CQM
Consultants in Quantitative Methods

UNIVERSITY OF TWENTE.

Efficient Design of Lifetime Tests to Estimate Product Reliability

Master Thesis Industrial Engineering and Management

15 September 2010

M.J. Pronk BSc
s0114219

Supervisors:

Dr. M.C. van der Heijden (University of Twente)

Dr. P.K. Mandal (University of Twente)

Drs. M.H. Schuld (CQM)

R. Wijgers MSc (CQM)

Management Summary

The subject of this thesis is product reliability. During the development process of a product the engineers want to predict the reliability of a product and verify that reliability requirements are met. But, they also have to deal with a strong pressure to develop the products in a short time. This leads to a need for fast and accurate lifetime testing methods. Accelerated lifetime testing, lifetime testing by using stress (e.g. temperature or voltage), is such a fast testing method that can be used. To obtain accurate results, the stress factors, the stress levels, the amount of products, and the test time must be chosen strategically. These factors are described in a test plan. Scientific literature only provides a few methods to design accurate accelerated lifetime test plans. Hence, the goal of this research is:

To develop a procedure for the design of efficient test plans to be able to make an accurate estimation of the reliability of a product and to be able to determine the influence of design parameters on this, by accelerating the occurrence of applicable failure modes.

In the development of test plans, we focus on different scenarios, based on the number of failure modes (i.e. the causes of failure), the number of stress factors, and the number of product design variants. Product design variants are largely similar, but differ on some specific points, such as size, type of material or amount of chemicals included; differences are represented by design parameters.

The research consists of three phases to finally fulfill the research goal. During the first phase we use a literature study to collect information about the transformation of lifetimes under stress to lifetimes under normal use conditions, methods to design test plans and methods to measure the accuracy of estimations. During the second phase, we use this information to develop procedures to design test plans for different scenarios. We start the research with the development of a test plan for the most basic scenario, namely one failure mode, one stress factor and one product design, and extend this stepwise until the most complex scenario; two or more failure modes, two or more stress factors and two or more product design variants. We describe these methods, but also make a program to be able to generate the plans. The last phase is the verifying and testing phase. We test the developed procedures based on cases from practice and compare the performance of different types of test plans within a scenario. Based on these results we conclude which procedure must be used to design test plans for lifetime testing in different situations.

According to our developed procedure, test plans are based on assumed situations. These assumptions can be based on knowledge of previous products, expert opinions, or earlier experiments. Dependent on the uncertainty about the situation a type of test plan must be selected. A test plan type determines the number of stress levels and the way of allocating the items over these stress levels. For the selected type of test plan, different configurations need to be evaluated during a simulation based on the assumed situation. The test configuration that gives the most accurate estimation of the lifetime under normal use conditions must be selected as the best test plan. To make the best test plan more robust against inaccurate assumptions, close alternative test plan configurations need to be compared for different situations. Based on those results the overall best test plan can be selected.

We examine the performance of the developed procedure during two case studies. In the first case 100 lamps and 3,000 hours of testing time are available. The goal of the test of the first case study is to estimate the product reliability and the influence of the design (i.e. the color of the light) with use of the stress factor(s) temperature and/or voltage. Based on the results of this case study we conclude for our developed procedure that:

- Using two independent stress factors instead of one stress factor reduces the needed sample size to obtain certain estimation accuracy up to 80%
- If multiple designs are compared in one single stress test instead of one stress test per design the needed sample size to obtain certain estimation accuracy is reduced by approximately 40%.

The second case concentrates on the development of a new type of aquarium lamp. For this test, 80 lamps and 4,000 hours of testing time are available to estimate the reliability and select the best design. The different designs can be described by four design parameters, namely type of emitter, filling pressure, amount of neon, and type of spiral. The influence of the design parameters on the reliability is unknown. We designed a test plan and based on the case study results we conclude for our developed procedure that:

- The developed procedure is flexible enough to be adjusted to practical scenarios. This became clear from the fact that although our research only assumed two independent design parameters, it can also deal with four interacting design parameters. Small changes in the procedure made it possible to generate a test plan for this “new” scenario.
- Accurate prior information is very important. The difference between the expected accuracy of the lifetime estimation based on the test plan configurations and the accuracy of the lifetime estimation based on the test results is approximately 8%. This difference can be reduced if assumptions about the influence of the design parameters on the reliability are available.

To make it possible for CQM to implement the developed procedure and to use the simulation program, four steps are needed. First, the developed procedure and the simulation program must be understood by the responsible employees and one person must be responsible for the maintenance. Second, the procedure must be used during a pilot project to evaluate the performance and identify improvement areas. Third, after validating the method, CQM must emphasize the importance of smart reliability testing at their customers and the developed procedure can be integrated into the standard working process. Last, the procedure must be extended or adapted based on the experiences from practice to be able to use it in more various situations. For example, a step to optimize the needed test time to reach certain estimation accuracy can be included.

Preface

This report is the result of my master graduation project, the final part of my master program Industrial Engineering and Management of the University of Twente. Six months I worked at CQM in Eindhoven to develop a procedure to design lifetime tests. Lifetime tests are tests to determine the lifetime of a new product, often a short period of time is available and stresses must be used to accelerate the failure process. This report is the result of the research.

I would like to thank a couple of people who supported me during this project. First of all, I would like to thank the CQM employees and the interns for the nice cooperation, and the pleasant and informative period. Especially, Bert Schriever for the projects he gave me to work on and his support and Roel Wijgers and Marc Schuld for the feedback, the questions and the suggestions to make the research as good as possible. From the University of Twente I would like to thank Pranab Mandal for the mathematical view and knowledge he gave me, and Matthieu van der Heijden for his feedback and his support. Finally, I would like to thank Simon for his comments and the confidence he gave me during the entire project.

Marije Pronk

Eindhoven, 06-08-2010

Table of Contents

Appendices

1. Introduction

Customers expect that the products they buy perform their intended function without failing. The probability that a product will perform its intended function until a specified point in time under encountered use conditions is called the reliability of a product (Escobar & Meeker, 2004). During the development process of a product, it is important to test the reliability of the product and improve it if necessary. Section 1.1 introduces the subject of reliability and the relevance of reliability testing. Section 1.2 describes the problems this research concentrates on. Section 1.3 introduces the initiator of this research, CQM (Consultants in Quantitative Methods). It describes the company profile and the relation towards reliability, the subject of this thesis. Section 1.4 gives an explanatory case description, based on a practical situation, which we use in this thesis to illustrate the methods that are described and developed.

1.1 Problem Introduction

During the development process of a product the engineers want to predict the reliability, verify that reliability requirements are met, identify weak points in design, and compare design variants to identify which design is best (Lydersen & Rausand, 1987). Insight in the reliability of the product can minimize cost of warranty and product recalls, can minimize customer dissatisfaction and can help to manage risks. This leads to more upfront testing of the products. But, engineers have to deal with strong pressure to develop high technological products in a short time, while improving quality, productivity and product reliability (Escobar & Meeker, 2006). This leads to a need for fast and accurate testing methods.

Based on results of lifetime tests, the life cycle of a product can be described. Figure 1 presents a general life cycle of a product, also called the rollercoaster curve. In the first-phase the failure rate is high, because of failures caused by manufacturing errors. Often, these failures are observed before the product arrives at the customer. During the second phase, the failure rate increases because failures occur due to design errors or wrong usage of customers, a subgroup of the population deals with these problems. During the third phase the failure rate is almost constant. Random failures occur due to use or environmental factors. In the fourth phase, the failure rate increases, because the product reaches the end of its life (Breyfogle, 2003). The shape and scale of this curve determine the mean life of the product or the moment in time where p% of the products is failed. Engineers try to develop a product such that the percentage of failures caused by early failures is as low as possible and the mean life of the product is higher than the pre specified lifetime of the product. A bad understanding of the life cycle of a product can cause product recalls. Appendix B gives some examples of product recalls based on causes of failures that are not detected during the development of the product. To obtain the desired lifetime curve, reliability must be considered during the complete development process of a product.

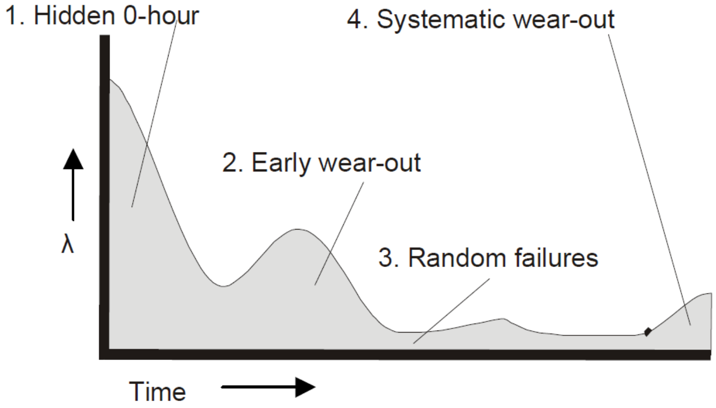


Figure 1: Four Phase Rollercoaster Reliability Curve (Brombacher, 1992)

Figure 2 describes the development process of a product based on the Design for Six Sigma phases (ReliaSoft, 2007). The activities with respect to reliability are summarized in this figure. The identification phase and the optimization phase are the most important phases for this research and are described in more detail. The reliability target is set in the identification phase, based on the expectations of the customer, a benchmark study to competitive products and the requirements of the developers. The modeling of the lifetime takes place in the optimization phase. During this phase, the failure modes, i.e. causes of failure, of the product are identified and their impact on the lifetime of the product is determined. Also the impact of design changes of the product on the reliability is important; more design variants can be available for testing. Tests are executed to obtain the information to establish these lifetime models. These lifetime tests are difficult; often there is not enough time to observe the whole life of a product, because this can take for years. Another problem is that often a limited amount of products is available for such tests; the products are not in production yet. This research focuses on the design of these lifetime tests. The next section describes the problems with respect to these tests in more detail.

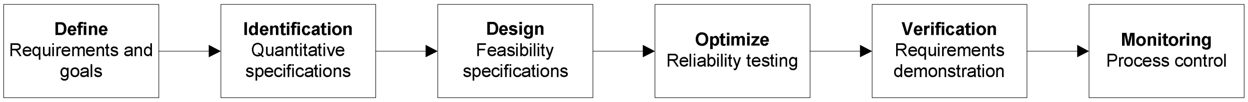


Figure 2: Reliability in the Product Development Process

1.2 Problem Description

As described in the problem introduction, tests must be executed to model the lifetime of a product in order to estimate the reliability. A problem that often occurs with these lifetime tests is that there is not enough time to observe the whole product lifetime, which can take years. For example the minimum lifetime of a LED lamp, Light Emitting Diode, is 20.000 hours; this would take almost three years to confirm (Philips: LED Lighting Systems). During the development process this amount of time is not available; engineers strive for a short time-to-market to introduce a new product earlier than the competitors do. The available time to test a LED lamp will be around one month (700 hrs). Therefore, ways to speed up this failure process must be used. This can be done via accelerated stress testing. The items are tested under more severe stresses than those encountered during normal use, to quickly obtain data, which yield desired information about the lifetime of the product under normal use conditions. Examples of these stresses are temperature, humidity, voltage, mechanical load and thermal cycling (Nelson, 1990). The levels of the stresses must be chosen such that there is a very low probability that failure modes occur that are not feasible under normal conditions, for example melted or burned components. But on the other hand enough failures must occur to obtain statistically significant results (Clark, Garganese, & Swarz, 1997). Also the amount of products that is available for testing as well as the testing equipment can be limited. For example for testing lighting products often 30 till 50 testing places are available per product type. Different designs of the product can be compared to examine the influence of a specific product design parameters on the reliability, for example size, chemicals included or thickness. The lifetime tests must be designed such that the time and the number of products needed are minimized, but still be able to make an accurate estimation of the reliability.

The controllable terms that must be determined for the lifetime tests are the stresses that must be used, the values of these stresses, the way of increasing the stress levels, the sample size, the number of replications per stress level and the test termination time. After testing, the observed lifetimes must be translated to lifetimes under normal use circumstances and the accuracy of the estimation must be determined. A procedure to design an efficient way of executing lifetime tests is needed, but not available yet. In the scientific literature methods can be found to analyze lifetime data, for example, Lawless (2003) and Elsayed (2003) describe different methods for these analyses. Some other authors, for example Nelson(1990), Meeker & Escobar(1998) and Guo & Pan(2007), describe the optimization problem for the design of a test plan for specific situations, but no clear description of a complete procedure that solves the optimization problem for different situations is available. This research focuses on the development of such a complete procedure to design lifetime tests.

1.3 Company Description

CQM is a consultancy firm located in Eindhoven, focused on decision problems, planning, and process improvements. The company exists over 30 years and originates from Philips. CQM supports Philips during a lot of projects, but also has customers like ASML, Océ and Prorail. Quantitative methods are used to analyze and improve the complex problems and processes of the customers. CQM consists of 30 consultants, divided over three groups; chain management, planning, and product and process improvement. This research is done for the latter group. This group enables companies to tackle and validate all relevant steps in industrial innovation projects from marketing to production. The concept Design for Six Sigma is used as method to shorten the time to market, make the design robust to input and process variances and improve the quality. Customers are supported with statistical methods to improve their decisions based on quantitative facts. The last few years, customers recognize the importance of reliability testing during the development of new products and ask for methods for reliability engineering. This thesis will describe the methods that can be used to design lifetime tests, such that the reliability of the products can be estimated in an early stage.

1.4 Case Introduction

---

2. Research Design

This chapter includes the research goal of this thesis and the strategy that will be followed to reach this goal. Section 2.1 states the research goal based on the problems described in chapter 1 and section 2.2 describes the corresponding research questions. Section 2.3 describes the scope of the research and section 2.4 presents the research strategy to be followed during this research. Finally, section 2.5 describes the outline of this thesis.

2.1 Research Goal

As described in chapter 1.2, the problem this research deals with is that there exists no complete procedure to design lifetime tests to estimate the reliability of a product. To solve this problem, the goal of this research is:

To develop a procedure for the design of efficient test plans to be able to make an accurate estimation of the reliability of a product and to be able to determine the influence of design parameters on this, by accelerating the occurrence of applicable failure modes.

- *Efficient tests* are tests that are executed in a minimum amount of time, using a minimum amount of products, containing purposeful changes in the input variables so that changes may be observed in the output response. The response must be such that based on this an accurate estimation of the reliability can be made (Montgomery, 2005).
- *An accurate estimation* is in this context a calculated approximation, using a statistical method, which estimates the value of a parameter as close as possible to the true parameter value.
- *Reliability* is the probability that a unit will perform its intended function until a specified point in time under encountered use conditions (Escobar & Meeker, 2004).
- *Design parameters* are physical or functional component or product characteristics represented by a variable, these characteristics can be for example material or size.
- A*ccelerating* means speeding up the failure process by making use of stress factors. Examples of stress factors are temperature, humidity or power.
- *Applicable failure modes* are the most important causes of failure, also called failure mechanisms (Nelson, 1990). A product can have several failure modes, but we focus on the failure modes that have the largest influence on the lifetime of the product. Every failure mode has its own lifetime distribution. To model the lifetime distribution of a product with multiple failure modes, a combined distribution must be used. Part of this research will focus on products with multiple failure modes.

Appendix A lists the most important definitions we use in this report.

This research is scientifically relevant as well as relevant for CQM and other companies in the same industry. The scientific literature is strongly focused on the analysis of lifetime data. Based on tests or observations during use, there is information available about the lifetime of a product. Based on these data, a statistical distribution can be fit and an expected lifetime can be determined. About the design of the lifetime tests, less literature is available. There exist scientific articles about the design of lifetime tests, but they are appropriate for specific situations, focus on comparisons of two product designs, or determine one of the experimental factors, given the other ones. However, they do not include the

effect of the design parameters and the reliability estimation. Some authors describe parts of the design, but a complete procedure to generate the test values is missing. This research focuses on these issues and will provide new insights into the fields of and develops methods to design lifetime tests.
Thereby, the research has not only scientific value, but is also relevant for CQM. Their customers express their need for methods to test product reliability, but currently these complete methods are not available. To give these customers full support, this research is necessary and the result will be a more complete procedure that describes how to design lifetime tests. The focus of companies on reliability shifts from production to the product development phase, so also other companies with R&D departments can profit from this research.

2.2 Research Questions

To meet the research goal, several subquestions have to be answered. This section presents these questions and describes the purpose of each question. Figure 3 shows the relation between the different research questions.

1. *Which models are used to describe the lifetime of a product and how can failure modes, stress factors and design parameters be included?*
   The goal of this question is to describe the most common lifetime models and the situations in which they can be applied; examples are the Normal and the Weibull distribution. But also models to combine lifetimes for products with different failure modes, data under stress or different design parameters are needed. Examples of such models are the Arrhenius model and Proportional hazard models. These models can be fitted through lifetime data obtained by tests; the model parameters are chosen such that the data fits the model as good as possible. These models are necessary, because based on them the reliability of a product can be estimated.
   We answer this question based on a literature study; the most common models for different situations are explored.

