Bridging the Gap: Selected Problems in Model Specification, Estimation, and Optimal Design from Reliability and Lifetime Data Analysis

Caleb Bridges King

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Statistics

Yili Hong, Chair
Stephanie P. DeHart
Xinwei Deng
G. Geoffrey Vining

March 23, 2015
Blacksburg, Virginia

Key Words: Accelerated lifetime testing, ADDT, Competing risks, Fatigue testing, Lifetime data, Product reliability.

Bridging the Gap: Selected Problems in Model Specification, Estimation, and Optimal Design from Reliability and Lifetime Data Analysis

Caleb Bridges King

Abstract

Understanding the lifetime behavior of their products is crucial to the success of any company in the manufacturing and engineering industries. Statistical methods for lifetime data are a key component to achieving this level of understanding. Sometimes a statistical procedure must be updated to be adequate for modeling specific data as is discussed in Chapter 2. However, there are cases in which the methods used in industrial standards are themselves inadequate. This is distressing as more appropriate statistical methods are available but remain unused. The research in Chapter 4 deals with such a situation. The research in Chapter 3 serves as a combination of both scenarios and represents how both statisticians and engineers from the industry can join together to yield beautiful results.

After introducing basic concepts and notation in Chapter 1, Chapter 2 focuses on lifetime prediction for a product consisting of multiple components. During the production period, some components may be upgraded or replaced, resulting in a new "generation" of component. Incorporating this information into a competing risks model can greatly improve the accuracy of lifetime prediction. A generalized competing risks model is proposed and simulation is used to assess its performance.

In Chapter 3, optimal and compromise test plans are proposed for constant amplitude fatigue testing. These test plans are based on a nonlinear physical model from the fatigue literature that is able to better capture the nonlinear behavior of fatigue life and account for effects from the testing environment. Sensitivity to the design parameters and modeling assumptions are investigated and suggestions for planning strategies are proposed.

Chapter 4 considers the analysis of ADDT data for the purposes of estimating a thermal index. The current industry standards use a two-step procedure involving least squares regression in each step. The methodology preferred in the statistical literature is the maximum likelihood procedure. A comparison of the procedures is performed and two published datasets are used as motivating examples. The maximum likelihood procedure is presented as a more viable alternative to the two-step procedure due to its ability to quantify uncertainty in data inference and modeling flexibility.

Dedication

To my beautiful wife, Lindsay, whose personal sacrifice and overwhelming support made this work possible.

Acknowledgments

There are so many people to whom I am grateful for their support and assistance on my journey. I'd like to thank my chair, Dr. Yili Hong, for the kindness and helpful advice given during my many years of research. I'd also like to thank the other members of my advisory committee: Dr. Stephanie P. DeHart for her insightful comments regarding statistical practice in industry, and Drs. Xinwei Deng and G. Geoffrey Vining for providing wisdom from experience. Acknowledgement is certainly required for my fellow graduate students at all levels, who in attempting to understand my work have provided me many opportunities for understanding it myself. I would especially like to thank my family, who continually encouraged me throughout the long process. I would also like to acknowledge Advanced Research Computing (ARC) at Virginia Tech (URL: http://www.arc.vt.edu), whose computational resources allowed for a speedy completing of this dissertation. A portion of this work was funded by NSF Grant CMMI-1068933 and the 2011 DuPont Young Professor Grant. I am exceedingly grateful for their financial support. A portion of this work was Finally, saving the best for last, I'd like to thank my wonderful new wife Lindsay, who sacrificed so much of herself to come alongside and serve as my motivation and the greatest companion one could ever have, and my God, Who was with me from the beginning and saw me through to the end.

Contents

List of Figures

List of Tables

Chapter 1 General Introduction

In this chapter, a brief introduction will be given to the field of lifetime data analysis with specific emphasis on reliability methodology. Any necessary terminology and concepts required for understanding the material in the following chapters will also be introduced and discussed.

1.1 Lifetime Data Analysis and Reliability

Lifetime data analysis is the branch of statistics and probability theory in which the primary variable of interest is the time until some event. This event is most commonly terminal, such as death or failure, but need not always be as is the case in recurrent events. Furthermore, "time" need not necessarily be associated with clock or calendar time as it can also be associated with steps in a sequence or runs of a process. Thus, lifetime data analysis is actually quite flexible in its application. Lifetime data analysis can be especially useful in obtaining reliability information on industrial products and systems. A general and useful introduction to the methods used in analyzing reliability data can be found in the textbooks by Meeker and Escobar (1998) and Lawless (2003).

1.1.1 Accelerated Failure Time and Log-Location Scale Models

For the purposes of this dissertation, T will denote the continuous time to failure of a randomly selected product. This random variable has a survival function $S(t;\mathbf{x},\boldsymbol{\theta}) = \Pr(T > t|\mathbf{x},\boldsymbol{\theta})$ which may be dependent on external factors $\mathbf{x}$ and governed by parameters $\boldsymbol{\theta}$. The cumulative distribution function (cdf) of T is the complement $F(t;\mathbf{x},\boldsymbol{\theta}) = 1 - S(t;\mathbf{x},\boldsymbol{\theta})$ of the survival function and the probability density function (pdf) of T is given by $f(t;\mathbf{x},\boldsymbol{\theta}) = \mathrm{d}F(t;\mathbf{x},\boldsymbol{\theta})/\mathrm{dt} = -\mathrm{d}S(t;\mathbf{x},\boldsymbol{\theta})/\mathrm{dt}$.

There are several classes of models used in analyzing lifetime data (see the monograph

by Bagdonavicius and Nikulin (2010) for a full discussion). In this dissertation, the class of models will be restricted to that of accelerated failure time (AFT) models. These models assume that any external factors that affect the lifetime of a product do so by "accelerating" the lifetime. This can expressed in the survival function as

$$S(t; \mathbf{x}, \boldsymbol{\theta}) = G[r(t, \mathbf{x}); \boldsymbol{\theta}] \tag{1.1}$$

where G is a base survival function and r is some specified function. In this dissertation, G is specified to be a member of the log-location-scale family of distributions. These models are very common in reliability due to their ease of use and interpretation. Specifically, a lifetime distribution is log-location-scale if its pdf or cdf can be written as

$$f(t) = \frac{1}{\sigma t}\phi\left[\frac{\log(t) - \mu}{\sigma}\right] \tag{1.2}$$

$$F(t) = \Phi\left[\frac{\log(t) - \mu}{\sigma}\right] \tag{1.3}$$

where ϕ and Φ are the standard pdf and cdf of a specified location-scale distribution. Two popular examples of log-locations-scale distributions are the lognormal distribution, with $\phi = \phi_{\text{NORM}}$ and $\Phi = \Phi_{\text{NORM}}$, and the Weibull distribution, with $\phi(z) = \exp[z - \exp(z)]$ and $\Phi(z) = 1 - \exp[-\exp(z)]$, where $z = [\log(t) - \mu]/\sigma$. Each log-location-scale model has two parameters: a location parameter μ and a scale parameter σ. If the lifetime is dependent on external factors, this relationship is most commonly expressed through the location parameter as a function $\mu(\mathbf{x})$. Thus, in the terminology of (1.1), $r(t, \mathbf{x}) = [\log(t) - \mu(\mathbf{x})]/\sigma$. A generous discussion of location-scale and log-location-scale distribution models can be found in (Meeker and Escobar, 1998, Chapter 4) and (Lawless, 2003, Chapter 1).

1.1.2 Competing Risks

Often an industrial product or system consists of multiple components with the overall system lifetime determined by the lifetimes of each component. If the system lifetime corresponds to the minimum component lifetime, the situation is referred to as "competing risks" as the components are in essence "competing" for the lifetime of the system. Several approaches exist in the literature for modeling competing risks data. The classical approach (e.g., Chiang, 1961; David and Moeschberger, 1978; Crowder, 2001) assumes the risks are independent with individual lifetime distributions. Alternative methods include the use of sub-hazard functions (Prentice et al., 1978) and mixture models (Maller and Zhou, 2002). Specification of AFT log-location-scale models in the context of classical competing risks is discussed in Section 2.2. An excellent reference for competing risks and other multivariate data analysis techniques in reliability is the recent book by Crowder (2012).

1.2 Accelerated Testing

Many industrial systems and products are highly reliable, with lifetimes consisting of several years. While such reliability is by no means undesirable, in terms of testing not every company and/or manufacturer has the luxury of waiting such long periods for results. As such, it is common for experimenters to impose greater levels of stress on their samples than would usually be encountered so as to accelerate the time to failure. After selecting varying levels of stress or "accelerating" factors and performing the desired number of runs, a model relating the stress to the lifetime is estimated and then used to extrapolate to the usual or "use" stress levels to retrieve the predicted lifetime. This procedure is known as accelerated testing.

1.2.1 Lifetime Testing

The most common form of accelerated testing is accelerated lifetime testing (ALT), in which the test under the accelerating factor level is run until the product fails. Much of the

research in this area has focused on the design of optimal ALTs with respect to a specific criterion. Examples of optimality criterion include minimizing the variance of parameter estimates, the variance of a lifetime predictor, or the variance of a quantile estimator. A good introduction to ALT methodology can be found in Nelson (1990) with a comprehensive bibliography on the ALT literature given by Nelson (2005).

In this dissertation, the ALT scenario we consider is a special case known as fatigue testing. This form of testing is used in the material sciences and involves subjecting a material sample to varying levels of tensile (stretching) or compressive (squeezing) stress. The typical form of fatigue testing is cyclic constant amplitude testing, which involves specifying a range of stresses though which to cycle until failure is achieved. The maximum stress is often considered the accelerating factor for modeling purposes. Some introductory texts for fatigue testing include Harris (2003) and Vassilopoulos and Keller (2011). Degrieck and Van Paepegem (2001) presents a thorough review of some of the models and methods used in analyzing fatigue in composite materials, which are the materials considered here. In regards to optimal ALT designs for fatigue testing, a good example is Pascual (2003), who derives several plans based on different optimality criterion for constant amplitude testing.

1.2.2 Degradation Testing

An alternative form of accelerated testing is accelerated degradation testing (ADT), in which repeated observations are made on a test subject under an accelerating factor. The measurements on the test subject are referred to as a "degradation" measure and it is usually assumed that they follow a degradation path which is a function of the accelerating factor and time. Once the degradation measures drop below a pre-specified level, a "soft" failure is said to have occurred. The research presented in this dissertation is a special case of ADT in which the observations result in the destruction of the subject. This type of ADT is known as accelerated destructive degradation testing (ADDT). Due to the destructive nature, greater care is needed in the planning of the test and analysis of the resulting data. A useful reference for accelerated degradation testing and ADDT is the article by Meeker et al. (2011).

1.2.3 Censoring

When collecting lifetime data, it is often the case that some of the test subjects have portions of their lifetimes fall outside of the observation period. This may be due to the experiment being terminated before the event occurred or the event occurring prior to the official start of the experiment. These partially observed lifetimes are said to be censored. Unlike truncation, in which no information is known about the subject, censored observations contain partial information on the lifetime distribution. Thus, it is beneficial to include censored observations in the modeling and prediction procedure, though doing so may lead to a slightly more complex analysis. Methods for handling censored data are discussed in Meeker and Escobar (1998) and Lawless (2003). Procedures for making predictions with censored data are given in Escobar and Meeker (1999).

1.3 Motivation

1.3.1 Genealogy Information for Product Components

For products that contain multiple components, it is common during the production period for these components to receive upgrades or be replaced by components from a different manufacturer due to financial or reliability reasons. In such instances, the components can be said to belong to a particular generation. This genealogical information is readily available in many product databases, but has not been used in standard competing risks model predictions. Through incorporation of this information into these models, it is believed that warranty prediction accuracy will improve significantly. In Chapter 2, methods are provided for including this genealogical information into a competing risks framework. Extensions of the prediction procedures given in Hong and Meeker (2010) are derived and simulations are used to compare the accuracy of predictions to those under a standard competing risks model.

1.3.2 Optimal Test Plans for Constant Amplitude Fatigue Testing

The current standards for fatigue testing use balanced designs for testing of any criterion and log linear models for analysis. The nonlinear physical models derived in the fatigue literature are more appropriate for such data and test plans can be derived from such models that are optimal for a desired criterion. In Chapter 3, optimal and compromise test plans based on the fatigue model for composite materials proposed by Epaarachchi and Clausen (2003) are derived for cyclic constant amplitude testing. The criterion of interest involves a weighted sum of the asymptotic variances of a quantile estimator on a range of design stress values. The effects of the design parameters and modeling assumptions on these plans are investigated and suggestions are made regarding appropriate test planning. Simulations are used to validate the designs.

1.3.3 Comparison of Industrial and Statistical Methodology for ADDT Model Estimation

The methodology presented in the industry standards for estimating quantities from an ADDT involves a two-step procedure of least squares estimation. This is in contrast with the accepted method of maximum likelihood in the statistical literature. In Chapter 4 a comparison of both procedures is presented. The advantages and disadvantages of each procedure are discussed with emphasis given on the maximum likelihood procedure as a more viable option. Two public datasets are used to assist in presenting and motivating each procedure.

1.4 Overview

The remainder of the dissertation is organized as follows. In chapter 2, methods for incorporating genealogical information on product components into current competing risks methodology is presented. Chapter 2 is mainly based on King, Hong, and Meeker (2015), a paper submitted to Quality and Reliability Engineering International for review. Chapter 3 discusses optimum test planning for constant amplitude fatigue testing based on a new fatigue

model. Chapter 3 is based on King et al. (2015), a paper in progress. Chapter 4 presents a comparison of two procedures for estimating models from ADDT data and proposes a viable alternative to the current industrial standard. Chapter 4 is based on King et al. (2015), a paper in progress.

1.5 Bibliography

Bagdonavicius, V. and M. Nikulin (2010). *Accelerated Life Models: Modeling and Statistical Analysis*. Boca Raton, Florida: CRC Press.

Chiang, C. L. (1961). A stochastic study of the life table and its applications. III. The follow-up study with the consideration of competing risks. *Biometrics 17*(1), 57–78.

Crowder, M. J. (2001). *Classical Competing Risks*. Boca Raton, Florida: CRC Press.

Crowder, M. J. (2012). *Multivariate Survival Analysis Competing Risks*. Boca Raton, Florida: CRC Press.

David, H. A. and M. L. Moeschberger (1978). *The Theory of Competing Risks*. London: Griffin.

Degrieck, J. and W. Van Paepegem (2001). Fatigue damage modelling of fibre-reinforced composite materials: Review. *Applied Mechanics Reviews 54*(4), 279–300.

Epaarachchi, J. A. and P. D. Clausen (2003). An empirical model for fatigue behavior prediction of glass fibre-reinforced plastic composites for various stress ratios and test frequencies. *Composites Part A: Applied Science and Manufacturing 34*(4), 313–326.

Escobar, L. A. and W. Q. Meeker (1999). Statistical prediction based on censored life data. *Technometrics 41*(2), 113–124.

Harris, B. (2003). *Fatigue in composites: science and technology of the fatigue response of fibre-reinforced plastics*. Boca Raton, Florida : Cambridge, England: CRC Press ; Woodhead Pub. Ltd.

Hong, Y. and W. Q. Meeker (2010). Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics 52*(2), 148–159.

King, C., Y. Hong, S. P. DeHart, and P. A. DeFeo (2015). Planning fatigue tests for polymer composites. *In Progress*.

King, C., Y. Hong, and W. Q. Meeker (2015). Product component genealogy modeling and warranty return prediction. *Unpublished Manuscript*.

King, C., Y. Hong, Y. Xie, J. H. Van Mullekom, S. P. DeHart, and P. A. DeFeo (2015). A comparison of least squares and maximum likelihood approaches to estimating thermal indices for polymeric materials. *In Progress*.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Hoboken, New Jersey: John Wiley & Sons.

Maller, R. A. and X. Zhou (2002). Analysis of parametric models for competing risks. *Statistica Sinica 12*(3), 725–750.

Meeker, W., Y. Hong, and L. Escobar (2011). Degradation models and analyses. *Encyclopedia of Statistical Sciences*.

Meeker, W. Q. and L. A. Escobar (1998). *Statistical Methods for Reliability Data*. New York: Wiley-Interscience.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. New York: Wiley-Interscience.

Nelson, W. (2005). A bibliography of accelerated test plans. *IEEE Transactions on Reliability 54*(2), 194–197.

Pascual, F. G. (2003). Theory for optimal test plans for the random fatigue-limit model. *Technometrics 45*(2), 130–141.

Prentice, R. L., J. D. Kalbfleisch, A. V. Peterson, N. Flournoy, V. T. Farewell, and N. E. Breslow (1978). The analysis of failure times in the presence of competing risks. *Biometrics 34*(4), 541–554.

Vassilopoulos, A. P. and T. Keller (2011). *Fatigue of Fiber-reinforced Composites*. London: Springer.

Chapter 2 Product Component Genealogy Modeling and Field-Failure Prediction

2.1 Introduction

2.1.1 Background

Most products or systems consist of multiple components that are necessary for maintaining proper product functionality. For example, a laptop computer needs a fully functioning keyboard, monitor screen, CPU, and memory drive. If any one of these components were to fail, it would result in the failure of the entire system. A manufacturer is responsible for the repairs or replacements of components that fail during the warranty period. It is required that manufacturers maintain cash reserves to cover warranty costs for their sold products. More broadly, companies are often interested in the longer-term performance of their products, after the warranty period is over. In some applications, there are more important concerns such as safety. Inaccurate predictions could lead to financial penalties or safety concerns. Thus it is always desirable to have accurate field-failure predictions.

Because the failure of components leads to product failure, the problem of modeling and predicting the failure time of such a product often involves a competing risk analysis. Classical competing risk models are often used in field-failure prediction for products with multiple failure modes (the failure of a particular component often corresponds to a particular failure mode). Most of the current prediction methods for competing risks data assume that the behavior of each risk is constant over the production period of the product. Specifically, while the hazard function for a component may change over time, the hazard function used in the prediction model is treated to be the same for units manufactured over different production periods. It should be noted, however, that a product's design often evolves over a period of time with some components being replaced by alternative lower-cost or higher-reliability

components, especially in certain technological areas where product components are evolving rapidly. For example, a manufacturer may choose to make component replacements due to customer feedback, to improve their product's performance relative to a competitor, to incorporate new technological developments or for various other reasons regarding the performance of their product, or to reduce cost.

Many common household products' (e.g., televisions, laptop computers, and printers) system components are constantly being updated over time. Let us consider the laptop computer example given earlier. Even for the same laptop computer model from a specific manufacturer, different generations of hard drives may be used during the production period to improve the reliability or performance or to lower the cost of the laptop model. Figure 2.1 gives a graphical example of a product that has four major components, which we will call Product G. Over time, some of the component part numbers are changed while others stay the same. The number of changes and the change times are generally not the same for all the components.

Products manufactured at different points in time will have different configurations, although consumers usually would not notice such changes. The overall reliability of the product, however, may change due to component part-number changes. Thus, the return rate of these products could depend on such changes. Product configuration information is generally available but is usually ignored in currently-used product-failure prediction models. Incorporating the product configuration information into field-failure prediction can provide important improvements over the classical prediction methods. A prediction model that does not consider product generation information is likely to be biased. One simple approach for handling the genealogy problem is to stratify the data according to the configuration groups and to perform separate analyses on each group. Doing so would, however, usually result in small amounts of data for some groups and is an inefficient use of available information. A better strategy is to "borrow strength" across different groups corresponding to particular part numbers by estimating common parameters. The primary objective here is to use generational information to make field-failure predictions. We will also compare different strategies and quantify the advantages of using generational information in predictions.

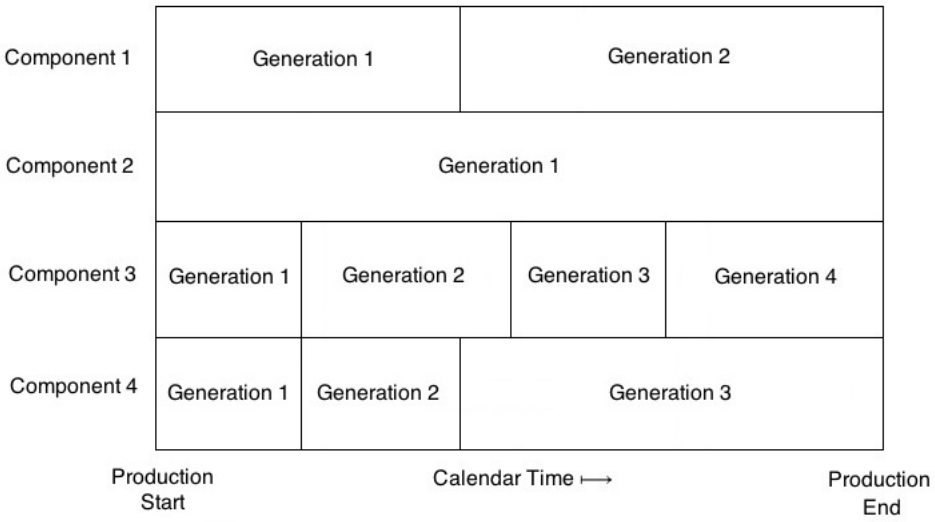


Figure 2.1: Example of product with multiple components of possibly multiple generations.

2.1.2 The Motivating Application

This research was motivated by a prediction problem that is similar to product D in Hong and Meeker (2010). Due to sensitive proprietary information, the actual data cannot be used here and a simulated dataset, similar to the original data, is used for illustration. The scenario, however, is close to the real application.

The dataset contains records of 6,000 system units, which entered into service at different times from March 2010 to March 2011, according to Figure 2.2. The failure of Product G is primarily caused by the failures of four key components. Based on early field returns, component four underwent several changes during production while component three was updated based on improvements in its manufacturing procedure. It was decided in mid 2010 to switch providers of component one for financial and quality reasons. Component two remained unchanged throughout the product's production. In March 2012, the dataset was frozen for analysis.

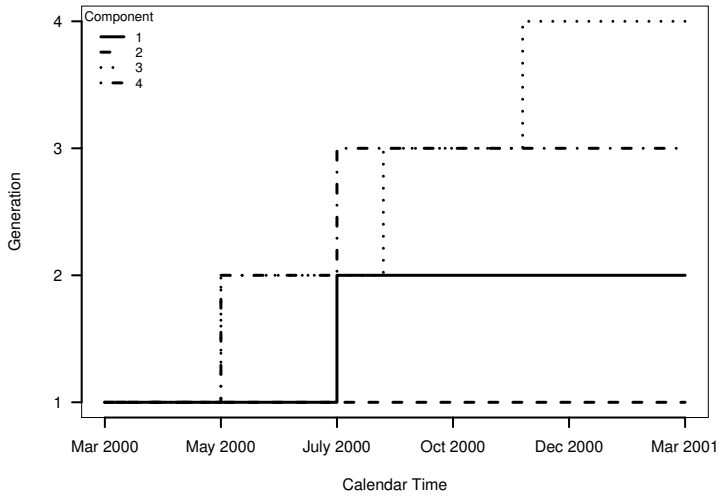


Figure 2.2: Generation functions for Product G.

2.1.3 Related Literature

A review of classical competing risks theory using latent response variables is given in Chiang (1961), Moeschberger and David (1971), Gail (1975), David and Moeschberger (1978), and Crowder (2001). Park and Kulasekera (2004) use this approach to model and compare competing risks among different groups of samples from different populations. A second method of using a cause-specific hazard rate to model competing risks data is discussed in Prentice et al. (1978). A third approach using mixture models is explored in Larson and Dinse (1985) and most recently in Maller and Zhou (2002). An excellent review of competing risks methods is given in Crowder (2012).

A general introduction to warranty data analysis is given by Lawless (1998) and Kalbfleisch et al. (1991). Methods for prediction in the presence of censored data are discussed in Escobar and Meeker (1999). Hong and Meeker (2010) considered field-failure prediction with multiple failure modes. The most common tool used to quantify potential field-failure prediction error is the prediction interval (PI), with much of the literature focused on calibration methods for the naive "plug-in" procedure. A theoretical perspective is taken in Komaki (1996), Barndorff-Nielsen and Cox (1996), and Vidoni (2009), while Beran (1990) and Escobar and Meeker

(1999) use simulation/resampling methods. The predictive distribution used in Lawless and Fredette (2005) provides a useful method of presentation of the prediction interval procedure that is equivalent to calibration of the naive intervals. Hong and Meeker (2010) adapted this approach for warranty-prediction applications involving multiple failure modes.

2.1.4 Overview

The rest of this chapter is organized as follows. Section 2.2 describes the data and model used in this paper. Section 2.3 gives the maximum likelihood (ML) procedure for estimating the model parameters. Section 2.4 gives the procedure for predicting the cumulative number of field failures at a future time for products in the risk set. Finally, Section 2.5 evaluates the advantage of incorporating component genealogy information relative to methods that ignore such information through simulation and discussion.

2.2 Data and Model

2.2.1 Product G Data

The failure-time data are denoted by $\{\tau_i, t_i, g_{ij}, \delta_{ij}; j = 1, \ldots, J\}$, $i = 1, 2, \ldots, n$. The value J denotes the number of components and n is the total number of observations. For the illustrative dataset, $J = 4$ and $n = 6{,}000$. Here, τ_i is the time of installation on the calendar time scale for system i and t_i is the observed time to failure (time in service) for the failed (censored) system i. The quantity δ_{ij} is the observed failure indicator for component j of system i. Here $\delta_{ij} = 1$ and $\delta_{ik} = 0$ $(k \neq j)$ if system i failed due to component j, and $\delta_{ij} = 0$ for all j if system i is censored. Finally, $g_{ij} = g(\tau_i, j)$ is the generation information function that gives the generation of component j based on the installation time τ_i for the system units. For example, for $j = 4$ and τ_i in June 2010 for the ith unit of Product G, $g_{i4} = 3$. We denote by G_j the maximum number of possible generations for component j. An example of the generation functions for Product G for all four components is presented in Figure 2.2. Of the 6,000 system units followed, 609 failed prior to the data freeze date (DFD). Table 2.1

Table 2.1: Failure proportion (# of failures / total # system units), the number of failures, and number of system units surviving of Product G by component and generation at the DFD. Here "C", "Prop.", "Fail.", and "RS" stand for, component, failure proportion, number of failures, and the number of system units in the risk set, respectively. The total number of system units is 6,000.

C	Generation												Total	
	1			2			3			4				
	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.
1	0.0055	33	1838	0.0038	23	3553		-			-		0.0093	56
2	0.0240	144	5391		-			-			-		0.0240	144
3	0.0048	29	777	0.0030	18	1520	0.0013	8	1426	0.0008	5	1668	0.0099	60
4	0.0528	317	777	0.0100	60	1061	0.0010	6	3553		-		0.0638	383

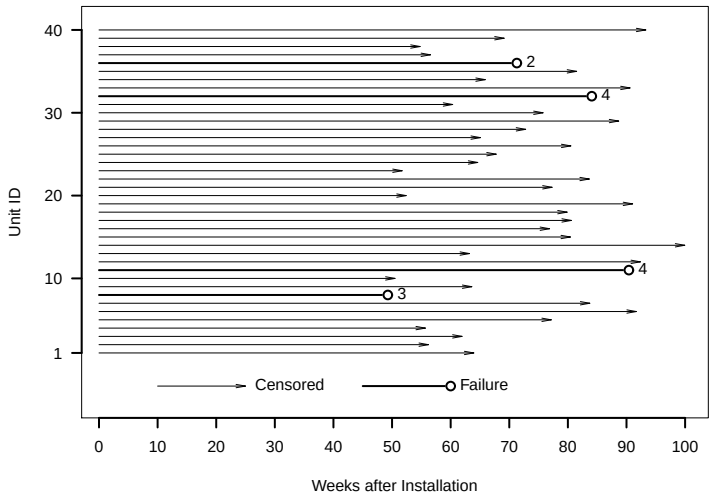


Figure 2.3: Event plot showing the lifetimes for the first 40 system units of Product G. The numbers next to the failure arrows indicate the component that caused failure.

shows the failure proportion (the number of failures divided by the total number of system units), the number of failures, and number of system units surviving (the risk set) of Product G by component and generation at the DFD.

