ij}, b_e, s_e) \\ = \frac{\pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m)}{\iint_{s_m, b_m} \pi_0(b_m, s_m) \cdot L(D_{e,ij}, D_{m,ij}, b_e, s_e | b_m, s_m) db_m ds_m} \end{aligned} \quad (7.60)$$

where $\pi_0(b_m, s_m)$ is the joint prior distribution of b_m and s_m , which is characterized with a prior distribution as lacking of information.

The true crack length model prediction then can be estimated by:

$$Y_t = F_m \cdot D_m \quad (7.61)$$

which can be estimated by a lognormal distribution as

$$Y_t \sim LN(\ln(D_m) + b_m, s_m) \quad (7.62)$$

Figure 7.3 illustrates the modified prediction results and the corresponding uncertainty bounds. The results show that the developed AE-model reasonably estimates the true crack length. While the results show that the developed model slightly over-predicts the crack lengths, such small bias can be addressed with the inclusion of the model error factor F_m . Therefore, the developed model can be used to appropriately get an unbiased estimate of the true crack length distribution.

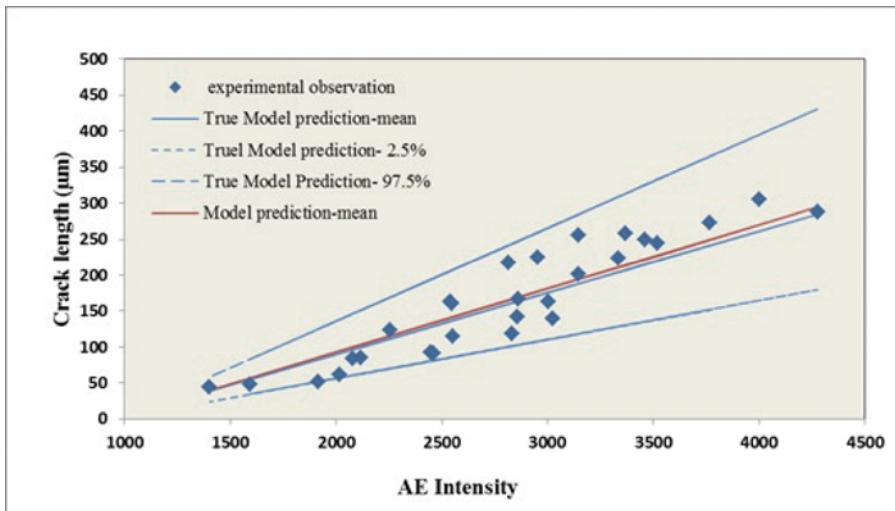


Figure 7.3: Model prediction with multiplicative error

REFERENCES

- Aitkin, M. Posterior Bayes Factors. *Journal of the Royal Statistics Society*, 1991, 1(B), 111-142.
- Box, G. Sampling and Bayes Inference in Scientific Modelling and Robust-ness. *Journal of the Royal Statistics Society*, 1980, B(143), 383-430.
- Carrol, R. and Stefanski, L. Approximate quasi-likelihood estimation in models with surrogate predictors. *JASA*, 1990(September), 652-663.
- Cheeseman, P., ed. Probabilistic versus fuzzy reasoning. (Elsevier, Dordrecht, 1986).
- Darby, S. and Feam, T. Berkson error model. In Kotz, S., CB, R. and DL, B., eds. *Encyclopedia of statistical sciences updates* (Wiley, New York, 1999).
- Fuller, W. Measurement error models. (Wiley, New York, 1987).
- Geweke, J. *Contemporary Bayesian Econometrics and Statistics*. (Wiley, Hoboken, 2005)
- Jeffrey, R. *The logic of decision*. (University of Chicago Press, Chicago, 1965).
- Jeffreys, H. *Theory of Probability*. (Oxford University Press, Oxford, 1961).
- Keshtgar A. and Modarres, M., Acoustic Emission-Based Fatigue Crack Growth Prediction, Proceedings of the 2013 Reliability and Maintainability Symposium (RAMS), Jan. 2013, Orlando, FL.
- Mallick, B., Hoffman, F. and Carroll, R. Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada test site. *Biometrics*, 2002, 58(1), 13-20.
- Tan, Z. and Xi, W. Bayesian analysis with consideration of data uncertainty in a specific scenario. *Reliability Engineering & System Safety*, 2003, 79(1), 17.
- Geisser, S. and Eddy, W. A predictive approach to model selection. *Journal of American Statistics Association*, 1979(74), 153-160.
- Gelfand, A.E. and Dey, D.K. Bayesian Model Choice: Asymptotics and Exact Calculations. In Solomon, P.H., ed (Department of Statistics, Stanford University, Stanford, California, 1993).
- Hubbard, R. and Lindsay, R. Why P Values Are Not a Useful Measure of Evidence in Statistical Significance Testing. *Theory & Psychology*, 2008, 18(1), 69-88.