2. *How can the accuracy of the lifetime model and the estimation methods be measured?*
   The lifetime model is fitted using the experimental data, but a criterion is needed to find the best fit and measure the deviation of the model from the reality. The objective of this question is to describe the criterion that can be used to fit the experimental data through the statistical model and describe a method that can be used to measure the possible deviation from the true value, also called accuracy. The accuracy of the model is measured when the developed procedure is used in the simulation study, research question 5, and when it is tested in a practical situation, research question 6. This method to measure accuracy is obtained from the scientific literature.

3. *Which procedures available in the scientific literature are appropriate to design tests to model the lifetime of a product, and when are these methods applicable?*
   The purpose of this question is to give an overview of the existing methods to design lifetime experiments. The methods are described based on available scientific literature. The methods are categorized according to the scenarios for which the methods can be used and the situations for which they are appropriate. Scenarios are based on the number of failure modes, the number of stress factors and the number of different product design variants for which a test plan must be designed. Situations are based on the assumptions made to design the test plan on, for example uncertainty about the family of distributions, or uncertainty about the failure probabilities. From this, missing methods can be recognized. These methods will be developed in answering research question 4, and together with the available methods they are used to create a complete procedure to design lifetime tests.

4. *Which complete procedure can be followed to design tests to model the lifetime of a* product?
   Answering this question results in a recipe for designing lifetime tests. Methods for the situations wherefore no methods are available in the scientific literature are developed. Combinations of these methods are used to develop the complete procedure. Based on a given situation, it must be clear how the sample size, stress levels, stress loadings, number of runs and run settings can be determined. To come up with this procedure, different scenarios must be included. These different scenarios are created based on the number of failure modes, the number of stress factors and the existence of different product designs. More details about these scenarios and the research approach can be found in section 2.4. Figure 4 gives a schematically overview of the different scenarios. The methods described by answering research question 3 are used for the development of the procedures, but also new methods are derived.

5. *How appropriate is the designed procedure to well known situations?*
   Via computer based simulation the procedures designed by answering question 4 are tested. A situation description is made and based on this, with use of the designed procedure, a test plan is developed. Subsequently, the test plan is executed and results are generated by simulation. In this step, the failures are simulated with use of a lifetime distribution and a life-stress relation. At the end, the appropriateness of the procedure can be determined based on the accuracy of the estimation, and the difference of the estimated values and the underlying model.
   Answering question 4 & 5 is an iterative procedure; after the development of methods for one of the scenarios this question must be answered and based on the test results the procedure can be adapted and tested again. The goal of this research question is to improve the developed procedure and determine the quality of the developed procedure.

6. *How appropriate is the designed procedure in a practical case?*
   After validation of the procedure via simulation, it can be tested in a practical situation. We describe the development of a test plan for an aquarium lamp and describe the results. During this test practical preferences and restrictions from the company that is involved are also incorporated. The appropriateness of the test plan procedure is measured based on the accuracy and the adaptability of the procedure.

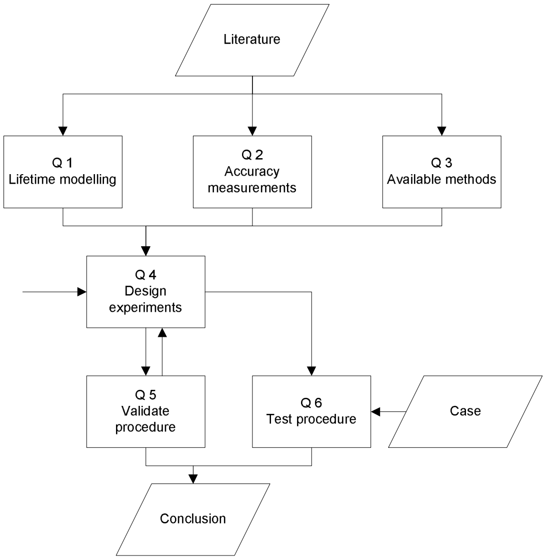


Figure 3: Relation Research Questions

2.3 Scope of the Research

Reliability and life testing are both very broad concepts; this research focuses on a specific part of reliability. We can describe this part as:

- *Overstress testing*
  Different kinds of accelerated life testing exist, namely usage rate tests, overstress tests and, degradation tests. Products with a low usage rate per day or week can be tested by high usage rate tests. These tests are executed under normal circumstances, but the products are used more and for longer periods than normal. During overstress tests the products are running at higher than normal stress levels to shorten the product life. Degradation tests are almost similar to overstress tests, but the performance is observed instead of the lifetime (Nelson, 1990). This research focuses on overstress testing. The high usage tests do not deal with the time problem and the degradation tests are only possible when the degradation process can be measured.
- *Quantitative methods*
  Both qualitative and quantitative accelerated life tests are used for reliability testing during the optimization phase. In the scientific literature both kinds of methods are described, but they have their own purpose. The main purpose of the quantitative methods is to obtain information about the failure time distribution at specified use conditions of the stress variables. The purpose of the qualitative methods is to identify product weaknesses caused by flaws in the product's design or

manufacturing process (Escobar & Meeker, 2006). The focus of this research is to estimate the lifetime and reliability of a product based on the failure time distribution, so the quantitative methods are used.

- *Highly reliable consumable items*

  High reliable products are defined as products with an expected lifetime much larger than the available test time, for example LED lamps with an average lifetime of 20.000 hours and a test time of 700 hours. Accelerated stress testing is used in such situations, but this is not necessary for products with a relatively short lifetime, therefore only high reliable products are taken into account.
- *Single component, nonrepairable item*

  The reliability is estimated for a single component or product, so serial systems are not taken into account. Based on individual reliabilities, the reliability of a system can be determined; this part is out of scope of this research. In general no new tests have to be done to examine this system reliability; it can be based on the experimental data of the components. Because we focus on the component level, the items that are failed cannot be repaired. Examples of products are electrical insulations, battery cells or lighting products.
- *One or more product design variants available*

  To make it possible to compare different designs and measure the influence of specific design parameters on reliability, we also include situations with one or more product design variants.
- *Time censoring, termination time given*

  There are different types of stress tests available. We concentrate on the situation where there is a fixed amount of time available for testing. This is the most common situation in practice. The consequence of this fixed amount of time is that the number of observed failures is not known exactly beforehand. In most cases not all the products are failed when the test terminates and this leads to censored data. Appendix C describes the different types of censoring in more detail.

Assumptions that we take into account are:

- *Family of lifetime distribution known*

  The lifetime of the products is assumed to be parametric. We use the Normal and Weibull distribution to describe the lifetime because they have totally different characteristics and can be used for different type of products. The assumption about the type of distribution and the parameters is based on knowledge from previous tests and physics of the product. When no parametric distribution can be applied, a nonparametric analysis can be used. This nonparametric method is only applied when a parametric method gives infeasible results.
- *Constant shape parameter of lifetime distribution for different stress levels and design variants*

  The fundamental principle of overstress testing is that the unit under test will exhibit the same behavior in a short time at high stress that it will exhibit in a longer time at low stress. This means that the shape parameter of the life distribution is assumed to be constant and only the scale parameter changes (Condra, 2001). Also, the design parameters only influence the scale parameter.
- *Failure modes to test known and independent*

  We assume that a qualitative analysis is done before the quantitative analysis, so the applicable failure modes are identified. Examples of qualitative methods are HALT, High Accelerated Life

Testing and MEOST, Multiple Environment Over Stress Testing. For more details about these methods see Bhote and Bhote (2004).

- *Applicable stress factors known with their normal use and highest level*
  We also assume that the stress factors with their normal use and maximum level are known. With the maximum stress level, also called high stress level, we mean the maximum level that we can use to stress the product without stressing the product that much that other failure modes occur that do not appear during normal use. The stress factors can be known based on the physics of the product, or are derived from the qualitative analysis.
- *After failure, the corresponding failure mode is known*
  To be able to treat different failure modes independently, and finally combine them to estimate the reliability of the product, the reason of failure of a product must be known when the product fails.

2.4 **Research Strategy**

The research can be split into three phases. Figure 3 displayed the research questions and the relations between them; the three middle layers of this figure correspond with the phases of the research. The first phase can be classified as literature study. The goal of this phase is to understand the current way of working with respect to modeling lifetime, estimating reliability and the design of tests. Based on scientific literature, methods to model reliability are described, we do this while answering research question 1. We describe methods to estimate the parameters of the lifetime model and measure the accuracy of these estimations, such as Maximum Likelihood Estimation and confidence intervals, by answering research question 2. Important references for the first two research questions are Meeker and Escobar (1998), Nelson (1990), Lawless (2003) and Deshpande and Purohit (2005). Subsequently, we describe methods currently used in the scientific literature to design lifetime tests and categorize them based on their assumptions and applications; we do this by answering research question 3. Based on this literature review, we can recognize lacking methods to design lifetime tests.

The second phase of the research can be classified as the design phase; we develop methods to design lifetime tests based on combinations of existing methods described by answering question 3 and derivations of methods for similar situations. Different scenarios are created dependent on the number of failure modes, the number of stress factors and the number of product designs to be tested. The product designs are for a large part the same, but differ on one or a few points. It can be efficient to test these designs, which have much in common, together. This is because it can increase the accuracy of the estimation and make it possible to see the influence of a design parameter on the reliability. This makes the decision for choosing the best design easier. Figure 4 shows the different scenarios schematically. Starting with the simplest scenario, we develop a procedure, this procedure is used to come up with procedures for more complex scenarios, and this process continues until the most complex scenario is reached. After developing a procedure for one scenario, this procedure is tested via a simulation study and adapted; this is described by research question 5. For the first two layers, we develop the generation of the test plans completely. For the other two layers, we describe the idea behind the method to design a test plan, but do not give the complete description.

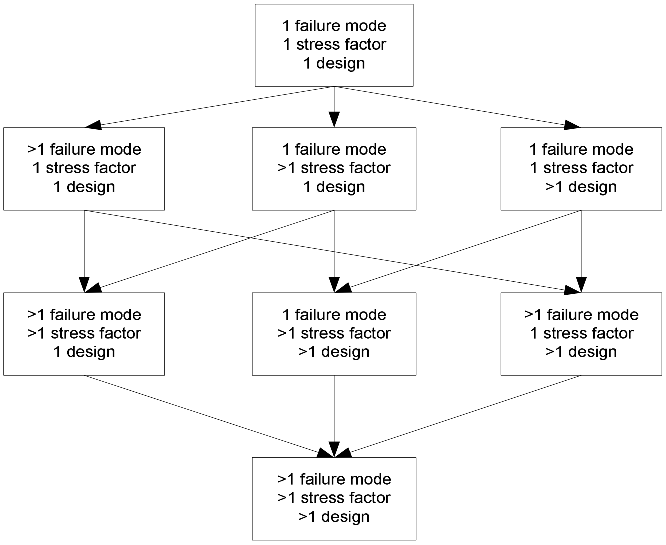


Figure 4: Research Approach

The last phase of this research can be classified as the validation and testing phase. The new developed procedures are validated through a simulation study. We design tests based on the new methods and simulate failures using Monte Carlo simulation based on a lifetime distribution and a life-stress relation. After this, we fit a lifetime distribution based on the data obtained by the lifetime tests and measure the difference between the estimated and the real reliability. This validation must be done for every scenario of Figure 4. This means that the design of the procedure and the validation follow an iterative procedure, we make adjustments to the procedures based on the test results and test the adapted procedures again. After validation, we test the procedure with use of a practical case. This test shows if the assumptions used during this research are realistic for practical situations and which adjustments have to be made to be able to implement this procedure in practice. At the end, we conclude how efficient lifetime tests can be designed to make accurate estimations of the reliability of products based on the scientific literature and the test results and we give suggestions for further research.

2.5 Outline Thesis

Chapter 3 presents different lifetime models, these models describe the reliability of a product on a specific moment. For every product, the parameters of these models can be estimated based on the experimental data. Chapter 4 describes methods to determine the accuracy of the parameter estimations, given the experimental data. To obtain these accurate results, Chapter 5 describes methods based on the scientific literature that can be used to set up lifetime tests for the one failure mode, one stress factor and one product design scenario. These methods are classified based on the situation they can be applied on, and are illustrated with use of an example. Chapter 6 describes procedures to set up lifetime tests for the other scenarios, based on the methods described in Chapter 5 and new designed methods to fill the gaps. Chapter 7 gives the results of the new designed procedure during a practical case. Chapter 8 describes the steps CQM has to execute to implement the developed. Based on the procedure and the test results, Chapter 9 answers the research questions and gives suggestions for further research.

3. Lifetime Modeling and Analysis

This chapter describes the models that can be used for lifetime modeling and reliability estimation. Section 3.1 describes the models that are employed to describe the lifetime and the reliability of a product. Section 3.2 describes the lifetime model of a product with several failure modes, consisting of a combination of individual models. Section 3.3 discusses the techniques to translate failure times of products under severe stress conditions back to normal use conditions. The output data of tests are used to choose an appropriate lifetime model and to estimate the parameters of this model. These parameters are estimated based on Maximum Likelihood Estimation; this method is described together with the methods to determine accuracy in chapter 4. Finally, section 3.4 describes the models to analyze different design variants. To be able to design lifetime tests, knowledge of the lifetime models is important.

3.1 Lifetime Modeling

The most widely used measure for reliability of a product is the failure time distribution; this distribution is constructed based on the failure time data of the product. The lifetime or failuretime of a product can be described by a nonnegative, continuous random variable T. The probability distribution for the lifetime T can be characterized by a cumulative distribution function, a probability density function, a reliability function or a hazard function. The reliability function gives the probability of an item surviving up to time t. The hazard rate function, also called failure rate function, describes the probability of failure in the small interval [t, t+Δ], given survival up to time t (Meeker & Escobar, 1998).

As described in the scope of the research, we focus on the Normal and the Weibull distribution. The exponential distribution is a special form of the Weibull distribution, so it is also included in the research. Appendix D describes the different characteristics of the Normal and the Weibull distribution. In the situation where no prior information is available about the distribution or the number of failures is very small the nonparametric approach can be used (Lawless, 2003). Section 3.1.3 describes this approach.

3.1.1 Normal Distribution

The normal distribution is applicable for products with wear out failure; this is an increasing failure rate. The older the product is, the higher the probability to fail becomes. Many consumable items can be described by the normal distribution. The disadvantage of the normal distribution is that its range is defined from -∞ to ∞, but in reality lifetimes cannot be negative. Therefore, the normal distribution can only be used to model the lifetime if the mean lifetime is large and the standard deviation is small. In this situation the fraction below zero is very small.

3.1.2 Weibull Distribution

The Weibull distribution is the most widely used lifetime distribution model (Nelson, 1990). The shape parameter gives the distribution the flexibility to fit many types of data. The Weibull distribution can describe an increasing, decreasing and monotone hazard rate as function of age. Figure 5 shows the hazard rate for different values of the shape parameter, β. If $\beta=1$, the hazard rate is monotone and the Weibull distribution equals the exponential distribution. If $\beta<1$, the hazard rate is decreasing over the time and if β>1 the hazard rate is increasing over time. In conclusion, the shape parameter determines the type of hazard rate (Meeker & Escobar, 1998).

We only use the two-parameter Weibull distribution. The three-parameter distribution also has a location parameter, this can be used to include a failure-free period, but this distribution is seldom used for accelerated life testing (Nelson, 1990).

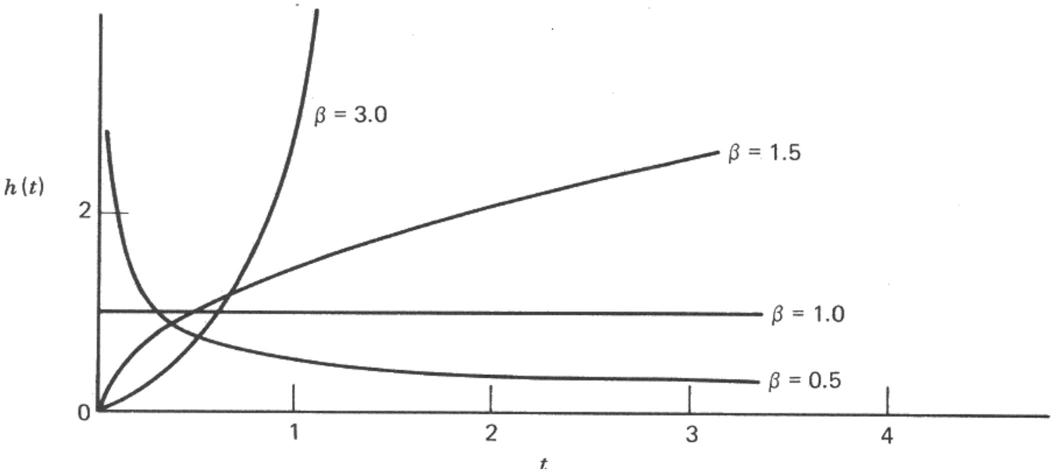


Figure 5: Weibull Hazard Functions (Pal, 2005)

The Weibull distribution can be transformed to the smallest extreme value (SEV) distribution. The smallest extreme value distribution has the advantage that it belongs to the family of location-scale distributions, like the Normal distribution, and for this kind of distributions general rules can be applied during the lifetime analysis (Meeker & Escobar, 1998). For example the likelihood function for location scale distributions can be expressed in a simple generic form. The transformation of the Weibull distribution to the smallest extreme value distribution can be described as:

If T ~ Weibull (β,η) then Y=ln(T) ~ Smallest extreme value(μ, σ) with μ=ln(η) and σ=1/β.