2.2.2 Failure-time Model

The log-location-scale family of distributions, which includes the lognormal and Weibull distributions as special cases, will be used here to model the component lifetimes. Let T_{ij} denote the latent lifetime of component j of system i. Assuming the particular log-location-scale distribution does not change across generations, which is reasonable if the failure mode

is the same within each component, the cumulative distribution function (cdf) of T_{ij} is given by

$$F_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) = \Phi_{ij}\left[\frac{\log(t) - \mu_{jg_{ij}}}{\sigma_j}\right], \tag{2.1}$$

where $\mu_{jg_{ij}}$ and σ_j are the location and scale parameter respectively for $\log(T_{jg_{ij}})$, $\boldsymbol{\xi}_{jg_{ij}} = (\mu_{jg_{ij}}, \sigma_j)'$ is the parameter vector for component j of system i at the g_{ij}th generation, and $\Phi(\cdot)$ is the standard cdf for the log-location-scale family of distributions. Within the competing-risk model, F_{ij} is the marginal cdf for component j of system i. The probability density function (pdf) for $T_{jg_{ij}}$ is given by

$$f_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) = \frac{1}{\sigma_j t}\phi_{ij}\left[\frac{\log(t) - \mu_{jg_{ij}}}{\sigma_j}\right], \tag{2.2}$$

where $\phi(\cdot)$ is the standard pdf for the log-location-scale family of distributions.

The primary focus will be on models in which only generational changes in the location parameter are considered, which we will refer to as "location-change" generational (LCG) models. However, this model can easily be extended to allow for generational changes in the scale parameter as well. We will refer to these models as "extended location-change" generational (ELCG) models. Because the scale parameter is often associated with the failure mechanism, the LCG model tends to be more appropriate with respect to most field applications. However, the ELCG model is more flexible and may be appropriate in cases where an updated component does contain a different failure mechanism than the previous generation. An example would be transitioning from a traditional hard drive to a solid-state hard drive in a laptop.

It is important to note that the location and scale parameters are used here in connection with their usage in the location-scale part of the log-location-scale distribution. For example, in the lognormal distribution, μ represents the location parameter and σ the scale parameter of the normal distribution, describing the logarithms of the times. This is to be distinguished from the shape and scale parameters of the lognormal, which can be represented by $1/\sigma$ and $\exp(\mu)$, respectively. It is with respect to the former case that we will be referring to the

location and scale parameters of the distribution.

Many well known distributions used in lifetime modeling belong to the log-location-scale family. For illustration purposes, we will model the component lifetimes using the Weibull and lognormal distributions. The Weibull distribution cdf and pdf are given by replacing Φ and ϕ in (2.1) and (2.2) with the standard smallest extreme value distributions $\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$, and $\phi_{\text{sev}}(z) = \exp[z - \exp(z)]$, respectively. The lognormal distribution cdf and pdf are obtained similarly, but with Φ_{nor} and ϕ_{nor}, which are, respectively, the cdf and pdf for the standard normal distribution.

2.2.3 System cdf and Sub-distribution Function

Under the classical competing risks framework, the distribution of a system's lifetime is the same as the distribution of the minimum latent lifetime over the J components. Because the components are assumed independent of one another, the cdf of the system lifetime T_i is given by

$$F_i(t;\boldsymbol{\xi}_i) = \Pr[\min(T_{i1},\ldots,T_{T_{iJ}}) \leq t] = 1 - \prod_{j=1}^{J}\left[1 - F_{ij}(t;\boldsymbol{\xi}_{jg_{ij}})\right], \tag{2.3}$$

where $\boldsymbol{\xi}_i = (\boldsymbol{\xi}'_{1g_{i1}},\ldots,\boldsymbol{\xi}'_{Jg_{iJ}})'$ is the set of parameters for system i. An alternative expression for the system lifetime cdf is

$$F_i(t;\boldsymbol{\xi}_i) = \sum_{j=1}^{J} F^*_{ij}(t;\boldsymbol{\xi}_i). \tag{2.4}$$

Here, $F^*_{ij}(t;\boldsymbol{\xi}_i)$ is known as the sub-distribution function (e.g., see Moeschberger and David (1971)), which is defined as

$$F^*_{ij}(t;\boldsymbol{\xi}_i) = \Pr(T_i \leq t, \Delta_{ij} = 1;\boldsymbol{\xi}_i), \tag{2.5}$$

where $\Delta_{ij} = 1$ and $\Delta_{il} = 0$ for all $l \neq j$ if component j is the cause of system failure. The $F^*_{ij}(t;\boldsymbol{\xi}_i)$ functions are needed to generate predictions for the individual failure modes. Peterson Jr (1976) showed that sub-distribution functions are not necessarily the same as true distribution functions as it may hold that $\lim_{t\to\infty} F^*_{ij}(t;\boldsymbol{\xi}_i) < 1$. The sub-distribution function is related to the marginal distributions of the components in classical competing risks theory by

the following expression,

$$F_{ij}^{*}(t;\boldsymbol{\xi}_i) = \int_0^t f_{ij}(s;\boldsymbol{\xi}_{jg_{ij}}) \prod_{l\neq j}[1 - F_{il}(s;\boldsymbol{\xi}_{lg_{il}})]ds. \tag{2.6}$$

The sub-density function is

$$f_{ij}^{*}(t;\boldsymbol{\xi}_i) = \frac{dF_{ij}^{*}(t;\boldsymbol{\xi}_i)}{dt} = f_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) \prod_{l\neq j}[1 - F_{il}(t;\boldsymbol{\xi}_{lg_{il}})]. \tag{2.7}$$

2.3 Parameter Estimation

2.3.1 The Likelihood Function

The parameters of the generational model are estimated using ML estimation. In this section, we construct the likelihood function for the LCG and ELCG models. We will denote the set of distinct parameters to be estimated by $\boldsymbol{\theta}$.

2.3.1.1 Location-Change Generational Model

For the LCG model, the set of distinct parameters is denoted by $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_J')'$, where $\boldsymbol{\theta}_j = (\mu_{j1}, \ldots, \mu_{jG_j}, \sigma_j)'$. The likelihood is expressed as follows:

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{i=1}^{n} \left(\left\{ \prod_{j=1}^{J} f_{ij}^{*}(t_i;\boldsymbol{\xi}_i)^{\delta_{ij}} \right\} [1 - F_i(t_i;\boldsymbol{\xi}_i)]^{1-\sum_{j=1}^{J}\delta_{ij}} \right). \tag{2.8}$$

Note that $\boldsymbol{\xi}_i$ is defined in (2.3) as the set of parameters for system i, while $\boldsymbol{\theta}$ here is the set of distinct parameters in the LCG model. The construction of the likelihood function is based on the following facts. If system i failed due to component j, then the likelihood contribution is

$$\Pr(T_i = t_i; \Delta_{ij} = 1, \Delta_{il} = 0, l \neq j) = \Pr(T_{ij} = t_i; T_{il} > T_{ij}, l \neq j) = f_{ij}^{*}(t_i;\boldsymbol{\xi}_i).$$

If system i is censored, then the likelihood contribution is

$$\Pr(T_{ij} > t_i; j = 1, \ldots, J) = 1 - F_i(t_i;\boldsymbol{\xi}_i).$$

The likelihood function in (2.8) can be re-expressed as

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left\{ \prod_{i=1}^{n} f_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}} [1 - F_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}. \tag{2.9}$$

The likelihood in (2.9) is equivalent to treating failures caused by components $l \neq j$ as if they were censored for component j, which is a common practice in competing risks analysis and is also appropriate here because generational information is specific to a component and is not affected by other components under our assumptions. One can also maximize the likelihood in (2.9) separately for each component because there are no common parameters across components.

2.3.1.2 Extended Location-Change Generational Model

For the ELCG model, the set of distinct parameters is denoted by $\boldsymbol{\theta} = (\boldsymbol{\theta}_{11}^{*'}, \ldots, \boldsymbol{\theta}_{1G_1}^{*'}, \ldots, \boldsymbol{\theta}_{J1}^{*'}, \ldots, \boldsymbol{\theta}_{JG_J}^{*'})'$. Here $\boldsymbol{\theta}_{jg}^{*} = (\mu_{jg}, \sigma_{jg})'$, and μ_{jg} and σ_{jg} are location and scale parameters for the log lifetime of the gth generation of component j, respectively. The likelihood can be expressed as

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left(\prod_{g=1}^{G_j} \left\{ \prod_{i \in \text{DATA}_{jg}} f_{ij}(t_i; \boldsymbol{\theta}_{jg}^{*})^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\theta}_{jg}^{*})]^{1-\delta_{ij}} \right\} \right), \tag{2.10}$$

where DATA_{jg} is the portion of the sample that has component j belonging to generation g. Here, the focus is on the lifetime contribution for a particular generation of the component with failures due to other generations within the same component or in other components being treated as censored observations.

2.3.2 Maximum Likelihood Estimates and Information Matrix

The ML estimate, denoted by $\widehat{\boldsymbol{\theta}}$, is obtained by maximizing (2.9) for the LCG model or (2.10) for the ELCG model. Depending on the distribution of each component, the solution may have a closed-form expression, but more often a numerical optimization procedure is needed. For our analysis, the ML estimates were calculated using a quasi-Newton iterative procedure. When estimating the scale parameter, a log transformation was used so that the optimization was unbounded. Also, when there is heavy censoring, it is useful to replace μ by a small quantile, so that the likelihood is more well-behaved.

The local information matrix is obtained by evaluating the negative Hessian at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In

Table 2.2: Summary of ML estimation for Product G data. The time unit is weeks.

Comp.	Gen.	Param.	True Values	Estimate	Std. Error	95% Lower	95% Upper
1	1	μ	6.20	6.080	0.213	5.662	6.498
	2	μ	6.30	6.055	0.250	5.564	6.545
	All	σ	0.40	0.369	0.129	0.286	0.475
2	1	μ	5.00	4.991	0.035	4.923	5.059
		σ	0.30	0.297	0.060	0.264	0.334
3	1	μ	5.63	5.635	0.148	5.345	5.925
	2	μ	5.73	5.739	0.181	5.384	6.094
	3	μ	5.83	5.768	0.219	5.339	6.196
	4	μ	5.93	5.747	0.251	5.254	6.240
	All	σ	0.30	0.291	0.127	0.227	0.374
4	1	μ	4.68	4.670	0.010	4.651	4.690
	2	μ	4.78	4.751	0.018	4.716	4.785
	3	μ	4.88	4.810	0.033	4.746	4.875
	All	σ	0.20	0.186	0.042	0.171	0.202

particular, the local information matrix is

$$I(\widehat{\boldsymbol{\theta}}) = -\frac{\partial^2 \log[L(\boldsymbol{\theta}|\text{DATA})]}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.$$

The details of the calculation of the local information matrix for each model are given in Appendix 2.A. An estimate of the variance-covariance matrix of the parameter estimators, $\widehat{\Sigma}$, is obtained by evaluating $[I(\boldsymbol{\theta})]^{-1}$ at $\widehat{\boldsymbol{\theta}}$.

Wald confidence intervals (CIs) can be calculated for the ML estimates using $\widehat{\Sigma}$. As an example, a $100(1-\alpha)\%$ Wald CI for μ_{jg} in the LCG model is given by

$$\widehat{\mu}_{jg} \pm z_{1-\alpha/2}\sqrt{\widehat{\text{var}}(\widehat{\mu}_{jg})}, \tag{2.11}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)\%$ quantile of the standard normal distribution and $\widehat{\text{var}}(\widehat{\mu}_{jg})$ is the estimated variance of $\widehat{\mu}_{jg}$ taken from the corresponding diagonal element of $\widehat{\Sigma}$.

2.3.3 Estimation for Product G Data

We apply the ML method for the LCG model to the Product G data to estimate the unknown model parameters. Table 2.2 gives the ML estimates, standard errors, and the approximate 95% CIs for the model parameters for each component and each generation.

2.4 Field-Failure Predictions

2.4.1 Point Predictions and Intervals

For most applications, lifetime prediction involves calculation of point predictions with corresponding PIs. Both point and interval predictions can be derived using conditional predictive distributions (Lawless and Fredette, 2005), which are probability distributions for the remaining lifetime of a system given its current lifetime. These distributions are presented here in terms of the LCG model. The computing of the predictive distributions for the ELCG model are similar.

The conditional predictive distribution for failure mode j of system i is

$$\rho_{ij} = \rho_{ij}(s) = \Pr(\Delta_{ij} = 1, T_i \leq t_i + s | T_i > t_i) = \frac{F^*_{ij}(t_i + s; \boldsymbol{\xi}_i) - F^*_{ij}(t_i; \boldsymbol{\xi}_i)}{1 - F_i(t_i; \boldsymbol{\xi}_i)}, \tag{2.12}$$

for some $s > 0$. The ML estimate of ρ_{ij}, denoted by $\widehat{\rho}_{ij}$, can be obtained by evaluating (2.12) at $\widehat{\boldsymbol{\theta}}$. The conditional predictive distribution for system i is

$$\rho_i = \rho_i(s) = \Pr(T_i \leq t_i + s | T > t_i) = \frac{F_i(t_i + s; \boldsymbol{\xi}_i) - F_i(t_i; \boldsymbol{\xi}_i)}{1 - F_i(t_i; \boldsymbol{\xi}_i)}, \tag{2.13}$$

for some $s > 0$. The ML estimator of ρ_i, denoted by $\widehat{\rho}_i$, can be obtained by evaluating (2.13) at $\widehat{\boldsymbol{\theta}}$. Note that here $\rho_i = \sum_{j=1}^{J} \rho_{ij}$. The point prediction for the total number of failures is given by $\widehat{N} = \widehat{N}(s) = \sum_{i \in \text{RS}} \widehat{\rho}_i$, where RS is the risk set containing systems that have not yet failed by the DFD.

2.4.2 Prediction Results for Product G

Point predictions can be computed as a function of time after the DFD and PIs can also be generated for each point prediction. In particular, the point prediction is obtained as $\widehat{N} = \sum_{i \in \text{RS}} \widehat{\rho}_i$ at the ML estimate. The computing of PIs is based on an extension of the procedure given in Hong and Meeker (2010). The details of the calculations are discussed in Appendix 2.B. Figure 2.4 shows the point predictions and PIs for the cumulative number of failures for the at-risk system units as a function of time after the DFD.

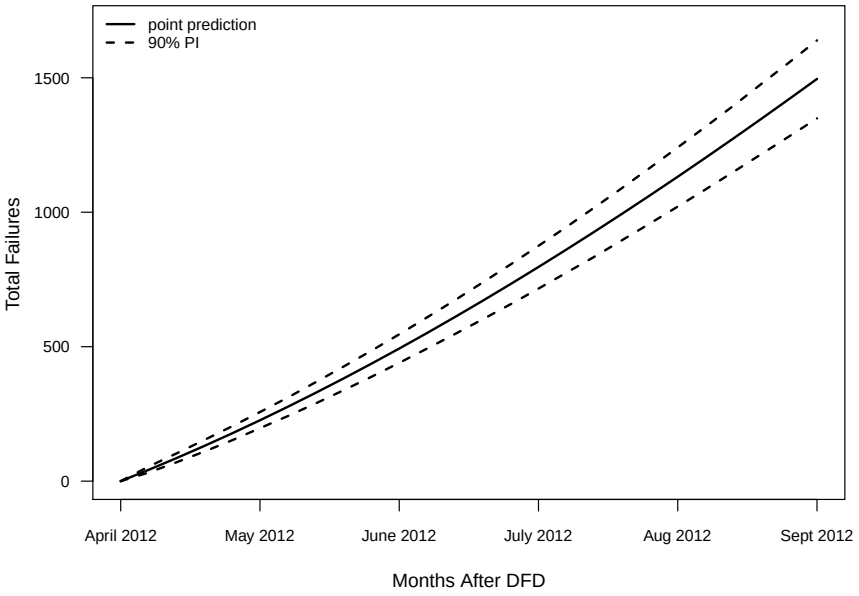


Figure 2.4: The predicted cumulative number of failures as function of time after the DFD, plotted in the calendar time scale.

2.5 Assessment of Model by Simulation

2.5.1 Prediction Models

In Section 2.2.2, we introduced the LCG and ELCG models. In this section, we will compare the standard competing risks (SCR) model, the LCG model, and the ELCG model. The SCR model corresponds to the commonly-used prediction procedure in which no generational changes are considered and all data are pooled to estimate a failure distribution for each component. Thus, the likelihood for the SCR model simplifies to

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left\{ \prod_{i=1}^{n} f_{ij}(t_i; \boldsymbol{\theta}_j^*)^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\theta}_j^*)]^{1-\delta_{ij}} \right\}, \tag{2.14}$$

where t_i is the lifetime of system i, $\boldsymbol{\theta}_j^* = (\mu_j, \sigma_j)'$ is the vector of location and scale parameters for the log of the lifetime of component j. The set of distinct parameters for the SCR model is $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^{*\prime}, \ldots, \boldsymbol{\theta}_J^{*\prime})'$. The derivation of the local information matrix is given in Appendix 2.A. The conditional predictive distributions used in the point prediction and PI calculation are

similar to those given in (2.12) and (2.13).

2.5.2 Comparison Criterion

The assessment of each model is based on the accuracy of its prediction of the cumulative total number of failures at a specified time point after the DFD. For each simulated dataset, the mean squared error (MSE) of its prediction will be estimated as a function of future time. The estimated MSE is calculated as follows:

$$\widehat{\mathrm{MSE}} = \frac{1}{B}\sum_{b=1}^{B}(\widehat{N}_b - \widehat{N})^2, \tag{2.15}$$

where $\widehat{N}_b$ is the estimated prediction based on bootstrap simulation b, B is the total number of bootstrap samples, and $\widehat{N} = \sum_{i\in \mathrm{RS}} \widehat{\rho}_i$ is the point prediction based on the simulated dataset. The bootstrap samples were generated based on the random-weight bootstrapping procedure. Details on this procedure are given in Appendix B. Here, we use bootstrapping to estimate the MSE for each dataset because each simulated dataset will have a different size risk set. Thus, the models were compared based on the average relative MSE (RMSE) of the MSEs calculated for each dataset at individual time points after the DFD using the LCG model as a basis. For example, the RMSE for SCR vs LCG is computed as the ratio of the MSE of prediction based on the SCR model and the MSE of prediction based on the LCG model for each dataset. A large number of datasets (e.g., 200) were simulated in order to ensure a stable average of the RMSE.

2.5.3 Simulation Setup

To investigate performance of our models over a wide range of situations, our simulation considers eight scenarios. In each scenario, lifetime data were simulated based on the competing risks model, in which each risk has a specified lifetime distribution and the overall lifetime of a system is determined by the minimum lifetime of the independent competing risks. Each dataset consisted of 6,000 system units with four competing risks, much like the original data. The lifetimes of the components were randomly generated using staggered entry and the dis-

Table 2.3: Summary of the simulation scenarios.

Component	Dist.	μ	σ	# Gens.	% Change in $t_{0.1}$ for eight different scenarios (S1-S8)							
					S1	S2	S3	S4	S5	S6	S7	S8
1	Weibull	6.20	0.40	2	0	5	11	28	65	5	5	11
2	Lognormal	5.00	0.30	1	-	-	-	-	-	-	-	-
3	Weibull	5.63	0.30	4	0	5	11	28	65	11	(5)(11)(28)	(11)(-28)(11)
4	Lognormal	4.68	0.20	3	0	5	11	28	65	28	(5)(11)	(11)(-5)

tribution form remained the same across generations with the parameters of the distribution being dependent on the date the product was manufactured. After a set period of time, no more lifetimes were generated, corresponding to the end of "production." The values of the parameters were chosen to make the total proportion failing similar to the original data.

Table 2.3 summarizes the generation scenarios, distribution information, and parameter values used in the simulation. In the simulation, we use the LCG model as the true model. That is, the generation changes only occur to the location parameter μ and the scale parameter σ remains constant over generations. To aid in understanding, we present the changes in location parameter in terms of changes in the 0.1 quantile of the component lifetime distribution. That is, we use the percentage $100 \times \exp(\mu + \Delta\mu + z_{0.1}\sigma)/\exp(\mu + z_{0.1}\sigma)\%$, where $\Delta\mu$ is the amount of change and the $z_{0.1}$ is the 0.1 quantile of the standard location-scale distribution. The percentage changes presented in Table 2.3 are rounded to the nearest hundredth.

For the first five scenarios, the change across generations (if it exists) is the same for all generations where the amount of change is specific to each scenario. This change in $t_{0.1}$ ranges from no change (a standard competing risks model) to an increase of 65%. The first five scenarios will be used to test the sensitivity of the LCG and ELCG models to the magnitude of the generational change.

The remaining three scenarios consider more general cases of changes in $t_{0.1}$. The sixth scenario allows the magnitude of changes to vary across components but remain the same within each component. The seventh scenario consists of changes in $t_{0.1}$ that gradually increase over time. Note that the change for component 3 in this scenario is represented by (5)(11)(28) in Table 2.3, which means the consecutive changes are 5%, 11%, and 28%, respectively. The last scenario implements a decrease in lifetime across two selected generations.

2.5.4 Simulation Results and Discussion

For each scenario, 200 datasets were simulated with estimation and point prediction performed for each dataset. Figure 2.5 gives the results for each scenario. The SCR and LCG models are nearly indistinguishable in the first scenario, which indicates that the LCG model is still adequate for situations in which no generational changes are present. In the remaining scenarios, the SCR model performs worse than the LCG when generational changes are introduced. This is due to an increase in both the variance and bias of the SCR model predictions. It is a clear indication of the inflexibility of modeling this type of data with the standard approach. The tapering effect in all of the plots is due to the number of system units in the risk sets decreasing as time goes on (eventually all system units will fail), meaning predictions are slowly becoming more stable. It is interesting to note that the behavior of the SCR model in Scenario 8 is not as poor as in the others. This is most likely due to the large decrease in the quantile $t_{0.1}$ compared to the increases, which yields data that is more similar to that generated by a situation with no generational changes.

It is surprising to see the ELCG model performing more erratically than the more restricted LCG. Further simulations were performed using progressively larger sample sizes while maintaining a consistent fraction failing to determine if the number of expected failures might be the cause. In the computing of the variance and bias, the normal approximation was used to simulate data, instead of using bootstrap, to reduce the time needed for simulations. The large size (e.g., 50,000) of several of these new simulations prevents us doing the bootstrap in a reasonable time frame. Figure 2.6 gives the variance and squared bias as a function of time after the DFD for sample size $n = 6{,}000, 10{,}000$, and 50,000. Both the variance and squared bias of the prediction were divided by the sample size n so that a fair comparison can be made across different sample sizes. Figure 2.6 reveals the prediction variance to be the dominant factor in the MSE. As the sample increases, the variance of the early predictions decreases as does the overall bias, confirming the sensitivity of the ELCG model to the number of failures. This is because estimation of the parameters in the ELCG model is performed at the genera-

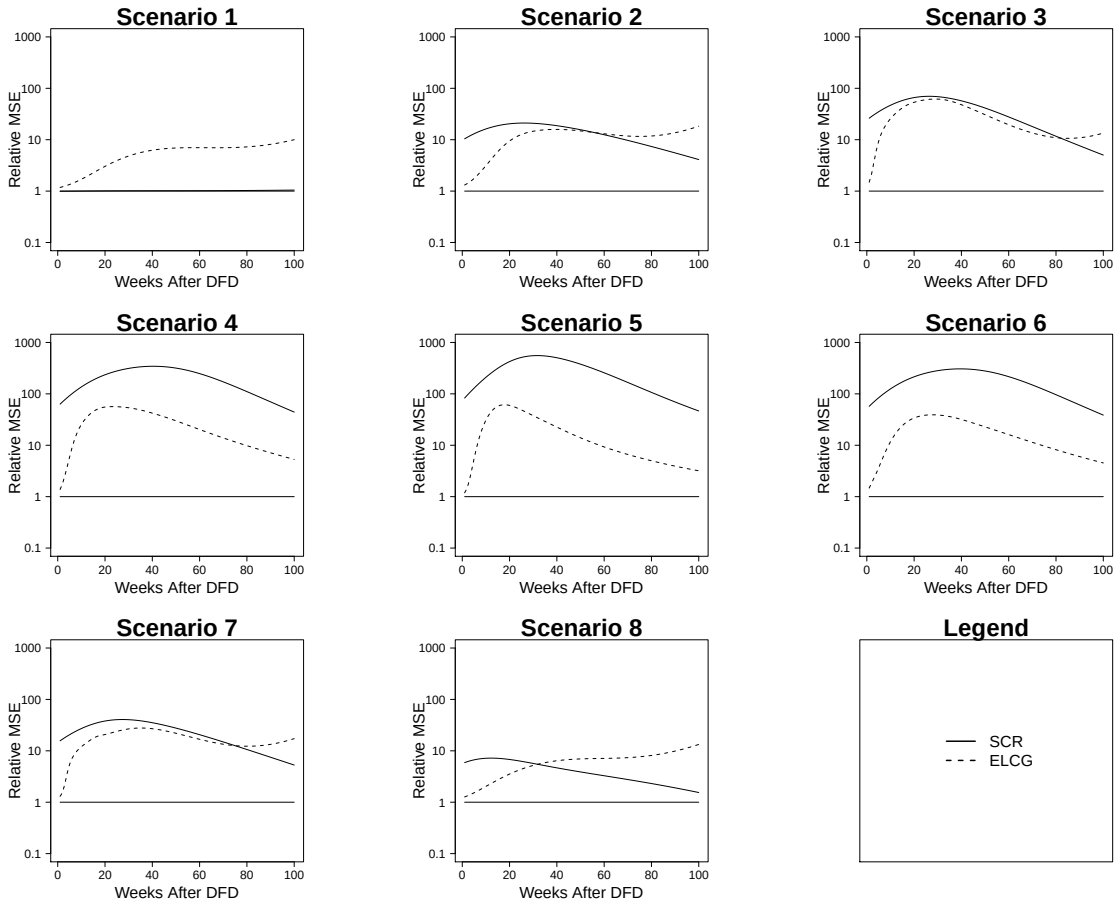


Figure 2.5: Comparison of SCR and ELCG models relative to the LCG model via relative MSE (RMSE) for each scenario. The horizontal line corresponds to RMSE= 1 for the LCG model.

tional level and so more failures are needed in every generation, especially the later ones, to yield more accurate predictions. At a sample size of $n = 6{,}000$, Table 2.1 shows the observed fraction failing for the fourth generation of component 3 is only 0.08% (5 failures), providing further insight into why the variance is large for the ELCG model. This simulation also shows an advantage of the LCG model in its ability to borrow strength across generations.

The approach to field-failure predictions presented in this chapter is general and can easily accommodate specific applications, such as warranty return prediction. One only needs to adjust the risk accordingly for those system units with expired warranty (e.g., typically one-year to three-year period). There are also situations when a company needs to predict warranty costs not just for system units that have already gone to the field, but also for system units

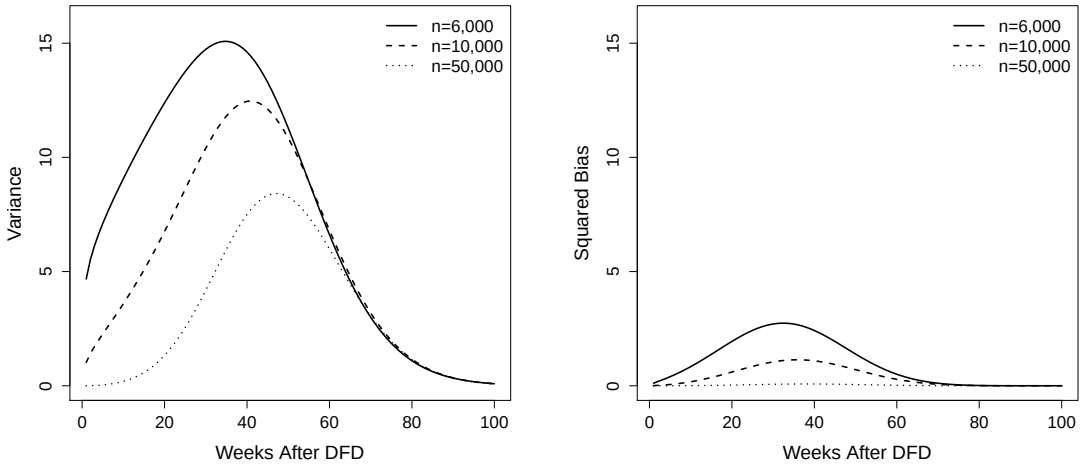


Figure 2.6: Variance and squared bias of ELCG model predictions adjusted for sample size.

that will be put into the field in the future. All that is needed is to add future system units to the risk set and the predictions can be obtained correspondingly.