Also of Interest

Check out these published and forthcoming titles in the Performability Engineering Series

Network Reliability

Measures and Evaluation

By Sanjay K. Chaturvedi

Published 2016. ISBN: 9781119223566

Machine Tool Reliability

By Bhupesh Kumar Lad, Divya Shrivastava and Makarand S. Kulkarni Published 2016. ISBN 978-1-119-03860-3

Binary Decision Diagrams and Extensions for Systems Reliability Analysis

By Suprasad Amari and Liudong Xing Published 2015. ISBN 978-1-118-54937-7

Quantitative Assessments of Distributed Systems

Methodologies and Techniques

By Dario Bruneo and Salvatore Distefano Published 2015. ISBN 978-1-118-59521-3

Fundamentals of Reliability Engineering

By Indra Gunawan

Published 2014. ISBN 978-1-118-54956-8

Building Dependable Distributed Systems

By Wenbing Zhao

Published 2014. ISBN 978-1-118-54943-8

Index

A

Accelerated Aging, 80
 Accelerated Corrosion
 Damage, 92
 Testing, 94
 Accelerated Degradation Modeling, 10
 Accelerated Test
 Conceptual, 9
 Data, 2, 117
 High Stress, 235
 Model, 239
 Multivariable, 121
 Qualitative, 10
 Accelerated Testing
 Stresses, 239
 Acceleration Factor
 Adjusted, 183
 Arrhenius model, 140
 Dual-Stress Life-Stress Model, 160
 Eyring Model, 149
 Inverse Power Law, 153
 Power-Exponential Model, 164
 Temperature-Humidity, 191
 Acceptance Test, 103
 Additive Error Model, 209, 210, 264
 agent of failure, 122, 124, 127
 Aging
 Failure Mechanisms, 1
 Archard's Wear Law, 60, 62, 63, 65, 79
 Arrhenius
 Acceleration Factor, 140
 eibull, 141
 Equation, 2, 80
 Exponential Life-Stress Model, 139
 Inverse Power Law, 169
 Least Squares Regression, 126
 Life Stress, 131
 Lognormal, 152
 Model, 125, 126, 128, 130, 139, 149, 169
 Relationship, 120, 126, 218, 240, 242
 Weibull, 142, 143
 Autocorrelation, 208, 209

B

Bathtub
 Curve, 103, 107
 Bayes factor

Posterior, 261
 Pseudo, 261
 Bayes' Theorem, 171, 174, 251, 253, 259
 Bayesian Approach, 175, 219, 239, 243
 Bayesian Estimation
 ADT Model Parameters, 219
 Method, 124, 172
 Prior Distribution, 172
 Probability Distributions, 172
 Bayesian Inference
 ADT Model Parameters, 219
 Expression, 199
 Method, 196
 Step-Stress Data Analysis, 196
 Bayesian Regression, 7
 Bearings
 Ball, 35, 69, 70, 72
 Life Models, 69
 Radial Ball, 70
 Roller, 69, 70
 Beta Distribution, 173, 262, 263
 Binomial Distribution
 Likelihood Function, 262
 Burn-in
 Process, 103
 Screening Test, 105
 Test, 103
 Burn-in Test, 103, 104

C

Censored Data
 lnterval, 132
 Left, 119
 Multiply, 119
 Right, 118, 131, 132, 137, 149, 154, 160, 193, 197, 198
 Coefficient of Friction, 60, 75, 78, 79
 Coefficient of Variation, 203
 Coffin-Manson Relationship, 28
 Combined Loading, 21
 Component Reliability, 10
 Conditional
 CDF, 192
 Combined Model, 140
 Distribution of Life, 123
 Dual-Stress-Exponential Life-Stress Model, 162
 Exponential Life-Stress, 140
 Eyring Life-Stress, 151
 Gibbs Sampling, 198