The characteristics of the SEV distribution are:

$$Cumulative\ distribution\ function: \quad F(t;\mu,\sigma) = \Phi_{SEV}\left[\frac{\log(t)-\mu}{\sigma}\right] \tag{1}$$

$$Probability\ density\ function: \quad f(t;\mu,\sigma) = \frac{1}{\sigma t}\phi_{SEV}\left[\frac{\log(t)-\mu}{\sigma}\right] \tag{2}$$

$$Hazard\ rate: \quad h(t;\mu,\sigma) = \frac{1}{\sigma\exp(\mu)}\left[\frac{t}{\exp(\mu)}\right]^{\frac{1}{\sigma}-1} \tag{3}$$

$$Percentile: \quad t_p = \exp[\mu + \mathit{\Phi}_{SEV}^{-1}(p)\sigma] \tag{4}$$

$With\ \Phi_{SEV}(z) = 1 - \exp[-\,exp(z)]\ and\ \phi_{SEV} = \exp[z - \exp(z)].$

$t: moment\ in\ time$

$\mu: location\ parameter$

$\sigma: scale\ parameter$

$p: percentile$

In the remaining part of this thesis, this transformation of the Weibull distribution is used because the normal distribution is also a location-scale distribution and in this way general expressions can be used that are appropriate for both distributions to make the methods more generic. For example during the parameter estimations via maximum likelihood estimation these expressions are used, see section 4.1. The value of the location parameter, μ, depends on the applied stress level.
In general the expression for μ equals:

$$\mu = \gamma_0 + \gamma_1 x, \qquad \gamma_0 \; and \; \gamma_1 \; are \; both \; parameters \; and \; x \; is \; the \; applied \; stress. \qquad (5)$$

The value of x depends on the life-stress relation; section 3.3 describes this in more detail.

3.1.3 Nonparametric Analysis

Both the Normal distribution and the Weibull distribution are parametric models. It is important that the assumptions about the type of distribution and parameters are accurate, because different assumptions can give quite different results. When no assumptions can be made about the distribution type, we can perform a nonparametric analysis. This method is more robust to wrong model estimations. The disadvantage of this method is that the confidence bounds of the model are much wider than with use of a parametric distribution and also predictions outside the range of observations are not possible. A useful nonparametric method is the Kaplan-Meier method. The reliability function can be estimated as follows:

$$\hat{R}(t_i) = \prod_{j=1}^{i} \frac{n_j - d_j}{n_j} \qquad i = 1..m \qquad (6)$$

$$With \;\; n_i = n - \sum_{j=0}^{i-1} d_j - \sum_{j=0}^{i-1} s_j \qquad i = 1..m$$

$\hat{R}(t_i)$: $Reliability\ at\ time\ t_i$

t_i: $time\ of\ the\ i^{th}\ observation\ moment$

m: $total\ number\ of\ observation\ moments$

n: $total\ number\ of\ units$

n_j: $number\ of\ units\ at\ risk\ before\ the\ j^{th}\ observation\ moment$

d_j: $number\ of\ failures\ within\ the\ j^{th}\ observation\ moment$

s_j: $number\ of\ censored\ observations\ within\ the\ j^{th}\ observation\ moment$

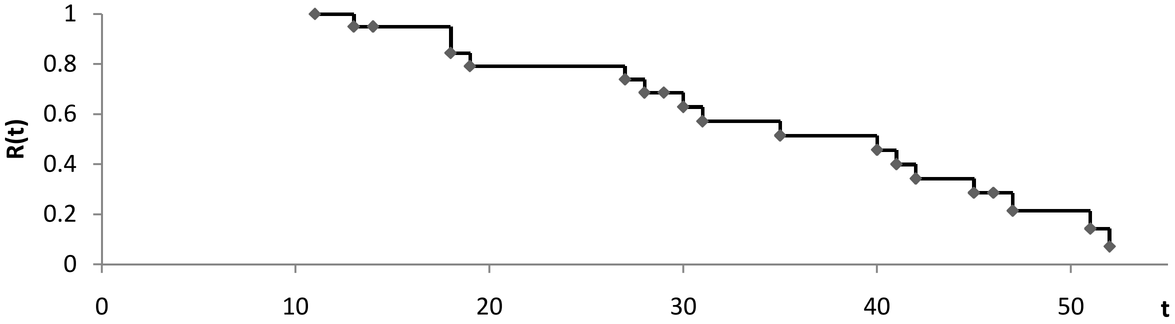


Figure 6: Kaplan-Meier Reliability Plot

Figure 6 shows the reliability plot of a product based on the Kaplan-Meier method. The Kaplan-Meier approach is only described to show a method to analyze data without assuming a family of distributions. For the design of a test plan we focus on the parametric approach, because then we can make some assumptions about the distribution of the lifetimes.

3.2 Competing Failure Modes

A product can fail due to different reasons; the reason why a product fails is called failure mode. During testing and analyzing the reliability it is important to make a distinction between these different failure modes, because they all have their own lifetime distribution with their own parameter values. This means that a model has to be fitted for each failure mode; subsequently they can be merged to the lifetime distribution of the product. This sequential approach has to be followed, because mixing different failure modes when fitting a distribution leads to high deviations and bad prognoses (Werner, 2009).

Products with statistically independent failure modes can be seen as independent systems and we can use the addition law for failure rates. For such products, the failure rates add (Nelson, 1990). Figure 7 shows this for a product with two failure modes in a graph.

The general equation for a combined failure rate function equals:

$$h(t) = h_1(t) + h_2(t) + \cdots + h_M(t) \tag{7}$$

$M: number\ of\ independent\ failure\ modes$

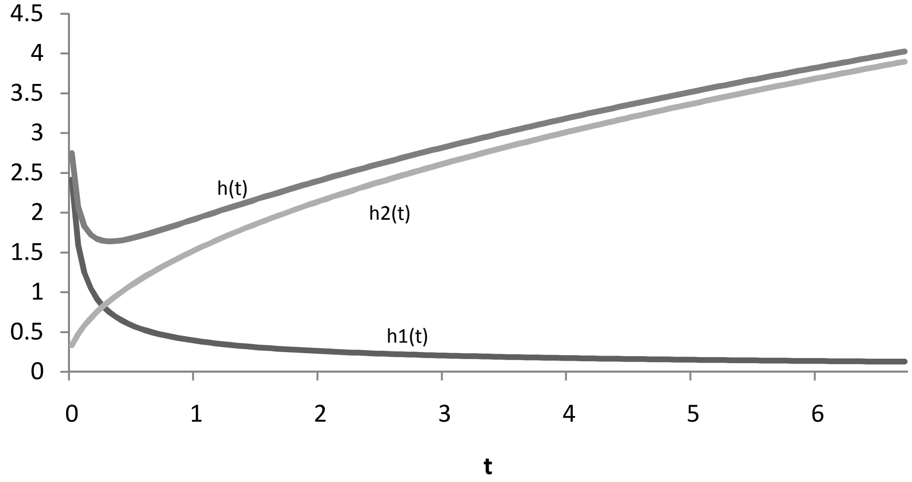


Figure 7: Hazard rate for Product with Two Failure Modes

Per failure mode, the failure probability distribution has to be determined. Based on these individual distributions, a function for the total failure rate or reliability can be established. These functions become:

$$R(t) = R_1(t) * R_2(t) * \ldots * R_M(t) \tag{8}$$

$$F(t) = 1 - [1 - F_1(t)] * \ldots * [1 - F_M(t)] \tag{9}$$

Sufficient failures per failure mode are necessary to make an accurate estimation for the individual reliabilities. We assume that independent of the stress level, all failure modes can be observed. The percentage in which they occur differ, but an initial estimate is available.

3.3 Life-Stress Relationships

The previous sections described models based on the normal use conditions of a product. To speed up the failure rates and obtain lifetime data faster, stress factors are used. The scientific literature describes different methods to relate the lifetime under stress to lifetime under normal conditions; most important and general applicable methods are described in this section. These models do not describe the shape of the distributions, but only the scale parameter (Condra, 2001).

Under the parametric models, for the Scale Accelerated Failure Time Models (SAFT) the assumption is made that the failure time distribution at stress level s_1 remains within the same family as on stress level s_2, s_3 etc. This means that the shape parameter of the distributions is the same for all stress levels, but the scale parameters can be different. Based on this assumption the general relation between life under normal use conditions and life under stress can be described as (Elsayed, 2003):

$$t_p(\boldsymbol{x}_u) = t_p(\boldsymbol{x}) * AF(\boldsymbol{x}) \qquad (10)$$

t_p: $\quad Time\ when\ p\%\ of\ the\ population\ is\ failed$

$\boldsymbol{x}_u$: $\quad Vector\ of\ stress\ factors\ under\ normal\ use\ condition$

$\boldsymbol{x}$: $\quad Vector\ of\ stress\ factors$

$AF(\boldsymbol{x})$: $Acceleration\ factor\ depending\ on\ vector\ \boldsymbol{x}$

$$F_u(t) = F_{\boldsymbol{x}}\left(\frac{t}{AF(\boldsymbol{x})}\right),\ f_u(t) = \left(\frac{1}{AF(\boldsymbol{x})}\right) f_{\boldsymbol{x}}\left(\frac{t}{AF(\boldsymbol{x})}\right),\ \ h_u(t) = \left(\frac{1}{AF(\boldsymbol{x})}\right) h_{\boldsymbol{x}}\left(\frac{t}{AF(\boldsymbol{x})}\right)$$

Beside the SAFT models, there exist also Proportional Hazard failure time models (PH). The main difference between these models is that the SAFT models assume that the effect of the variables is to multiply the predicted random variable, lifetime, by some constant and the PH assumes that the effect of the variables is to multiply the hazard rate, also called failure rate, by some constant (Lawless, 2003). The hazard rate of the SAFT models depends on the stress level, and the baseline hazard rate, the failure rate under normal use condition, of the PH models is independent of the stress levels. The general description of the proportional hazard model is:

$$h(t, \boldsymbol{x}) = \psi(\boldsymbol{x}) * h_0(t) \qquad (11)$$

$\boldsymbol{x}$: $\quad Vector\ of\ stress\ factors$

$h(t, \boldsymbol{x})$: $Hazard\ rate\ at\ time\ t, depending\ on\ vector\ \boldsymbol{x}$

$\psi(\boldsymbol{x})$: $\quad Relative\ risk\ ratio$

$h_0(t)$: $\quad Baseline\ hazard\ function$

Another type of stress models is based on step-stress. In this type of tests, the items are first tested at a pre-specified stress level for a specified period of time. The items that are not failed are tested again, subject to a higher stress level for another period of time. This process continues until the termination time is reached. The life-stress relation that can be used to analyze this process is the cumulative exposure model.

Several life-stress relations exist for different situations depending on the product, the stress factor, the stress loading, and the stress level. The most important relations are described in more detail. Based on the applied stress factors, previous experience with similar products or preliminary testing the correct relation should be chosen. Per life-stress relation the expression for the lifetime dependent on stress and the acceleration factor are described.

The life-stress relations can be included in the lifetime model by replacing the location (SEV/normal), or scale (Weibull) parameter by the expression for the lifetime dependent on the stress, T(V). In this way it is possible to make the lifetime distribution dependent of the stress factor. Examples are given in Appendix E.

3.3.1 Arrhenius Relationship

The Arrhenius relation originates from the thermodynamics. It is based on observations in chemical reactions, the more energy in the system, the more likely the reactors are to cross the energy barrier, and the shorter the life. This theorem can be applied when a thermal stress is used (Meeker & Escobar, 1998). Applications include electrical insulations, semiconductor devices, battery cells, plastics and lamp filaments (Nelson, 1990). The Arrhenius relation expresses the time for a failure to occur as an exponential function of the inverse of the applied absolute temperature (Condra, 2001).
It is described as:

$$T(V) = C * e^{\frac{B}{V}} \tag{12}$$

$T(V)\!: Stochastic\ variable\ for\ lifetime\ dependent\ on\ stress\ V$
$V\!: Temperature\ value\ in\ degrees\ Kelvin$
$C, B\!: Model\ parameters\ to\ be\ estimated$

T(V) describes the characteristic life of a product, for that reason it can replace the scale parameter η of the Weibull distribution.

Based on the relation for the characteristic lifetime based on temperature, the acceleration factor can be determined:

$$AF(V) = \frac{T(V_{use})}{T(V)} = e^{B*\left(\frac{1}{V_{use}} - \frac{1}{V}\right)} \tag{13}$$

B is based on the activation energy of the chemical reaction divided by the Boltzmann constant, or can be a parameter to be determined based on the experimental data. We will use the general expression, because the activation energy of the process is most of the time not known and has to be estimated too.

3.3.2 Inverse Power Relationship

The inverse power relationship is used when a stress factor analogous to pressure is applied; voltage is for example often used. Some applications are ball and roller bearings and flash lamps. The relation assumes that the life of a system is inversely proportional to the applied stress (Condra, 2001). The expression is:

$$T(U) = \frac{1}{K * U^n} \tag{14}$$

$U: Value\ non\ thermal\ stress\ factor$
$K, n: Model\ parameters\ to\ be\ estimated$

The acceleration factor that can be derived from this relation equals:

$$AF(U) = \left(\frac{U}{U_{use}}\right)^n \tag{15}$$

3.3.3 Eyring Relationship

The Eyring relationship can be used when temperature and one other stress factor are applied. The model consists of two constant stress terms and no interactions between those terms. Its general form is:

$$T(U,V) = A\left(\frac{1}{U}\right)^n Ce^{\frac{B}{V}} \tag{16}$$

$U: Non\ thermal\ stress$
$V: Thermal\ stress$
$A, B, C, n: Model\ parameters\ to\ be\ estimated$

Stress U can be used in a variety of transforms. It can represent humidity, but also voltage. In this way, also a combination of the Inverse Power Law and the Arrhenius law can be made. This combination equals the product of both individual relations (Condra, 2001). The expression for the acceleration factor becomes:

$$AF(U,V) = \left(\frac{U}{U_{use}}\right)^n * e^{B*\left(\frac{1}{V_{use}} - \frac{1}{V}\right)} \tag{17}$$

3.3.4 Generalized Log Linear Relationship

When a test involves more than one stress factor, the generalized log linear relationship can be used. It is possible to include multiple stress factors (Nelson, 1990).

$$T(\boldsymbol{x}) = e^{\alpha_0 + \sum \alpha_j x_j} \tag{18}$$

$$AF(\boldsymbol{x}) = e^{\sum \alpha_j x_{use,j} - \sum \alpha_j x_j} \tag{19}$$

With the use of transformations, we can show that the Arrhenius relationship, the Inverse Power Law and the Eyring relationship can be described by this generalized relationship.

Table 1 summarizes the transformations that must be made to the stresses to represent these relationships. Other stress factors that are often used are also included with their transformation.

Stress relation	X
Arrhenius	1/V
Inverse power	ln(U)
Eyring	1/V & ln(U)
Humidity	ln(RH) or ln(RH/(1-RH))
Size	ln(Thickness)
Linear	V

Table 1: Stress Transformations

When one stress factor, temperature, is applied, with x=1/V, the relation becomes:

$$T(\boldsymbol{x}) = e^{\alpha_0+\alpha_1\frac{1}{V}} = e^{\alpha_0}e^{\frac{\alpha_1}{V}} = C * e^{\frac{B}{V}}$$

The last expression equals the expression of the Arrhenius relationship.

For computational reasons, the transformed stress is often standardized to a value between 1 and 0. In this case, 1 corresponds to the lowest value of x, and 0 corresponds to the highest value of x. This makes it easier to compute the parameters that determine the location (scale) of the distribution. The standardized stress is called ξ. The standardized stress can be calculated based on the relation:

$$\xi = \frac{x_{high} - x}{x_{high} - x_{use}} \tag{20}$$

3.3.5 Proportional Hazard Regression Model

The Proportional Hazard regression model is a flexible model, which can be used to isolate effects of explanatory variables. This model assumes that the explanatory variables have a multiplicative effect on the hazard rate. The first applications were in the medical industry, but this model is also applicable to describe product reliability (Dale, 1985). The explanatory variables can be continuous variables like stress factors, as well as categorical variables, like design parameters, section 3.4 explains more about design parameters. The advantage of this model is that both kinds of variables can be included in one model. If stress factors such as temperature or voltage are used, the Arrhenius or Inverse Power Law relation can be included by defining variable x as transformed stress. Equation 11 described the Proportional Hazard model.

The Proportional Hazard model is a semi-parametric model, meaning that the distribution of the baseline hazard rate function does not has to be specified, while the relation of the explanatory variables does. The most common relations are the log linear and the linear relation (Deshpande & Purohit, 2005).

The relative risk ratios for both situations can be described as:

$\psi(\boldsymbol{x}) = \exp(\sum_{i=1}^{n} \alpha_i x_i)$ Log linear relation

$\psi(\boldsymbol{x}) = 1 + (\sum_{i=1}^{n} \alpha_i x_i)$ Linear relation

$h_0(t): baseline\ hazard\ rate$
$\alpha_i: Model\ parameter\ to\ be\ estimated$
$n: number\ of\ different\ explanatory\ variables$

If the baseline hazard function is chosen to be parametric, the model parameters can be estimated based on maximum likelihood estimation. This method is described in section 4.1. The baseline hazard rate function is often describes as a quadratic relation, $h_0(t) = \gamma_0 + \gamma_1 t + \gamma_2 t^2$. If the baseline hazard rate remains unspecified, the likelihood function is only based on the relative risk ratio.