2.6 Bibliography

Barndorff-Nielsen, O. E. and D. R. Cox (1996). Prediction and asymptotics. *Bernoulli 2*(4), 319–340.

Beran, R. (1990). Calibrating prediction regions. *Journal of the American Statistical Association 85*(411), 715–723.

Chiang, C. L. (1961). A stochastic study of the life table and its applications. III. The follow-up study with the consideration of competing risks. *Biometrics 17*(1), 57–78.

Crowder, M. J. (2001). *Classical Competing Risks*. CRC Press.

Crowder, M. J. (2012). *Multivariate Survival Analysis Competing Risks*. CRC Press.

David, H. A. and M. L. Moeschberger (1978). *The Theory of Competing Risks*. Griffin.

Escobar, L. A. and W. Q. Meeker (1999). Statistical prediction based on censored life data. *Technometrics 41*(2), 113–124.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. *Biometrics 31*(1), 209–222.

Hong, Y. (2013). On computing the distribution function for the poisson binomial distribution. *Computational Statistics & Data Analysis 59*, 41–51.

Hong, Y. and W. Q. Meeker (2010). Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics 52*(2), 148–159.

Kalbfleisch, J. D., J. F. Lawless, and J. A. Robinson (1991). Methods for the analysis and prediction of warranty claims. *Technometrics 33*(3), 273–285.

Komaki, F. (1996). On asymptotic properties of predictive distributions. *Biometrika 83*(2), 299–313.

Larson, M. G. and G. E. Dinse (1985). A mixture model for the regression analysis of competing risks data. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 34*(3), 201–211.

Lawless, J. F. (1998). Statistical analysis of product warranty data. *International Statistical Review / Revue Internationale de Statistique 66*(1), 41–60.

Lawless, J. F. and M. Fredette (2005). Frequentist prediction intervals and predictive distributions. *Biometrika 92*(3), 529–542.

Maller, R. A. and X. Zhou (2002). Analysis of parametric models for competing risks. *Statistica Sinica 12*(3), 725–750.

Moeschberger, M. L. and H. A. David (1971). Life tests under competing causes of failure and the theory of competing risks. *Biometrics 27*(4), 909–933.

Park, C. and K. Kulasekera (2004). Parametric inference on incomplete data with competing risks among several groups. *IEEE Transactions on Reliability 53*(1), 11–21.

Peterson Jr, A. V. (1976). Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks. *Proceedings of the National Academy of Sciences 73*(1), 11–13.

Prentice, R. L., J. D. Kalbfleisch, A. V. Peterson, N. Flournoy, V. T. Farewell, and N. E. Breslow (1978). The analysis of failure times in the presence of competing risks. *Biometrics 34*(4), 541–554.

Vidoni, P. (2009). Improved prediction intervals and distribution functions. *Scandinavian Journal of Statistics 36*(4), 735–748.

2.A Likelihood Information Matrix Computation

2.A.1 Location-Change Generational Model

The likelihood given in (2.9) can be divided into a product of sub-likelihoods at the component level. These sub-likelihoods are given by

$$L^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}}[1-F_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}, \tag{2.16}$$

where DATA_j is the portion of the sample associated with component j. These sub-likelihoods are useful in determining the local information matrix of the parameters.

The log-likelihood $\mathcal{L}(\boldsymbol{\theta}|\text{DATA}) = \log[L(\boldsymbol{\theta}|\text{DATA})]$ is the sum of the logarithms of the sub-likelihoods, which are given as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \sum_{i=1}^{n} \left\{ \delta_{ij}\log[f_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})] + (1-\delta_{ij})\log[1-F_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})] \right\}, \tag{2.17}$$

Because no system unit can belong to two different generations within the same component, (2.17) can be re-expressed as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \sum_{g=1}^{G_j} \sum_{i\in\text{DATA}_{jg}} \left\{ \delta_{ij}\log[f_{ij}(t_i;\boldsymbol{\xi}_{ig})] + (1-\delta_{ij})\log[1-F_{ij}(t_i;\boldsymbol{\xi}_{ig})] \right\}, \tag{2.18}$$

where the subscript g indicates the generation and DATA_{jg} is the portion of the sample associated with generation g of component j.

Because the log-likelihood can be expressed as the sum of sub-log-likelihoods, it follows that

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{jg'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_j \partial \sigma_{j'}} = 0$$

for all $j \neq j'$ and $g \neq g'$. The local information matrix is then a $(G+J)\times(G+J)$ block-diagonal matrix, where $G = \sum_{j=1}^{J} G_j$, with J sub-information matrices given by

$$I^{(j)} = \begin{bmatrix} I_{\mu_{j1}} & 0 & \cdots & 0 & I_{\mu_{j1}\sigma_j} \\ 0 & I_{\mu_{j2}} & \cdots & 0 & I_{\mu_{j2}\sigma_j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_{\mu_{jG_j}} & I_{\mu_{jG_j}\sigma_j} \\ I_{\mu_{j1}\sigma_j} & I_{\mu_{j2}\sigma_j} & \cdots & I_{\mu_{jG_j}\sigma_j} & I_{\sigma_j} \end{bmatrix}.$$

Using the notation given in (2.1) and (2.2) for log-location-scale distributions, we define the following functions for arbitrary μ and σ:

$$r_{ij}(z) = \frac{\phi'_{ij}(z)}{\phi_{ij}(z)}, \text{ and } h_{ij}(z) = \frac{\phi_{ij}(z)}{1 - \Phi_{ij}(z)}.$$

Here $z = [\log(t) - \mu]/\sigma$ and $\phi'_{ij}(z)$ is the first derivative of $\phi_{ij}(z)$ with respect to z. Using these functions, it follows that

$$I_{\mu_{jg}} = \frac{1}{\sigma_j^2} \sum_{i \in \text{DATA}_{jg}} \left[\delta_{ij} r'_{ij}(z_{ijg}) - (1 - \delta_{ij}) h'_{ij}(z_{ijg})\right]$$

$$I_{\sigma_j} = \frac{1}{\sigma_j^2} \sum_{g=1}^{G_j} \Bigg(\sum_{i \in \text{DATA}_{jg}} \Big\{ \delta_{ij} \left[z_{ijg}^2 r'_{ij}(z_{ijg}) + 2 z_{ijg} r_{ij}(z_{ijg}) + 1 \right] - (1 - \delta_{ij}) \left[z_{ijg}^2 h'_{ij}(z_{ijg}) + 2 z_{ijg} h_{ij}(z_{ijg}) \right] \Big\} \Bigg)$$

$$I_{\mu_{jg}\sigma_j} = \frac{1}{\sigma_j^2} \sum_{i \in \text{DATA}_{jg}} \left\{ \delta_{ij} \left[z_{ijg} r'_{ij}(z_{ijg}) + r_{ij}(z_{ijg}) \right] - (1 - \delta_{ij}) \left[z_{ijg} h'_{ij}(z_{ijg}) + h_{ij}(z_{ijg}) \right] \right\}$$

, where $z_{ijg} = [\log(t_i) - \mu_{jg}]/\sigma_j$.

2.A.2 Standard Competing Risk Model

In a similar manner for the LCG model, the SCR likelihood given in (2.14) can also be written as a product of J component-level sub-likelihoods given by

$$L^{(j)}(\boldsymbol{\theta}_j^*|\text{DATA}_j) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i;\boldsymbol{\theta}_j^*)^{\delta_{ij}} [1 - F_{ij}(t_i;\boldsymbol{\theta}_j^*)]^{1-\delta_{ij}} \right\}$$

with the corresponding log-sub-likelihoods given as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j^*|\text{DATA}_j) = \sum_{i=1}^{n} \left\{ \delta_{ij} \log\left[f_{ij}(t_i;\boldsymbol{\theta}_j^*)\right] + (1-\delta_{ij}) \log\left[1 - F_{ij}(t_i;\boldsymbol{\theta}_j^*)\right] \right\}.$$

Once again, it follows that

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_j \partial \mu_{j'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_j \partial \sigma_{j'}} = 0$$

for all $j \neq j'$, where $\mathcal{L}(\boldsymbol{\theta}|\text{DATA})$ is the logarithm of (2.14). The local information matrix is a $(2J) \times (2J)$ block-diagonal matrix of J sub-information matrices given by

$$I^{(j)} = \begin{bmatrix} I_{\mu_j} & I_{\mu_j\sigma_j} \\ I_{\mu_j\sigma_j} & I^*_{\sigma_j} \end{bmatrix},$$

where

$$\begin{aligned}
I_{\mu_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left[\delta_{ij} r'_{ij}(z_{ij}) - (1-\delta_{ij}) h'_{ij}(z_{ij}) \right] \\
I^*_{\sigma_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left\{ \delta_{ij} \left[z_{ij}^2 r'_{ij}(z_{ij}) + 2 z_{ij} r_{ij}(z_{ij}) + 1 \right] - (1-\delta_{ij}) \left[z_{ij}^2 h'_{ij}(z_{ij}) + 2 z_{ij} h_{ij}(z_{ij}) \right] \right\} \\
I_{\mu_j\sigma_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left\{ \delta_{ij} \left[z_{ij} r'_{ij}(z_{ij}) + r_{ij}(z_{ij}) \right] - (1-\delta_{ij}) \left[z_{ij} h'_{ij}(z_{ij}) + h_{ij}(z_{ij}) \right] \right\}.
\end{aligned}$$

Here, $z_{ij} = [\log(t_i) - \mu_j]/\sigma_j$.

2.A.3 Extended Location-Change Generational Model

For the ELCG model, the likelihood given in (2.10) can be divided at the generation level with the sub-likelihoods given as

$$L^{(jg)}(\boldsymbol{\theta}_{jg}^*|\text{DATA}_{jg}) = \prod_{i \in \text{DATA}_{jg}} f_{ij}(t_i;\boldsymbol{\xi}_{ig})^{\delta_{ij}} [1 - F_{ij}(t_i;\boldsymbol{\xi}_{ig})]^{1-\delta_{ij}}. \tag{2.19}$$

and the corresponding sub-log-likelihoods as

$$\mathcal{L}^{(jg)}(\boldsymbol{\theta}^*_{jg}|\text{DATA}_{jg}) = \sum_{i\in\text{DATA}_{jg}} \delta_{ij}\log\left[f_{ij}(t_i;\boldsymbol{\xi}_{ig})\right] + (1-\delta_{ij})\log\left[1-F_{ij}(t_i;\boldsymbol{\xi}_{ig})\right]. \tag{2.20}$$

As with the previous models, it follows that

$$\begin{aligned}\frac{\partial^2\mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial\mu_{jg}\partial\mu_{j'g}} &= \frac{\partial^2\mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial\mu_{jg}\partial\mu_{jg'}} = \frac{\partial^2\mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial\mu_{jg}\partial\mu_{j'g'}} \\ = \frac{\partial^2\mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial\sigma_{jg}\partial\sigma_{j'g}} &= \frac{\partial^2\mathcal{L}}{\partial\sigma_{jg}\partial\sigma_{jg'}} = \frac{\partial^2\mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial\sigma_{jg}\partial\sigma_{j'g'}} = 0\end{aligned}$$

for all $j \neq j'$ and $g \neq g'$, where $\mathcal{L}(\boldsymbol{\theta}|\text{DATA})$ is the logarithm of (2.10). Thus, the local information matrix is then a $(2G)\times(2G)$ block-diagonal matrix with G sub-information matrices given by

$$I^{(jg)} = \begin{bmatrix} I^*_{\mu_{jg}} & I_{\mu_{jg}\sigma_{jg}} \\ I_{\mu_{jg}\sigma_{jg}} & I_{\sigma_{jg}} \end{bmatrix},$$

where

$$\begin{aligned} I^*_{\mu_{jg}} &= \frac{1}{\sigma^2_{jg}} \sum_{i\in\text{DATA}_{jg}} \left[\delta_{ij} r'_{ij}(z^*_{ijg}) - (1-\delta_{ij}) h'_{ij}(z^*_{ijg})\right] \\ I_{\sigma_{jg}} &= \frac{1}{\sigma^2_{jg}} \sum_{i\in\text{DATA}_{jg}} \Big\{\delta_{ij}\left[z^{*2}_{ijg} r'_{ij}(z^*_{ijg}) + 2z^*_{ijg} r_{ij}(z^*_{ijg}) + 1\right] \\ &\qquad -(1-\delta_{ij})\left[z^{*2}_{ijg} h'_{ij}(z^*_{ijg}) + 2z^*_{ijg} h_{ij}(z^*_{ijg})\right]\Big\} \\ I_{\mu_{jg}\sigma_{jg}} &= \frac{1}{\sigma^2_{jg}} \sum_{i\in\text{DATA}_{jg}} \left\{\delta_{ij}\left[z^*_{ijg} r'_{ij}(z^*_{ijg}) + r_{ij}(z^*_{ijg})\right] - (1-\delta_{ij})\left[z^*_{ijg} h'_{ij}(z^*_{ijg}) + h_{ij}(z^*_{ijg})\right]\right\}. \end{aligned}$$

Here, $z^*_{ijg} = [\log(t_i) - \mu_{jg}]/\sigma_{jg}$.

2.B Prediction Procedure

Here we briefly describe the procedure for obtaining PIs in the context of the LCG model. The procedure is similar to Hong and Meeker (2010) but is extended for the LCG model.

2.B.1 Sampling Distribution

1. Simulate random values $Z_i, i = 1,\ldots,n$ that are independent and identically distributed with a distribution that has the property $\text{E}(Z_i) = \sqrt{\text{Var}(Z_i)}$. For our simulation, the

distribution of choice was the exponential distribution with $\mathrm{E}(Z_i) = 1$.

2. The weighted likelihood is computed as

$$L^*(\boldsymbol{\theta}|\text{DATA}) = \prod_{i=1}^{n} [L_i(\boldsymbol{\xi}_i|\text{DATA})]^{Z_i},$$

where

$$L_i(\boldsymbol{\xi}_i|\text{DATA}) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}^{Z_i}.$$

3. Obtain the ML estimate $\widehat{\boldsymbol{\theta}}^*$ by maximizing $L^*(\boldsymbol{\theta}|\text{DATA})$.

4. Repeat steps 1-3 B times to get B bootstrap samples $\widehat{\boldsymbol{\theta}}^*_b, b = 1, 2, \ldots, B$.

2.B.2 Prediction Intervals

1. Simulate I_i^* from Bernoulli$(\widehat{\rho}_i), i \in \text{RS}$, and compute $N^* = \sum_{i \in RS} I_i^*$.

2. Repeat step 1 B times to get $N_b^*, b = 1, 2, \ldots, B$.

3. Obtain $\widehat{\boldsymbol{\theta}}^*_b$ by using the procedure described in B.1

4. Compute $U_b^* = F_N(N_b^*; \widehat{\boldsymbol{\theta}}^*_b), b = 1, 2, \ldots, B$, where $F_N(\cdot)$ is the cdf of the Poisson-Binomial distribution with probability vector $\widehat{\rho}_i$'s, as given in Hong (2013).

5. Let u_N^l, u_N^u be, respectively, the lower and upper $\alpha/2$ sample quantiles of $U_b^*, b = 1, 2, \ldots, B$. Compute the $100(1-\alpha)\%$ calibrated PI by solving for N_l and N_u in $F_N(N_l; \widehat{\boldsymbol{\theta}}) = u_N^l$ and $F_N(N_u; \widehat{\boldsymbol{\theta}}) = u_N^u$, respectively.

Chapter 3 Planning Fatigue Tests for Polymer Composites

3.1 Introduction

3.1.1 Motivation

All materials, regardless of their physical makeup, exhibit some form of weakening after long periods of field use. This weakening is known as "fatigue" in the material sciences and is of great importance for material manufacturers. Knowledge of fatigue behavior can be very helpful for marketing promotions, pricing, differentiating, making claims and offering warranties. It may also help manufacturers improve upon existing products or even allow for the development of new and improved offerings. In light of these benefits, a large body of research exists pertaining to the study of fatigue behavior in material products. Recently, the focus of this research has shifted from homogeneous materials, such as metals and alloys, to the study of heterogeneous materials, such as polymer composites. Composite materials have become more commonplace due to their ability to remain lightweight while still retaining comparable levels of strength and endurance, a property that makes them more energy efficient. This has resulted in polymer composites becoming a key component in several large industries related to energy consumption, including transportation manufacturing and alternative energy production.

Acquiring knowledge of polymer composite fatigue requires proper experimentation. A well-designed experiment can yield invaluable information about a manufacturer's product. The majority of testing performed in this field is in accordance with the standards provided in ASTM E739 (2010) for testing and analysis of polymer fatigue. All of the test plans discussed within these standards are of a balanced nature with equal replication and spacing of the test samples. While balanced plans are very desirable, research in the statistical literature on optimal test planning for fatigue testing (e.g., Nelson and Kielpinski 1976; Nelson and Meeker

1978; Pascual 2003) has shown that these plans may not result in the most precise information. Optimal test planning involves the design of experiments in which the data that are collected yield the maximum amount of information regarding a particular quantity of interest, such as the lifetime by which at least a specified proportion of the product is expected to fail. These optimal test plans tend to be more unbalanced and yet yield more precise estimates of a selected quantity than could be achieved under a balanced design. As such, they may be better suited for the task of gathering information on polymer composite fatigue behavior.

The most common form of fatigue testing is known as constant amplitude fatigue testing. In this form of testing, a sample of material is either stretched or compressed (or in some cases a combination of the two) to a maximum stress $\sigma_{\max}$ and then cycled between that stress and a lower level of stress $\sigma_{\min}$ until failure is achieved. See Figure 3.1 for a graphical example of this procedure. Due to the resilient nature of polymer composite materials, the stress levels used in testing are usually higher than those experienced in the field in order to induce failures on a shorter timescale. This is an example of accelerated life testing (ALT), in which varying levels of higher than normal stress are selected for testing and the resulting data are used to extrapolate out to the normal use levels about which inference is desired. Much of the optimal design literature on ALTs consider a single use stress level as part of the design criterion. However, it is more likely that a material will experience several varying levels of stress over its lifetime. For example, the polymer composite materials in motor vehicles are subject to a variety of stresses during the lifetime of the vehicle, such as changes in the terrain (potholes, train tracks, etc.) or weather conditions to name a few. Thus, we propose an optimality criterion consisting of a weighted sum of asymptotic variances over a distribution of design stress levels with the weights corresponding to the frequency with which each stress level is expected to be encountered. Such optimal test plans would be more relevant to fatigue researchers.

In addition to specifying more relevant criteria, we also seek to incorporate existing fatigue knowledge through specification of an appropriate model. Many of the optimal designs proposed for fatigue testing are based on linear models, one of the few exceptions being the

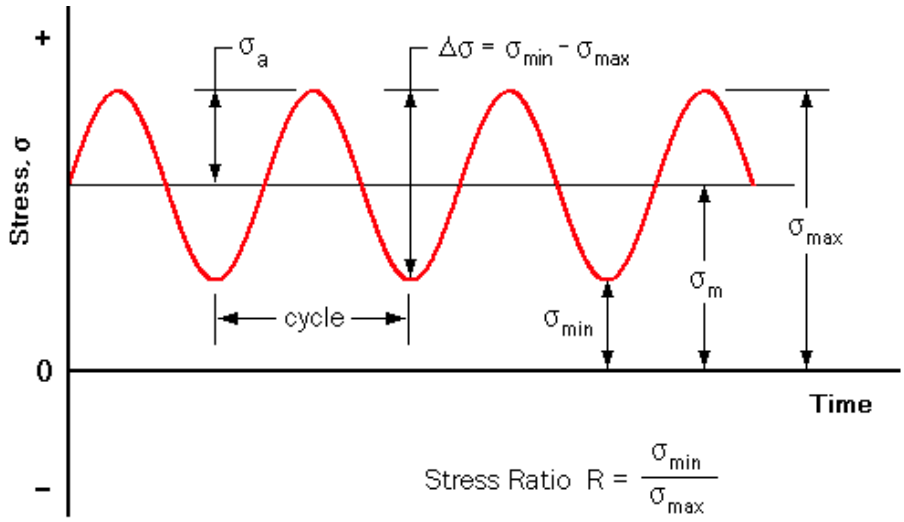


Figure 3.1: Example of constant-amplitude fatigue testing. Retrieved 2 April, 2014 from ETBX Stress-Life Fatigue Analysis Module `http://www.fea-optimization.com/ETBX/stresslife_help.html` (Used under fair use 2015)

designs proposed by Pascual (2003). This is also reflected in the choice of a linear model for analyzing fatigue data proposed by the standards (ASTM E739, 2010). A casual observance of the models proposed in the fatigue literature (see Section 3.1.2 for a small selection) reveals the inadequacy of a linear model for explaining the relationship between stress and lifetime for polymer composites. Instead, we propose using a physical model taken from the fatigue literature as the basis for our optimal design. This will provide stronger justification for our proposed test plans through the incorporation of subject-matter expertise. We also consider compromise designs that utilize more design points so as to further increase robustness to any assumption violations.

3.1.2 Related Literature

Our discussion of the related literature is divided into three parts. First, we present relevant literature on knowledge regarding fatigue behavior and its relationship to the testing procedure. Second, we review a selection of the most important models used to examine fatigue behavior. Lastly, we discuss fatigue testing from the perspective of the statistical literature, with an emphasis on optimal designs for accelerated lifetime testing.

3.1.2.1 Fatigue Damage Theory

It is well known in the fatigue literature that the fatigue of metals and other homogeneous materials is primarily due to the initiation and propagation of a dominant crack. However, the same is not necessarily true for polymer composite materials. Due to their heterogenous makeup, the fatigue of polymer composites may consist of one or more failure modes such as fiber breakage, matrix cracking, debonding, and delamination (Dharan, 1975; Reifsnider, 1980; Talreja, 1981; Konur and Matthews, 1989; Harris, 2003; Vassilopoulos and Keller, 2011). In-depth investigation of fatigue due to each failure mode has been performed for several materials; for example, Owen and Howe (1972) and Karlak et al. (1974) investigated failure of glass fiber-reinforced plastics (GFRPs) due to interface debonding. Furthermore, the presence of a particular failure mode is often dependent on several factors related to fatigue testing.

For example, in unidirectional testing of GFRPs, high stress/low cycle fatigue is primarily due to catastrophic fiber breakage while intermediate stress fatigue is due to matrix cracking that leads to delamination (Dharan, 1975; Owen, 1980). The dominance of a particular failure mode is also influenced by the material and layup used (Dharan, 1975; Owen, 1980; Reifsnider, 1980; Chen and Harris, 1993), the loading frequency (Stinchcomb et al., 1974), the testing direction (Talreja, 1981), and the ratio of minimum to maximum stress (Gamstedt and Sjögren, 1999; Niu and Talreja, 2000). Because of this sensitivity, some attempts have been made to find a standard by which various materials can be compared, such as the "characteristic damage state" proposed by Reifsnider and Talug (1980). For the most part though, research has been relatively specific to a certain material. Several reference materials for the fatigue of polymer composites include Konur and Matthews (1989), Harris (2003) and Vassilopoulos and Keller (2011).

3.1.2.2 Fatigue Modeling

There is an abundance of literature on models for polymer composite materials. Some are specific to a particular material while others are more general. Here, the focus will be on accelerated failure time (AFT) models for the stress-life (S-N) curve under cyclic constant amplitude stress testing. In particular, models relating the central tendency of lifetime to stress will be considered. Distributional models, which attempt to describe the conditional distribution of lifetime given stress, are an alternative to AFT models and were first considered by Bastenaire (1972). Examples of their usage can be found inCastillo et al. (1985); Castillo and Hadi (1995); Castillo and Fernández-Canteli (2009), and Pascual and Meeker (1999).

Table 3.1 summarizes some of the important models in the fatigue regression literature. With regards to the terminology, N is the lifetime in loading cycles, σ_M is the maximum stress applied during testing, σ_e is the endurance limit below which lifetime is theoretically infinite, and σ_u is the ultimate strength of the material, which represents the maximum stress the material can withstand before instantaneous failure. All other parameters are model specific and details regarding their function are given in the following descriptions.

Table 3.1: Regression models for material fatigue from the literature.

Model Reference	**Formula**
Wöhler (1870); ASTM E739 (2010)	$\log(N) = A - B\sigma_M$
Basquin (1910)	$\log(N) = A - B\log(\sigma_M)$
Strohmeyer (1914)	$\log(N) = A - B\log(\sigma_M - \sigma_e)$
Palmgren (1924)	$\log(N + D) = A - B\log(\sigma_M - \sigma_e)$
Weibull (1949)	$\log(N + D) = A - B\log\left(\frac{\sigma_M - \sigma_e}{\sigma_u - \sigma_e}\right)$
Stüssi (1955)	$\log N = A - B\log\left(\frac{\Delta\sigma_M - \Delta\sigma_e}{\Delta\sigma_u - \Delta\sigma_e}\right)$
Spindel and Haibach (1981)	$\log\left(\frac{N}{N_e}\right) = A\log\left(\frac{\sigma_M}{\sigma_e}\right) - B\log\left(\frac{\sigma_M}{\sigma_e}\right) + B\left\{\frac{1}{C}\log\left[1 + \left(\frac{\sigma_M}{\sigma_e}\right)^{-2C}\right]\right\}$
Kohout and Vechet (2001)	$\log\left(\frac{\sigma_M}{\sigma_e}\right) = A\log\left(\frac{N + N_1}{N + N_2}\right)$
Epaarachchi and Clausen (2003)	$\log(N) = \frac{1}{B}\log\left\{\left(\frac{B}{A}\right) f^B \left(\frac{\sigma_u}{\sigma_M} - 1\right) \times \left(\frac{\sigma_u}{\sigma_M}\right)^{\gamma(\alpha)-1} [1 - \psi(R)]^{-\gamma(\alpha)} + 1\right\}$

The earliest models (Wöhler, 1870; Basquin, 1910) were simple linear regression models on either semi-log or log-log data. These are the models that are presented in ASTM E739 (2010) for fatigue data analysis. However, as Castillo et al. (1985) noted, fatigue data is non-linear by nature and the possible existence of an endurance limit rules out standard linear regression techniques. The first attempt at expanding the original models came when Strohmeyer (1914) introduced an endurance limit into the Basquin model. A further adjustment was made by Palmgren (1924) with the introduction of an offset term D, which represents the point where the lifetime changes from the low-cycle to the intermediate stage and is primarily a curve-fitting tool. Weibull (1949) then tweaked this new model with the inclusion of ultimate tensile strength. The model attributed to Stüssi (1955) is actually just a special case of the Weibull model with the maximum stress replaced by the stress ranges $\Delta\sigma_M = \sigma_M - \sigma_m$ and $\Delta\sigma_e = \sigma_e - \sigma_m$, where σ_m is the minimum stress. The model proposed by Spindel and Haibach (1981) incorporates the endurance limit through a smooth splicing of two lines. Kohout and Vechet (2001) decided to expand on the Palmgren model by incorporating the transitions from low-cycle (N_1) or high-cycle (N_2) lifetimes to intermediate lifetimes simultaneously.

All of the previous models have attempted to enhance a linear model with known physical properties of the fatigue process, namely the inclusion of an endurance limit and/or an ultimate strength parameter. However, they are not necessarily derived primarily from a physical understanding of the process. An example of a physically motivated model is given by Epaarachchi and Clausen (2003). This model is derived from assumptions on the accumulation of fatigue damage in polymer composite materials. It also considers the effect of variables involved in the testing procedure based on previously known experimental data. This, combined with its parsimonious nature (only parameters A and B are unknown), provided us with sufficient evidence to consider it as the basis for our optimal designs. Further details on its parameterization are discussed in Section 3.2.

3.1.2.3 Fatigue Testing

The industry standards for creating balanced test plans are given in ASTM D3479 (2012). Discussions of data analysis are also presented there. Sample sizes for the test plans are calculated separately based on the methods given in ASTM E122 (2009). With regards to optimal fatigue testing, there has been substantial research in the field of accelerated life testing, of which fatigue testing can be considered a special case. The earliest ALT plans were based on linear models relating lifetime to stress. Chernoff (1962) proposed an asymptotically optimal design with r stress levels, where r is the number of parameters to be estimated. This design was based on an exponential distribution of the lifetime and was constructed so as to minimize the variance of the maximum likelihood (ML) estimates of the parameters. Nelson and Kielpinski (1976) derived plans based on normal and lognormal distributions with the goal of minimizing the ML estimate of the mean/median lifetime. A similar approach was taken in Nelson and Meeker (1978) except with Weibull and extreme value distributions and the goal of minimizing the ML estimate of a specified quantile. Later, Meeker (1984) compared plans based on Weibull and lognormal distributions and concluded that compromise plans may be more suitable as they are more robust to departures of model assumptions.