- Life-Stress Model, 149
 pdf, 133
 Power Life-Stress, 154
 Power-Lognormal, 156
 Probability, 256
Confidence Interval
 Eyring-Weibull Model, 150
 Local Fisher, 162
 Model Parameters, 164
 One-sided, 135
 Two-sided, 135, 141, 142, 148, 150, 153, 155, 157, 161, 163, 165, 167
 Confidence Level, 134, 135, 146, 233
 Confidence Limits, 135
 Correlation, 7, 57, 124, 209
Corrosion
 Crevice, 13, 16, 91, 92
 Current Density, 92, 93
 Damage, 91, 92
 Erosion, 92
 Failure, 91, 106
 Fatigue, 58
 Galvanic, 16, 91, 92
 Life Prediction Model, 91
 Mass Loss, 95
 Pitting, 13, 16, 91, 92
 Rate, 91, 92, 93, 94, 95, 96
 Stress Corrosion Cracking, 91
 Testing, 91
 Uniform, 16, 92
Covariance
 Matrix, 134, 135, 207, 209
Crack
 Initiation, 15, 23, 47, 48, 56, 58, 59, 237
 Propagation, 15, 23, 47, 49, 51, 59, 85, 92, 110, 237, 259
 Crack Growth, 2, 23, 47, 48, 49, 51, 52, 54, 56, 57, 92, 110, 113, 180
Creep
 Behavior, 81, 82, 83, 86
 Data, 82
 Deformation, 87
 Failure, 109
 Logarithmic, 86
 Mechanisms, 2
 Parabolic, 86
 Range, 57
 Strain, 81, 82, 83, 85, 86, 87, 89
 Test, 81
Cumulative Damage
 Behavior, 109
 Distribution, 110
 Effect, 179, 182
 Cycles to Failure, 17, 18, 19, 28, 34, 35, 47, 50, 92, 113, 137
- D**
- Damage
 Agents, 15
 Distribution, 5, 121
 Models, 35, 37
 Values, 5
 Damage-Endurance Model, 5, 109
Degradation
 Mechanism, 4, 6, 7, 90, 236
 Process, 7, 178, 205, 237
 Testing, v, 8, 204, 205
Degrees of Freedom, 243
Destruct Limit
 Upper and Lower, 10
Detection
 of Degradation, 203
 Probability of, 203
 Screens, 106
Distribution
 Common Life, 124, 136
 Gamma Prior, 243
 Joint Posterior, 174, 198
 Lognormal, 112, 125, 126, 137, 138, 146, 151, 155, 156, 158, 161, 165, 174, 212, 258, 265, 267
 Posterior, 171, 172, 198, 255, 256, 260, 263
 Prior, 171, 172, 173, 174, 198, 239, 240, 241, 242, 247, 267
 Representative Life, 1
 Stress and Strength, 109, 111
 Uniform, 174, 225, 263
 Weibull, 117, 135, 136, 137, 140, 149, 153, 154, 164, 169, 170, 172, 183, 184, 189, 192, 196, 212, 242
- E**
- Empirical**
 Data, 38
 Exponents, 63
 Model of Impact, 67
 Relationships, 60
 Wear Models, 64
Environmental
 factors, 3, 6, 57, 95
 inter- and intra-, 3
 Stress Screen, 104
Error
 Berkson, 251, 253, 256, 257, 258, 259
 Classical, 251, 252, 255, 256, 258, 259
 factor, 267
Estimation
 Bayesian, 124, 171, 172, 175, 207, 227, 266
 Maximum likelihood, 124
 Model Parameter, 123
 of the ADT Parameters, 219
Expert

Bias, 258
 Elicitation, 239, 245
 Imperfect, 258
 Opinion, 173, 258
 Perfect, 258
 Exponential Distribution, 133, 136
 Extreme Value Distribution, 247
 Eyring
 Life-Stress Model, 148
 Lognormal Model, 151
 Model, 149, 151, 159, 169
 Relationship, 120, 148, 149, 159
 Weibull Model, 149