3.3.6 Cumulative Exposure Model

The cumulative exposure model is needed when step stress is applied. During a step-stress experiment, different stress levels are applied to the same item, to obtain failures quickly and to minimize the total amount of products needed (Liao & Tseng, 2006). Within a step a constant stress level is applied, thus the item will fail according to the distribution at the current step but with a starting age corresponding to the total accelerated time up (ReliaSoft, 2007). The lifetime distribution becomes:

$$F_i(t;V_i) = 1 - e^{-[KV_i{}^n((t-t_{i-1})+\varepsilon_{i-1})]^\beta} \qquad (21)$$

$$With\ \varepsilon_{i-1} = (t_{i-1} - t_{i-2})\left(\frac{V_{i-1}}{V_i}\right) n + \varepsilon_{i-2}$$

$\varepsilon_i: Accumulated\ exposure$

$V_i: Value\ stress\ factor\ on\ level\ i, i = 1 \ldots m$

$K, n, \beta: Model\ parameters\ to\ be\ estimated$

More information about these formulas and the analysis of step stress tests is described by Liao and Tseng (2006).

3.4 Design Parameters

Different settings of a product can lead to different product reliabilities. These settings can be represented by discrete or continuous design parameters. Discrete parameters, represented by binary variables, describe the presence or absence of a given setting, for example material, supplier, batch or type of test equipment. Continuous parameters can also describe a given setting, for example size, thickness, or amount of chemical included, but uses the exact value.

To compare different design settings, a one-factor-at-a-time analysis can be done per setting and the results can be compared afterwards. But whenever applicable, a more efficient test should be used to be able to establish a combined regression model. The advantage of the combined regression model is that by combining the test to determine the reliability and the test to determine the influence of the different design parameters on reliability, the number of samples needed during the test can be decreased and some information about the best design choices can be obtained. Based on the assumption that the total sample size per design variant is equal, the best design is the one with the highest lower bound of the confidence interval of the performance indicator. The focus is on these lower bounds, because the variance of the performance indicator is taken into account in these bounds. If we only compare the lifetime estimate, it can be possible that the one with the highest lifetime also has the largest variance, so it is not sure that the true value of this design is better than the true value of the other design. The performance indicator can be for example the value of the p^th^ percentile or the reliability at time t. Percentiles are often used in practice, because it can help setting the guarantee period of the product, and it gives information about the amount of failures at a certain point in time

(Kuo & Wan, 2007). For these reasons, we also use percentiles as performance indicator during this research.

The one-factor-at-a-time analysis follows the method for a single product analysis as described in chapter 5, and does not improve the efficiency of the tests so is not described in more detail. On the contrary the regression based analyses needs other methods, and can improve the efficiency. The remaining part of this section focuses on this method.

As stated in the scope of the research, we assume that the design parameters only influence the scale parameter of the lifetime model; the family of the distribution remains the same. For the life-stress relations this assumption also holds, therefore we can use some earlier described regression models to model different designs under stress conditions. These are the generalized log linear model; see section 3.3.4, and the proportional hazard model as described in section 3.3.5. Both models include weights and variables that can be used to represent the influence of the specific product design variables. The generalized log linear model can be used in the parametric case and the proportional hazard model in the nonparametric case. When the baseline hazard rate of the proportional hazard model is Weibull or Standard Extreme Value distributed, the reliability functions of the generalized log linear model and the proportional hazard model are equal and the stress coefficients can be transformed to each other. Appendix F shows this relation between the Proportional Hazard and the Generalized Log Linear model.

The first question in testing different designs is which combinations need to be tested in order to be able to fit an accurate regression model. The theory of Design of Experiments (DoE) can be used in this situation; instead of testing all possible combinations, fractional factorial experiments can be executed. Fractional factorial experiments only focus on the main effects and do not include all individual interactions. See for example Condra (2001) for more information about DoE.

3.5 Conclusion

This chapter described models that can be used to model the lifetime of a product under normal circumstances, but also including different failure modes, stress factors, or design parameters. The families of distributions we consider are the Normal and the Weibull distribution. Table 2 gives a summary of the most important life-stress relations and the situations wherein they can be applied.

In the remaining part of this thesis, we focus on the generalized log linear relationship, because via stress transformations the Arrhenius, Inverse Power Law and the Eyring relation can be represented by this relation. The second reason is that this relation can also be used in situations where more than one stress factor is applied. The Proportional Hazard model equals the Generalized Log Linear relationship in case of the Weibull distribution, so is not used during this research. If the focus is on nonparametric models, this model must be used. Cumulative exposure models ask for other methods for the design of tests and analysis, so they are out of scope for this research. The next chapter describes how the model parameters of these life-stress relations can be estimated based on experimental data.

Life-stress relation	Stress factor	Stress load	# stress factors	Lifetime model	Design parameters
Arrhenius	Thermal stress	Constant stress	1 stress factor	Parametric	No
Inverse Power Law	Non thermal stress	Constant stress	1 stress factor	Parametric	No
Eyring	(Non) thermal stress	Constant stress	2 stress factors	Parametric	No
Generalized Log Linear relation	(Non) thermal stress	Constant stress	≥ 1 stress factor	Parametric	Yes
Proportional Hazard	(Non) thermal stress	Constant stress	≥ 1 stress factor	(Non) parametric	Yes
Cumulative Exposure	(Non) thermal stress	Step-stress	1 stress factor	Parametric	No

Table 2: Summary Life-Stress Relations

4. Parameter Estimation and Accuracy Measurement

Chapter 3 described different lifetime models and life-stress relations, but to be able to estimate the lifetime according to a specific percentile, the unknown parameters of the models have to be estimated. This parameter estimation is often done based on the experimental data via Maximum Likelihood Estimation. The experimental data consist of the time until failure of the items at a specific stress level, or the test termination time for the items that are not failed at the end of the test, named censored data. Section 4.1 describes Maximum Likelihood Estimation to estimate the unknown model parameters. These parameter estimations have certain accuracy; this accuracy can be described by confidence intervals. Section 4.2 describes methods to determine these intervals for parametric cases and section 0 for nonparametric cases. Section 4.3 illustrates the use of MLE and the confidence intervals with an example. Finally, section 4.5 summarizes the methods and describes which variables determine the accuracy of the estimations and have to be optimized in an efficient test plan.

4.1 Maximum Likelihood Estimation

To determine the parameters of the life (stress) distribution, Maximum Likelihood Estimation (MLE) can be used. For large samples, MLE has beneficial mathematical properties such as asymptotical normality, asymptotic unbiasness, and consistency; this makes MLE preferable to, for example, least square method (Pascual, 2008). The unknown parameters are estimated based on the likelihood function, a function that contains the unknown parameters and expresses the “likelihood” of the data, given the values of the parameters. Maximizing this expression, by changing the parameter values, results in the parameter values that make the data most likely. The likelihood function can be described as the product of the underlying probability density function evaluated for each data point. A general expression for the likelihood of a sample is:

$$L(\boldsymbol{p}) = L(\boldsymbol{p}; DATA) = \prod_{i=1}^{n} L_i(\boldsymbol{p};\ data_i) = \prod_{i=1}^{n} f(data_i; \boldsymbol{p}) \qquad (22)$$

$L_i(\boldsymbol{p};\ data_i)$: $Likelihood\ of\ observation\ i$
$data_i$: $Lifetime\ of\ observation\ i$
$\boldsymbol{p}$: $Vector\ of\ parameters\ to\ be\ estimated, described\ by\ \boldsymbol{\theta}\ for\ paramteric\ estimates$

For the location-scale distributions with exact right-censored observations, which means that some items that are not failed at the end of the test time, the likelihood function can be written as:

$$L(\mu,\sigma) = \prod_{i=1}^{n}[f(y_i;\mu,\sigma)]^{\delta_i}[1-F(y_i;\mu,\sigma)]^{1-\delta_i} = \prod_{i=1}^{n}\left[\frac{1}{\sigma}\phi\left(\frac{y_i-\mu}{\sigma}\right)\right]^{\delta_i} * \left[1-\Phi\left(\frac{y_i-\mu}{\sigma}\right)\right]^{1-\delta_i} \qquad (23)$$

$$\delta_i = \begin{cases} 1 & if\ y_i\ is\ an\ exact\ observation \\ 0 & if\ y_i\ is\ a\ right\ censored\ observation \end{cases}$$

$$y_i = \begin{cases} t_i & if\ Normal\ distribution\ is\ used \\ \log(t_i) & if\ Weibull\ distribution\ is\ used \end{cases}$$

$$\phi, \Phi = \begin{cases} \phi_{normal}, \Phi_{\mathrm{normal}} & if\ Normal\ distribution\ is\ used \\ \phi_{sev}, \Phi_{\mathrm{sev}} & if\ Weibull\ distribution\ is\ used \end{cases}$$

t_i: $Lifetime\ observation\ i$

Maximizing the likelihood function gives estimations of the unknown parameters. In practice, it is often computationally easier to maximize the log likelihood function to find the vector of the maximum likelihood estimators. The log likelihood function for the smallest extreme value distribution equals:

$$\mathcal{L}(\theta) = \log\big(L(\theta)\big) = \textstyle\sum_{i=1}^{n} \mathcal{L}_i(\theta)$$
$$= \textstyle\sum_{i=1}^{n} \delta_i * \left(-\ln(\sigma) + \frac{y_i-\mu}{\sigma} - \exp\left(\frac{y_i-\mu}{\sigma}\right)\right) + (1-\delta_i) * (ln\, (\exp\left(-\exp\left(\frac{y_i-\mu}{\sigma}\right)\right)))$$

The maximum of the log likelihood function, if exists, occurs at the same value of the parameters as the maximum of the likelihood function (Meeker & Escobar, 1998). Setting all partial derivatives equal to zero and solving the equations simultaneously lead to the maximum and give the estimations of the parameters. To check if the solution is indeed the maximum, the second derivative has to be negative for the estimated parameters (Larsen & Marx, 2006).

The range of the true parameter value with a certain confidence can be given by a confidence interval. Section 4.2 describes these intervals.

4.2 **Confidence Intervals Parametric Case**

A confidence interval is a range around a given statistical estimate within the true value is located with some special degree of confidence (Condra, 2001). There are different types of confidence intervals based on the type of model and the sample size that is used. Important types are Normal approximation confidence intervals and Bootstrap based confidence intervals. We use the normal approximation based confidence intervals, they are often used in reliability testing and can be computed analytically.

The normal approximation confidence interval assumes that $Z_{(\hat{\theta})} = \left[\hat{\theta} - \theta\right]/\sqrt{\widehat{var}_{\hat{\theta}}}$ can be approximated by the standard normal distribution. The 100(1-α)% confidence interval for parameter θ is:

$$\left[\hat{\theta} - z_{\left(\frac{\alpha}{2}\right)} * \sqrt{var\big(\hat{\theta}\big)}, \qquad \hat{\theta} + z_{\left(\frac{\alpha}{2}\right)} * \sqrt{var\big(\hat{\theta}\big)}\right] \tag{24}$$

With $var\big(\hat{\theta}\big) = E\left[\frac{d^2\mathcal{L}(\theta)}{d\theta^2}\right]^{-1}$ and $z_{\left(\frac{\alpha}{2}\right)}\ meaning\ P(Z \geq z) = \frac{\alpha}{2}, with\ Z \sim N(0,1)$

The variance of the estimated parameters can be calculated based on the second derivative of the Likelihood function. The variance can be extracted from the variance-covariance matrix, which is the inverse of the Fisher Information Matrix (Meeker & Escobar, 1998).

$$\begin{bmatrix} var(\widehat{\theta_1}) & \cdots & cov(\widehat{\theta_1}, \widehat{\theta_n}) \\ \vdots & \ddots & \vdots \\ cov(\widehat{\theta_n}, \widehat{\theta_1}) & \cdots & var(\widehat{\theta_n}) \end{bmatrix} = I^{-1} = \begin{bmatrix} E\left\{-\dfrac{\partial^2\mathcal{L}}{\partial^2\theta_1}\right\} & \cdots & E\left\{-\dfrac{\partial^2\mathcal{L}}{\partial\theta_1\partial\theta_n}\right\} \\ \vdots & \ddots & \vdots \\ E\left\{-\dfrac{\partial^2\mathcal{L}}{\partial\theta_n\partial\theta_1}\right\} & \cdots & E\left\{-\dfrac{\partial^2\mathcal{L}}{\partial^2\theta_n}\right\} \end{bmatrix}^{-1}$$

Because of the complexity of the computation of this matrix, especially for functions with more than two unknown parameters, a numerical method is used to solve this. We use the statistical package STATA to calculate the Maximum Likelihood Estimators and the Fisher Information Matrix. STATA converts the maximization problem to a root finding problem and solves this numerically using the Newtons Raphson Method. This is an iterative procedure and stops after convergence. Gould and Sribney (1999) describe the methods STATA uses in more detail. Appendix G shows the likelihood function and the partial derivatives for a Weibull model without stress. Also the likelihood function of a Weibull-Arrhenius model is presented in this appendix to show the complexity of this kind of functions and the need for numerical methods to solve these problems. In the remaining part of this thesis STATA is used for the MLE and variance calculations.

If the parameter to be estimated is a positive parameter, the log transformation generally improves the normal approximation confidence interval in accuracy. The log transformation also ensures that the lower endpoint of the confidence interval will be positive, which is not always the case with the standard confidence intervals. The log transformed confidence interval assumes that $Z_{log(\hat{\theta})} = \left[\log(\hat{\theta}) - \log(\theta)\right] / \sqrt{\widehat{var}_{\log(\hat{\theta})}}$ can be approximated by the standard normal distribution (Meeker & Escobar, 1998).

The 100(1-α)% confidence interval for parameter θ is:

$$\left[\hat{\theta} / \exp\left(\frac{z_{\frac{\alpha}{2}} * \sqrt{var(\hat{\theta})}}{\hat{\theta}}\right), \qquad \hat{\theta} * \exp\left(\frac{z_{\frac{\alpha}{2}} * \sqrt{var(\hat{\theta})}}{\hat{\theta}}\right)\right] \tag{25}$$

The confidence interval for a function of parameters $g(\boldsymbol{\theta})$, for example the pth percentile, $g(\boldsymbol{\theta}) = t_p$ can be described as:

$$\left[\exp\left(\ln\left(g(\hat{\boldsymbol{\theta}})\right) - z_{\frac{\alpha}{2}} \frac{\sqrt{var\left(g(\hat{\boldsymbol{\theta}})\right)}}{g(\hat{\boldsymbol{\theta}})}\right), \qquad \exp\left(\ln\left(g(\hat{\boldsymbol{\theta}})\right) + z_{\frac{\alpha}{2}} \frac{\sqrt{var\left(g(\hat{\boldsymbol{\theta}})\right)}}{g(\hat{\boldsymbol{\theta}})}\right)\right] \tag{26}$$

With $var\left(g(\hat{\boldsymbol{\theta}})\right) = \sum_{i=1}^{n} \left(\frac{\partial g}{\partial \theta_i}\right)^2 var(\widehat{\theta_\iota}) + \sum_{i=1}^{n-1} \sum_{j=\iota+1}^{n} (2 * \frac{\partial g}{\partial \theta_i} \frac{\partial g}{\partial \theta_j} cov(\widehat{\theta_\iota}, \widehat{\theta_J}))$ (27)

The variances of the parameters are calculated based on the Fisher Information Matrix as equation 24 describes.

The variance of the parameter or function of interest determines the spread of the confidence interval. During the development of a test plan, the variance of the pth percentile estimate under normal use conditions must be minimized to make the estimation as accurate as possible.

To be able to compare the variance of the pth percentile estimate for different test plans, a confidence interval of the variance must be constructed. The variance of the variance of the pth percentile estimate can be calculated based on:

$$var\left(var\left(g(\hat{\boldsymbol{\theta}})\right)\right) = var\left(g(\hat{\boldsymbol{\theta}})\right) * \sqrt{\frac{2}{n}} \tag{28}$$

Based on the variance of the pth percentile and a target precision factor, the sample size needed to reach certain estimation accuracy can be computed. A target precision factor of 1.5 means an approximate estimated deviation of 50% between the estimation and the lower (or upper) confidence bound. It is also possible to use the bound ratio as target; the target precision becomes in that situation the square root of the bound ratio. The bound ratio is the upper bound of the confidence interval divided by the lower bound. The formula for the sample size becomes:

$$n^* = \frac{z_{\frac{\alpha}{2}}^2 * \ V_{\log(\hat{g})}}{[\log(R_T)]^2} \tag{29}$$

$$With\ V_{\log(\hat{g})} = \frac{Var(\hat{g})}{\hat{g}^2} * n$$

$\hat{g}: Function\ or\ parameter\ of\ interest$

$R_T: Target\ precision\ factor$

$n: Sample\ size\ at\ which\ Var(\hat{g})\ is\ evaluated$

For the Weibull distribution the value of beta has a large influence on the needed sample size. A small value of the shape parameter beta is related to more relative variability (Zhang & Meeker, 2005). This can also be observed from the shape of the failure rate function. In the case of a small beta value, the failure rate is almost stable and this causes a lot of variability in failure times.