An important weakness of the previous designs is their assumption of a linear model. In light of this, there have been attempts to derive test plans for situations where certain modeling assumptions are violated. Meeter and Meeker (1994) derive plans in which the scale parameter is non-constant and Pascual (2006) discusses plans that are robust to model misspecifications. In regards to fatigue testing specifically, Pascual (2003) derives several plans based on his random fatigue limit model. A useful reference for ALT literature is the bibliography compiled by Nelson (2005) and his book (Nelson, 1990) is considered a classic reference in the field.

3.1.3 Overview

The remainder of this chapter is organized as follows. Section 3.2 discusses further details on the model used to build the test plans. Section 3.3 presents the settings and notation for derivation of the test plans along with a description and example of each type of test plan. The construction of the optimality criterion is also presented. In Section 3.4 we discuss the effects of the design parameters and modeling assumptions on the test plans. A simulation study is performed using selected test planning values to compare the asymptotic variance with the actual variance under the test plans. Finally, Section 3.5 presents some suggestions for fatigue test planning and a final summary of the results.

3.2 The Statistical Model

3.2.1 Model

For our test planning purposes, the fatigue data are assumed to come from a log-location-scale family of distributions with the probability density function (pdf) and cumulative density function (cdf) given as

$$f(t) = \frac{1}{\nu}\phi\left(\frac{\log(t) - \mu}{\nu}\right)$$

and

$$F(t) = \Phi\left(\frac{\log(t) - \mu}{\nu}\right),$$

respectively. Here, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard forms of the pdf and cdf, respectively. Common examples of the log-location-scale family of distributions include the lognormal and Weibull distribution, which is very popular in fatigue modeling. The scale parameter ν is assumed to be constant while the location parameter $\mu = \mu(\sigma_M)$ is specified to be a function of the stress. As mentioned in Section 3.1.2, we specify the relationship according to the Epaarachchi and Clausen (2003) model as

$$\mu(\sigma_M; A, B) = \frac{1}{B} \log \left\{ \left(\frac{B}{A}\right) f^B \left(\frac{\sigma_u}{\sigma_M} - 1\right) \left(\frac{\sigma_u}{\sigma_M}\right)^{\gamma(\alpha)-1} [1 - \psi(R)]^{-\gamma(\alpha)} + 1 \right\} \tag{3.1}$$

The unknown parameters are $\boldsymbol{\theta} = (A, B, \nu)'$, where A represents environmental effects on the material fatigue and B is a material-specific parameter which represents effects from the material itself. The remaining parameters are known from the testing procedure. The ultimate strength σ_u is the same as described previously and f is the frequency of the cyclic testing. The parameter ψ is a function of the ratio $R = \sigma_m/\sigma_M$, where

$$\psi(R) = \begin{cases} R & -\infty < R < 1 \\ \\ \frac{1}{R} & 1 < R < \infty \end{cases},$$

and $\gamma(\alpha) = 1.6 - \psi|\sin(\alpha)|$ is a function of the smallest angle α between the testing direction and the fiber direction. These relations are based on previous experimental results. There are several reasons why this specific model was chosen as the basis for our optimal test plans:

- The model is physically motivated as opposed to a curve-fitting technique or an extension thereof. It is derived on the assumption of a power law relationship for damage within the material structure (D'Amore et al., 1996).

- It incorporates current knowledge regarding the effects of the testing parameters on the relationship between stress and lifetime in the material. None of the other models discussed in Section 3.1.2 even considered these effects.

- It is parsimonious in terms of the number of unknown parameters to be considered. Linear models have at least two parameters to be estimated and the same goes for this model.

- Each of the unknown parameters has a physical meaning. In a linear or linearized model, the slope and intercept may not have a physical meaning or may yield values that are physically impossible.

One primary weakness of this model in comparison to a linear model is that it is highly nonlinear in its structure, resulting in a more complex estimation procedure. However, current computing power and technology can easily be used to overcome this complexity.

3.3 Test Plan Settings

Consider a fatigue test with s stress levels and k_i number of samples tested at stress levels $i = 1, \ldots, s$. The total number of units under test is $k = \sum_{i=1}^{s} k_i$. Let σ_{Mi} be the maximum stress for level i. We define the design points for the test plan as $q_i = \sigma_{Mi}/\sigma_u$; the ratio of maximum stress to the ultimate tensile strength. Note that $0 < q_i \leq 1$. In practice, a planning range of q_i is specified prior to experimentation. We denote the planning range by $q_i \in [q_L, q_U]$, where q_L and q_U are the lower and upper bounds, respectively. An example of a planning range could be $q_i \in [0.35, 0.75]$. Denote the proportion of samples allocated to level i by $\pi_i = k_i/k$. Note that $0 \leq \pi_i \leq 1$ and $\sum_{i=1}^{s} \pi_i = 1$. Here, we enforce π_i to be a valid proportion of the total sample size. That is, $\pi_i = p/q$ for some positive integers p and q. This is done to ensure that the resulting test plans yield sensible proportions of the total sample making for easier implementation. Finally, our test plan consists of the vector $\boldsymbol{\eta} = (q_1, \cdots, q_s, \pi_1, \cdots, \pi_s)'$ containing the design points where samples are to be allocated and the proportion of the sample allocated to each point. These are the parameters that can be adjusted to create the desired test plan.

3.3.1 Traditional Test Plan

The test plan we will use to represent the current methodology, which we will denote as the "traditional" test plan, is represented by $\boldsymbol{\eta}_{\text{trad}} = (0.35, 0.50, 0.75, 1/3, 1/3, 1/3)'$. This plan is representative of what is commonly used in fatigue testing and will serve as the basis for comparison.

3.3.2 Optimal Test Plan

We will denote as the "optimal" test plan the vector $\boldsymbol{\eta}_{\text{opt}} = (q_1^*, q_2^*, \ldots, q_s^*, \pi_1^*, \pi_2^*, \ldots, \pi_s^*)'$, where q_i^* and π_i^* are the design points and sample allocation values that optimize a desired criterion. A common optimality criterion used in reliability known as C-optimality is to minimize the variance of the p quantile of the product lifetime distribution at a specified use level (e.g., Nelson and Kielpinski 1976, Nelson and Meeker 1978, Nelson 1990 and Pascual 2003). This quantity is commonly used for marketing purposes or as part of industrial standards to ensure products are performing to proper specifications. Here, we extend this criterion to minimizing the weighted sum of the large-sample approximate variance of an estimator of the p quantile of the lifetime distribution at a vector of specified use levels. Specifically, let $\mathbf{q}_{\text{use}} = (q_{\text{use},1}, q_{\text{use},2}, \ldots, q_{\text{use},n})'$ denote a vector of use levels and $\mathbf{w} = (w_1, w_2, \ldots, w_n)'$ denote a vector of weights, where $\sum_{i=1}^n w_i = 1$. The vectors $\mathbf{q}_{\text{use}}$ and $\mathbf{w}$ are chosen to best represent the stresses and associated frequencies the material under consideration would experience in the field. Thus, our optimal design can be represented as

$$\boldsymbol{\eta}_{\text{optim}} = \arg\min_{\boldsymbol{\eta}} \sum_{i=1}^{n} w_i \text{AVar}\left\{\log\left[\widehat{T}_p(q_{\text{use},i})\right]\right\} \tag{3.2}$$

where $\widehat{T}_p$ is the estimator of the p quantile of the lifetime distribution. This criterion is subject to the following constraints:

(C1) The optimal design points must fall within the pre-specified planning range ($q_i^* \in [q_L, q_U]$).

(C2) The optimal sample allocations must be valid reasonable proportions ($\pi_i^* \in [0,1], k\pi_i^* = k_i$ and $\sum_{i=1}^{s} \pi_i^* = 1$).

(C3) The expected total time of the fatigue test must not exceed a pre-specified total test time $T_{\text{tot.}}$ ($k\sum_{i=1}^{s} \pi_i^* \mathbf{E}[T_i] \leq T_{\text{tot.}}$).

The expected lifetime $\mathbf{E}[T_i]$ at stress level i is computed according to the assumed distribution. For example, in the case of the lognormal distribution $\mathbf{E}[T_i] = \exp(\mu_i + \nu^2/2)$ and in the case

of the Weibull distribution $\mathbf{E}[T_i] = \exp(\mu_i)\Gamma(1+\nu)$, where μ_i is the location parameter for stress level i.

3.3.2.1 The Likelihood and Fisher Information Matrix

For our test plans, the estimator $\log(\widehat{T}_p)$ will be derived based on the maximum likelihood estimates of the model parameters in (3.1). Specifically, we proceed as follows. Let the data from a fatigue test be denoted by $\{T_{ij}, d_{ij}\}, j = 1, \ldots, k_i$, where T_{ij} is the lifetime of the jth sample at stress level i and d_{ij} is a censoring indicator. In particular,

$$d_{ij} = \begin{cases} 1 & T_{ij} \leq T_M \\ \\ 0 & T_{ij} > T_M \end{cases},$$

where T_M is a pre-specified maximum test time. It is assumed for now that T_M is the same for each stress level. The total time under test can be computed as $T_{\text{tot.}} = \sum_{i=1}^{s} k_i T_M$. The likelihood function based on this data and model (3.1) is

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{i,j} \left\{ \frac{1}{\nu T_{ij}} \phi \left[\frac{\log(T_{ij}) - \mu_i(A,B)}{\nu} \right] \right\}^{d_{ij}} \left\{ 1 - \Phi \left[\frac{\log(T_M) - \mu_i(A,B)}{\nu} \right] \right\}^{1-d_{ij}}, \tag{3.3}$$

where $\mu_i(A,B) = \mu(\sigma_{Mi}; A, B)$ is the Epaarachchi and Clausen (2003) model as defined in (3.1). The log-likelihood function is

$$l(\boldsymbol{\theta}|\text{DATA}) = \log[L(\boldsymbol{\theta}|\text{DATA})] = \sum_{ij} l_{ij}(\boldsymbol{\theta}),$$

where $l_{ij}(\boldsymbol{\theta})$ is the log-likelihood contribution from observation j at stress level i and $\boldsymbol{\theta} = (A, B, \nu)'$ is the vector of model parameters. In particular,

$$l_{ij}(\boldsymbol{\theta}) = d_{ij} \left\{ -\log(\nu) - \log(T_{ij}) + \log[\phi(z_{ij})] \right\} + (1 - d_{ij}) \log[1 - \Phi(z_{Mi})], \tag{3.4}$$

where $z_{ij} = [\log(T_{ij}) - \mu_i(A,B)]/\nu$ and $z_{Mi} = [\log(T_M) - \mu_i(A,B)]/\nu$. The maximum likelihood (ML) estimates are those values of $\boldsymbol{\theta}$ that maximize (3.4). The ML estimator $\log\left[\widehat{T}_p(q_{\text{use},i})\right]$ is then derived as

$$\log\left[\widehat{T}_p(q_{\text{use},i})\right] = \mu_{\text{use},i}(\widehat{A}, \widehat{B}) + z_p \widehat{\nu}, \tag{3.5}$$

where $\widehat{A}, \widehat{B}$ and $\widehat{\nu}$ are the ML estimators of A, B and ν, respectively, and $z_p = \Phi^{-1}(p)$.

3.3.2.2 The Large-sample Approximate Variance

The large-sample approximate variance of the estimator in (3.5) is derived from the Fisher information matrix

$$I_{\boldsymbol{\theta}} = \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right].$$

The details of the calculation of $I_{\boldsymbol{\theta}}$ for model (3.1) are given in Appendix 3.A. Specifically, under the standard regularity conditions, the large-sample approximate variance-covariance matrix of the ML estimators of A, B, and ν is

$$\begin{aligned}\Sigma_{\boldsymbol{\theta}} &= \begin{bmatrix} \text{AVar}\left(\widehat{A}\right) & \text{ACov}\left(\widehat{A}, \widehat{B}\right) & \text{ACov}\left(\widehat{A}, \hat{\nu}\right) \\ \text{ACov}\left(\widehat{A}, \widehat{B}\right) & \text{AVar}\left(\widehat{B}\right) & \text{ACov}\left(\widehat{B}, \hat{\nu}\right) \\ \text{ACov}\left(\widehat{A}, \hat{\nu}\right) & \text{ACov}\left(\widehat{B}, \hat{\nu}\right) & \text{AVar}\left(\hat{\nu}\right) \end{bmatrix} \\ &= \begin{pmatrix} \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A^2}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A \partial B}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A \partial \nu}\right] \\ \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A \partial B}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial B^2}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial B \partial \nu}\right] \\ \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A \partial \nu}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial B \partial \nu}\right] & \mathbf{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \nu^2}\right] \end{pmatrix}^{-1}. \end{aligned}$$

Note that $\Sigma_{\boldsymbol{\theta}} = I_{\boldsymbol{\theta}}^{-1}$. Then $\text{AVar}\left\{\log\left[\widehat{T}_p(q_{\text{use},i})\right]\right\}$ is computed as

$$\text{AVar}\left\{\log\left[\widehat{T}_p(q_{\text{use},i})\right]\right\} = \left[\frac{\mu_{\text{use},i}(A,B)}{\partial A}, \frac{\mu_{\text{use},i}(A,B)}{\partial B}, z_p\right] \Sigma_{\boldsymbol{\theta}} \left[\frac{\mu_{\text{use},i}(A,B)}{\partial A}, \frac{\mu_{\text{use},i}(A,B)}{\partial B}, z_p\right]'. \tag{3.6}$$

The exact variances and covariances of the ML estimates of the parameters or functions of parameters from nonlinear models can be very difficult to obtain in closed form so $\Sigma_{\boldsymbol{\theta}}$ or functions thereof are often used as a viable substitute. The design of optimal test plans usually involve minimizing a function of $\Sigma_{\boldsymbol{\theta}}$ primarily because of this constraint.

Because the quantities in $\Sigma_{\boldsymbol{\theta}}$ usually depend on the true value of the model parameters, planning values for the parameters are assumed so as to yield numeric values for the optimality criterion. The use of planning values is quite controversial as the optimality of a test plan depends strongly on the assumed model with any deviations from this model resulting in

plans that produce suboptimal results. Great care should be taken in their selection. We suggest that planning values for any experimental design project should be chosen based on expert knowledge of the fatigue process or on the results of experiments performed on similar materials. For the purpose of exposition, our planning values will be selected based on ML estimates derived from data collected on a polymer composite material. Due to the sensitive nature of the experiment and the resulting dataset, it is not presented here and has been rescaled and relabeled for use with our test plans.

3.3.3 Compromise Test Plans

While optimal test plans are desired due to their ability to extract the greatest amount of information from data, their reliance on an assumed model structure is a prominent weakness. Furthermore, these test plans tend to be rather economical in terms of the number of design points needed to implement them. Though this is certainly advantageous, it often does not leave much room for experimenters to determine if their modeling assumptions were correct. In order to alleviate this risk, many researchers propose a "compromise" plan in which additional design points and/or more allocation of the sample to these points are included. These additions help make the test plan more robust to possible misspecification of the model. In a similar manner, we consider a compromise plan $\boldsymbol{\eta}_{\text{comp}} = (q_1^{**}, q_2^{**}, \ldots, q_s^{**}, \pi_1^{**}, \pi_2^{**}, \ldots, \pi_s^{**})'$ in which additional constraints are imposed on the optimality criterion given in (3.2). Specifically, the optimality criterion is now subject to the following constraints:

(C'1) The optimal design points must fall within the pre-specified planning range ($q_i^{**} \in [q_L, q_U]$).

(C'2) The optimal sample allocations must be valid reasonable proportions and allocate no less than a specified minimum $\pi_{\min}^{**}$ to each design point. ($\pi_i^{**} \in [\pi_{\min}^{**}, 1], k\pi_i^{**} = k_i$ and $\sum_{i=1}^{s} \pi_i^{**} = 1$).

(C'3) The expected total time of the fatigue test must not exceed a pre-specified total test

time $T_{\text{tot.}}$ $(k \sum_{i=1}^{s} \pi_i^{**} \mathbf{E}[T_i] \leq T_{\text{tot.}})$.

(C'4) The number of design points must be no less than a pre-specified number s^{**} $(s \geq s^{**})$.

The compromise plans we consider here have $s^{**} = 3$ for constraint (C'4) as the inclusion of more design points would result in test plans that are larger than the traditional plan rendering them less efficient and so less desirable.

3.4 Assessment of Test Plans

The test plans will be assessed in two areas: the effects of the design parameters (total sample size, total time on test and use stress level distribution) and the effects of the modeling assumptions (planning values and lifetime distribution). The results of these assessments will help us understand the behavior of our proposed test plans under a variety of possible circumstances and make appropriate suggestions accordingly. Unless otherwise specified, the default settings for the design parameters are $k = 12$ and $T_M = 1.0 \times 10^7$ cycles. The planning values are based on those estimated from the rescaled data provided to us and a lognormal distribution is used as the default lifetime distribution.

Optimization over the test plan parameters was performed using the optimization procedure given in Appendix 3.B. Optimization according to the constraints given in Section 3.3.2 consistently resulted in test plans with only two design points at each end of the planning range. This is a common occurrence in ALT plans as the optimal number of design points is usually equal to the number of parameters to be estimated in the model, of which there are only two here. In light of these results, a minimum distance of 0.1 was imposed between any two design points for optimization of the compromise test plans to ensure they did not revert to the optimal two design points.

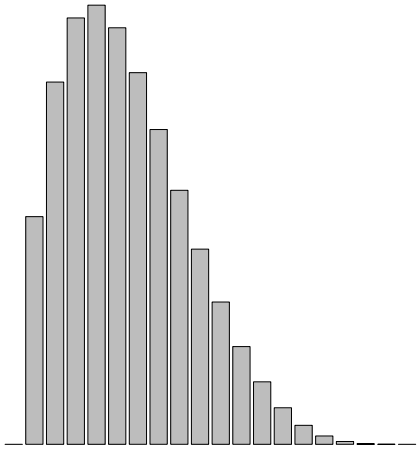
3.4.1 Design Parameters

The range of sample sizes consists of $k \in \{9, 12, 15, 18, 21, 24, 27, 30\}$. As each of these sample sizes is a multiple of three, this makes for a more direct comparison with the traditional test plans. The effect of the total time on test is expressed through the individual censoring time T_M. The range of censoring times considered are $T_M \in \{(0.5, 0.75, 1.0, 1.5, 2.0, 3.0, 5.0) \times 10^6\}$ cycles and was chosen to ensure fair comparison among the test plan types. More restrictive censoring times would lead to poor test plans and less restrictive censoring times would have no effect. The range of use stress levels consisted of 20 stress levels ranging from 5% to 25% of the ultimate strength of the material. Four distributions representing different possible use profiles were selected and are presented in Figure 3.2. The effect of the design parameters will be assessed under each use profile.

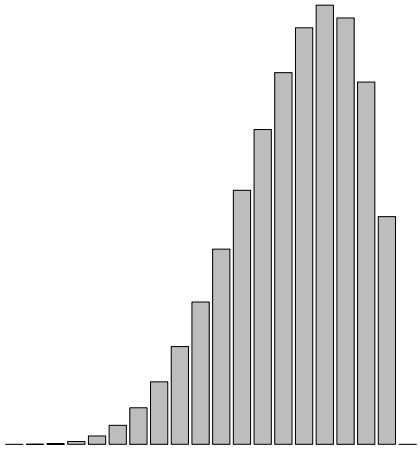
3.4.1.1 Sample Size

The configuration of each type of test plan under varying sample sizes and use profiles is presented in Table 3.2. Figure 3.3 summarizes the effect of the sample size on each type of test plan under each use distribution. As expected, all of the test plans yield more precise estimates as the sample size increases. However, it is clear that the optimal and compromise test plans yield more precise estimators across the board. Furthermore, they are both much more efficient than the traditional test plan. For example, in order to obtain a more precise estimator than that resulting from an optimal test plan with $k = 9$ samples under the first use profile, it would require $k = 15$ samples for a traditional test plan, a roughly 67% increase in the required number of samples.

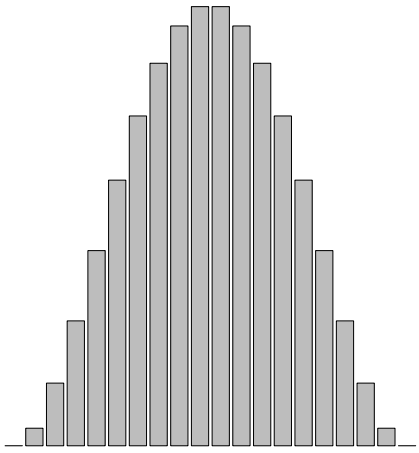
When looking at the test plans in Table 3.2, one can see a marked difference in the allocation of samples between the traditional and alternative plans. For both the optimal and compromise test plans, the majority of the samples are allocated to the lower end of the planning range. For the optimal test plan, this allocation is roughly 2:1 while for the compromise plan it is roughly 6:1:3. This is also common in ALT plans as data collected in



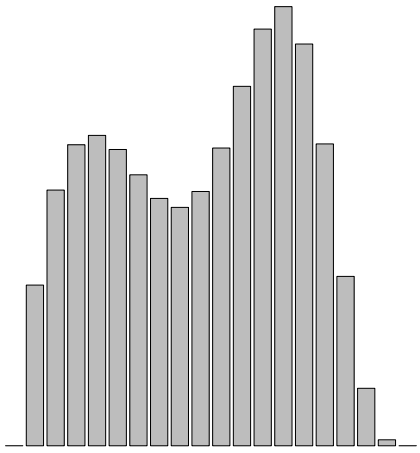
(a) Profile 1: Skewed right



(b) Profile 2: Skewed left



(c) Profile 3: Unimodal symmetric



(d) Profile 4: Bimodal asymmetric

Figure 3.2: Distribution profiles of use stress levels for the test plan assessment. Stress levels increase towards the right.

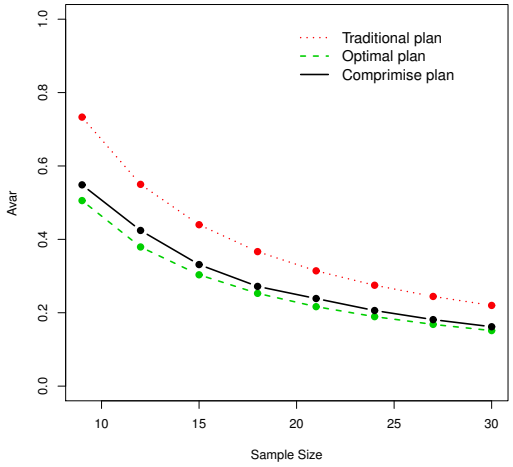
the planning range must then be used to extrapolate to the range of interest. Thus, optimal test plans tend to assign more samples to the lower end of the planning range as these design points are closest to the range of interest and so will yield more information regarding the lifetime behavior there.

One notable result in terms of the sample size is the decreasing gap in the asymptotic variance between the optimal and compromise plans. Specifically, for sample sizes $k > 25$, the difference between the two plan types is almost negligible. Thus, if the sample size is large enough, the penalty of choosing a compromise test plan over an optimal one when the underlying assumptions are in fact true are much less severe. However, the overall discrepancy between the asymptotic variances of the optimal and compromise plans is not that large even with small samples.

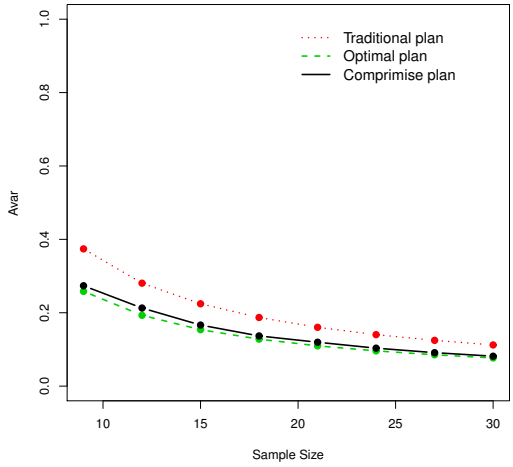
With regards to the effect of the use profile, for the second use profile, which put most of the weight on stress levels further away from the lower bound of the planning range, the test plans yield less precise estimators. This is to be expected as the farther away the use levels are from the planning range, the more uncertainty there is in the extrapolation. Furthermore, when the majority of the weight is on use stress levels closer to the planning range, more of the sample tends to be allocated to the lower stress levels in the planning range. Once again, because these regions of the planning range are closer to the use levels (in these situations, much closer), it is intuitive to assign even more samples to these regions as they are sufficiently close to yield the maximum amount of information.

3.4.1.2 Censoring Time

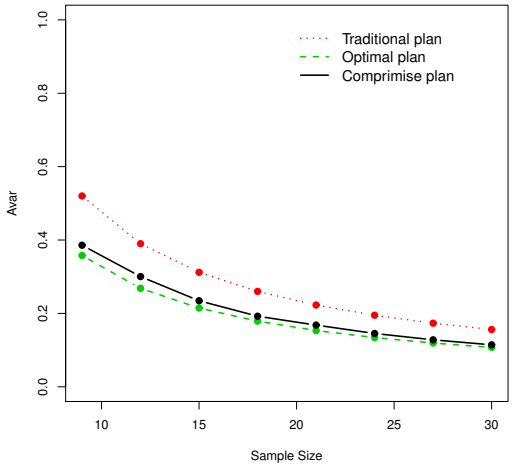
The effect of the censoring time is summarized in Figure 3.4. The configuration of each type of test plan under varying censoring times and use distributions is presented in Table 3.3. The efficiency of the optimal and compromise plans remains resistant to severe time constraints. This is most evident when $T_M = 500{,}000$ cycles. The traditional plan is untenable as the location of the design points and allocation of samples yields a test plan that will require a longer average time of completion than is possible under this constraint. Both the optimal



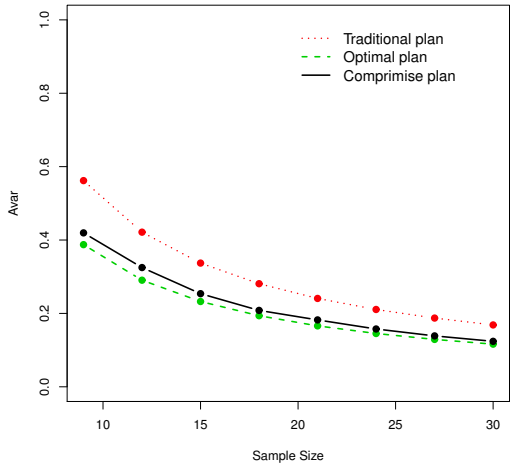
(a) Profile 1: Skewed right



(b) Profile 2: Skewed left



(c) Profile 3: Unimodal symmetric



(d) Profile 4: Bimodal asymmetric

Figure 3.3: Sample size effect for each plan under each use profile. The lines connecting the observation points only serve as a visual reference and have no physical meaning.

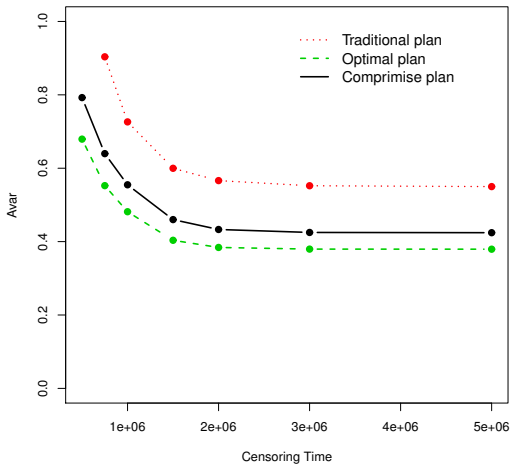
Table 3.2: Test plans by sample size and use profile.