F

Failure Analysis, 178
 Failure Mechanisms
 Aging, 1
 Fatigue, 57
 Modeling, 1
 PPoF model of, 6
 Failure Mode
 High-Stress Induced, 103
 Masked, 234, 235
 of Interest, 118, 234
 Principal, 103
 Real-Life, 106
 Secondary, 131
 Time Dependent, 106
 Failure Model
 Mechanistic-Based, 1
 Failure Models
 Time-to-, 173
 Failure rate
 Actual, 76
 Instantaneous, 103
 Failure Rate
 Base, 76
 Parameter, 136
 fatigue
 High-Cycle, 57, 59
 Fatigue
 Low-Cycle, 57, 58, 98, 120, 157
 Fatigue test
 Long Life, 38
 Fatigue Test
 Accelerated, 237
 Axial Tension, 57
 Cost Effective, 57
 of Smooth Specimens, 41
 Fisher Information Matrix
 Local, 141, 148, 150, 152, 155, 157, 161, 162, 164
 Fracture Mechanics
 Linear Elastic, 48

G

Gaussian Distribution, 114, 258
 Gerber Relationship, 20
 Goodman Relationship, 20, 23
 goodness of fit tests, 117, 124
 Graphical Method, 124
 G_{rms} , 102

H

HALT, 10, 102, 104, 105, 106, 107, 181
 Hardware
 Components, 15
 Units, 103
 HASS
 Test, 106
 Hazard Function, 169
 Hazard Rate
 Baseline, 169
 Function, 140, 154, 162

I

Infant Mortality
 Failures, 10
 Hazard Rate, 103
 Inherent Uncertainty, 251

J

Joint Distribution
 Life-Stress, 242
 of Parameters, 174, 265
 Joint pdf, 175, 209

L

Larson-Miller theory, 82
 LEFM, 48, 49, 53
 Life-Stress Method, 17, 18, 19, 23
 Likelihood Function, 131, 132, 133, 136, 154, 172, 174, 197, 198, 219, 239, 252, 255, 256, 264, 265
 Linear Regression, 94, 124, 125, 128, 129
 Location Parameter, 247
 Log-Likelihood Function
 Eyring-Weibull, 149
 Power-Lognormal, 156
 Power-Weibull, 154
 Lognormal Distribution
 Log-Likelihood Function, 138, 156
 Median, 112, 146, 151, 155, 161, 165
 Standard Deviation, 138
 Lower Destruct Limit (LDL), 10
 Lubricated Contact, 79

M

- Manson-Haford Theory, 82, 83
 Maximum Likelihood Estimation
 For Acceleration Degradation and Model Parameters, 209
 For ALT Data Analysis, 131
 For Step-Stress Data Analysis, 192
 Method, 124, 131, 192
 Point Estimate, 134
 Miner's Rule, 35, 37
 Miscellaneous Factor, 59
 Mission Profile, 6
 Model
 Selection, 259
 Uncertainties, 6, 8, 234
 Morrow Relationship, 20
 Multiplicative Error Model, 264, 265

N

- Normal Distribution
 As Additive Error Model, 209
 CDF, 134
 Cumulative Standard, 109
 For Large Sample Sizes, 134, 141, 142, 148, 150, 153
 Multivariate, 207
 Notch
 Effect, 37, 38, 46
 Factor, 38, 39, 41, 42

O

- Observation, 29, 68, 105, 118, 196, 197, 208, 239, 251, 252, 256, 257, 258, 259
 Operating Limit, 15, 105, 106
 Oxidation Life, 80

P

- Parameter Estimation
 Approach to ALT Models, 171
 Based on Expert Elicitation, 245
 MLE-Based, 136, 139
 Model, 123
 Paris Equation, 49
 Performance Measurements, 113, 204
 Performance-Requirement Model, 110
 Peterson Equation, 38
 PH Model, 169, 175
 Physics of Failure, 2, 117, 130, 178
 Planning
 ADT, 246
 ALT, 118, 237, 238
 Polarization Resistance, 93
 Power Law Model, 56, 184