4.3 Example Analysis of Lifetime Data

4.4 **Confidence Interval Nonparametric Case**

In case of a nonparametric situation, the Fisher Information Matrix cannot be used because there are no parameter estimations. The nonparametric confidence interval for the cumulative failure probability can be constructed based on the Greenwood rule:

$$\left[\hat{F}(t_i) - z_{1-\alpha/2} * \sqrt{\widehat{var}\left[\hat{F}(t_i)\right]}\,; \qquad \hat{F}(t_i) + z_{1-\alpha/2} * \sqrt{\widehat{var}\left[\hat{F}(t_i)\right]}\right] \tag{31}$$

$$\widehat{var}\left[\hat{F}(t_i)\right] = \widehat{var}\left[\hat{R}(t_i)\right] = \left[\hat{R}(t_i)\right]^2 \sum_{j=1}^{i} \frac{\hat{p}_j}{n_j\left(1-\hat{p}_j\right)} \tag{32}$$

$With\ \hat{p}_i = \frac{r_i}{n_i} \quad and \quad n_i = n - \sum_{j=0}^{i-1} r_j - \sum_{j=0}^{i-1} s_j$

$m: total\ number\ of\ observation\ moments$

$n: total\ number\ of\ units$

$r_j: number\ of\ failures\ within\ the\ jth\ observation\ moment$

$s_j: number\ of\ censored\ observations\ within\ the\ jth\ observation\ moment$

If the sample size is small, the normal approximation does not hold. A better confidence interval in that situation is:

$$\left[\frac{\hat{F}(t_i)}{\hat{F}(t_i) + (1-\hat{F}(t_i)) * w}\,; \quad \frac{\hat{F}(t_i)}{\hat{F}(t_i) + (1-\hat{F}(t_i))/w}\right] \tag{33}$$

$With\ w = \exp\left(\frac{z_{1-\alpha/2} * \sqrt{\widehat{var}[\hat{F}(t_i)]}}{\hat{F}(t_i)\left(1-\hat{F}(t_i)\right)}\right)$

4.5 **Conclusion**

Maximum Likelihood Estimation can be used to estimate the unknown parameters of the lifetime models. Because of the complexity of the likelihood formula, STATA is used to calculate the parameters. The variance of the parameter estimations can be derived from the inverse Fisher Information Matrix. Based on these variances and the normal approximation, the confidence interval of the percentiles can be computed. The variance determines the spread of the interval, to make estimations as accurate as possible the expected variance of the parameters or function of interest must be minimized during the design of a test plan. The standard error is the square root of the variance and has the same dimension as the pth percentile; this makes the standard error easier to interpret as the variance. So, in the remaining part of this thesis, we focus on minimizing the standard error instead of the variance. The case shows that an equal allocation of the items over three stress levels did not lead to very accurate percentile estimates; the confidence interval is very wide. Hence, in the next chapters we describe methods to design test plans that lead to more accurate estimations.

5. Test Design

The goal of lifetime experiments is to make an accurate estimation of the reliability or lifetime of a product. This chapter describes how such test plans can be designed for the scenario with one stress factor, one failure mode and one product design. Section 5.1 describes properties of a test plan in general; section 5.2 presents the setup of a test plan and gives a description how to design a test plan for the one stress factor, one failure mode and one product design variant scenario. Chapter 6 describes the extension of this method to methods for the other scenarios as shown in Figure 4. Section 5.3 illustrates the method with an example based on the case introduced in section 1.4. Section 5.4 compares the test configurations of the different types of test plans and evaluates the robustness of the test plans with respect to parameter deviations and model departures. Finally, section 0 concludes this chapter by describing in which situation which type of test plan can be used best and gives some directions for the extensions of the methods to the more complex scenarios.

5.1 Test Plan Properties

Important information and capabilities to design an effective accelerated test plan are:

- Knowledge of the functional- and reliability requirements of a product
- Understanding of a product's normal use conditions, applicable stress levels, and the stress range. This means that some assumptions only hold for this limited range of stress or that new failure modes can occur outside this limit that do not occur under normal use circumstances.
- Expected failure modes.
- Life-stress relations for each of the failure modes in order to design the lifetime test and to analyze the test results.
- The capability to analyze physical failures and identify due to which failure mode the product has failed.

We assume that this information and these capabilities are available before designing the test plans. This can be obtained from qualitative analyses, such as MEOST and HALT, and experiences from previous products. See Bhote and Bhote (2004) for more details about such qualitative methods.

A test plan includes information about the configurations and the number of different test runs. This must be chosen such that an accurate estimation of the p^{th} percentile can be made and practical constraints are met. These practical constraints can be test time or sample size restrictions. A test plan consists of the following information:

- The total number of test units, n
  This number is often based on product availability or resource capacity, but the number can also be determined based on a target precision of the percentile estimate. If the last one is the case, the total number of test units is determined based on the expected standard error of the p^{th} percentile estimate for the best test plan configurations, see equation 29.

- Test termination time, t_c
  Based on the time available for testing or other practical constraints, a decision about the duration of the lifetime test has to be made.
- Stress factors that can be used to accelerate the failure process
  Based on the physics of the failure process of the product, applicable stress factors should be chosen. Also the number of different stress factors that can be used depends on the physics of the product.
- Number of different stress levels
  The number of different stress levels determines the accuracy of the percentile estimation. If the behavior of the failure process is uncertain more test levels have to be used than if the behavior is assumed to be known. Section 5.2.2 describes rules to determine the number of test levels.
- Value of the stress levels per stress factor, x_{ij} (i: stress factor indicator, j: stress level indicator)
  The highest stress level is fixed for all test plan designs; it is set at the limit of the stress range. This highest stress level is located there to generate as many failures as possible; the probability of non failed items is the smallest at this level. The other stress level values are decision variables during the optimization of the test plan configurations.
- Allocation of the units over the stress levels , π_{ij} (i: stress factor indicator, j: stress level indicator)
  The allocation of the units over the different stress levels can be given based on a specific test plan type or must be optimized. A given allocation makes the test plan design faster, because fewer alternatives have to be evaluated, but the results are less accurate. Different types of test plans are described in section 5.2.2. In most of the cases, the allocation is a decision variable of the optimization problem. On the one hand the goal is to generate as much failures as possible at the lowest stress level, because this leads to an accurate estimation of the pth percentile for that specific stress level. But on the other hand, if a large amount of the items is allocated to the lowest stress level, a lot of these items will be censored, because the failure probability at the lowest stress level is low. This large amount of censoring leads to less accurate results than when these items were allocated to a higher stress level. In the second case, most of these items were failed and a more accurate extrapolation over stress can be made. A trade-off between the accuracy due to the number of failures and the extrapolation over stress is made during the optimization.

The optimization criterion to obtain the test configurations is minimizing the standard error of the pth percentile estimate under normal use conditions. The decision variables are the values of the lowest stress levels and, depending on the test plan type, the allocation of the units over the stress levels. The remaining part of this chapter focuses on methods to solve this optimization problem.

5.2 General Setup for Test Plans

Independent of the situation or the type of test plan that is used, a general setup can be described that can be followed during the design of a test plan. First of all we give a global description of this setup, and after that, the different steps are described in more detail.

The general steps during the development of a test plan are:

1. Obtain prior information and initialize the life-stress model

   Prior information is necessary to make assumptions about the lifetime model and the corresponding parameter values. We need this assumed model to evaluate different test configurations and to select the best one.

   Section 5.2.1 describes the prior information that must be obtained. This information can be obtained from knowledge of previous products, earlier tests or physics of the product. The parameter values can be estimated based on the linear life-stress relation, section 5.2.1 also describes this method. The optimization needs the assumed lifetime model with parameter values to determine the expected amount of failures per stress level and to find the values of the decision variables that minimize the standard error of the p^{th} percentile estimate.
2. Choose type of test plan and design the test plan

   The scientific literature describes different types of test plans for the one failure mode, one stress factor and one product design variant scenario. Based on the uncertainty of the prior information the most appropriate type can be selected. The type of test plan defines the number of stress levels that must be used and the determination of the allocation of the items over the stress levels. Some types of plans specify this allocation explicitly, and in some type of plans this allocation has to be chosen during the optimization. Section 5.2.2 describes the test plan types for the one stress factor, one failure mode and one product design variant scenario. Chapter 6 describes the test plant types for the other scenarios.

   In this step we select the best test plan, the plan that scores the best on the optimization criterion, via optimization. The optimization problem has a nonlinear form, so it is difficult to solve. Instead of solving this nonlinear optimization problem analytically, we use simulation and select the plan with the best performance. Section 5.2.3 describes this optimization via simulation in more detail.
3. Test robustness of test plan by changing prior estimations

   To design a test plan that is robust to changes of the prior information due to uncertainty about these values, we established a robustness test. The best test configurations obtained in step 2 are used for this test. Not only the best test plan is used, also the plans that do not differ significant from this best test plan. We measure the significance of the difference with use of a confidence interval for the standard error of the p^{th} percentile estimate. We change the prior information assumptions that can affect the best test configurations one by one and evaluate the effect on the standard error of the p^{th} percentile estimate for each of the selected test configurations. Section 5.2.4 describes which assumptions have the largest impact on the test configurations and gives a method to execute these tests. Based on the test results, we choose the final test plan configurations to incorporate the uncertainty about the prior information.

5.2.1 Prior Information

In order to describe the expected standard error of the percentile estimate that results from a particular test plan, it is necessary to have some “prior information” about the lifetime distribution. This information is needed to assess the effect of sample size, stress levels and allocation of the items over the different stress levels on the outcome of a test plan. Such information can be obtained from design specifications, expert opinion, or previous experiences (Meeker & Escobar, 1998). Tang and Liu (2010) describe a sequential accelerated life test procedure. Before the real lifetime test is executed a part of the items is used to obtain the needed prior information. These items are tested at the highest stress level and based on the results of this lifetime test the prior information assumptions are made. This can be an efficient way to obtain information if this is not available from previous experiences or expert opinions.

The amount of prior information that is necessary depends on the scenario for which the test plan has to be designed. This section describes the information that is needed in every scenario. When a range of possible values is available instead of one specific value, test plans have to be constructed for the limits and the expectation of this value, the best test configurations per value can be compared, and a plan that is generally satisfactory should be chosen (Meeker W. , 1984).

The necessary prior information is:

- (Assumed) family of lifetime distribution (Weibull/(Log)Normal)
  This information is needed to construct the assumed lifetime model, and to be able to estimate the model parameter values with their variance and covariances.
- Normal use level & highest level of stress
  These values define the experimental region of the stress levels that can be used during the lifetime test. The standard error of the p^{th} percentile estimate is minimized at the normal use stress level, because we are interested in the reliability under normal use conditions. In general this stress level is too low to generate failures in a short period of time, so most of the time it is not used as test level. The highest stress level is always used as test level during the lifetime tests, the failure probability is the largest at this level and the amount of censoring is the lowest. This increases the accuracy of the percentile estimate.
- Life-stress relation (Arrhenius, Inverse power etc., see section 3.3)
  Possible stress relations for different stress factors are described in chapter 3. The expected relation is used to transform the stress to be able to use a Generalized Log Linear lifetime model.
- Assumed percentage of failures at end of test at normal use stress and at highest stress level
  These two percentages define, together with the stress relation, the acceleration factor of the applied stress. The parameters of the life-stress relation can be estimated based on these values to construct the expression for the location parameter dependent on stress.
- Available test time
  Based on the available test time and the failure probabilities, the expected number of failures per stress level can be calculated. This expected number of failures per stress level is necessary to define the range of the lowest stress level. To generate accurate results the probability of failure at the lowest level must be as least as large as the percentile of interest, because otherwise extrapolation over time is needed, and this reduces the accuracy extremely. Extrapolation over time means that the percentile point for a lower percentage than the percentage of interest can be determined based on the experimental data, so the percentile point of interest is estimated outside the range of observations.

- Assumed value of the scale parameter (beta (Weibull) or standard deviation ((Log)Normal))
  This estimation gives, together with the expression for the location parameter the complete expression for the estimated lifetime, dependent on the stress. This information is used to make an estimation of the standard error of the percentile of interest and to determine the sample size needed to obtain certain accuracy.
- Performance indicator
  The indicator should be chosen based on the goal of the lifetime test and the information one would like to obtain with the test. In practice often the 10th percentile is of interest, so during this research we use the 10th percentile as performance indicator. We minimize the standard error of the 10th percentile estimate during the development of a test plan, to obtain an accurate estimation.

We use the case to illustrate the need of the prior information and to illustrate the method to initialize the life-stress model.

The next section presents different type of test plans, and after that, section 5.2.3 describes how to come up with a test plan based on the initialized lifetime model and the chosen type of test plan.

5.2.2 Type of Test Plans

Different methods to design a test plan for the one failure mode, one stress factor, one product design scenario are available in the scientific literature; this section describes the most appropriate ones. We examined the appropriateness based on the reasoning behind the methods and the assumptions made. If an improved method is developed based on exactly the same situation and assumptions, we only describe the improved method. The assumptions of the methods must match the assumptions made during this research, as described by section 2.3. For each of the test plans holds that the highest stress level is assumed to be fixed, and equals the highest stress level as defined by the prior information. Table 3 summarizes the different test plan types and describes the main differences.

1a. *Two Level Statistical Optimum Plan*

This method minimizes the standard error of the p^{th} percentile, $SE(t_p)$, by choosing the lowest stress level and the proportion of units allocated to the lowest stress level. In total two stress levels are used. This method is appropriate if the prior information is expected to be true, because it is very sensitive to model departures (other family of distributions or stress relationship) and parameter deviations (Meeker & Escobar, 1998). Meeker and Nelson (1975) describe this method in more detail and use their own figures to come up with optimum stress and allocation values, based on the prior information. Methods to create these figures are not described, so we cannot use exactly the same method. We use our own method to optimize $SE(t_p)$ to generate the best test plan configurations for this test plan type. Section 5.2.3 describes the method we use for optimization in more detail.

1b. Three Level Best Compromise Plan

The idea behind this method is that the resulting test plan is robust to wrong parameter estimates and model assumptions while the statistical optimum plan, plan 1a, is not. A middle stress level is introduced to ensure this robustness. This method minimizes SE(t_p) by choosing the lowest stress level and the proportion of units allocated to the lowest stress level as the optimum plan does, but allocates a fixed percentage of the items to the middle stress level. The lowest stress level that result from this optimization can be lower than that of the statistical optimum plan, plan 1a, because also a middle stress level is used to generate failures. The middle level is located at the normalized average of the chosen low stress level and the high stress level. The percentage of units allocated to the middle stress level is fixed, often 20%, because otherwise by optimization the units will be allocated to the low and the high stress levels only. The minimal SE(t_p) will be higher than that of the statistical optimum plan, plan 1a, but the advantage of this method is the robustness to wrong parameter estimates (Meeker & Hahn, 1977).

1c. Three Level 4:2:1 Allocation Plan

To make the extrapolation from the test stress level to the normal use stress level as accurate as possible, an equal number of failures per stress level is best. This idea is based on the theory of Design of Experiments (DoE). But in case of lifetime testing, this leads to a very high allocation of the items to the lowest stress level, where the probability of failure is low, and the amount of censored observations is high. While fewer items are allocated to the highest stress levels this leads to less failures on these levels. To overcome this problem, but use the idea of DoE Nelson (1990) constructed a compromise plan with this same idea. A high fraction, 4/7th, of the items is allocated to the lowest stress level, 2/7th is allocated to the middle stress level and 1/7th is allocated to the high stress level. The lowest stress level is chosen such that SE(t_p) is minimized, the middle stress level is located at the average of the normalized low and the normalized high stress level.

Other Plans and Guidelines

Some other type plans are described in the scientific literature. They are used for specific situations and are difficult to generalize, but some ideas can be used to design new types of test plans. Appendix I gives a short summary of these methods with references to the complete descriptions. We developed a new method based on the ideas of these other plans and this type of test plan is described below.

1d. New Proposed Plan

We develop the new proposed plan as variant of the best compromise plan, plan 1b, with use of the ideas of Meeker and Hahn (1985) and Tang and Yang (2002). They state that the middle stress level is only used to check model departures and to make the test plan less sensitive to wrong parameter estimations, so a low fraction of the items can be allocated to this level. It is important to ensure that at least p% of the items, with a minimum of 5 items, will fail at the middle stress level. At least p% of the items must fail because this reduces the variance for the p^{th} percentile estimate, otherwise extrapolation over time has to be done. The minimum of 5 items is needed to be able to estimate the parameter values with some accuracy. This constraint is easily met for large sample sizes, but for small sample sizes this can be a problem. So, for the generation of the test plans we allocate that amount of items to the middle stress level that cause p% of the total items to fail. If the total sample size available for the test is known beforehand, the settings for this small sample sizes must be optimized based on the minimum of 5 failures restriction. In summary, the goal is to optimize SE(t_p) by choosing the lowest stress level and the fraction allocated to the lowest stress level as the best compromise plan, only with the extra constraint on the middle stress level to improve the accuracy.

Table 3 summarizes the different test plan types and describes their special properties. The next section describes how the test configurations for the different types can be generated.