Profile	k	Traditional Plan							Optimum Plan					Compromise Plan						
		AVar	Stress Level			Allocation			AVar	Stress Level		Allocation		AVar	Stress Level			Allocation		
			q_1	q_2	q_3	k_1	k_2	k_3		q_1	q_2	k_1	k_2		q_1	q_2	q_3	k_1	k_2	k_3
1	9	0.73	0.35	0.50	0.75	3	3	3	0.51	0.35	0.75	6	3	0.55	0.35	0.45	0.75	5	1	3
	12	0.55	0.35	0.50	0.75	4	4	4	0.38	0.35	0.75	8	4	0.42	0.35	0.65	0.75	7	2	3
	15	0.44	0.35	0.50	0.75	5	5	5	0.30	0.35	0.75	10	5	0.33	0.35	0.65	0.75	9	2	4
	18	0.37	0.35	0.50	0.75	6	6	6	0.25	0.35	0.75	12	6	0.27	0.35	0.65	0.75	11	2	5
	21	0.31	0.35	0.50	0.75	7	7	7	0.22	0.35	0.75	13	8	0.24	0.35	0.65	0.75	12	3	6
	24	0.27	0.35	0.50	0.75	8	8	8	0.19	0.35	0.75	15	9	0.21	0.35	0.65	0.75	14	3	7
	27	0.24	0.35	0.50	0.75	9	9	9	0.17	0.35	0.75	17	10	0.18	0.35	0.65	0.75	16	3	8
	30	0.22	0.35	0.50	0.75	10	10	10	0.15	0.35	0.75	19	11	0.16	0.35	0.65	0.75	18	3	9
2	9	0.37	0.35	0.50	0.75	3	3	3	0.26	0.35	0.75	6	3	0.27	0.35	0.65	0.75	6	1	2
	12	0.28	0.35	0.50	0.75	4	4	4	0.19	0.35	0.75	9	3	0.21	0.35	0.65	0.75	8	2	2
	15	0.22	0.35	0.50	0.75	5	5	5	0.15	0.35	0.75	11	4	0.17	0.35	0.65	0.75	10	2	3
	18	0.19	0.35	0.50	0.75	6	6	6	0.13	0.35	0.75	13	5	0.14	0.35	0.65	0.75	12	2	4
	21	0.16	0.35	0.50	0.75	7	7	7	0.11	0.35	0.75	15	6	0.12	0.35	0.65	0.75	14	3	4
	24	0.14	0.35	0.50	0.75	8	8	8	0.10	0.35	0.75	17	7	0.10	0.35	0.65	0.75	16	3	5
	27	0.12	0.35	0.50	0.75	9	9	9	0.09	0.35	0.75	19	8	0.09	0.35	0.65	0.75	18	3	6
	30	0.11	0.35	0.50	0.75	10	10	10	0.08	0.35	0.75	21	9	0.08	0.35	0.65	0.75	20	3	7
3	9	0.52	0.35	0.50	0.75	3	3	3	0.36	0.35	0.75	6	3	0.39	0.35	0.65	0.75	6	1	2
	12	0.39	0.35	0.50	0.75	4	4	4	0.27	0.35	0.75	8	4	0.30	0.35	0.65	0.75	7	2	3
	15	0.31	0.35	0.50	0.75	5	5	5	0.21	0.35	0.75	10	5	0.23	0.35	0.65	0.75	9	2	4
	18	0.26	0.35	0.50	0.75	6	6	6	0.18	0.35	0.75	12	6	0.19	0.35	0.65	0.75	11	2	5
	21	0.22	0.35	0.50	0.75	7	7	7	0.15	0.35	0.75	14	7	0.17	0.35	0.65	0.75	13	3	5
	24	0.20	0.35	0.50	0.75	8	8	8	0.13	0.35	0.75	16	8	0.15	0.35	0.65	0.75	15	3	6
	27	0.17	0.35	0.50	0.75	9	9	9	0.12	0.35	0.75	18	9	0.13	0.35	0.65	0.75	17	3	7
	30	0.16	0.35	0.50	0.75	10	10	10	0.11	0.35	0.75	20	10	0.11	0.35	0.65	0.75	19	3	8
4	9	0.56	0.35	0.50	0.75	3	3	3	0.39	0.35	0.75	6	3	0.42	0.35	0.65	0.75	6	1	2
	12	0.42	0.35	0.50	0.75	4	4	4	0.29	0.35	0.75	8	4	0.32	0.35	0.65	0.75	7	2	3
	15	0.34	0.35	0.50	0.75	5	5	5	0.23	0.35	0.75	10	5	0.25	0.35	0.65	0.75	9	2	4
	18	0.28	0.35	0.50	0.75	6	6	6	0.19	0.35	0.75	12	6	0.21	0.35	0.65	0.75	11	2	5
	21	0.24	0.35	0.50	0.75	7	7	7	0.17	0.35	0.75	14	7	0.18	0.35	0.65	0.75	13	3	5
	24	0.21	0.35	0.50	0.75	8	8	8	0.15	0.35	0.75	16	8	0.16	0.35	0.65	0.75	15	3	6
	27	0.19	0.35	0.50	0.75	9	9	9	0.13	0.35	0.75	18	9	0.14	0.35	0.65	0.75	17	3	7
	30	0.17	0.35	0.50	0.75	10	10	10	0.12	0.35	0.75	20	10	0.12	0.35	0.65	0.75	19	3	8

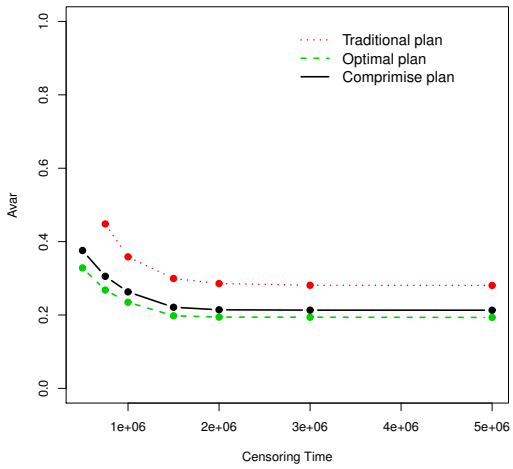
and compromise plans are able to be run within this time constraint as they allow for more flexible locations of the design points while still maintaining the mantra of "more samples at the low points". The degree of change during transitions to less restrictive time constraints is fairly consistent for all but use profile two, where there is a change in sample allocation for the optimal plan and the transition in the middle design point for the compromise plan is much sooner. As with the previous results, more weight on use stresses closer to the lower design point yield test plans that attempt to place more of the data at the closest stress level as soon as is feasible.

3.4.2 Modeling Assumptions

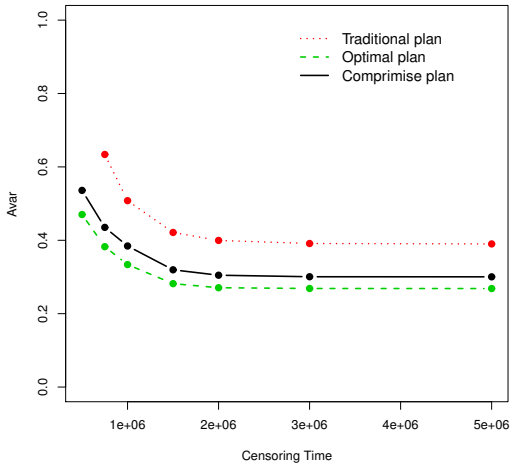
The range of planning values were chosen to roughly cover a distance of one standard deviation away from the estimated values from the rescaled motivating data. In addition to the default lognormal lifetime distribution, the Weibull distribution will be considered as an alternative due to the popularity of the related Weibull distribution in modeling fatigue data. For the purposes of this portion of the assessment, the use profile was fixed at Skewed Right



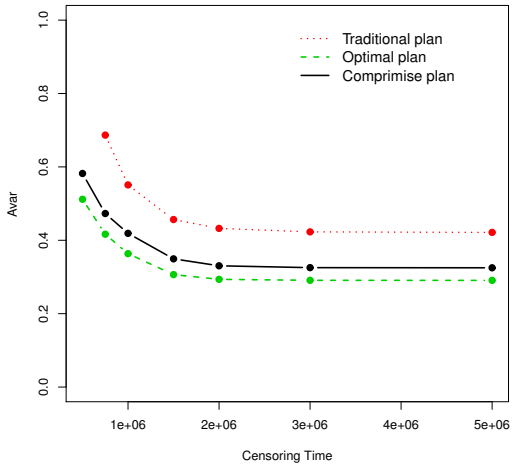
(a) Profile 1: Skewed right



(b) Profile 2: Skewed left



(c) Profile 3: Unimodal symmetric



(d) Profile 4: Bimodal asymmetric

Figure 3.4: Censoring time effect for each plan under each use profile.

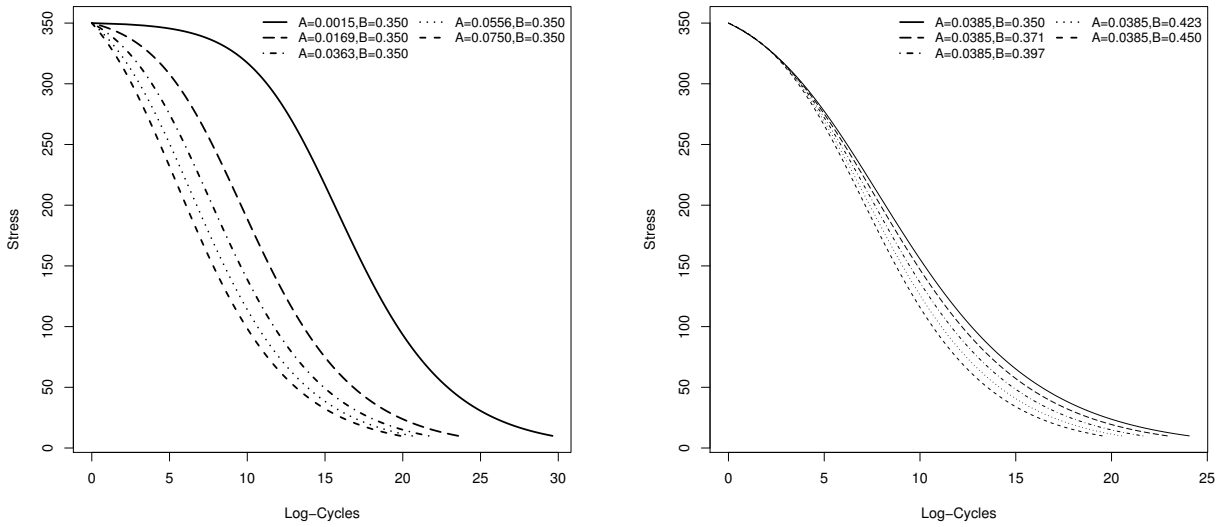
Table 3.3: Test plans by censoring time and use profile. T_M is in millions of cycles. For T_M = 500,000 cycles, the traditional plan was untenable, requiring more than the allotted total time to implement.

Profile	T_M	Traditional Plan							Optimum Plan					Compromise Plan						
		AVar	Stress Level			Allocation			AVar	Stress Level		Allocation		AVar	Stress Level			Allocation		
			q_1	q_2	q_3	k_1	k_2	k_3		q_1	q_2	k_1	k_2		q_1	q_2	q_3	k_1	k_2	k_3
1	0.50	-	0.35	0.50	0.75	4	4	4	0.68	0.41	0.75	8	4	0.77	0.41	0.51	0.75	6	2	4
	0.75	0.90	0.35	0.50	0.75	4	4	4	0.55	0.39	0.75	8	4	0.63	0.39	0.48	0.75	6	2	4
	1.00	0.73	0.35	0.50	0.75	4	4	4	0.48	0.37	0.75	8	4	0.54	0.36	0.46	0.75	6	2	4
	1.50	0.60	0.35	0.50	0.75	4	4	4	0.40	0.35	0.75	8	4	0.46	0.35	0.45	0.75	6	2	4
	2.00	0.57	0.35	0.50	0.75	4	4	4	0.38	0.35	0.75	8	4	0.43	0.35	0.65	0.75	7	2	3
	3.00	0.55	0.35	0.50	0.75	4	4	4	0.38	0.35	0.75	8	4	0.42	0.35	0.65	0.75	7	2	3
	5.00	0.55	0.35	0.50	0.75	4	4	4	0.38	0.35	0.75	8	4	0.42	0.35	0.65	0.75	7	2	3
2	0.50	-	0.35	0.50	0.75	4	4	4	0.33	0.41	0.75	8	4	0.38	0.40	0.50	0.75	6	2	4
	0.75	0.45	0.35	0.50	0.75	4	4	4	0.27	0.38	0.75	8	4	0.31	0.39	0.65	0.75	8	2	2
	1.00	0.36	0.35	0.50	0.75	4	4	4	0.23	0.37	0.75	8	4	0.26	0.37	0.65	0.75	8	2	2
	1.50	0.30	0.35	0.50	0.75	4	4	4	0.20	0.35	0.75	9	3	0.22	0.35	0.65	0.75	8	2	2
	2.00	0.29	0.35	0.50	0.75	4	4	4	0.19	0.35	0.75	9	3	0.21	0.35	0.65	0.75	8	2	2
	3.00	0.28	0.35	0.50	0.75	4	4	4	0.19	0.35	0.75	9	3	0.21	0.35	0.65	0.75	8	2	2
	5.00	0.28	0.35	0.50	0.75	4	4	4	0.19	0.35	0.75	9	3	0.21	0.35	0.65	0.75	8	2	2
3	0.50	-	0.35	0.50	0.75	4	4	4	0.57	0.41	0.75	8	4	0.54	0.41	0.51	0.75	6	2	4
	0.75	0.63	0.35	0.50	0.75	4	4	4	0.38	0.38	0.75	8	4	0.44	0.38	0.48	0.75	6	2	4
	1.00	0.51	0.35	0.50	0.75	4	4	4	0.33	0.37	0.75	8	4	0.38	0.36	0.46	0.75	6	2	4
	1.50	0.42	0.35	0.50	0.75	4	4	4	0.28	0.35	0.75	8	4	0.32	0.35	0.65	0.75	8	2	2
	2.00	0.40	0.35	0.50	0.75	4	4	4	0.27	0.35	0.75	8	4	0.30	0.35	0.65	0.75	7	2	3
	3.00	0.39	0.35	0.50	0.75	4	4	4	0.27	0.35	0.75	8	4	0.30	0.35	0.65	0.75	7	2	3
	5.00	0.39	0.35	0.50	0.75	4	4	4	0.27	0.35	0.75	8	4	0.30	0.35	0.65	0.75	7	2	3
4	0.50	-	0.35	0.50	0.75	4	4	4	0.51	0.41	0.75	8	4	0.58	0.41	0.51	0.75	6	2	4
	0.75	0.69	0.35	0.50	0.75	4	4	4	0.42	0.39	0.75	8	4	0.47	0.38	0.48	0.75	6	2	4
	1.00	0.55	0.35	0.50	0.75	4	4	4	0.36	0.37	0.75	8	4	0.41	0.36	0.46	0.75	6	2	4
	1.50	0.46	0.35	0.50	0.75	4	4	4	0.31	0.35	0.75	8	4	0.35	0.35	0.45	0.75	6	2	4
	2.00	0.43	0.35	0.50	0.75	4	4	4	0.29	0.35	0.75	8	4	0.33	0.35	0.65	0.75	7	2	3
	3.00	0.42	0.35	0.50	0.75	4	4	4	0.29	0.35	0.75	8	4	0.33	0.35	0.65	0.75	7	2	3
	5.00	0.42	0.35	0.50	0.75	4	4	4	0.29	0.35	0.75	8	4	0.33	0.35	0.65	0.75	7	2	3

from Figure 3.2.

3.4.2.1 Planning Values

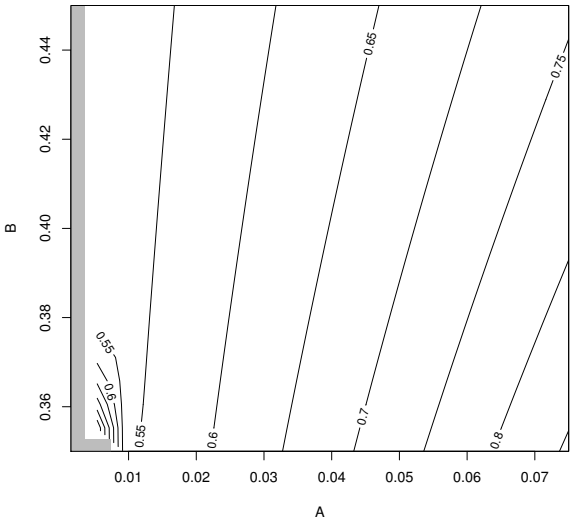
Prior to beginning our assessment of the effect of the planning values, we believe it important to take a quick detour to examine the effect of the unknown model parameters A and B on the Epaarachchi and Clausen (2003) model. In light of their descriptions given in Section 3.2, we will for the remainder of this section refer to A as the "environmental" parameter and B as the "material" parameter. The effects of each of these parameters on the model curve are summarized in Figure 3.5. From this, we can clearly see that perturbing the environmental parameter causes a much greater change in the overall curve than perturbations in the material parameter. This is in agreement with what one would expect with small changes in the environmental factors of the testing procedure as opposed to small changes in the material makeup. Note that the effect is inversely related to the direction of increase for both parameters. This knowledge will prove useful in evaluating the effects of changes in the planning values for the test plans.



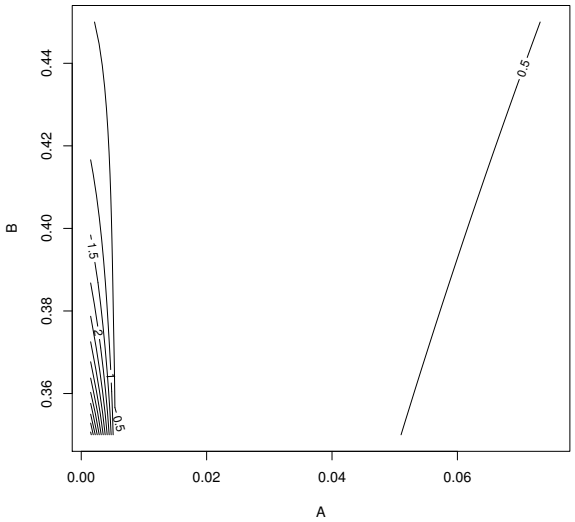
(a) Effect of environmental parameter A (b) Effect of material parameter B

Figure 3.5: Parameter effects on Epaarachchi and Clausen (2003) model.

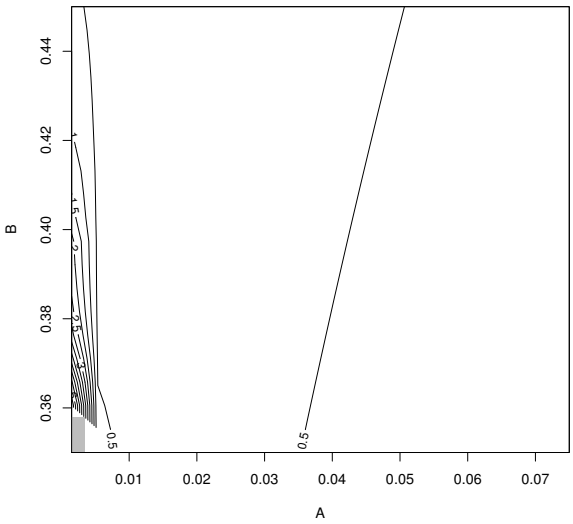
Figure 3.6 gives contour plots of the weighted sum of asymptotic variances for each combination of parameter values for the three test plan designs. Figures 3.7 and 3.8 give shade plots for the design points and sample allocation for the optimal and compromise plans, respectively. For both plans, the location of the upper design point was fixed at 0.75 for all combinations of the planning values. Both the optimal and compromise test plans are more robust to small changes in the planning values, much more so than the traditional plan. Only for very small values of the environmental parameter do the optimal and compromise plans become less robust and yield results that are worse than the traditional test plan. The reason for this is due to the inverse effect discussed previously. As the planning value of the environmental parameter deviates to smaller values, the lifetime increases much more dramatically resulting in more time required to evaluate the test plans. Thus, both the optimal and compromise plans must adjust accordingly. However, for small changes in the environmental parameter (< 0.03) beyond these small values, both the optimal and compromise plans are robust in terms of outcome and setup.



(a) Traditional Plan

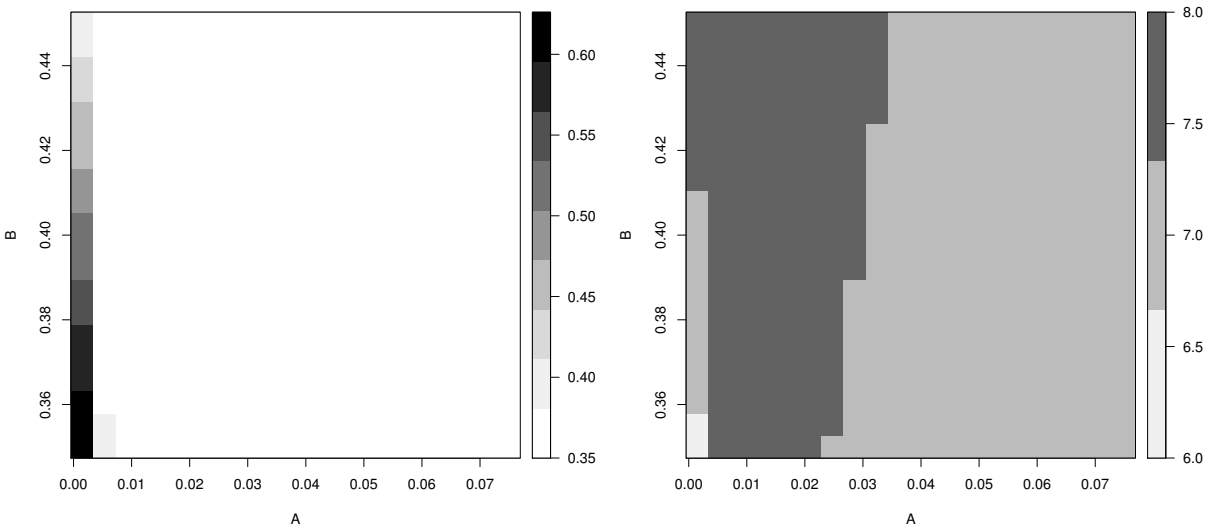


(b) Optimal Plan



(c) Compromise Plan

Figure 3.6: Planning value effect on optimality criterion for each design type. The shaded regions correspond to values that yield untenable plans due to the time restrictions.



(a) Lower design point q_L

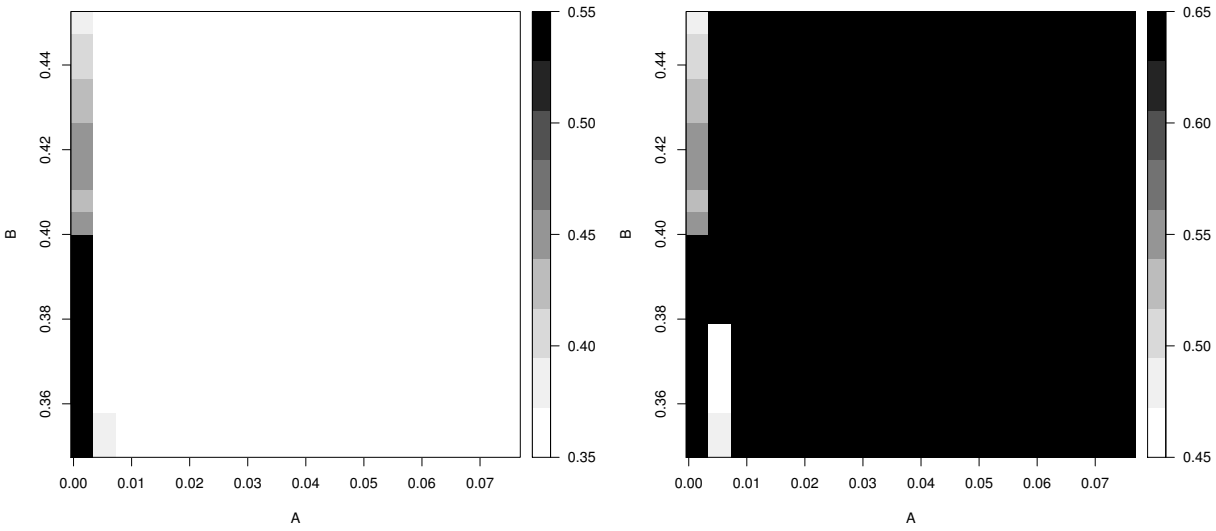
(b) Sample allocation k_L to q_L

Figure 3.7: Changes in selected design parameters for the optimal design. The upper design point q_U was always fixed at 0.75.

3.4.2.2 Distribution Assumption

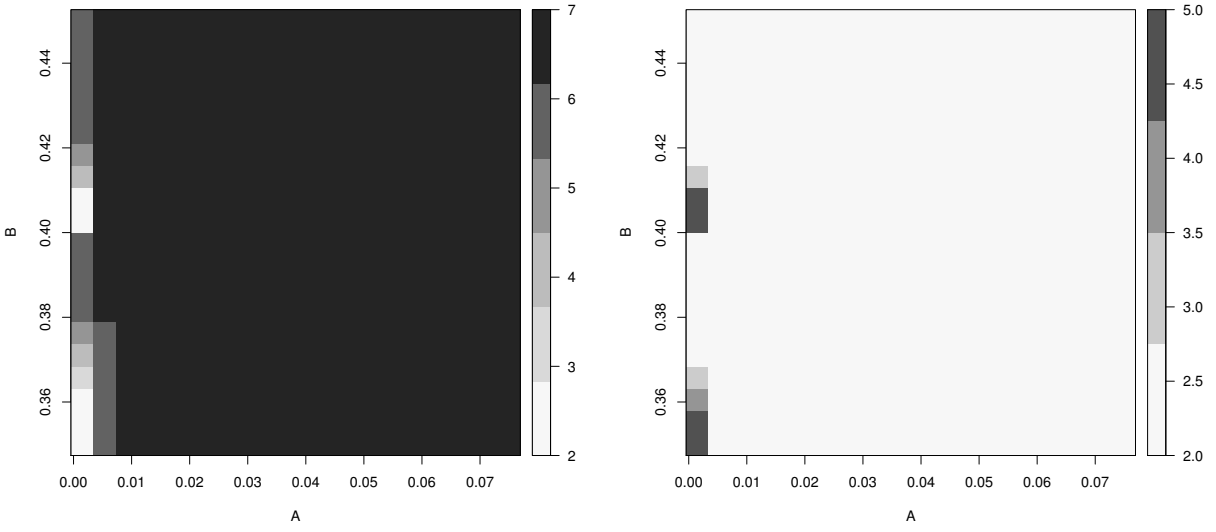
The effect of choice of lifetime distribution is summarized in Figure 3.9 and the specific layout of each test plan is presented in Table 3.4. It was decided to compare lifetime distributions across a variety of sample sizes both for ease of comparison and to detect any changes in sample size effect. We can see that though there is a large discrepancy in the asymptotic variances under each lifetime distribution, the layout of the test plans is not that much different, with the use of the Weibull distribution resulting in more samples allocated to the lowest design point only in a few instances.

The previous results would seem to indicate that using a Weibull lifetime distribution would result in more precise estimates. However, it is important to note that our criterion is based on asymptotic values. The accuracy of the criterion may be dependent on the choice of lifetime distribution, especially when that lifetime distribution is far from symmetric. We performed a simulation study to ascertain the accuracy of our criterion under each of the selected lifetime distributions and sample sizes. For each combination of distribution and



(a) Lower design point q_L

(b) Middle design point q_M



(c) Sample allocation k_L to q_L

(d) Sample allocation k_M to q_M

Figure 3.8: Changes in selected design parameters for the compromise design. The upper design point qL was always fixed at 0.75.

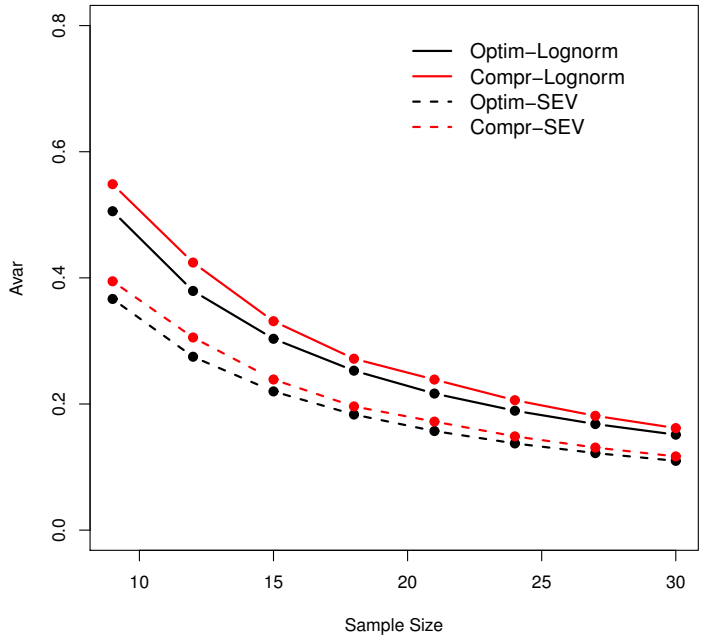


Figure 3.9: Effect of distributions on optimum and compromise plans. Lines are for visual representation only.