Prior Distribution

- Joint, 198, 241, 247, 267
 Negative-Log Gamma, 243
 Non-Informative, 242
 Parameters, 171, 225
 Uniform, 172, 173, 262
Prior Information, 172
Probabilistic Physics of Failure, 1
Probability Density Function, 109
Probability Distribution, 7, 8, 117, 119, 124, 131, 134, 136, 137, 171, 172, 175, 252
Probability Plots, 125, 128, 130

Q

- Qualitative Tests**, 10, 101, 107, 181
Quantitative Tests, 10, 101

R

- Random**
 Environmental Variations, 6
 Loading, 113
 Sample, 242
 Vibration, 6
Random Stress, 108
Random Variable, 1, 5, 111, 112, 133, 137, 203, 209, 243
Regression Analysis, 130
Reliability Assessment, 6, 234
Reliability Engineering, 137
Reliability Function
 Expression, 151, 154, 156
 Lognormal, 147, 151, 156, 162, 166
 PH-Weibull Model, 170
 Power-Lognormal, 156
Reliability Growth, 101
Root Cause Analysis, 106

S

- Safety**
 Factors, 108, 109
 Margin, 105, 109, 110
 Margins, 47
Sample Size
 Normal Distribution for Large, 134, 141, 142, 148, 150, 153
Scale Parameter
 Weibull, 140, 149, 153, 160, 164, 196, 241
 Weighted Average, 125, 128
Scenario
 Alternate Notch, 54
 Amplitude Loading, 37
 Common Basic Fatigue, 38
Scenarios
 Amplitude Loading, 33

- Seals, 16, 60, 73, 74
 Sensitivity Analysis, 234
 Sequence Effect, 34
Shape Parameter
 Common, 125, 128
 Confidence Interval, 136
 Generic, 238
 Weibull, 144
Shock
 Mechanical, 102
 Thermal, 75
Smith-Watson-Topper, 20, 29
S-N
 Diagram, 18, 20, 21
 Line, 18
 Test Data, 18
Soderberg Relationship, 20
Standard Deviation
 Logarithmic, 265
Step-Stress
 Accelerated, 179
 ALT Analysis, 182
 Continuum, 180
 Data Analysis, 182
 Pattern, 181
 Test, 179, 180, 182, 183, 192, 237
Stern-Gary Model, 93
Strain
 Elastic, 25, 28
 Engineering, 24
 Plastic, 23, 28, 29, 44, 46, 57, 81
 True, 24, 25
Strain-Life Model, 46
Stress
 Concentration Factor, 37, 38, 39, 40, 41, 42, 158
 Corrosion Cracking, 91, 92
 Effective, 21, 22, 23, 36
 Intensity Factor, 2, 49, 51, 53, 92, 159
 Mean, 19, 20, 21, 23, 29, 32, 33, 36, 56
 True, 17, 24, 25, 43
Stress Amplitude, 20, 21, 23, 34, 59, 242
Stress-Strength Model, 4, 108
Stribeck Curve, 79
- T**
- Temperature Cycling, 9, 103, 104
 Temperature Stress, 139
 Test Duration, 179, 180
 Test Equipment, 179
 Thermal Expansion, 3
 Truncated Distribution
 Normal, 113
 Type 1 Censoring, 119
- Type 2 Censoring, 119
- U**
- Uncertainty
 Aleatory, 6
 Epistemic, 6
Uniform Prior Distribution, 172, 173, 262
Unknowns of Interest, 175, 251, 252, 255, 259, 262
- V**
- Variable Amplitude, 33, 34, 35, 37, 56
Variance
 Asymptotic, 239
 Posterior, 241, 242, 249
 Prior, 245
 Sample, 242
Vibration
 Cycling, 104
 Harmonic, 102
 Induced Fatigue, 3
 Physical, 104
 Random, 6, 102
 Stress, 234
Voltage
 Margining, 105
 Mean, 5
 Sequencing, 105
 Stress, 183
- W**
- Walker Relationship, 20
Warranty
 Of Time, 134
 Return, 101
Warranty Levels, 123
Wear
 Abrasive, 16
 Adhesive, 61, 64, 66, 74, 75, 79, 80
 Impact, 16
 Rolling, 68
 Sliding, 64
Weibull distribution
 PH, 170
Weibull Distribution
 Hazard Function, 169
 Joint Conditional Exponential Life-Stress, 140
 Joint Life-Stress Model, 164
 PH, 170