Plan	Author(s)	Optimization criterion	Decision Variable(s)	# Test levels	Remark(s)
Statistical Optimum Plan	Meeker & Escobar/ Ahmad	SE(t_p)	x_{Low}, π_{Low}	2	Very sensitive to wrong estimations
Best Compromise Plan	Meeker & Hahn	SE(t_p)	x_{Low}, π_{Low}	3	Compromise between statistical optimality and robustness
4:2:1 allocation	Nelson	SE(t_p)	x_{Low}	3	Based on DOE principles, equal # failures per level
New Proposed Plan	-	SE(t_p)	x_{Low}, π_{Low}	3	Increase accuracy by allocating less items to the middle stress level

(x_{low}: lowest stress level, π_{Low}: allocation fraction to lowest stress level)

Table 3: Overview of Methods to Design Test Plans for the Basic Scenario

5.2.3 Generation of Test Plans

Section 5.2.2 described methods to design test plans, but a nonlinear optimization is necessary to obtain the values of the decision variables and no method to solve this is described in the scientific literature. The variances of the parameter estimations are necessary for this optimization, because the variance and the standard error of the percentile estimate can be calculated based on these variances. Normally, the variance can be computed based on the observations via MLE, but no observations are available before the lifetime test is executed. To overcome this problem, we choose to use Monte Carlo simulation to generate lifetime observations based on prior information and to make estimating the variance of the parameters possible. Based on the prior information, we assume a life-stress model, and based on this model the lifetimes are simulated. We use this simulated data to estimate the variance of the model parameter estimates via MLE and based on this we calculate the standard error of the percentile estimate of interest. Finally we select the test configurations with the minimum standard error as best test plan configurations. This section explains the methods and ideas behind the optimization via simulation; section 5.3 illustrates this with use of an example.

Based on the type of test plan, we create different test plan configurations and thereafter we generate experimental results per test plan configuration based on simulated lifetimes. We select the test plan configuration with the smallest standard error as the best one. To obtain accurate results, the simulation consists of some replications. The average standard error per test plan configuration over the replications is used as optimization criterion. Figure 8 presents the generation process schematically.

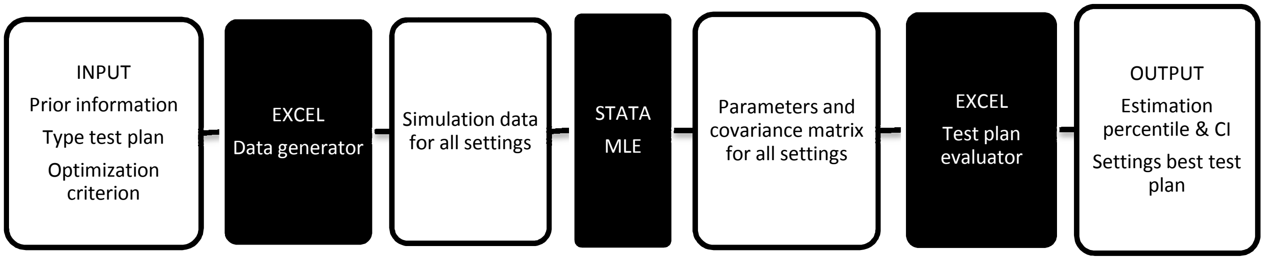


Figure 8: Test Plan Generation Process

The first step of the generation process is the generation of simulation data; we use Microsoft Office Excel for this step. We made a program called “Data Generator” to generate the failuretime data based on the prior information. We use a large sample size per test to be able to make use of the asymptotical unbiasness and the asymptotical normal properties of MLE. Experiments have shown that a sample size of 2,000 can be used to generate unbiased estimations of t_p. Appendix J shows the experimental results. To make the comparison between test plans fair, we use the same lifetimes for all test plan configurations, this means that we make use of common random numbers to generate the lifetimes per observation for the different test configurations. To generate stable results, we use replications of the simulation. Experiments have shown that 40 replications generate stable results, see also Appendix J for these experimental results

We use the prior information, type of test plan, and the optimization criterion to come up with different test plans configurations. To generate these test plan configurations, we established the following rules based on experimental results and rules of thumb:

- The lower bound of the lowest stress level is chosen such that the failure probability is at least as high as the percentile that is used for the optimization.
  $\xi_{\text{Low}} \geq \frac{\ln(t_c)-\Phi_{SEV}^{-1}(p)*\sigma-\alpha_0}{\alpha_1}$
  This minimum is necessary to reduce variance, because otherwise extrapolation over time is needed. Only the percentile point for which failures are observed can be determined based on the sample data, the remaining points must be estimated based on extrapolation of this data and this increases the variance.
- The upper bound of the lowest stress level is the stress level for which the failure probability is two times as large as the percentile used for optimization, with a maximum of 100%.
  $\xi_{\text{Low}} \leq \frac{\ln(t_c)-\Phi_{SEV}^{-1}(\min(2p,1))*\sigma-\alpha_0}{\alpha_1}$
  An upper bound is needed to limit the possible stress configurations. Several experiments have shown that a stress level near to the lower bound is chosen, so the upper bound needs not to be

that large. To give the optimization some flexibility, we choose to set the upper bound at the level for which the failure probability is two times as large as the percentage of interest. The value can be changed by the user of the simulation program.

- A finite number of stress levels must be examined. We take only integer values in our example.
- The lower bound of the allocation fraction to the lowest level is 1/k.

  $\pi_{Low} \geq \frac{1}{k}$, k is the number of different test levels

  The reason for this is that the fraction allocated to the lowest stress level must be at least as large as the fraction allocated to the other stress levels to obtain accurate results. (For some plans the fractions are given, the lower and upper bound become the pre specified fraction in this case)

  The upper bound of the allocation fraction to the lowest stress level equals 1-(k/10).

  $\pi_{Low} \leq 1 - \frac{k}{10}$

  The other stress levels must also have a fraction of units to create failures, so not all units can be allocated to the lowest stress level. We use this rule because if two stress levels are used at least 2/10th is allocated to the highest stress level, this is necessary for accurate model and parameter estimates on that level. For three stress levels, 2/10th is allocated to the middle stress level and at least 1/10th must be allocated to the highest stress level, so at most 7/10th can be allocated to the lowest stress level. We developed several three level test plans, and the allocation percentage was in all cases between the 40% and the 60%. Based on this, we conclude that this rule can be used.
- The step size of the allocation fractions is finite; we choose a step size of 1/10th in the example.

Per test plan configuration the allocation of the items over the stress levels, or the lowest stress level value is different. For example suppose we want to create a statistical optimum plan, plan 1a. This means that two stress levels are used and both the lowest stress level and the allocation fraction to the lowest stress level are decision variables. The allocation fraction at the lowest stress levels can vary from 0.5-0.8 according to the rules described above; this leads to 4 possible allocation fractions. The lowest stress level can take the values 96W-101W based on the rules above and the expression for the failure probability dependent on stress, this leads to 6 different lowest stress levels. This means that 24 different test plan configurations have to be compared (4*6). We simulate failuretime data based on the prior information and per test plan configurations this failure time is accelerated according to the applied stress. If the accelerated failuretime for that stress level is smaller than the total test time, the item fails during the experiment, otherwise the item is censored. We generate random failuretimes based on the assumed lifetime distribution, common random numbers per test plan, and the life-stress relation. Per test configuration we use the obtained failure data to estimate the parameters of the lifetime model based on MLE, and we use the covariance-matrix of these parameters for further analysis. Based on this covariance-matrix, the variance and standard error of the pth percentile estimate can be calculated. The test plan configuration with the lowest average standard error over all replications for the pth percentile under normal use conditions becomes the best test plan. Appendix J, describes the simulation process and the program we created to determine the variance and standard error in more detail. Section 5.3 illustrates which type of test plan fits the example and what the test plan settings become in this situation.

5.2.4 Test Robustness

To ensure that the designed test gives accurate results, even if the true lifetime model deviates from the assumed lifetime model, the robustness of the test plan that results from the optimization must be evaluated. We evaluate the best test plan and the test plans that are not significantly different from this test plan, during this robustness test. Per test configuration, we construct the 95% confidence interval of the standard error of the p^{th} percentile estimate with use of equation 28. If the confidence interval of a test plan overlaps with the confidence interval of the standard error of the best test plan, we consider this test plan not significantly different from the best test plan, and this plan must be taken into account during the robustness test. Because this leads to a lot of test plans that must be evaluated, we only focus on the plans for which $SE(t_p)$ lies within the confidence interval of the best test plan. We insert new values of the prior information to check which of these selected test plan settings can deal best with the deviations, and a final plan is chosen based on these test results. The assumed prior information can be changed per assumption, but also in couples if it is more feasible that they change together. The assumed value of the scale parameter is not taken into account, because this value does not influence the best test configurations, it only influences the accuracy proportional to the parameter change. Appendix K describes the influence of the scale parameter in more detail.

Assumptions within the prior information that have a large impact on the resulting test plan are:

- Percentage of failure at the end of the test under normal use conditions

  This percentage determines the acceleration factor of the stress. An overestimate of this percentage leads to a too low lowest stress level, less failures than expected will occur, and the expected accuracy cannot be reached.
- Life-stress relation

  The stress relation is chosen based on the physics of the product. If there is some uncertainty about the relationship, test plans for other relationships can be generated and the differences can lead to a compromise plan. Which plan has to be selected depends on the differences in test plan settings and the probability that the different assumptions will be true.
- Family of distributions

  The family of distributions is assumed based on expert knowledge or knowledge from previous products. If one is not sure about this assumption, different families of distributions can be evaluated and a test plan that scores well for the most reasonable families must be selected.

5.3 Example Generation of a Test Plan

5.4 **Comparison of Test Plan Types**

To give an overview of the test plan types and the situations in which they are most appropriate, we compare the settings and the standard error of the different type of test plans for different situations. A deviation of the scale parameter (Weibull: (1/beta), (log)normal: sigma)) is not taken into account during this analysis because this parameter has only influence on the sample size needed to reach certain accuracy and not on the stress level settings. Appendix K describes the influence of beta on the variance of the percentile and the sample size needed to reach certain accuracy.

Situation 1. Assumed model

The first situation for which we compare the performance of the different test plan types is the normal situation, the situation for which the assumptions about the model and the stress relations are true. We compare the standard error of the 10th percentile estimates and the test plan type with the smallest standard error is the most appropriate type of test plan for this situation.

Table 4 summarizes the test plan configurations for the four different test plan types for the example described in section 5.2.2. Based on situation 1 we conclude that test plan type 1a, the statistical optimum plan, is the best test plan type if the assumed situation seems to be the true situation.

Table 4: Test Plan Configurations per Test Plan Type

To check the robustness of these test plans types, the ability to deal with parameter deviations and model departures, we use the test plan settings as given in situation 1. We change some parameters of the model that simulates the failuretimes, such that the true model with parameter values deviates from the assumed model with parameter values. We do not change all the parameters at once, so different situations are used for this comparison. These situations are described below and Table 10 presents the scores.

Situation 2. Deviation of the parameters of the acceleration model

In this situation we change the parameter values of the acceleration model, $\mu(\xi) = \alpha_0 + \alpha_1 \xi$. This deviation can be obtained easily by changing the initial estimates of the failure probabilities at the end of the test, because the parameters are derived from these estimates.

So the changes we make in this situation are:
- Increase the failure probability under normal use condition
- Decrease the failure probability under normal use condition
- Increase the failure probability at the highest stress level
- Decrease the failure probability at the highest stress level

Several deviation percentages can be chosen for the comparison. For our example we choose deviations of 25%, 50% and a factor 10, but also other values can be used, depending on the situation.

Situation 3. Other stress relation

In the example we chose to use the power relation to relate the lifetimes under stress to the lifetimes under normal use condition. This choice is based on the physics of the product and the applied stress. In practice it can be true that the failure process can be described better by another stress relation. So in this situation we generate the lifetimes based on other stress relations than initial assumed. For the example we evaluate the Arrhenius and the Linear relation instead of the Power relation, and we examine the effect on the standard error of the estimate for each type of test plan.

Situation 4. Other family of distributions

In this situation we change the assumed family of distributions. In the example the Weibull distribution is used to describe the lifetime, so we have to evaluate the performance of the different test plan types when the true family of distribution is Lognormal or Normal. The scale parameter of the "new" distribution is chosen such that the expected lifetime of the product remains the same.

Table 5 summarizes the score of each of the four types of test plans, as described during this chapter, on these situations. This table describes the standard error of the 10th percentile estimate per situation and per test plan type. We also use the median and the spread of the standard errors between replications in examining which test plan is the best with use of histograms. For most of the situations the best test plan type scores best on all three criteria, but in some situations not. The reviewer has the responsibility to make the end decision based on the performance on the three criteria. In this case we are the reviewers and prefer the test plan type with the smallest average SE(t_p) unless the difference in average is smaller than 0.5% with respect to another test plan which has a smaller spread and a smaller median. The gray cells in the table indicate per situation the test plan that scores the best according to our criteria. Only the cell marked with a * is chosen based on the best score on the criteria besides the average. We conclude that the "Statistical Optimum Plan" seems to have the best overall performance, but the "Best Compromise Plan" performs better if model departures occur. This plan can also deal best with large parameter deviations, especially parameter increases. The "New Proposed Plan" can also deal with parameter deviations, but performs best if these deviations are small. In the other situations, the "New Proposed Plan" performs second best. From the comparison results, we derive suggestions for the appropriateness of the different types in different situations. Figure 9 summarized the results of the comparison study. With fast optimization we mean that the best test configurations can be generated fast, because a small optimization is used. The 4:2:1 test plan chooses only the stress levels; the allocation of the units over the levels is pre-specified, so the time needed to obtain the best settings is small. The optimization speed is the advantage of this type of test plan, but this can result is a less accurate test plans than with a full optimization. If no decision can be made about the situation of Figure 9 the "New Proposed Plan" can be used best, because it gives always the best or the second best test plan.

Table 5: Comparison of SE(t_{10}) for Different Types of Test Plans in Different Situations

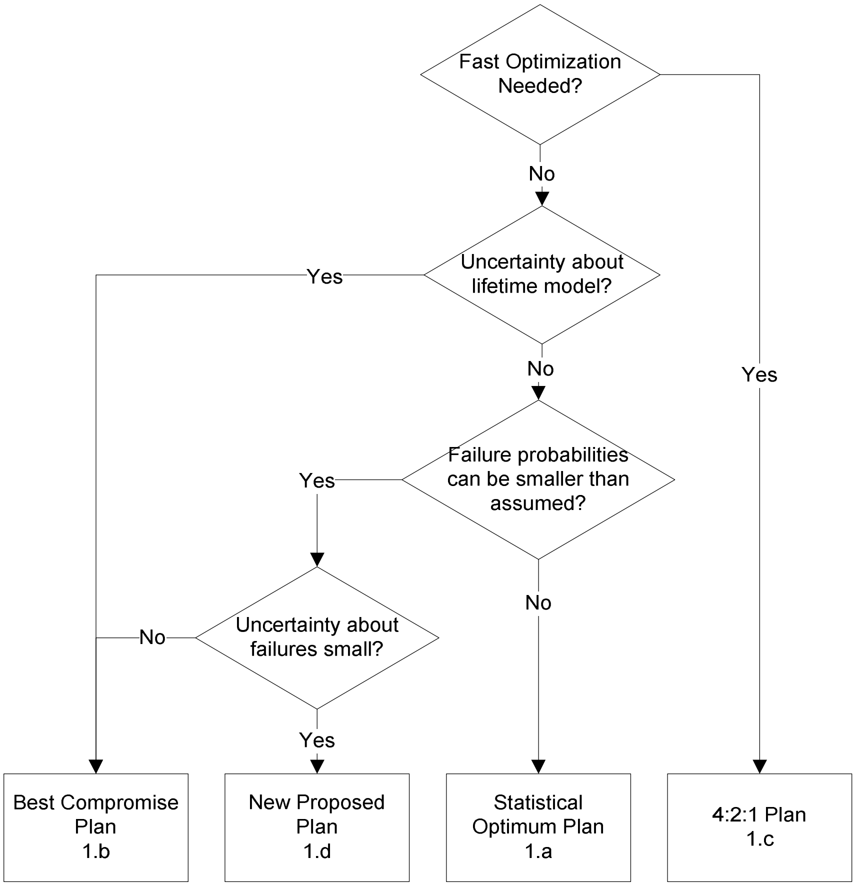


Figure 9: Most Appropriate Test Plan Type per Situation- One Failure Mode, One Stress Factor and One Product Design

5.5 Conclusion

This chapter described several methods available in the scientific literature to develop tests plans for the one failure mode, one stress factor and one product design scenario based on the Smallest Extreme Value (Weibull) distribution. The described methods can also be applied to the Normal and the Lognormal distribution if some small changes are made. To make the methods appropriate for the Lognormal distribution, instead of using the Standard Extreme Value distribution, the Normal distribution must be used as basic model. For the Normal distributed lifetime also the Normal distribution must be used, and instead of calculating the log failure times, immediately the normal failuretimes are calculated. So, no logarithmic transformations have to be made in this situation. The simulation program we developed can deal with these three families of distributions.