Table 3.4: Test plans by sample size and lifetime distribution.

Distribution	k	Optimum Plan					Compromise Plan						
		AVar	Stress Level		Allocation		AVar	Stress Level			Allocation		
			q_1	q_2	k_1	k_2		q_1	q_2	q_3	k_1	k_2	k_3
Lognormal	9	0.51	0.35	0.75	6	3	0.55	0.35	0.45	0.75	5	1	3
	12	0.38	0.35	0.75	8	4	0.42	0.35	0.65	0.75	7	2	3
	15	0.30	0.35	0.75	10	5	0.33	0.35	0.65	0.75	9	2	4
	18	0.25	0.35	0.75	12	6	0.27	0.35	0.65	0.75	11	2	5
	21	0.22	0.35	0.75	13	8	0.24	0.35	0.65	0.75	12	3	6
	24	0.19	0.35	0.75	15	9	0.21	0.35	0.65	0.75	14	3	7
	27	0.17	0.35	0.75	17	10	0.18	0.35	0.65	0.75	16	3	8
	30	0.15	0.35	0.75	19	11	0.16	0.35	0.65	0.75	18	3	9
Weibull	9	0.37	0.35	0.75	6	3	0.39	0.35	0.45	0.75	5	1	3
	12	0.27	0.35	0.75	8	4	0.31	0.35	0.65	0.75	7	2	3
	15	0.22	0.35	0.75	10	5	0.24	0.35	0.65	0.75	9	2	4
	18	0.18	0.35	0.75	12	6	0.20	0.35	0.65	0.75	11	2	5
	21	0.16	0.35	0.75	14	7	0.17	0.35	0.65	0.75	13	3	5
	24	0.14	0.35	0.75	16	8	0.15	0.35	0.65	0.75	14	3	7
	27	0.12	0.35	0.75	17	10	0.13	0.35	0.65	0.75	16	3	8
	30	0.11	0.35	0.75	19	11	0.12	0.35	0.65	0.75	18	3	9

Table 3.5: Asymptotic and simulated variances for the lognormal and Weibull lifetime distributions for selected sample sizes.

k	Lognormal				Weibull			
	Optimum		Compromise		Optimum		Compromise	
	True	Asymp	True	Asymp	True	Asymp	True	Asymp
9	0.51	0.51	0.54	0.55	0.69	0.37	0.74	0.39
12	0.39	0.38	0.42	0.42	0.52	0.27	0.51	0.31
15	0.31	0.30	0.33	0.33	0.41	0.22	0.41	0.24
18	0.26	0.25	0.27	0.27	0.35	0.18	0.34	0.20
21	0.21	0.22	0.24	0.24	0.30	0.16	0.29	0.17
24	0.19	0.19	0.20	0.21	0.27	0.14	0.25	0.15
27	0.16	0.17	0.18	0.18	0.23	0.12	0.22	0.13
30	0.15	0.15	0.16	0.16	0.21	0.11	0.20	0.12

sample size, an optimal and compromise test plan as given in Table 3.4 was implemented and the resulting data used to estimate the unknown parameters A, B and ν. The estimated quantile at each of the default use stress levels was then computed. This simulation and estimation was performed 10,000 times and the final "exact" criterion was computed using the empirical variance of the estimates with the weights selected according to the Skewed Right use distribution. The results are summarized in Table 3.5. Note that the asymptotic criterion severely underestimates the true value of the criterion when the Weibull lifetime distribution is assumed.

3.5 Discussion

The results of this assessment lead us to make the following observations and suggestions regarding fatigue test planning:

- Assigning more samples to the design point closest to the use levels yields more precise estimates of quantiles at those use levels. Furthermore, if the use levels closest to the planning range are more likely to be encountered, than even more of the sample should be allocated to the lower end of the planning range.

- Both the optimal and compromise plans are robust to slight miscalculations in the planning values ($< 0.03 units$) for all but the smallest values of the environmental parameter

A. However, the compromise test plans are more consistent than the optimal test plans with regards to sample allocation.

- Great care must be taken when using asymptotic variances as a criterion for test planning. As the simulation study showed, the accuracy of the criterion depends on the choice of lifetime distribution. A quick simulation may be needed to determine the required number of samples for particular choices of lifetime distribution and planning values. We do not condone choosing lifetime distributions solely for the purpose of creating easy test plans.

Regarding the decision to use one type of test plan over another, we propose using the compromise test plan over the optimal test plan in general except when knowledge of the underlying model is very strong. In addition to its ability to detect more curvature than the optimal test plan, the difference in the criterion value between the two plans is rather small, especially for large samples. Even when the asymptotic criterion results in underestimation, the bias is roughly the same for both test plans for the distributions considered. This along with the robustness and consistency in deviations from the planning values is justification enough for us to advocate the use of compromise test plans for general test planning.

3.6 Bibliography

ASTM D3479 (2012). Test method for tension-tension fatigue of polymer matrix composite materials. Technical report, ASTM International.

ASTM E122 (2009). Practice for calculating sample size to estimate, with specified precision, the average for a characteristic of a lot or process. Technical report, ASTM International.

ASTM E739 (2010). Practice for statistical analysis of linear or linearized stress-life (S-N) and strain-life (σ-N) fatigue data. Technical report, ASTM International.

Basquin, O. H. (1910). The exponential law of endurance tests. *Proceedings of American Society of Testing Materials ASTEA 10*, 625–630.

Bastenaire, F. A. (1972). New method for the statistical evaluation of constant stress amplitude fatigue-test results. In *Probabilistic aspects of fatigue*, ASTM STP 511, pp. 3–28. American Society for Testing and Materials.

Castillo, E., A. F. Canteli, V. Esslinger, and B. Thurlimann (1985). Statistical model for fatigue analysis of wires, strands and cables. In *International Association for Bridge and Structural Engineering Proceedings P-82/85*, Zurich, pp. 1–40.

Castillo, E. and A. Fernández-Canteli (2009). *A unified statistical methodology for modeling fatigue damage*. Netherlands: Springer.

Castillo, E. and A. S. Hadi (1995). Modeling lifetime data with application to fatigue models. *Journal of the American Statistical Association 90*(431), 1041–1054.

Chen, A. S. and B. Harris (1993). Fatigue-induced damage mechanisms in carbon fibre-reinforced plastic composites. *Journal of Materials Science 28*(8), 2013–2027.

Chernoff, H. (1962). Optimal accelerated life designs for estimation. *Technometrics 4*(3), 381–408.

D'Amore, A., G. Caprino, P. Stupak, J. Zhou, and L. Nicholais (1996). Effect of stress ratio on the flexural fatigue behaviour of continuous strand mat reinforced plastics. *Composites Science and Technology 4*(1), 1–8.

Dharan, C. K. H. (1975). Fatigue failure in graphite fibre and glass fibre-polymer composites. *Journal of Materials Science 10*(10), 1665–1670.

Epaarachchi, J. A. and P. D. Clausen (2003). An empirical model for fatigue behavior prediction of glass fibre-reinforced plastic composites for various stress ratios and test frequencies. *Composites Part A: Applied Science and Manufacturing 34*(4), 313–326.

Gamstedt, E. and B. Sjögren (1999). Micromechanisms in tension-compression fatigue of composite laminates containing transverse plies. *Composites Science and Technology 59*(2), 167–178.

Harris, B. (2003). *Fatigue in composites: science and technology of the fatigue response of fibre-reinforced plastics*. Boca Raton, Florida : Cambridge, England: CRC Press ; Woodhead Pub. Ltd.

Hong, Y., H. Ma, and W. Q. Meeker (2010). A tool for evaluating time-varying-stress accelerated life test plans with log-location-scale distributions. *IEEE Transactions on Reliability 59*, 620–627.

Karlak, R. F., F. W. Crossman, J. J. Grant, and U. Meier (1974). Interface failures in composites. In J. N. Fleck and R. L. Mehan (Eds.), *Failure Modes in Composites II*, University of Pittsburgh, pp. 119–130. The Metallurgical Society of AIME.

Kohout, J. and S. Vechet (2001). A new function for fatigue curves characterization and its multiple merits. *International Journal of Fatigue 23*(2), 175–183.

Konur, O. and F. Matthews (1989). Effect of the properties of the constituents on the fatigue performance of composites: a review. *Composites 20*(4), 317–328.

Meeker, W. Q. (1984). A comparison of accelerated life test plans for Weibull and lognormal distributions and type I censoring. *Technometrics 26*(2), 157–171.

Meeter, C. A. and W. Q. Meeker (1994). Optimum accelerated life tests with a nonconstant scale parameter. *Technometrics 36*(1), 71–83.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. New York: Wiley-Interscience.

Nelson, W. (2005). A bibliography of accelerated test plans. *IEEE Transactions on Reliability 54*(2), 194–197.

Nelson, W. and T. J. Kielpinski (1976). Theory for optimum censored accelerated life tests for normal and lognormal life distributions. *Technometrics 18*(1), 105–114.

Nelson, W. and W. Q. Meeker (1978). Theory for optimum accelerated censored life tests for weibull and extreme value distributions. *Technometrics 20*(2), 171–177.

Niu, K. and R. Talreja (2000). Modeling of compressive failure in fiber reinforced composites. *International Journal of Solids and Structures 37*(17), 2405–2428.

Online. Typical cyclic loading parameters. `http://www.fea-optimization.com/ETBX/stresslife_help.html`.

Owen, M. J. (1980). Fatigue processes in fibre reinforced plastics. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 294*(1411), 535–543.

Owen, M. J. and R. J. Howe (1972). The accumulation of damage in a glass-reinforced plastic under tensile and fatigue loading. *Journal of Physics D: Applied Physics 5*(9), 1637.

Palmgren, A. (1924). Die lebensdauer von kugallagern. *Zeitschrift des Vereines Deutscher Ingenieure 68*(14), 339–341.

Pascual, F. G. (2003). Theory for optimal test plans for the random fatigue-limit model. *Technometrics 45*(2), 130–141.

Pascual, F. G. (2006). Accelerated life test plans robust to misspecification of the stress-life relationship. *Technometrics 48*(1), 11–25.

Pascual, F. G. and W. Q. Meeker (1999). Estimating fatigue curves with the random fatigue-limit model. *Technometrics 41*(4), 277–290.

Reifsnider, K. (1980). Fatigue behavior of composite materials. *International Journal of Fracture 16*(6), 563–583.

Reifsnider, K. and A. Talug (1980). Analysis of fatigue damage in composite laminates. *International Journal of Fatigue 2*(1), 3–11.

Spindel, J. E. and E. Haibach (1981). Some considerations in the statistical determination of the shape of S-N curves. In R. E. Little and J. C. Ekvall (Eds.), *Statistical Analysis of Fatigue Data*, ASTM STP 744, pp. 89–113. American Society for Testing and Materials.

Stinchcomb, W. W., K. Reifsnider, R. S. Williams, and L. A. Marcus (1974). Frequency dependent fatigue damage. In J. N. Fleck and R. L. Mehan (Eds.), *Failure Modes in Composites II*, University of Pittsburgh, pp. 131–143. The Metallurgical Society of AIME.

Strohmeyer, C. E. (1914). The determination of fatigue limits under alternating stress conditions. *Proceedings of the Royal Society of London. Series A 90*(620), 411–425.

Stüssi, F. (1955). *Die Theorie der Dauerfestigkeit und die Versuche von August Wöhler*. Verlag VSB.

Talreja, R. (1981). Fatigue of composite materials: Damage mechanisms and fatigue-life diagrams. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 378*(1775), 461–475.

Vassilopoulos, A. P. and T. Keller (2011). *Fatigue of Fiber-reinforced Composites*. London: Springer.

Weibull, W. (1949). A statistical report of fatigue failure in solids. In *Transactions*, Volume 27, Stockholm. Royal Institute of Technology of Sweden.

Wöhler, A. (1870). Über die Festigkeitsversuche mit Eisen und Stahl. *Zeitschrift für Bauwesen 20*, 73–106.

3.A Formulae for the Fisher Information Matrix

This appendix provides the derivatives of the log likelihood given in Section 3.3.2.1, and the details of the approach that we used to calculate needed expectations, and the large-sample approximate variance-covariance matrix of the ML estimators of $\boldsymbol{\theta} = (A, B, \nu)'$.

Let $c = (\sigma_u/\sigma - 1)\,(\sigma_u/\sigma)^{\gamma(\alpha)-1}\,[1-\psi(R)]^{-\gamma(\alpha)}$. Then

$$\mu(A,B) = \frac{1}{B}\log\left\{\left(\frac{B}{A}\right)f^B c + 1\right\} = \frac{\log(D) - \log(A)}{B},$$

where $D = A + Bf^B c$. Let

$$\begin{aligned}
\mu_1 &= \frac{\partial \mu(A,B)}{\partial A} = -\frac{1}{AB} + \frac{1}{BD} \\
\mu_2 &= \frac{\partial \mu(A,B)}{\partial B} = \frac{\log(A)}{B^2} - \frac{\log(D)}{B^2} + \frac{E}{BD} \\
\mu_{11} &= \frac{\partial^2 \mu(A,B)}{\partial A^2} = \frac{1}{A^2 B} - \frac{1}{BD^2} \\
\mu_{12} &= \frac{\partial^2 \mu(A,B)}{\partial A \partial B} = \frac{1}{AB^2} - \frac{1}{B^2 D} - \frac{E}{BD^2} \\
\mu_{22} &= \frac{\partial^2 \mu(A,B)}{\partial B^2} = -\frac{2\log(A)}{B^3} + \frac{2\log(D)}{B^3} - \frac{2E}{B^2 D} - \frac{E^2}{BD^2} + \frac{F}{BD}.
\end{aligned}$$

Here, $E = Bf^B \log(f)c + f^B c$, and $F = Bf^B[\log(f)]^2 c + 2f^B \log(f)c$.

Considering a log-likelihood for a single observation, the first partial derivatives with respect to the unknown parameters can be expressed as

$$\begin{aligned}
\frac{\partial l}{\partial A} &= -\frac{d}{\nu}\frac{\phi'(z)}{\phi(z)}\mu_1 + \frac{(1-d)}{\nu}\frac{\phi(z_m)}{1-\Phi(z_m)}\mu_1 \\
\frac{\partial l}{\partial B} &= -\frac{d}{\nu}\frac{\phi'(z)}{\phi(z)}\mu_2 + \frac{(1-d)}{\nu}\frac{\phi(z_m)}{1-\Phi(z_m)}\mu_2 \\
\frac{\partial l}{\partial \nu} &= -\frac{d}{\nu}\left[1 + \frac{\phi'(z)z}{\phi(z)}\right] + \frac{(1-d)}{\nu}\frac{\phi(z_m)z_m}{1-\Phi(z_m)}.
\end{aligned}$$

The second partial derivatives are

$$\begin{aligned}
\frac{\partial^2 l}{\partial A^2} =& \frac{d}{\nu^2}\left[\frac{\phi'(z)}{\phi(z)}\right]'\mu_1^2 - \frac{(1-d)}{\nu^2}\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'\mu_1^2 - \frac{d}{\nu}\frac{\phi'(z)}{\phi(z)}\mu_{11} + \frac{(1-d)}{\nu}\frac{\phi(z_m)}{1-\Phi(z_m)}\mu_{11} \\
\frac{\partial^2 l}{\partial B^2} =& \frac{d}{\nu^2}\left[\frac{\phi'(z)}{\phi(z)}\right]'\mu_2^2 - \frac{(1-d)}{\nu^2}\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'\mu_2^2 - \frac{d}{\nu}\frac{\phi'(z)}{\phi(z)}\mu_{22} + \frac{(1-d)}{\nu}\frac{\phi(z_m)}{1-\Phi(z_m)}\mu_{22} \\
\frac{\partial^2 l}{\partial \nu^2} =& \frac{d}{\nu^2}\left(1 + \frac{2\phi'(z)z}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z^2\right) - \frac{(1-d)}{\nu^2}\left(\frac{2\phi(z_m)z_m}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m^2\right) \\
\frac{\partial^2 l}{\partial A \partial B} =& \frac{d}{\nu^2}\left[\frac{\phi'(z)}{\phi(z)}\right]'\mu_1\mu_2 - \frac{(1-d)}{\nu^2}\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'\mu_1\mu_2 \\
& - \frac{d}{\nu}\frac{\phi'(z)}{\phi(z)}\mu_{12} + \frac{(1-d)}{\nu}\frac{\phi(z_m)}{1-\Phi(z_m)}\mu_{12} \\
\frac{\partial^2 l}{\partial A \partial \nu} =& \frac{d}{\nu^2}\left(\frac{\phi'(z)}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z\right)\mu_1 - \frac{(1-d)}{\nu^2}\left(\frac{\phi(z_m)}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m\right)\mu_1 \\
\frac{\partial^2 l}{\partial B \partial \nu} =& \frac{d}{\nu^2}\left(\frac{\phi'(z)}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z\right)\mu_2 - \frac{(1-d)}{\nu^2}\left(\frac{\phi(z_m)}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m\right)\mu_2.
\end{aligned}$$

Here,

$$\left[\frac{\phi'(z)}{\phi(z)}\right]' = \frac{\partial}{\partial z}\left[\frac{\phi'(z)}{\phi(z)}\right], \text{ and } \left[\frac{\phi(z)}{1-\Phi(z)}\right]' = \frac{\partial}{\partial z}\left[\frac{\phi(z)}{1-\Phi(z)}\right].$$

Detailed formulae of $\Phi, \phi, \phi/(1-\Phi), [\phi/(1-\Phi)]', \phi'/\phi$, and $[\phi'/\phi]'$ for commonly-used location-scale distributions (normal, smallest extreme value, largest extreme value, and logistic, corresponding to the lognormal, Weibull, Fréchet, and loglogistic log-location-scale distributions) can be found in Table I of Hong et al. (2010). The scaled expectations of the second derivatives are

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial A^2}\right] =& -\int_{-\infty}^{z_m} \mu_1^2\left[\frac{\phi'(z)}{\phi(z)}\right]'\phi(z) - \nu\mu_{11}\frac{\phi'(z)}{\phi(z)}\phi(z)dz \\
&+ \mu_1^2\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'[1-\Phi(z_m)] - \nu\mu_{11}\phi(z_m)
\end{aligned} \tag{3.7}$$

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial B^2}\right] =& -\int_{-\infty}^{z_m} \mu_2^2\left[\frac{\phi'(z)}{\phi(z)}\right]'\phi(z) - \nu\mu_{22}\frac{\phi'(z)}{\phi(z)}\phi(z)dz \\
&+ \mu_2^2\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'[1-\Phi(z_m)] - \nu\mu_{22}\phi(z_m)
\end{aligned} \tag{3.8}$$

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial \nu^2}\right] =& -\int_{-\infty}^{z_m}\left(1 + \frac{2\phi'(z)z}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z^2\right)\phi(z)dz \\
&+ \left(\frac{2\phi(z_m)z_m}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m^2\right)[1-\Phi(z_m)]
\end{aligned} \tag{3.9}$$

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial A\partial B}\right] =& -\int_{-\infty}^{z_m} \mu_1\mu_2\left[\frac{\phi'(z)}{\phi(z)}\right]'\phi(z) - \nu\mu_{12}\frac{\phi'(z)}{\phi(z)}\phi(z)dz \\
&+ \mu_1\mu_2\left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'[1-\Phi(z_m)] - \nu\mu_{12}\phi(z_m)
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial A\partial \nu}\right] =& -\int_{-\infty}^{z_m} \mu_1\left(\frac{\phi'(z)}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z\right)\phi(z)dz \\
&+ \mu_1\left(\frac{\phi(z_m)}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m\right)[1-\Phi(z_m)]
\end{aligned} \tag{3.11}$$

$$\begin{aligned}
\nu^2\mathbf{E}\left[-\frac{\partial^2 l}{\partial B\partial \nu}\right] =& -\int_{-\infty}^{z_m} \mu_2\left(\frac{\phi'(z)}{\phi(z)} + \left[\frac{\phi'(z)}{\phi(z)}\right]'z\right)\phi(z)dz \\
&+ \mu_2\left(\frac{\phi(z_m)}{1-\Phi(z_m)} + \left[\frac{\phi(z_m)}{1-\Phi(z_m)}\right]'z_m\right)[1-\Phi(z_m)].
\end{aligned} \tag{3.12}$$

3.B Optimization Procedure

The following algorithm represents the procedure used during optimization:

1. Given k and s, determine all P possible sample allocation schemes $\boldsymbol{\pi}^p = (\pi_1, \ldots, \pi_s)^t, p = 1, \ldots, P$ (t denotes the transpose operator). If $\pi_{\min}$ is specified, all sample allocations that include values of $\pi_i < \pi_{\min}$ are excluded. This first step takes advantage of the constraint on $k_i = k\pi_i$ being an integer value.

2. For each $\boldsymbol{\pi}^p$:

   (a) Divide the planning range into a grid of values separated by a specified grid size "$q-$step". Select all D possible combinations $\mathbf{q}^{pd} = (q_1, \ldots, q_s)^t, d = 1, \ldots, D$ of design points using these grid points. If a minimum distance between any two design points is specified, remove those combinations that contain grid points closer than this distance. Evaluate each plan $\boldsymbol{\eta}^{pd} = ((\mathbf{q}^{pd})^t, (\boldsymbol{\pi}^p)^t)^t$ and select the $\mathbf{q}^{pd}_{\min}$ that yields the minimum value of the optimality criterion. This step was required due to the sensitivity of the optimization to the initial values.

   (b) Using $\mathbf{q}^{pd}_{\min}$ as the initial values, optimize over the design point space to find the overall minimum value of the optimality criterion for the given $\boldsymbol{\pi}^p$. This was performed using the free statistical software package R using the BFGS algorithm option in the optim() function. To ensure stability in the optimization, the design points were first transformed by the following function, which maps values from the range $[a, b]$ to the real line. Specifically, given a value x the transformed value x^T is given by,

$$x^T = -\log\left[\frac{(b-x)}{(x-a)}\right].$$

   To enforce the time constraint, a constant penalty term $M = 10{,}000$ was added to the result if the expected total time exceeded the maximum total time.

3. Once a test plan $\boldsymbol{\eta}^p$ is obtained for each $\boldsymbol{\pi}^p$, select the test plan $\boldsymbol{\eta}_{\min}$ that yields the minimum value of the optimality criterion across all $\boldsymbol{\pi}^p$. This is the optimal/compromise plan.

Chapter 4 A Comparison of Least Squares and Maximum Likelihood Approaches to Estimating Thermal Indices for Polymeric Materials

4.1 Introduction

4.1.1 Motivation

Polymeric materials, including thermoplastic, thermosetting, and elastomeric materials, are widely used in many industrial applications. Accelerated destructive degradation testing (ADDT) is the most appropriate technique for evaluating long-term performance of such polymeric materials. For example, ADDT is often used to evaluate properties which degrade over time such as tensile strength. The destructive nature of the testing results from measuring tensile strength by stretching the material until failure. When the tensile strength of an aged material reaches a certain percentage (e.g., 50%) of the original tensile strength, a failure is said to have occurred. The time taken to reach this point is referred to as the time to failure. Since polymeric materials can last for long periods of time under standard use conditions, accelerated testing, of which ADDT is a special case, is used to reduce the amount of time required to observe a failure. For the tensile strength testing considered here, the acceleration factor is temperature. Since the test measurements are destructive, only one measurement can be taken from a single sample unit at a particular time and temperature. UL746B (2013) specifies the current industry guidelines for long-term property evaluation of polymeric materials via ADDT.

A thermal index (TI) or relative thermal index (RTI) is often used to characterize long-term performance of polymeric materials. According to UL746B (2013), a TI of a material serves as a measure of the material's ability to retain a specific property (physical, electrical, etc.) under exposure to elevated temperatures over a specified period of time (usually 100,000

hours). For example, a TI of 200°C may be interpreted as the material can reasonably be expected to maintain properties at a temperature of 200°C for 100,000 hours. The TI can be used for the purposes of determining if a material is acceptable for a particular application. It can also be used as a means of comparing multiple materials for an end use. A material with a higher TI possesses a stronger resistance to thermal damage and so provides a competitive advantage over similar materials with lower TIs. If the material is new to a particular field, the TI of this material can be compared to a list of known TI values, such as the table containing the RTI of common polymeric materials given in UL746B (2013).

There are two approaches to estimating the TI based on data collected from ADDT: the least squares (LS) procedure and the maximum likelihood (ML) procedure. The LS procedure is specified in the current industrial standards and is widely used in industrial applications. The ML procedure is frequently used in the statistical literature but rarely seen in industrial ADDT applications. At present, the ML procedure is not specified in current standards. In this paper, we will compare these two procedures using two motivating datasets from the literature and a simulation study. Due to its greater flexibility in uncertainty quantification and its use in other aspects of statistical analysis, we propose that the ML procedure is a viable alternative to the LS procedure. It is our intent that a comparison of the two procedures should be considered in the designation of statistical methods for industrial standards.

4.1.2 Related Literature

Numerous industrial standards address property testing using an applications approach. For example, UL746B (2013) gives the current guidelines for long term property evaluations for polymeric materials. Other guidelines exist that are specific to thermal testing. IEEE-STD-98-2002 (2002) is the IEEE standard for the preparation of test procedures for the thermal evaluation of solid electrical insulating materials. IEEE-STD-101-1987 (2010) is the IEEE guide for the statistical analysis of thermal life test data. Finally, IEEE-STD-421.1-2007 (2007) specifies the IEEE standard general principles for temperature limits in the rating of electric equipment and for the evaluation of electrical insulation.

In contrast to the industrial standards, the statistical literature describes reliability analysis and test planning using a statistical theory approach. Lu and Meeker (1993) first use degradation data to obtain product reliability information. Tseng and Wen (2000) propose an accelerated degradation analysis technique for highly reliable products. Meeker, Hong, and Escobar (2011, 2011a) give a review of commonly used degradation models and data analysis techniques. Recent development in test planning includes Shi et al. (2009), which proposes procedures for accelerated destructive degradation test planning. Shi and Meeker (2012) perform ADDT test planning using Bayesian methods. The applications of parametric ADDT analysis include Escobar et al. (2003), Tsai et al. (2013), and Li and Doganaksoy (2014). A general introduction to statistical methods for reliability data can be found in Meeker and Escobar (1998).

4.1.3 Overview

This paper will compare the methods presented in both the industrial and statistical literature and illustrate a proposed approach for estimating thermal indices with motivating examples. Section 4.2 provides an overview of ADDT for long-term material property evaluation including test designs and data. The motivating examples are also introduced in this section. Section 4.3 describes the LS and ML procedures for estimating the TI from ADDT data. Each procedure is illustrated using a published dataset from the literature. Section 4.4 will give a comparison of both methods via a simulation study. Section 4.5 will address additional considerations for the ML procedure including uncertainty quantification, prediction, and software implementation. In section 4.6, both procedures are applied to an additional published dataset based on an alternative test design to help illustrate further the discrepancies between the LS and ML procedures. Finally, section 4.7 will present conclusions and recommendations for ADDT analysis.

4.2 Accelerated Destructive Degradation Testing

4.2.1 Test Designs and Data

Industry standards give detailed recommendations on how ADDT is to be conducted for polymeric materials. Table 4.1 shows an example of a test plan in accordance with the UL standard. Specifically, this test plan was designed based on the "Fixed Time Sampling Method" laid out in Section 22 of UL746B (2013). According to this section, at least four temperature levels are required with a span of no less than 10°C between each temperature level. Measurements are made according to a predetermined time schedule, with the UL standard suggesting preferred time points. At least five samples are tested at each combination of time and temperature level with an additional ten samples tested at the initial time for a baseline measurement (usually tested at the lowest temperature level of the experiment at 48 hours). The UL standard uses 50% of the baseline measurement as the failure threshold though this may vary for other materials. A set of measurements is required after the degradation passes the 50% threshold in the LS procedure. The data resulting from the test are then used to obtain the TI of the material.