The methods described during this chapter use the standard error of the p^{th} percentile estimate as optimization criterion and use the low stress level and, in most cases, the allocation of units to the lowest stress level as decision variables. Methods to execute this optimization are not described in the scientific literature, some authors use their own tables, and other authors give the nonlinear optimization problem, but they do not describe generic methods to solve this. To generate the best test settings for different situations, we developed our own optimization method based on simulation. The procedure we developed to design a test plan consists of the steps:

1. Obtain prior information and initialize the life-stress model
2. Choose type of test plan and design the test plan with use of simulation
3. Test robustness of test plan by changing prior estimations

Based on a comparison of the different test plan types in different situations we conclude that the Statistical Optimum Plan seems to have the overall best performance, but the Best Compromise Plan, plan 1b, can deal best with wrong assumptions. The New Proposed Plan has the best or second best performance, so it can be chosen if one does not want to make a decision about the situation. Figure 9 suggests which test plan type can be used best in which situation; based on this type the test configurations can be established. Chapter 6 extends the methods described in this chapter and combines them with other methods available to describe the design of test plans for the other scenarios.

6. Development of Test Plans

The goal of this chapter is to describe a complete procedure to develop test plans appropriate for different scenarios based on the number of failure modes, the number of stress factors and the number of different product design variants. To come up with this procedure, we use the method described in chapter 5 and develop new methods. We also extend the simulation program used in chapter 5 for the generation of test plans. This chapter consists of detailed descriptions of methods to generate test plans for the scenarios with one factor larger than one, so for example two or more failure modes, one stress factor and one product design or one failure mode, two or more stress factors, one product design variant, descriptions of ideas to generate test plans for the remaining scenarios and a conclusion with a guidelines to generate test plans in general. Per scenario we describe appropriate test plan methods from the scientific literature, if they exist or methods of other scenarios. We describe the generation process not in detail, because it works largely the same as the generation process described in chapter 5. We only describe special results and comparisons of different plans within one scenario.

6.1 Two or More Failure Modes, One Stress Factor, One Product Design

For this scenario we assume that a product can fail for more than one reason. Examples of failure modes are cracking of the packaging, short circuit or wear. The failure modes that can occur during normal use of a product are taken into account. When one of the modes appears the product fails and the other modes cannot be observed. We assume that when a product fails, we can observe the corresponding failure mode and we have some initial estimates for all failure modes we want to include in the lifetime test. This section describes which types of test plans can be used for this scenario and how these test plans can be generated.

6.1.1 Type of Test Plans

The first type of test plan for this scenario is a test plan obtained from the scientific literature. We do not use this test plan type directly, because it is very situation specific and no methods to really design the plan are described, but we use the plan as theoretic background to develop new plans.

2a. D- and D_s- Optimal Plan

The test plan that can be found in the literature for this scenario is the D- optimal plan. A test plan is called D-optimal if it maximizes the determinant of the Fisher Information Matrix, $|F|$. When not all model parameters, but a subset is of interest, D_s-optimal test plans are more suitable. For example when the focus is only on failure mode *i*, the determinant of $|F_i|$ is maximized instead of the complete matrix. (Pascual, 2008). Pascual (2007) describes a simplification of the D-optimal method for situations where the shape parameter of the Weibull distribution (scale parameter for SEV distribution) is known and assumed to be equal for the different failure modes. In this case, an expression for the percentile can be established and the variance of the percentile estimation can be calculated with use of the Fisher matrix. The expression for the pth percentile for a product with *k* different failure modes becomes: $t_p(x) = \left\{\frac{-\ln(1-p)}{\sum_{i=1}^{k} \exp\left(\frac{-\mu_i(x)}{\sigma}\right)}\right\}^{\sigma}$. In practice it is uncommon that all the failure modes have the same scale parameter, so we also want to develop a procedure for different scale parameters. The prior estimates are difficult to obtain for the *k* failure mode case, because the estimations of the failure probability under normal use condition is often a combination of all failure modes. To get around this problem, we assume that independent estimations are available.

The simulation model we used for the generation of test plans for the one-one-one scenario optimizes SE(t_p) instead of the determinant of the Fisher Matrix. To make the test plan generation procedure as generic as possible, and because the additional step to compute SE(t_p) is not that difficult with use of the simulation program, we decide to use this criterion, min SE(t_p) also for this scenario. The literature does not describe how to come up with the test plan settings, so we have to develop such a method ourselves.

2b. Two Failure Modes Test Plan

The idea behind this test plan is to search for stress levels and allocation fractions that minimizes SE(t_p) for all the applicable failure modes. The highest stress level is chosen beforehand, just like for the one failure mode scenario. The lowest stress level must be chosen to minimize SE(t_p). Because different failure processes with their own lifetime distribution take place, we treat the failure processes separately for the parameter estimation and variance calculation. At the end of the test plan evaluation we establish the model for the overall failure process, so including all applicable failure modes. Equation 9 described how such combined models can be made if the failure modes are independent. If the scale parameter is not the same for all failure modes, no general expression for the p^{th} percentile as function of all individual failure modes can be established. The expression for the probability that an items with two failure modes fails before time *t* equals:

$$F(t,\xi) = 1 - \left[1 - \Phi_{SEV}\left(\frac{\ln(t)-\mu_1(\xi)}{\sigma_1}\right)\right] * \left[1 - \Phi_{SEV}\left(\frac{\ln(t)-\mu_2(\xi)}{\sigma_2}\right)\right] \tag{35}$$

Based on this expression we can conclude that an expression for the percentile point t can only be derived if $\sigma_1 = \sigma_2$, see Appendix L for the derivation. If this is not the case, a value for t must be searched via a search method. This also means that numeric calculation of the standard error of this percentile is not possible. For that reason, we determine the standard error of the percentile for the combined model with use of simulation. We use parameter values from a multivariate normal distribution, with means equal to the expected parameter values based on MLE of the simulated data and variances equal to the variances obtained from the Fisher Information Matrix. We use these new parameter values to generate lifetimes. We use the standard error of these new generated lifetimes as estimate for the standard error of the overall product. The steps needed to generate the best test plan are:

1. Obtain prior information per failure mode.
2. Choose a type of test plan based on the one failure mode, one stress factor, one product design variant scenario, Figure 9.
3. Determine the bounds of the lowest stress level based on the total failure probability, equation 34, as described in section 5.2.3.
4. Simulate lifetimes per failure mode per product. The failure mode that occurs first causes the failure of the product and determines the lifetime.

5. Estimate the life stress model parameters and their variances per failure mode. If a failure occurs due to another failure mode, assign a censored indicator to this failure mode and take this lifetime as censored observation into account.
6. Generate independent lifetimes per test plan for the pth percentile based on multivariate normal draws for the model parameters. With multivariate normal draws we mean a draw from the normal distribution, but with a mean equal to the estimated value of the parameter and a variance equal to the estimated variance. The parameters corresponding to the stress relation are allowed to take negative values, but sigma must always take a positive value. The lognormal distribution seems appropriate for this situation.

   The relation we use to generate these lifetimes t via a search method is:

   $$p = F(t,\xi) = 1 - \left[1 - \Phi_{\mathrm{SEV}}\left(\frac{\ln(\mathrm{t}) - (\hat{\gamma}_{10} + \hat{\gamma}_{11}\mathrm{x})}{\hat{\sigma}_1}\right)\right] * \left[1 - \Phi_{\mathrm{SEV}}\left(\frac{\ln(\mathrm{t}) - (\hat{\gamma}_{20} + \hat{\gamma}_{21}\mathrm{x})}{\hat{\sigma}_2}\right)\right]$$

   $\hat{\gamma}_{\mathrm{ij}}, \hat{\sigma}_{\mathrm{i}}: parameter\ estimations\ based\ on\ multivariate\ random\ normal\ draws$
   $p: percentile\ point\ used\ for\ evaluation$

   The value of *t* for which this relation holds becomes the simulated lifetime for a product for the pth percentile, *k* times a lifetime is generated per test plan.
7. Calculate the standard error of the pth percentile based on the generated percentile point estimates per test plan. The test plan with the configurations that give the smallest total standard error is the best test plan.

   $$SE(t_p) = \sqrt{\frac{\sum_{i=1}^{k}\left(t_{p,i} - \bar{t_p}\right)^2}{k-1}} \qquad (36)$$

   $t_{p,i}: percentile\ point\ for\ ith\ replication$, i=1…k

   $\bar{t_p}: mean\ percentile\ point, \quad \frac{\sum_{i=1}^{k}\left(t_{p,i}\right)}{k}$

Appendix M describes the application of this method to the case that is used as example in this report.

6.2 One Failure Mode, Two or More Stress Factors and One Product Design

If the lifetime of a product is much longer than the available test time, even when a certain stress factor is applied, it can be worthwhile to use more stress factors together. This section describes how the stress level combinations on which the items must be tested have to be chosen in this situation. We assume that the stress factors do not interact. The model can be extended to include interaction with an extra interaction component, but this is out of scope for this research. We describe the methods for a two stress factor situation, but the same idea holds for a k-stress factor situation. Section 6.2.1 describes some different types of test plans and the method to generate these plans, section 6.2.2 compares the different test plan types and describes in which situation which test plan type performs best. Section 6.2.3 compares the sample size needed to obtain certain accuracy for the one stress factor test with the two stress factor test.

6.2.1 Type of Test Plans

3a. Two Factor Optimum Plan

The idea behind this method is the same as that of the statistical optimum plan, plan 1.a of section 5.2.2. The two-factor problem can be transformed to a one-factor problem and solved by the method for the statistically optimum plan; a description of this transformation is given below. The plan that results according to this method is very sensitive to uncertain inputs and not capable of estimating the effect of the individual stress factors, because combined stress values are used (Escobar & Meeker, 1995). In practice this is a very impractical plan, but we use it is as starting point for the improved optimum split plan (plan 3b) and the best compromise plan (plan 3c).

The first step is to create the experimental region; this is a 2-dimensional area for a two-stress factor problem. We describe this method with use of an example based on the case introduced in section 1.4.

Figure 10 shows the experimental region for this example. Stress factor 1 represents power and stress factor 2 represents temperature. We transform both stresses based on respectively the Inverse Power Law and the Arrhenius relationship. After that, we standardize the stresses between 1 and 0 to make parameter estimation of the stress relation easier. The standardized stress becomes $\xi = \frac{x_{high} - x}{x_{high-} x_{use}}$. The bold lines in the picture are the borders of the experimental region and the dotted lines are probability lines. All combinations of the two standardized stress levels located on such a probability line correspond to the same failure probability. The probability lines can be constructed based on the failure probability function dependent on the test time and stress level, equation 34. As described in section 5.2.1 the values of the model parameters can be determined based on the failure probability estimations at the normal use and high level for each stress. The relation for the example becomes:

$$Failure\ probability\ (\boldsymbol{\xi}, 2000) = \Phi_{\mathrm{SEV}}\left(\frac{\ln(2000) - (5.98 + 2.15\xi_1 + 1.70\xi_2)}{0.294}\right)$$

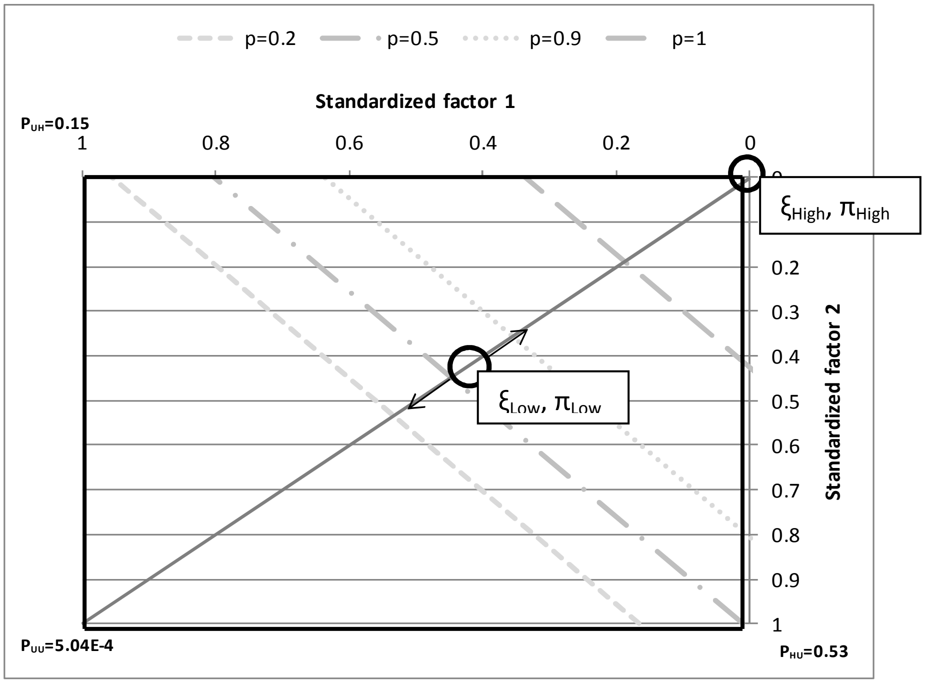


Figure 10: Experimental Region Two Stress Factors

The total experimental region is feasible if all stress combinations can be tested. This means that there are no equipment constraints on certain stress combinations and the failure modes that occur are the failure modes of interest. We assume that in the example the total experimental region is feasible, the constrained plan, plan 3d, focuses on situations where this is not the case.

Escobar and Meeker (1995) describe a method to transform the two stress level problem to a one stress factor problem. Methods to solve the one stress factor problem are available, so this problem can be solved too. The idea of the transformation is that if all the points within the experimental region are feasible, the optimum combined stress levels can be located on the line from the standardized use stresses $(\xi_1 = \xi_2 = 1)$ through the standardized high stresses $(\xi_1 = \xi_2 = 0)$. This makes the standardized stress level for both stresses equal and the problem can be transformed to a one-factor problem; $\mu(x) = \alpha_0 + \alpha_1 \xi_1 + \alpha_2 \xi_2 = \alpha_0 + (\alpha_1 + \alpha_2)\, \xi$. This problem can be optimized via the method described at the statistical optimum plan, plan 1a. ξ_L, the standardized low stress level and π_L, the proportion allocated to the lowest stress level, should be chosen such that the standard error of the p^{th} percentile estimate is minimized.

We use optimization by simulation, as described in chapter 5, with the adjustments described above to generate the stress level and the allocation fractions. If we apply this, the standardized stress level combinations that result from the optimization are (0.449, 0.449) and (0, 0) with an allocation percentage of respectively 70% and 30% . If we transform these values back to the original stress, test combination 1 becomes (96W, 118°C) and test combination 2 becomes (120W, 200°C) Because two test levels are chosen, the test plan is sensitive to wrong assumptions and the individual effects of the stress factors cannot be determined. Test plan 3b and test plan 3c are extensions of this method and take these disadvantages into account.

3b. Two Factor Optimum Split Plan (3 levels)

The optimum split plan has the same asymptotic variance as the two factor optimum plan, plan 3a, but is also possible to estimate the effects of the individual stress factor based on the test results. The disadvantages of this type of plan are that it is not possible to estimate the parameters of an interaction model and it is sensitive to uncertain inputs, like the one factor two-level statistical optimum plan, plan 1a. (Escobar & Meeker, 1995)

The starting point of this method is the stress level combinations obtained from the two-factor optimum plan, plan 3a. Based on the probability of failure of the lowest stress level combination, $(\xi_{Low}) = (\xi_{1,Low}, \xi_{2,Low})$, obtained from this plan, two stress combinations at the borders of the experimental region can be derived, for which the probability of failure remains the same.

$$\Phi_{\text{SEV}}\left(\frac{\ln(t_c)-(\alpha_0+\alpha_1\xi_{1,low}+\alpha_2\xi_{2,low})}{\sigma}\right)=\Phi_{\text{SEV}}\left(\frac{\ln(t_c)-(\alpha_0+\alpha_1\xi_{1,low1}+\alpha_2*0)}{\sigma}\right)=\Phi_{\text{SEV}}\left(\frac{\ln(t_c)-(\alpha_0+\alpha_1\xi_{1,low1}+\alpha_2*1)}{\sigma}\right)$$

The new stress level combinations are $(\xi_{1,Low1}, \xi_{2,Low1})$ and $(\xi_{1,Low2}, \xi_{2,Low2})$. Figure 11 shows how we split the low stress level of test plan 3a into two new stress level combinations on the border of the experimental region. The proportion of units allocated to the lowest level, π_{Low}, of the two factor optimum plan, 3a, is also split over the two new combinations of levels, such that $\pi_{Low1}\xi_{1,Low1} + \pi_{Low2}\xi_{1,Low2} = \pi_{Low}\xi_{1Low}$ and $\pi_{Low1}\xi_{2,Low1} + \pi_{Low2}\xi_{2,Low2} = \pi_{Low}\xi_{2,Low}$ to maintain the optimality. Detailed information about the maintained optimality is described by Escobar and Meeker (1995).