Table 4.1: Illustration of a test plan that satisfies the requirements of UL746B (2013). The table shows the number of samples allocated to each combination of temperature levels and measuring time points.

Temperature	Time Points (Hours)					
Levels (°C)	48	552	1008	2016	3528	5040
-	10					
240		5	5	5	5	5
250		5	5	5	5	5
260		5	5	5	5	5
270		5	5	5	5	5

In practice, material scientists often have prior knowledge about products that lead to an unbalanced test plan as illustrated in Table 4.2. The plan in Table 4.2 eliminates test combinations beyond which samples will show little or no degradation. In addition, combinations at

which samples will excessively degrade, even to a potentially non-measurable state, are also eliminated in Table 4.2. These types of combinations do not contribute additional information to the calculation of a TI.

It is important for practitioners to note that the statistical procedures described in this chapter can be used for either test plan. As will be seen in Section 4.2.2, the first motivating example consists of a balanced test plan similar to Table 4.1. For simplicity, the simulations in Section 4.4 will also be based on a balanced test plan.

Table 4.2: Illustration of a test plan that satisfies the requirements of UL746B (2013) with elimination of non-informative points. The table shows the number of samples allocated to each combination of temperature levels and measuring time points.

Temperature	Time Points (Hours)					
Levels (°C)	48	552	1008	2016	3528	5040
-	10					
220						5
230					5	5
240				5	5	5
250			5	5	5	5
260		5	5	5	5	5
270		5	5	5	5	
280		5	5	5		
290		5	5			
300		5				

To help with the presentation of each procedure, it is necessary to first introduce some notation. Let y_{ijk} be the degradation measurement (tensile strength) for the kth sample for temperature level i at time point j. Let TempC_i be the temperature in degrees Celsius for temperature level i and let t_j be the time in hours up to time point j. Here, $k = 1, \ldots, n_{ij}$, where n_{ij} is the number of samples tested under temperature level i and time point j. For example, $i = 3, j = 4$, and $n_{ij} = 5$ in Table 4.1. The failure time t_f is typically defined as the time to reach $100p\%$ of the original strength, where p is a pre-determined level corresponding to material failure. For example, the UL standard specifies a failure at 50% of the original strength so $p = 0.5$. In some cases, the failure time is defined as the time to reach a fixed

strength. This can easily be seen as equivalent to the previous definition.

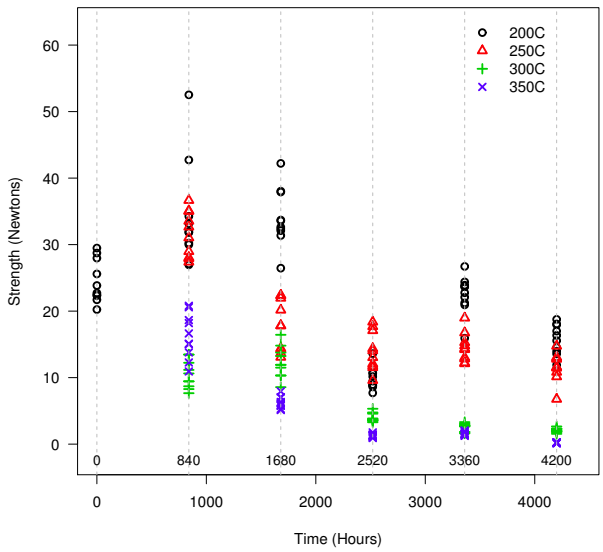
4.2.2 Motivating Examples

The motivation for this line of research is related to real world applications which cannot be disclosed publicly. Consequently, two published datasets will be used to illustrate and compare the LS and ML procedures. The first example is the Seal Strength dataset presented in Li and Doganaksoy (2014). The data are summarized in both Table 4.3 and Figure 4.1(a). The experiment consisted of testing the strength of a new seal. The test plan is similar to that discussed in the UL standard. In the original application in Li and Doganaksoy (2014), the failure is defined when the strength dropped below 0.5 Newtons (roughly 2% of the initial strength). For the purposes of comparison, we use a 50% failure threshold here. This dataset will be used to illustrate each of the procedures as they are introduced in Section 4.3.

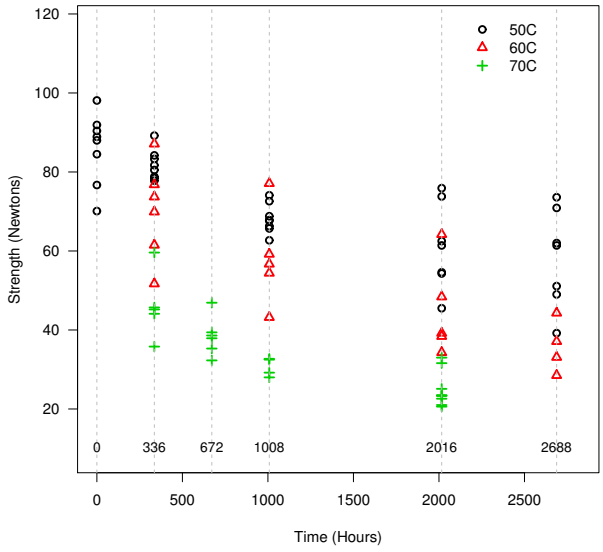
The second example is the Adhesive Bond B data from Escobar et al. (2003) containing the results of an ADDT on the strength of an adhesive bond. The data are summarized in both Table 4.4 and Figure 4.1(b). The response for the Adhesive Bond B data is the material tensile strength. Once again, for the purposes of comparison the failure threshold for the Adhesive Bond B material is set at 50% of the baseline measurement. The ADDT experiment discussed in Escobar et al. (2003) that generated this data is more similar to the "Fixed Temperature Sampling Method" discussed in Sections 13-21 of the UL standard UL746B 2013 as the data are not as balanced. For this method, the number of temperature levels is predetermined and fixed while the selection of time points is more flexible. This dataset will be used in the context of an application in Section 4.6 to further illustrate the differences between the LS and ML procedures.

Table 4.3: Data for Seal Strength Li and Doganaksoy (2014).

Temperature (°C)	Hours	Seal Strength (Newtons)									
100	0	28.74	25.59	22.72	22.44	29.48	23.85	20.24	22.33	21.70	27.97
200	840	52.52	30.23	31.90	33.15	34.26	31.82	27.10	30.00	26.96	42.73
	1680	31.37	37.91	38.03	42.21	32.64	32.10	32.37	33.59	26.46	33.69
	2520	9.47	13.61	8.95	8.61	10.16	8.82	8.84	10.73	10.63	7.70
	3360	26.72	21.24	22.76	24.39	15.93	23.90	22.09	23.69	23.67	20.94
	4200	14.53	17.95	11.90	17.00	15.56	18.07	13.96	13.57	16.35	18.76
250	840	28.97	35.01	27.39	36.66	27.91	31.03	32.65	35.08	28.05	33.54
	1680	14.29	20.16	22.35	21.96	13.67	14.40	22.37	13.08	17.81	17.82
	2520	9.59	14.37	12.08	11.79	17.69	14.05	17.08	11.52	13.03	18.37
	3360	14.23	12.83	13.02	16.74	12.11	12.24	18.97	15.29	14.38	14.80
	4200	14.75	11.54	11.57	10.83	12.78	10.14	11.45	12.91	13.06	6.76
300	840	10.63	8.28	13.46	13.47	9.44	7.66	11.16	8.70	9.44	12.23
	1680	10.34	13.24	8.57	11.93	13.76	16.44	14.81	11.50	11.92	10.30
	2520	3.86	4.76	5.32	3.74	4.58	3.62	3.58	3.47	3.29	3.63
	3360	2.89	3.31	1.81	1.61	2.65	2.83	2.70	2.79	1.83	3.08
	4200	1.95	1.55	2.19	2.00	2.00	2.33	1.80	2.34	1.88	2.66
350	840	13.79	15.10	20.58	18.20	16.64	10.93	12.28	18.65	20.80	15.04
	1680	5.78	5.90	6.99	7.94	7.06	5.13	5.80	6.20	5.30	6.34
	2520	1.34	0.92	1.31	1.76	1.30	1.47	1.11	1.25	1.02	1.30
	3360	1.24	1.57	2.06	1.56	1.94	1.39	1.91	1.44	1.61	1.50
	4200	0.27	0.20	0.26	0.26	0.27	0.18	0.13	0.20	0.13	0.21



(a) Seal Strength



(b) Adhesive Bond B

Figure 4.1: Scatter plots of the data for the two motivating examples.

Table 4.4: Data for Adhesive Bond B Escobar et al. (2003).

Temperature (°C)	Hours	Strength (Newtons)								
50	0	70.1	76.7	84.5	88.0	88.9	90.4	91.9	98.1	
50	336	77.8	78.4	78.8	80.5	81.7	83.3	84.2	89.2	
	1008	62.7	65.7	66.3	67.7	67.8	68.8	72.6	74.1	
	2016		45.5	54.3	54.6	61.4	62.5	73.8	75.9	
	2688		39.2	49.0	51.1	61.4	62.0	70.9	73.6	
60	336	51.7	61.5	69.9	73.7	76.8	87.1			
	1008		43.2	54.4	56.7	59.2	77.1			
	2016		34.3	38.4	39.2	48.4	64.2			
	2688			28.5	33.1	37.1	44.3			
70	336		35.8	44.1	45.2	45.7	59.6			
	672	32.3	35.3	37.9	38.6	39.4	46.9			
	1008			28.0	29.2	32.5	32.7			
	2016	20.6	21.0	22.6	23.3	23.4	23.5	25.1	31.6	33.0

4.3 Estimation Procedures

4.3.1 The Least Squares Procedure

The LS procedure is the name that will be given to the methodology described in UL746B (2013) as the primary method for estimation is based on the least squares approach. This procedure is the current accepted standard for ADDT data analysis in industry. The procedure is outlined as follows:

Least Squares Procedure

1. Compute the averages of the sample measurements at each combination of time and temperature level.

2. Fit a least-squares polynomial curve to the averages for each temperature level.

3. Interpolate the failure time for each temperature level using the estimated polynomial

curve.

4. Fit a least-squares line to the failure time data.

5. Use the fitted line to estimate the TI for the material.

Note that if the fitted mean degradation level from the polynomial regression has not yet reached the failure threshold by the last time point of a particular temperature level, all data collected from that temperature must be excluded from the analysis as there is insufficient information to interpolate the failure time.

A third-order polynomial is typically used in steps 2 and 3 to interpolate the failure time for each temperature level. Specifically, for temperature level i, we seek t_{fi} satisfying

$$a_0 + a_0 t_{fi} + a_0 t_{fi}^2 + a_0 t_{fi}^3 = p\alpha, \tag{4.1}$$

where α is the initial strength of the material. Figure 4.2 gives an example of the polynomial interpolation for the Seal Strength data. In step 4, a least-squares line is then fit to the failure time data according to the following model based on the Arrhenius relationship:

$$\log_{10}(t_{fi}) = \beta_0 + \beta_1 x_i, \tag{4.2}$$

where $x_i = 1/(\text{TempC}_i + 273.16)$. The estimated coefficients for the Seal Strength data based on the LS procedure are $\widehat{\beta}_0 = 0.568$ and $\widehat{\beta}_1 = 1487.999$. Finally, in step 5, the TI at a specific value of t_d (e.g., 100,000 hours), which is used in the UL standard, is computed as

$$T = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16\,. \tag{4.3}$$

The estimate $\widehat{T}$ is found by evaluating (4.3) at the least squares estimates of β_0 and β_1. The estimated TI for the Seal Strength data based on the LS procedure is 63°C and the fitted time-temperature line is shown in Figure 4.4.

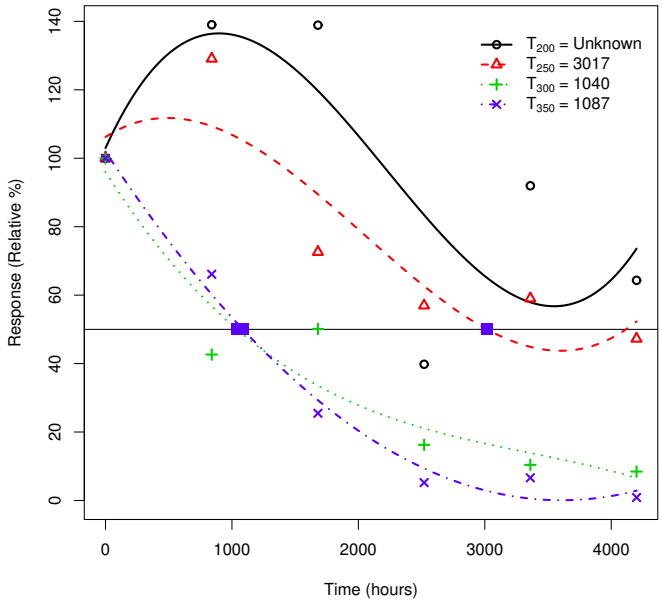


Figure 4.2: Polynomial interpolation from the LS procedure for the Seal Strength data. The labels give the interpolated failure time for each temperature, as marked by a square on the failure threshold (50%) line.

The primary advantage of the LS procedure is the computational simplicity of each step. However, it suffers from several drawbacks. The procedure uses polynomial regression to model the degradation curves, which may serve well as a curve-fitting technique, but is known to have issues on the boundaries of the data, as can be seen in Figure 4.2. While this may not be as strong a disadvantage due to the lack of interest in extrapolation beyond the data, these polynomial models have no physical basis in the process generating the data and so can serve only as crude approximations. Furthermore, by averaging the observations at each time and temperature level, the procedure ignores information that could be gleaned from the variation in and between batches.

Another issue related to the first stage of the procedure lies in imposing the constraint of requiring observations below the failure threshold. While this might ensure more accurate estimates of the failure time, it severely restricts the usage of the procedure. Note that in Figure 4.2, even though the mean observation at 250°C and 40,000 hours is below the 50%

threshold, the estimated mean from the polynomial regression is not below the threshold resulting in the data for this temperature being discarded. The constraint cannot be resolved by attempting to extrapolate to the failure threshold due to the aforementioned boundary instability of the polynomial regression. This can become an issue when time constraints are present or measurements are hard to obtain below the failure threshold.

A prominent weakness in the procedure involves its two-step nature for estimating the TI. Not only does this make the quantification of uncertainty regarding the estimate rather difficult, but it also makes the procedure susceptible to errors. The error in model-fitting in the first stage will pass on into the second stage propagating throughout the entire procedure.

4.3.2 The Maximum Likelihood Procedure

The ML procedure is the name that will be given to the methodology used in the majority of the statistical literature on ADDT. This is because the primary method for estimation and inference is based on a parametric model and maximum likelihood theory. The procedure is outlined as follows:

Maximum Likelihood Procedure

1. Based on the degradation measurement under investigation (e.g., tensile strength), use a parametric model and probability distribution that is adequate for describing the degradation path and construct the likelihood function.

2. Estimate the parameters of the degradation function by finding those parameter values that maximize the likelihood of the data.

3. Use the fitted function to estimate the TI for the material.

The ML procedure assumes the underlying degradation path is some function $\mu(t, x)$ of

time t and acceleration level x. The model for the degradation measurement is

$$y_{ijk} = \mu(t_j, x_i) + \epsilon_{ijk}, \tag{4.4}$$

where $x_i = 1/(\text{TempC}_i + 273.16)$ and ϵ_{ijk} is an error term. Because of this model specification, the ML procedure is able to estimate the complete model in a single step.

The function $\mu(t, x)$ is typically specified to be a decreasing function of t because the tensile strength should be decreasing over time. It is also specified to be an increasing function of x (i.e., decreasing function of temperature) because higher temperature should lead to a higher rate of degradation. For polymer materials, the following parametric form is suitable for the form of $\mu(t, x)$ Vaca-Trigo and Meeker e.g., 2009:

$$\mu(t, x) = \frac{\alpha}{1 + \left[\dfrac{t}{\eta(x)}\right]^{\gamma}}, \tag{4.5}$$

where $\eta(x) = \exp(\nu_0 + \nu_1 x)$ is based on the Arrhenius relationship.

Here, the error term is modeled as $\epsilon_{ijk} \sim \text{N}(0, \sigma^2)$ and is used to describe unit to unit variability. Since measurements of samples tested from the same batch could possibly be correlated, we model the correlation as $\text{Corr}(\epsilon_{ijk}, \epsilon_{ijk'}) = \rho$. The unknown parameters $\boldsymbol{\theta} = (\nu_0, \nu_1, \alpha, \gamma, \sigma, \rho)'$ are estimated by the maximum likelihood method. The likelihood is given as

$$L(\boldsymbol{\theta}) = \prod_{i,j} f(\boldsymbol{y}_{ij}; \boldsymbol{\theta}) = \prod_{i,j} (2\pi)^{-\frac{n_{ij}}{2}} |\Sigma_{ij}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} [\boldsymbol{y}_{ij} - \boldsymbol{\mu}(t_j, x_i)]' \Sigma_{ij}^{-1} [\boldsymbol{y}_{ij} - \boldsymbol{\mu}(t_j, x_i)] \right\} \tag{4.6}$$

where $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, \ldots, y_{ijn_{ij}})'$ is the corresponding vector of degradation measurements and $f(\boldsymbol{y}_{ij}; \boldsymbol{\theta})$ is the multivariate normal density with mean vector $\boldsymbol{\mu}(t_j, x_i)$, a $n_{ij} \times 1$ column vector of $\mu(t_j, x_i)$'s, and variance-covariance matrix Σ_{ij}, a $n_{ij} \times n_{ij}$, a matrix with σ^2 on the diagonal entries and $\rho\sigma^2$ on the off-diagonal entries. The parameter estimates $\widehat{\boldsymbol{\theta}}$ consist of those values of the parameters that maximize this likelihood function. The parameter estimates for the Seal Strength data based on the ML procedure are $\widehat{\boldsymbol{\theta}} = (0.695, -8901.925, 3.421, 0.504, 1.057, 1.556)'$. Figure 4.3 gives an example of the degradation paths estimated for the Seal Strength data using (4.5) as the underlying path model.

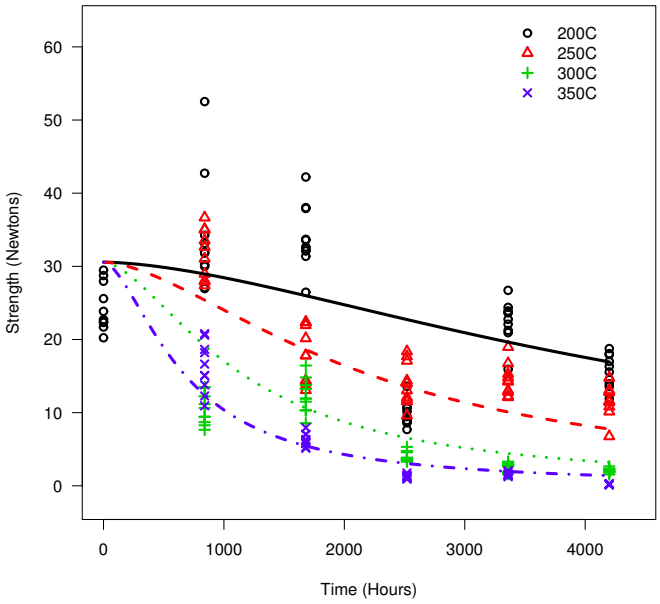


Figure 4.3: Parametric fit of the Seal Strength data using the ML procedure. The lines with different styles show the fitted mean functions for each temperature level.

The time to failure is obtained by solving $\mu(t_f, x) = p\alpha$ for t_f, leading to the model

$$\log_{10}(t_f) = \frac{\nu_0}{\log(10)} + \frac{\nu_1 x}{\log(10)} + \frac{1}{\gamma \log(10)} \log\left[\frac{1-p}{p}\right]. \tag{4.7}$$

When $p = 0.5$ (i.e., the failure threshold is 50%), as is the case in the UL standard, (4.7) reduces to

$$\log_{10}(t_f) = \beta_0 + \beta_1 x \tag{4.8}$$

where $\beta_0 = \nu_0/\log(10)$, and $\beta_1 = \nu_1/\log(10)$. The TI at t_d can be computed as

$$T = \frac{\beta_1}{\log_{10}(t_d) - \beta_0 - \frac{1}{\gamma}\log_{10}\left[\frac{1-p}{p}\right]} - 273.16 \tag{4.9}$$

If $p = 0.5$, (4.9) reduces to (4.3). The estimate $\widehat{T}$ is obtained by evaluating (4.9) at $\widehat{\boldsymbol{\theta}}$. The estimated TI for the Seal Strength data is 72°C and the fitted time-temperature line is shown in Figure 4.4.

The ML procedure has several advantages over the LS procedure. It is able to perform estimation in a single step rather than the two steps required in the LS procedure, making it

more robust to error propagation. Also, while a polynomial regression curve is quite flexible in data modeling, the ML procedure allows the analyst to incorporate subject-matter knowledge into the modeling procedure in both the choice of degradation path and distribution of the response. No restrictions are placed on the need to have degradation levels below the failure threshold as knowledge of the failure time is extracted from the estimated degradation path model. The procedure also considers the information contribution from each individual observation, resulting in estimation of within- and between-batch variation. Accounting for these two quantities allows for greater accuracy in determining the underlying degradation path, thus yielding more accurate measures of the TI. Furthermore, as will be discussed in Section 4.5, the ML procedure is able to quantify the uncertainty in TI estimates or estimates of any other desired quantity with relative ease.

The ML procedure is not without its faults, however. One disadvantage of the ML procedure is the greater complexity required to estimate the path model parameters compared to estimation in the LS procedure. This disadvantage is primarily related to the complexity present in the model selection and could be avoided by choosing a simpler model; however, current computing power and statistical software packages can readily handle such complexity as discussed in Section 4.5. Another potential weakness lies in the procedure's strong dependence on the specification of the degradation path and response distribution. However, this too can be remedied by consulting the subject-matter expertise of those involved in the material production. If there is still any uncertainty in the model selection, appropriate measures can be taken to assess any deviation from the assumed model via residual analysis and other similar procedures.

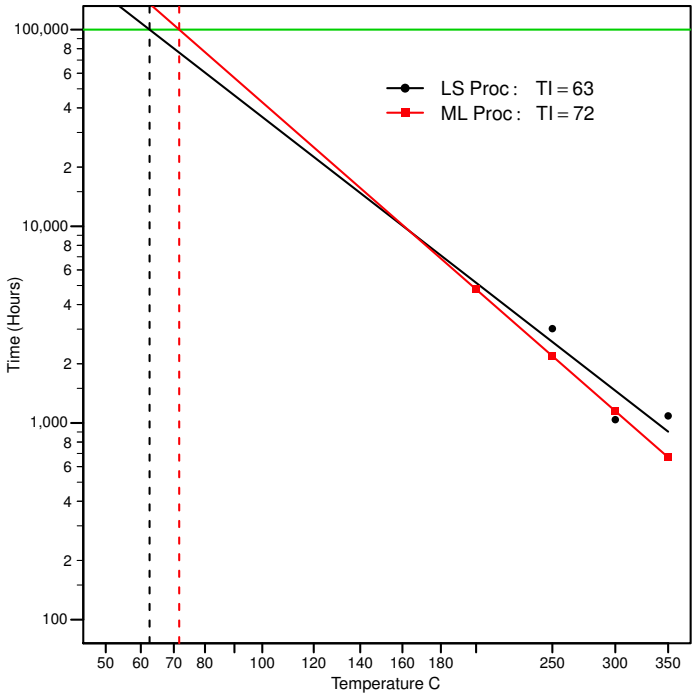


Figure 4.4: Fitted temperature-time relationship lines from both procedures and the corresponding estimated TIs for the Seal Strength data.

4.4 Simulation Study for Estimation Performance

In this section, we conduct a simulation study to compare the performance of the LS and ML procedures in terms of estimating the TI.

4.4.1 Simulation Setting

The data were simulated according to the model given in (4.5) with $\alpha = 9000, \nu_0 = -16\log(10), \nu_1 = 12500\log(10)$, and $\gamma = 2$. The failure threshold was the same as that in the UL standard, $p = 0.5$. The true TI at 100,000 hours is 181°C. Eight scenarios with varying numbers of time and temperature levels were considered: four with four time points and four with five time points. Consideration of fewer time points than those suggested by the UL standard would allow us to test each procedure's sensitivity to time constraints. The number of temperature levels ranged from three to five to assess sensitivity to constraints on accelerating factors. Two scenarios were considered for the case of four temperature levels,

with the range of levels consisting of either higher or lower temperature levels. This was done to determine the effect of the levels of accelerating factor in terms of distance from use levels. Details for each scenario are given in Table 4.5. There were 600 runs for each scenario. The choice of time schedule and temperature levels were based on the test plan from the original motivating data.

Table 4.5: The temperature levels and measuring time points for the eight simulation scenarios.

Scenarios	Temperature Levels (°C)					Time Points (Hours)				
1: Temp. 3, Time 4		250	260	270		552	1008	2016	3528	
2: Temp. 4, Time 4		250	260	270	280	552	1008	2016	3528	
3: Temp. 4, Time 4	240	250	260	270		552	1008	2016	3528	
4: Temp. 5, Time 4	240	250	260	270	280	552	1008	2016	3528	
5: Temp. 3, Time 5		250	260	270		552	1008	2016	3528	5040
6: Temp. 4, Time 5		250	260	270	280	552	1008	2016	3528	5040
7: Temp. 4, Time 5	240	250	260	270		552	1008	2016	3528	5040
8: Temp. 5, Time 5	240	250	260	270	280	552	1008	2016	3528	5040

4.4.2 Simulation Results

Table 4.6 shows the estimated mean, bias, standard deviation (SD), and root mean square error (RMSE) of the TI for the LS and ML procedures from the simulation study. Figure 4.5 plots the estimated mean and RMSE of the TI estimators for both procedures for each scenarios. The LS procedure appears to be sensitive to the number of time points whereas the ML procedure is relatively stable. As the number of temperature levels increases, the RMSE of the ML procedure decreases, as would be expected, but for the LS procedure the change in RMSE is not as monotone. The LS procedure also appears to be more sensitive to the selection of temperature levels than the ML procedure. Since higher levels of acceleration and increased observation times give a greater chance of yielding failures, the LS procedure's improved behavior in these scenarios is intuitive as less failures yields less data for this procedure. Thus, a good TI estimate from the LS procedure would require more temperature levels and more time points while the ML procedure is more stable across varying numbers

Table 4.6: Estimated mean, bias, SD, and RMSE of the TI estimators for the LS and ML procedures from the simulation study.

Scenarios	True TI	Mean		Bias		SD		RMSE	
		LS	ML	LS	ML	LS	ML	LS	ML
1: Temp. 3, Time 4	181	167	178	14	3	26	15	29	15
2: Temp. 4, Time 4	181	165	181	16	0	26	9	31	9
3: Temp. 4, Time 4	181	165	181	16	0	26	9	31	9
4: Temp. 5, Time 4	181	174	181	7	1	15	7	16	7
5: Temp. 3, Time 5	181	177	178	4	3	17	14	17	14
6: Temp. 4, Time 5	181	181	179	0	2	10	10	10	10
7: Temp. 4, Time 5	181	175	180	6	1	13	8	14	8
8: Temp. 5, Time 5	181	179	180	3	1	8	7	9	7

of time and temperature levels. In summary then, the ML procedure has superior estimation properties to the LS procedure.

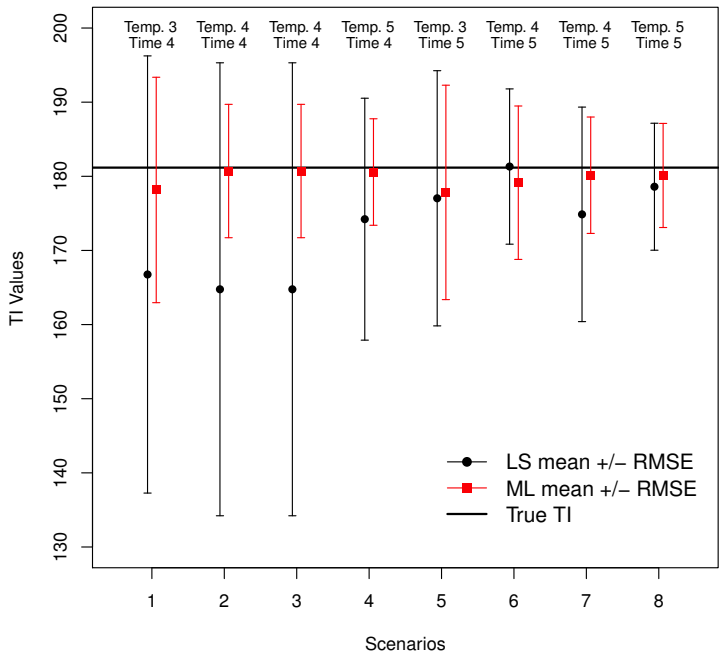


Figure 4.5: Plot of the estimated mean and RMSE of the TI estimators for the LS and ML procedures from the simulation study. The labels across the top correspond to the particular scenario represented.

4.5 Advantages of the ML Procedure

4.5.1 Uncertainty Quantification

As mentioned in Section 4.3, one important advantage of the ML procedure is its ability to quantify uncertainty regarding $\widehat{T}$. This can be accomplished through the construction of an approximate confidence interval based on the asymptotic normality of the maximum likelihood estimates. Specifically, the ML estimator $\widehat{\boldsymbol{\theta}}$ is asymptotically distributed as $\mathrm{N}(\boldsymbol{\theta}, \Sigma_{\widehat{\boldsymbol{\theta}}})$. Here, $\Sigma_{\widehat{\boldsymbol{\theta}}}$ is the large-sample variance-covariance matrix obtained as the inverse of the local information matrix $I_{\widehat{\boldsymbol{\theta}}}$ given by

$$I_{\widehat{\boldsymbol{\theta}}} = -\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.$$

Let $\widehat{T}$ be the ML estimator of the TI as given in (4.9). Then, by the delta method, it follows that

$$\widehat{T} \dot{\sim} \mathrm{N}\left(T, \left[\frac{\partial T(\widehat{\boldsymbol{\theta}})}{\partial \widehat{\boldsymbol{\theta}}}\right]' \Sigma_{\widehat{\boldsymbol{\theta}}} \left[\frac{\partial T(\widehat{\boldsymbol{\theta}})}{\partial \widehat{\boldsymbol{\theta}}}\right]\right), \tag{4.10}$$

where $\partial T(\widehat{\boldsymbol{\theta}})/\partial \widehat{\boldsymbol{\theta}} = \partial T(\boldsymbol{\theta})/\partial \boldsymbol{\theta}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$. A large-sample approximate $(1-q) \times 100\%$ confidence interval for the TI is then given by

$$\widehat{T} \pm z_{1-q/2} \sqrt{\left[\frac{\partial T(\widehat{\boldsymbol{\theta}})}{\partial \widehat{\boldsymbol{\theta}}}\right]' \Sigma_{\widehat{\boldsymbol{\theta}}} \left[\frac{\partial T(\widehat{\boldsymbol{\theta}})}{\partial \widehat{\boldsymbol{\theta}}}\right]}.$$

The approximation of the interval is with regards to the coverage probability, which will approach the desired confidence as the sample size increases. A 95% confidence interval for the Seal Strength data is $[14.942°\mathrm{C}, 128.420°\mathrm{C}]$. Note that this interval contains the LS point estimate of the TI (63°C). There is a large amount of variation in the data leading to a wide confidence interval for this data. As will be seen in Section 4.6, such wide intervals are not the norm.

4.5.2 Material Comparisons

The UL standards also define a relative thermal index (RTI) for the comparison of materials of different brands. For example, if a control material has an established TI of 180°C, an RTI,

denoted by R, can be obtained for the candidate material. There are two situations in which this comparison can be made with the ML procedure:

a). The only sample tested is from the candidate material and the established TI is treated as a fixed constant.

b). Samples from both the candidate and control materials are tested and the TI of the control material is forced to be equal to the established value for purposes of comparison.

For situation (a), inference on the comparison can be made using the confidence interval procedure outlined in Section 4.5.1. If the established TI is contained within the interval, then there is evidence to conclude that the candidate TI is not significantly different from the established value.

For situation (b), one can use the model in (4.5) with a separate set of parameters, $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_a$ to describe the data collected from the control and candidate materials, respectively. Using the ML procedure, we will have lines $\log_{10}(t_f) = \beta_{c0} + \beta_{c1}x$ and $\log_{10}(t_f) = \beta_{a0} + \beta_{a1}x$ for both the control and candidate materials, respectively, with the coefficients estimated from the data. The RTI for the candidate material can be obtained from the following expression:

$$R = R(\boldsymbol{\psi}) = \frac{\beta_{a1}(T + 273.16)}{(\beta_{c0} - \beta_{a0})(T + 273.16) + \beta_{c1}} \ , \tag{4.11}$$

where $\boldsymbol{\psi} = (\boldsymbol{\theta}_c', \boldsymbol{\theta}_a')$. The expression in (4.11) results from equating the control and candidate lines together at T and solving for R. For example, suppose $\beta_{c0} = -39.316, \beta_{c1} = 20082.24, \beta_{a0} = -60.151$, and $\beta_{a1} = 30082.24$ with the control material having an established TI of 180°C. Then, using equation (4.11) the candidate material's RTI is 188.57°C. Figure 4.6 gives a graphical view of this example.

The estimate $\widehat{R}$ can be obtained by substituting in the ML estimates. Then, by the delta method, it follows that

$$\widehat{R} \dot{\sim} \mathrm{N}\left(R, \left[\frac{\partial R(\widehat{\boldsymbol{\psi}})}{\partial \widehat{\boldsymbol{\psi}}}\right]' \Sigma_{\widehat{\boldsymbol{\psi}}} \left[\frac{\partial R(\widehat{\boldsymbol{\psi}})}{\partial \widehat{\boldsymbol{\psi}}}\right]\right). \tag{4.12}$$

Assuming the data for the control and candidate materials are independent of one another, let $\Sigma_{\widehat{\boldsymbol{\theta}}_c}$ and $\Sigma_{\widehat{\boldsymbol{\theta}}_a}$ be the large-sample covariance matrix for the control and candidate ma-

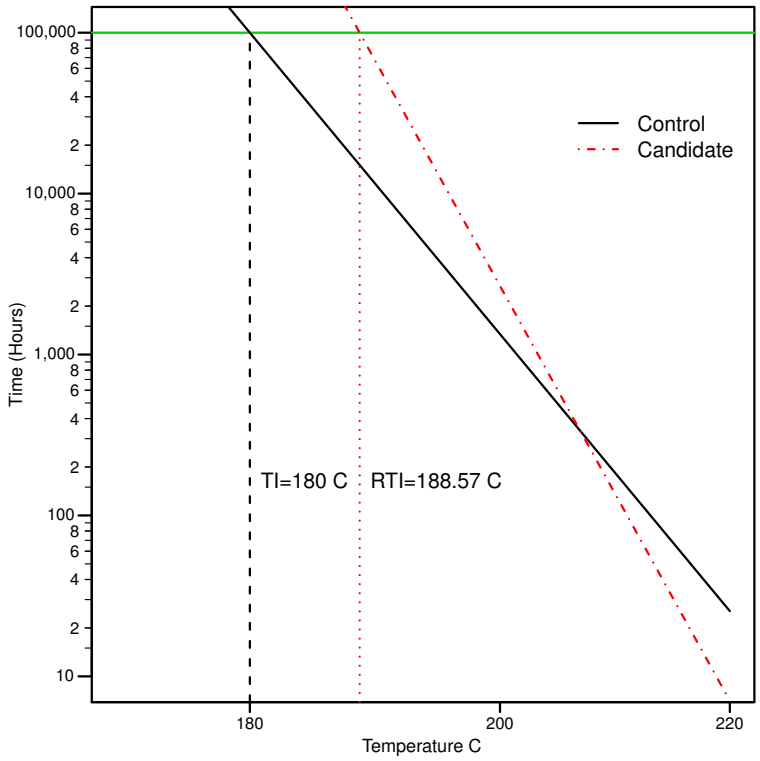


Figure 4.6: Graphical illustration of the calculation of RTI in material comparisons.

terial parameters, respectively. The large-sample covariance matrix is then given by $\Sigma_{\widehat{\boldsymbol{\psi}}} = \text{Diag}(\Sigma_{\widehat{\boldsymbol{\theta}}_c}, \Sigma_{\widehat{\boldsymbol{\theta}}_a})$. An approximate $(1-q)\times 100\%$ confidence interval for the RTI is then given by

$$\widehat{R} \pm z_{1-q/2}\sqrt{\left[\frac{\partial R(\widehat{\boldsymbol{\psi}})}{\partial \widehat{\boldsymbol{\psi}}}\right]' \Sigma_{\widehat{\boldsymbol{\psi}}} \left[\frac{\partial R(\widehat{\boldsymbol{\psi}})}{\partial \widehat{\boldsymbol{\psi}}}\right]}.$$

If the confidence interval contains the known TI of the control material, it will imply there is no significant difference between the candidate and control materials or vice versa. Of course, one must always consider power and the effect of sample size when making statements about significant differences. However, such calculations are beyond the scope of this paper.

4.5.3 Conditional Prediction

Prediction of a future observation based on currently observed data could also be of interest when studying material performance with ADDT. For example, based on data collected up to 3528 hours after the start of a test, we may want to predict the expected tensile strength

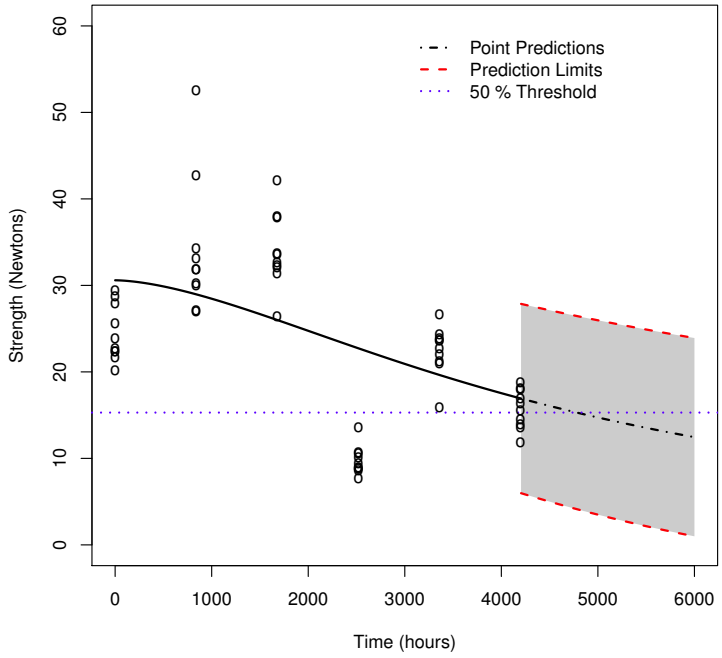


Figure 4.7: Point predictions and corresponding pointwise prediction intervals for the Seal Strength data at 200°C, based on the ML procedure.

of observations at the 5040 hour mark and/or derive a prediction interval for the observations at that time point. A point prediction can be achieved as $\widehat{y}_{ij^*} = \widehat{\mu}(t_{j^*}, x_i)$, where t_{j^*} is some future time point. To create a prediction interval with the ML procedure, we can use the variance of the residual

$$\text{Var}(y_{ij^*} - \widehat{y}_{ij^*}) = \sigma^2 \left[\rho + \frac{1-\rho}{n_{ij^*}}\right] + \left[\frac{\partial \mu(t_{j^*}, x_i)}{\partial \boldsymbol{\theta}}\right]' \Sigma_{\widehat{\boldsymbol{\theta}}} \left[\frac{\partial \mu(t_{j^*}, x_i)}{\partial \boldsymbol{\theta}}\right], \tag{4.13}$$

where ρ is the within-batch correlation of the sample. An asymptotic $(1-q)\times 100\%$ prediction interval for y_{ij^*} has the form

$$\widehat{y}_{ij^*} \pm z_{1-q/2} \sqrt{\widehat{\text{Var}}(y_{ij^*} - \widehat{y}_{ij^*})},$$

where $\widehat{\text{Var}}(y_{ij^*} - \widehat{y}_{ij^*})$ is (4.13) evaluated at the ML estimate $\widehat{\boldsymbol{\theta}}$. Figure 4.7 gives an example of prediction based on the ML procedure for a selected temperature level of the Seal Strength data.

4.5.4 Software Implementation

In terms of computation, the ML procedure is not as straightforward as the LS procedure. However, computation of the ML estimates can easily be performed with modern computing power and current statistical software packages. We have implemented the computing algorithms used in this paper in R, a popular and free statistical software package. An R package `ADDT` (Hong et al., 2014) containing several of the functions used in this paper has been developed to handle computing of the LS and ML procedures and an R GUI is currently in development. The main function of the package is `addt.fit()`, which provides model fitting and basic inference. For model-fitting of a single material, one can use the following command:

```
addt.fit(response~time+temperature, data=data, proc="ML")
```

For the comparison of two materials with samples of each material having been collected, the `addt.fit()` command can be adapted as:

```
addt.fit(response~time+temperature+materials, data=data, proc="ML")
```

The `summary()` and `plot()` functions can then be used to obtain the results.

4.6 Application

In this section, we apply both the LS and ML procedures to the Adhesive Bond B data mentioned in Section 4.2.2. Even though only two of the three temperature levels result in a failure, as seen in Figure 4.8, it is still possible to perform the analysis using the LS procedure. For the ML procedure, the same model assumptions used for the Seal Strength data are used here and the resulting degradation paths are shown in Figure 4.9. The estimated time-temperature line for both procedures are shown in Figure 4.10.

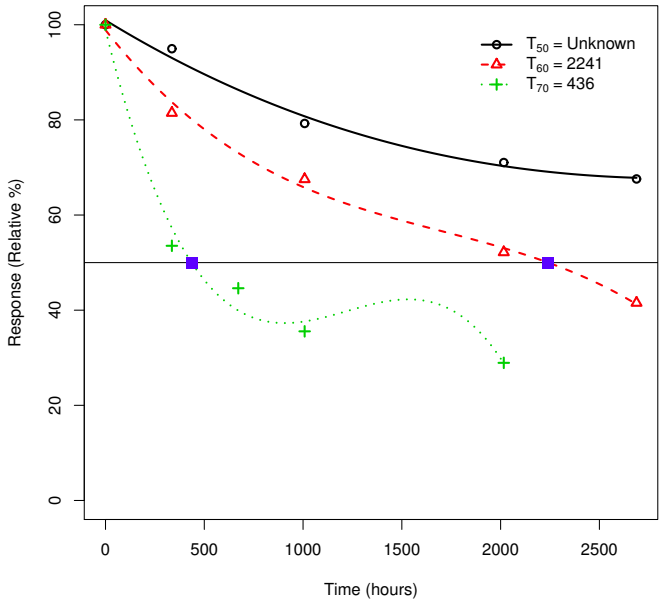


Figure 4.8: Polynomial interpolation from the LS procedure for the Adhesive Bond B data. The labels give the interpolated failure time for each temperature, as marked by a square on the failure threshold (50%) line.

The estimated TI based on the ML procedure is 33°C, which is not too far from the value of 39°C obtained by the LS procedure. Note that the 95% confidence interval for the estimated TI, roughly given as $[27°C, 39°C]$, contains the estimated TI from the LS procedure. This interval is also much smaller than that for the Seal Strength data as there is less variability in the data. The ML procedure is also able to provide predictions and their associated uncertainties for the temperature level that did not result in failure. Figure 4.11 shows pointwise prediction intervals for time points after 2688 hours at temperature 50°C for the Adhesive Bond B data. The plot is helpful in determining the time at which the mean of 5 test samples will pass the 50% threshold at the 50°C level.

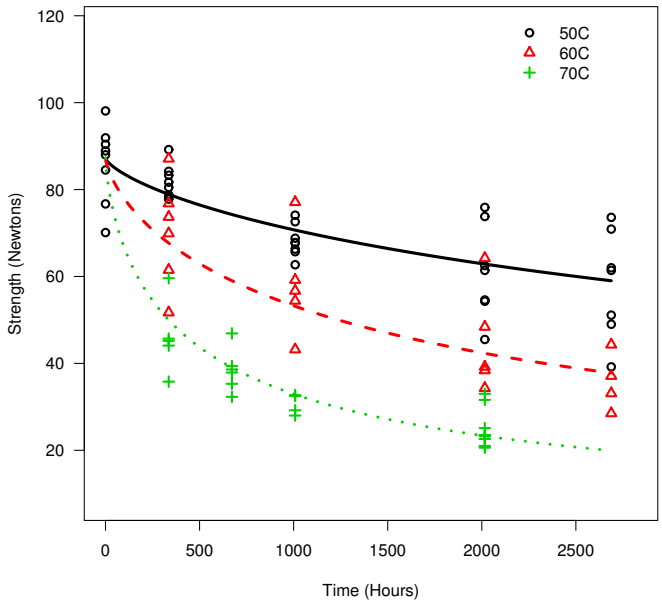


Figure 4.9: Parametric fit of the Adhesive Bond B data using the ML procedure.

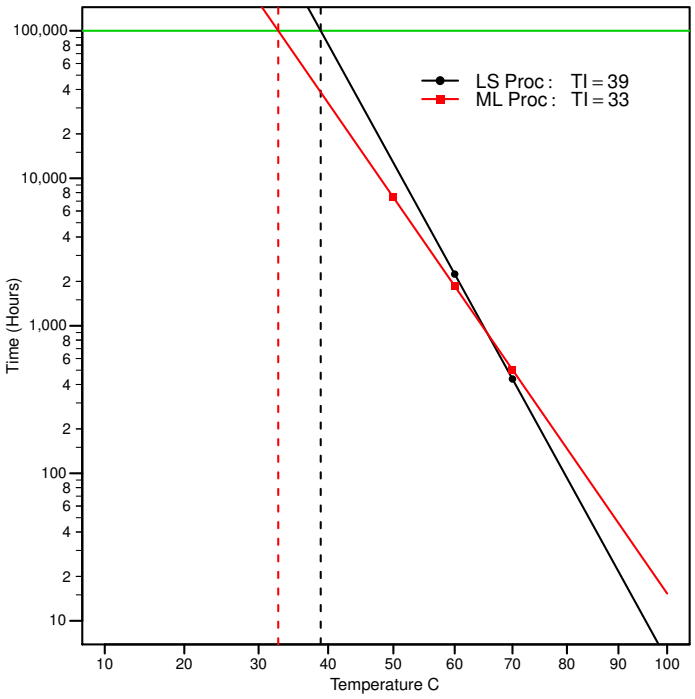


Figure 4.10: Fitted temperature-time relationship lines from both procedures and the corresponding estimated TIs for the Adhesive Bond B data.

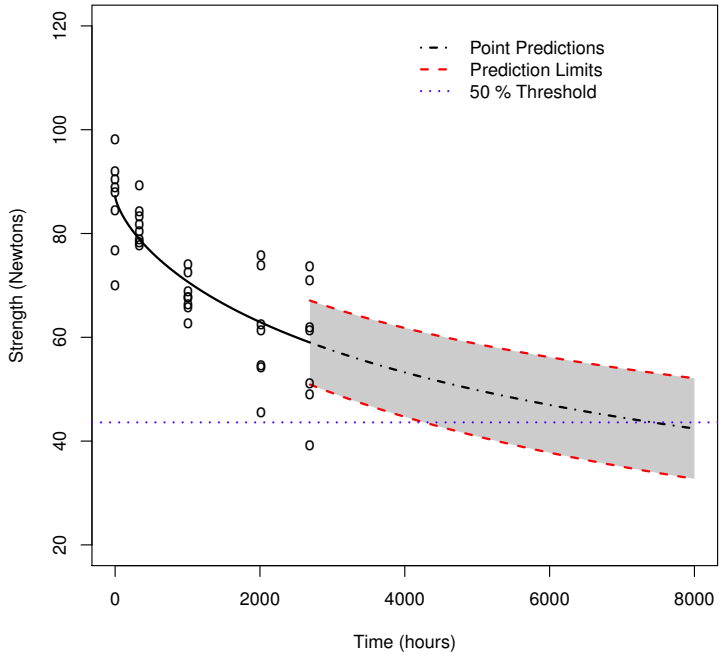


Figure 4.11: Point predictions and corresponding pointwise prediction intervals for the Adhesive Bond B data at 50°C, based on the ML procedure.

4.7 Discussion

Table 4.7 summarizes the comparison of the LS and ML approaches performed in this chapter. Both methods give similar conclusions about the TI, but the ML procedure has several advantages over the LS procedure. The ML procedure can be used to analyze data that have not yet reached the failure threshold whereas the LS procedure is unable to do so. The ML procedure is also less sensitive to constraints on time and temperature levels. It can also be used to quantify uncertainties in the estimation of the TI or in the comparison of TIs between two different materials. Since the LS procedure requires two steps to estimate the TI, it makes it very difficult to quantify uncertainties. While the ML procedure is slightly more computationally intensive than LS, it should be noted that this can be easily overcome with basic software packages, some of which are free to download. Thus, we believe the ML procedure is more appropriate for analyzing ADDT data than the LS procedure. Industry

Table 4.7: Comparison of the LS and ML procedures.

Feature	LS Procedure	ML Procedure
Specified in current standard	Yes	No
Uses data beyond the failure threshold	No	Yes
Requires parametric model	No	Yes
Considers within /between batch variations	No	Yes
Provides uncertainty quantification	No	Yes
Provides significance testing	No	Yes
Provides prediction capability	No	Yes
Requires statistical/computational software	No	Yes

procedures must be aligned with the most applicable and versatile statistical methods that are easily implemented in software. We strongly encourage further consideration of the ML procedure for setting performance indices for advanced materials.

4.8 Bibliography

Escobar, L. A., W. Q. Meeker, D. L. Kugler, and L. L. Kramer (2003). Accelerated destructive degradation tests: Data, models, and analysis. In B. H. Lindqvist and K. A. Doksum (Eds.), *Mathematical and Statistical Methods in Reliability*. Singapore: World Scientific Publishing Company.

Hong, Y., Y. Xie, and C. King (2014). ADDT: A package for analysis of accelerated destructive degradation test data. *R Package 1.0*.

IEEE-STD-101-1987 (2010). *IEEE STD 101-1987(R2010), IEEE Guide for the Statistical Analysis of Thermal Life Test Data*. IEEE.

IEEE-STD-421.1-2007 (2007). *IEEE STD 421.1-2007, IEEE Standard Definitions for Excitation Systems for Synchronous Machines*. IEEE.

IEEE-STD-98-2002 (2002). *98-2002 - IEEE Standard for the Preparation of Test Procedures for the Thermal Evaluation of Solid Electrical Insulating Materials*. IEEE.

Li, M. and N. Doganaksoy (2014). Batch variability in accelerated-degradation testing. *Journal of Quality Technology 46*(2), 171–180.

Lu, C. J. and W. Q. Meeker (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics 34*, 161–174.

Meeker, W. Q. and L. A. Escobar (1998). *Statistical Methods for Reliability Data*. New York: John Wiley & Sons, Inc.

Shi, Y., L. A. Escobar, and W. Q. Meeker (2009). Accelerated destructive degradation test planning. *Technometrics 51*, 1–13.

Shi, Y. and W. Q. Meeker (2012). Bayesian methods for accelerated destructive degradation test planning. *IEEE Transactions on Reliability 61*, 245–253.

Tsai, C.-C., S.-T. Tseng, N. Balakrishnan, and C.-T. Lin (2013). Optimal design for accelerated destructive degradation tests. *Quality Technology & Quantitative Management 10*, 263–276.

Tseng, S.-T. and Z.-C. Wen (2000). Step-stress accelerated degradation analysis for highly reliable products. *Journal of Quality Technology 32*, 209–216.

UL746B (2013). *Polymeric Materials - Long Term Property Evaluations, UL 746B*. Underwriters Laboratories, Incorporated.

Vaca-Trigo, I. and W. Q. Meeker (2009). A statistical model for linking field and laboratory exposure results for a model coating. In J. Martin, R. A. Ryntz, J. Chin, and R. A. Dickie (Eds.), *Service Life Prediction of Polymeric Materials*, Chapter 2. NY: New York: Springer.

Chapter 5 Conclusion and Areas for Future Research

This dissertation considered several research problems related to the field of reliability. Specifically, each of these problems represented the manner in which statistics and industry can each benefit from the other, using expertise from one field to provide a solution to a deficiency of methodology in the other. Chapter 3 discussed a special case in which both fields brought something to the table, a true example of the collaboration that is hoped will continue long into the future. In the remainder of this chapter, we summarize the research presented in each of the preceding chapters and consider areas in which the research can be extended further.

Chapter 2 presented an extension of the standard competing risks model to include the possibility of product components changing over the production period. Methods for estimating parameters and calculating conditional predictive distributions were proposed. An assessment of the generational model and an extended generational model was performed using cumulative predictions under eight different scenarios. It was found that the SCR model performed poorly relative to the LCG model when generational changes were present in the data. The models presented in Chapter 2 were basic in their construction and several restrictions were imposed in order to obtain a suitable foundation for the idea. In future work, these restrictions can be removed so as to formulate more complicated models. Allowing for covariates, especially dynamic covariates, in a regression setting could also be considered. Methods for lifetime predictions using dynamic covariates have been discussed in Hong and Meeker (2013) and similar procedures could be adapted here. Heterogeneities in operating environments (e.g., Ye et al. 2013) can also be considered with generational changes. One further restriction involved the assumption of independence among the components. This assumption, however, is not always true for products in the field. Although components may be grouped together to allow for independence among groups, incorporating dependence among the components and

investigating how it varies across generations and affects predictions may be useful in some applications.

In Chapter 3, we sought to derive optimal test plans for cyclic constant amplitude fatigue testing of polymer composites based on the model proposed by Epaarachchi and Clausen (2003). The optimality criterion consisted of the weighted sum of asymptotic variances of a quantile estimator for a vector of desired use stress levels. The model and criterion were both chosen to better incorporate subject-matter knowledge into the test planning procedure. Optimal and compromise plans were derived and assessed in terms of the effects of the design parameters and modeling assumptions. The optimum and compromise plans were more efficient than the traditional plans and their flexibility in their configuration allowed them to maintain that efficiency in the face of various constraints. In terms of the modeling assumptions, both plans were robust to slight changes in the planning values and provided similar results under differing lifetime distributions, though the accuracy of the optimality criterion wavered strongly under skewed log-lifetime distributions as confirmed by simulations. The plans we presented here are specific to cyclic constant amplitude fatigue testing, which is the most popular form of fatigue testing used in the industry. However, it is well known that such testing procedures do not adequately represent the stress that materials experience in the field. As such, procedures have been developed that allow the stress pattern to vary over the testing of the material. Epaarachchi (2006) present an extension of their model for the case of block testing, a simple extension of constant amplitude testing in which a sample is subjected to blocks of cyclic testing with the amplitude varying among blocks. The number of blocks and the amplitude for each block can be chosen to suit the experimenter's need. A logical next step would be to derive optimal test plans for fatigue testing under this scheme. Future work could then include consideration of spectral testing, which is a more complex extension of block testing. We can also consider a Bayesian approach to test planning (e.g., Hong et al. 2014), which can allow for the incorporation of uncertainty in the choice of planning values.

Finally, in Chapter 4, we compared two procedures for estimating the TI for a specified material: the LS procedure given in the UL standard for ADDT testing and the ML procedure

used in the statistical literature. The ML procedure possessed several advantages over the LS procedure, especially with regards to uncertainty quantification. Both procedures were applied to two different sets of data to further assist in illustrating the differences between the two procedures. Though it did not necessarily present any new methodology, this chapter helped to showcase the outstanding discrepancies that remain in current industrial standards which current statistical methodology can help to fill. It is our intention to continue to promote as well as contribute to such methodology in the future.

5.1 Bibliography

Epaarachchi, J. A. (2006). A study on estimation of damage accumulation of glass fibre reinforce plastic (GFRP) composites under a block loading situation. *Composite Structures 75*(1-4), 88–92.

Epaarachchi, J. A. and P. D. Clausen (2003). An empirical model for fatigue behavior prediction of glass fibre-reinforced plastic composites for various stress ratios and test frequencies. *Composites Part A: Applied Science and Manufacturing 34*(4), 313–326.

Hong, Y., C. King, Y. Zhang, and W. Q. Meeker (2014). Bayesian life test planning for the log-location scale family of distributions. *Journal of Quality Technology*.

Hong, Y. and W. Q. Meeker (2013). Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics 55*(2), 135–150.

Ye, Z.-S., Y. Hong, and Y. Xie (2013). How do heterogeneities in operating environments affect field failure predictions and test planning? *Annals of Applied Statistics 7*, 2249–2271.