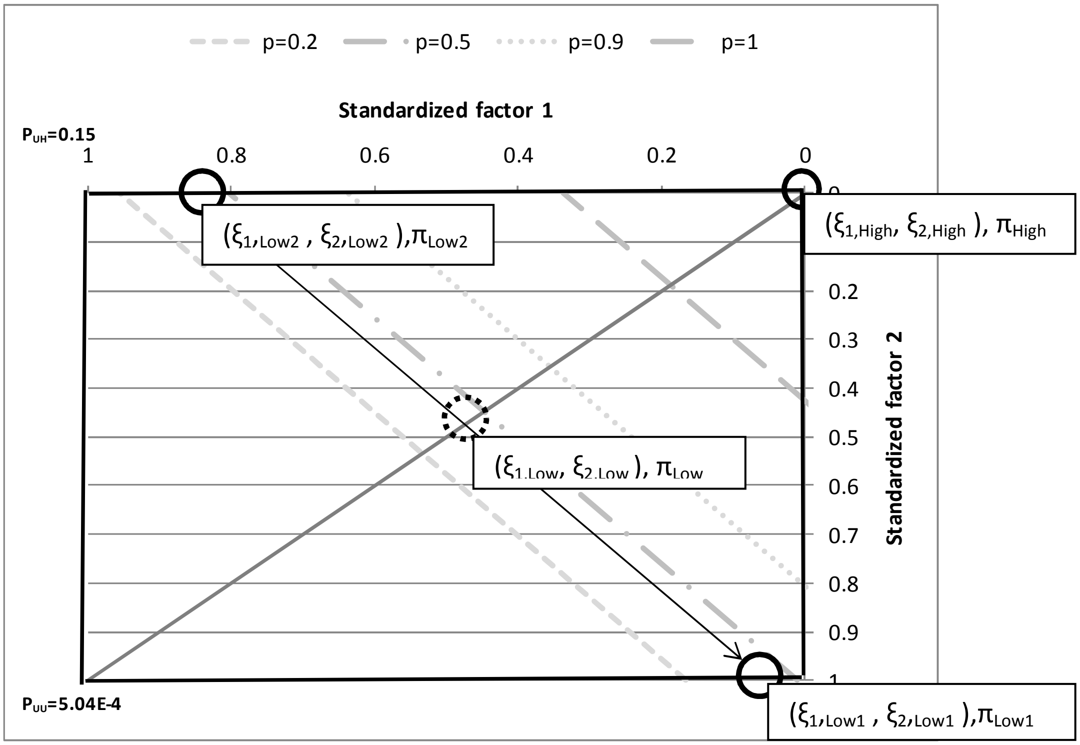


Figure 11: Two Factor Optimum Split Plan

If we apply this method to the example as described at the two factor optimum plan, we have to split the lowest stress level (96W, 118°C) or (0.449, 0.449) into two new stress combinations on the borders of the experimental regions for which the probability of failure remains the same, 49.5%. Stress 2 is the stress with the smallest effect on the probability of failure if it increases from use level to the highest level, so for the first new low stress combination we set stress factor 2 at the normal use level and for the second new stress combination we set stress factor 2 at the highest level. Based on these values, we choose values for stress factor 1 such that the probability of failure remains 49.5%.

Test combination low 1:

$Failure\ probability\ \left(\xi_{1,Low1}, 1; 2000\right) = 0.495 = \Phi_{\mathrm{SEV}}\left(\frac{\ln(2000)-(5.98+2.15\xi_1+1.70*1)}{0.294}\right) \rightarrow \xi_1 = 0.015$

Test combination low 2:

$Failure\ probability\ \left(\xi_{1,Low2}, 0; 2000\right) = 0.495 = \Phi_{\mathrm{SEV}}\left(\frac{\ln(2000)-(5.98+2.15\xi_1+1.70*0)}{0.294}\right) \rightarrow \xi_1 = 0.83$

If we transform these standardized stresses back to the original stresses, the stress combinations on the borders of the experimental region that give the same failure probability are (119W, 50°C) and (80W, 200°C). Sometimes it can be the case that the failure probability cannot be reached without increasing the use stress level of the second stress factor. If this is the case, we set the first stress factor at the highest stress level and choose a value for the second stress level such that the failure probability is reached. Instead of the lower border of the experimental region, the right border is used in this situation.

We split the allocation of 70% of the items to the lowest stress level into two parts based on the formulas described before:

$\pi_{Low1} * 0.015 + \pi_{Low2} * 0.83 = 0.7 * 0.449$ and $\pi_{Low1} * 1 + \pi_{Low2} * 0 = 0.7 * 0.449$

Based on these equations, we allocate 31% of the items to test combination low 1 and 39% of the items to test combination low 2. The remaining fraction, 30% is allocated to the high stress level combination, (120W, 200°C).

3c. Two Factor Compromise Split Plan (5 levels)

The compromise split plan can be developed using the same method as the optimum split plan, but is based on the one factor best compromise plan, plan 1b, instead of the one factor statistical optimum plan, plan 1a. Besides the low stress level, also the middle stress level is split according to the same method as described at the two factor optimum split plan, plan 3.b. Figure 12 shows how the best compromise plan is split into test values for the two factor compromise split plan. As example the case is used.

This compromise split plan increases the standard error of the p$^{\text{th}}$ percentile estimate, but is able to estimate the parameters more accurately and can also estimate the interaction (Escobar & Meeker, 1995). It is also possible to generate a five level, two-factor test plan based on the 4:2:1 allocation method, plan 1c., and use these stress level combinations to split the test plan. The construction of these 5 level split plans are not illustrated with use of the case, because they can be constructed in the same way as the two factor optimum split plan, only with one more stress combination that has to be split.

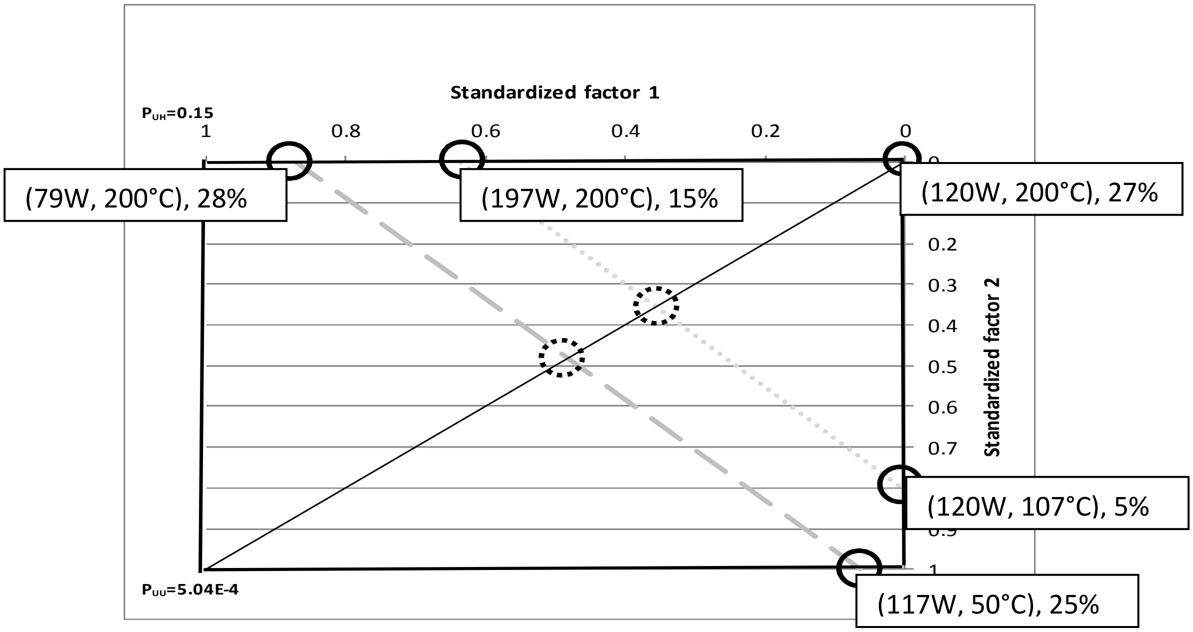


Figure 12: Compromise Split Plan

For the example, the one factor best compromise plan, plan 1b, becomes:

	(Stress1 (ξ), Stress2 (ξ))	Fraction	Prob. failure
Low 1	(0.47, 0.47)	53%	40%
Middle 1	(0.24, 0.24)	20%	99.9%
High	(0, 0)	27%	100%

We use this plan to generate the two factor best compromise split plan. We split the levels and the allocation fractions; Figure 12 shows the resulting test plan graphically. The difference in probability of failure between the middle and the highest stress level is small, 99.9% vs. 100%. This happens because in a two stress factor scenario, a 100% probability of failure line is reached much earlier than in the one stress factor scenario. So there can be a lot of stress combinations that have a failure probability of 100%. The middle stress level is always located at the normalized average of the low and the high stress level, and the parameters of the life-stress relation can be estimated based on the failure times on the three stress levels. The optimal percentage of failure is chosen during the optimization of SE(t_p).

3d. Constrained Split Plan

The plans described above can only be used if all stress level combinations are feasible. In practice it can be the case that certain stress combinations lead to other failure modes due to the high stress, or certain combinations are not possible due to practical restrictions, such as equipment constraints, for example it is not possible to use the equipment under very high temperature with a high voltage or there are environmental constraints. To incorporate this, we developed a new type of test plan. We use the concept of the stress level combinations on a straight line from use conditions for both factors, but instead of the highest stress levels for both factors, the reference point of the line becomes the maximum feasible stress combination. Based on the slope of this line and the intersection with the border of the experimental region, we construct a new model. Figure 13 shows the constrained experimental region and the line we use for test plan generation.

On the straight line through (1, 1) and the maximum stress combination holds: $\xi_2 = a\xi_1 + b$, with $a = slope\ of\ the\ line$ and $b = intersection\ of\ the\ line\ with\ the\ right\ experimental\ border$.
This leads to the expression: $\mu(x) = \alpha_0 + \alpha_1\xi_1 + \alpha_2(slope * \xi_1 + intersection)$

Based on these relations, the constrained model becomes:

$$\ln(tp) = (\alpha_0 + intersection * \alpha_2) + (\alpha_1 + slope * \alpha_2)\xi_1 + \Phi_{SEV}^{-1}(p) * \sigma \qquad (37)$$

With $intersection = 1 - slope \quad and \quad slope = \frac{1-\xi_{2,max}}{1-\xi_{1,max}}$

$\xi_{i,max}$: $Maximum\ standardized\ level\ of\ stress\ i\ in\ combination\ with\ the\ other\ stresses$

We use this model in the same way as the model to develop the plans for a nonconstrained situation as described before. We illustrate this with use of an example based on the case.
Suppose that the maximum feasible stress combination is 109W and 142°C, this means a normalized stress combination (0.2, 0.3).
The highest test stress combination changes from (120W, 200°C) to (109W, 142°C). To make it possible to estimate the effect of the individual stresses, this high stress level must also be split into two separate stress level combinations on the border of the experimental region. The lowest stress level combinations have to be chosen such that SE(t_p) is minimized. We generate the stress level values with use of the adapted model based on the optimum split plan (plan 3b) or the compromise split plan (plan 3c).

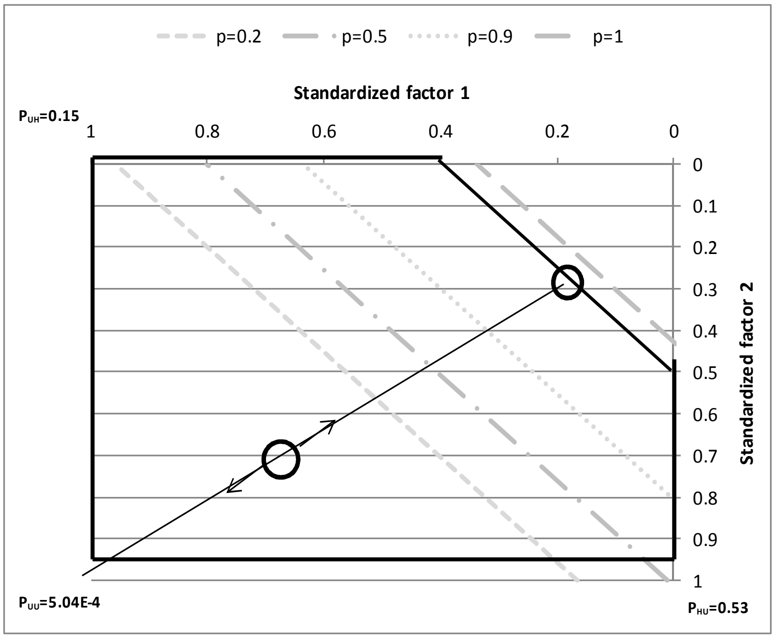


Figure 13: Experimental Region Including Stress Restriction

Based on the slope and intersection of the stress combination line, we adjust the parameters of the standard model, and use these values in the simulation. The test plan settings for the constrained optimum split plan become:

	Stress1 (W)	Stress2 (C)	Fraction	Prob. failure
Low 1	116	50	31.4%	36.78%
Low 2	78	200	38.6%	36.78%
High1	120	103	13.5%	99.99%
High 2	96	200	16.5%	99.99%

6.2.2 Comparison of Test Plan Types

To describe the applicability of the test plan types in different situations, we use a comparison study as we used for the one stress factor scenario in section 5.4. We add one new situation to examine the differences in performance if there is some interaction between the stress factors. For some applications it can be important that the test plan can deal with interaction.

We do not compare all the types of test plans described for this scenario. Plan 3a is not taken into account because it gives no information about the influence of the individual stress factors and the results cannot be used in practice, it describes only the basics of the other plans. Also plan 3d is not taken into account because it is generally the same as the other plans, but appropriate for situations where constraints about the stress level values are given. The test plans for this two stress factor scenario are based on the plans for the one stress factor scenario. The 4:2:1 allocation plan, plan 1c., is not used as basic plan during the comparison, because this type of plan has the worst performance, but the advantage is the fast optimization. Therefore we only include it in the schematically categorization, Figure 14. Based on the comparison of the standard error for the 10th percentile estimate for all situations, as displayed in

---

Table 6, we can conclude which type of test plan is most appropriate for which situation. Remarkable is that the Optimum Split Plan can deal better with model deviations and wrong parameter assumptions as the Statistical Optimum Plan of scenario 1 can. For this two stress factor scenario, the optimum split plan performs best in all situations, except the situation in which there is some interaction. Figure 14 summarizes this categorization schematically.

Table 6: SE(t_{10}) per Situation and Type of Test Plan for the Two Factor Plans

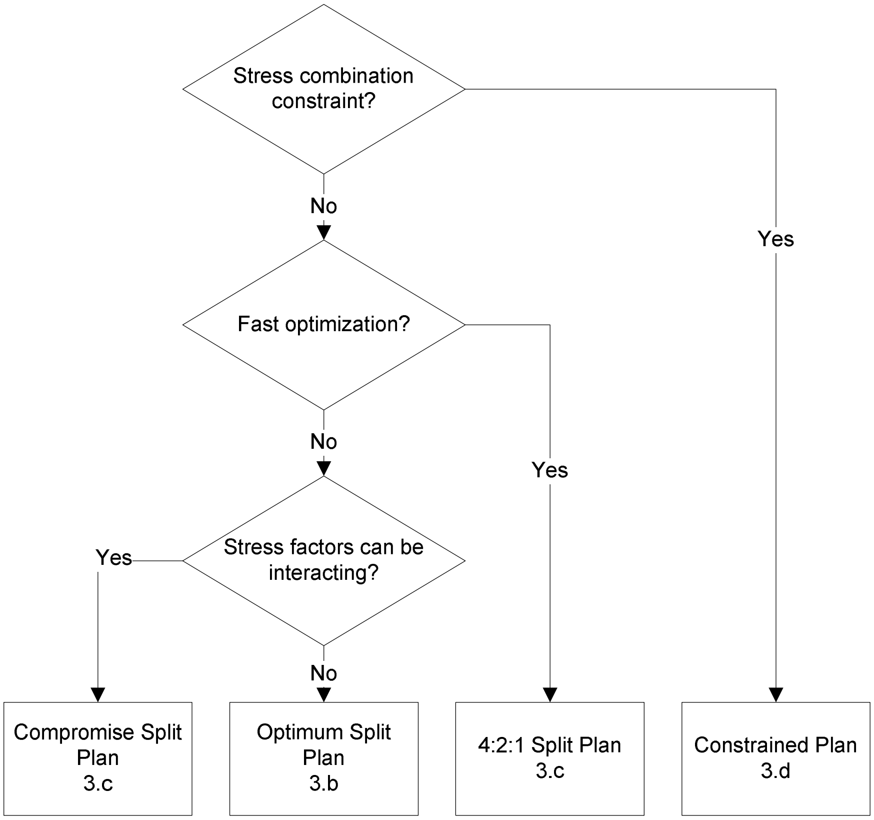


Figure 14: Most Appropriate Test Plan Types per Situation - One Failure Mode, Two Stress Factors and One Product Design

6.2.3 Comparison Performance One Stress Factor Test vs. Two Stress Factors Test

We introduced methods to design test plans for the two stress factor scenario. These methods increase the complexity of the design, but also improve the accuracy, so a smaller amount of products can be used to reach the same estimation accuracy. For the example we use in this research, we compare the needed sample size for both kind of tests. The available test time for both situations is 3,000 hours.

In the first situation, we design a test plan based on the best compromise plan, plan 1b. In the second situation, we also design a test plan based on the best compromise plan but we use two stress factors, namely power and temperature. So in fact we use the two factor best compromise split plan, plan 3c. Table 7 describes the test plan settings for both situations, the standard error of the p^{th} percentile estimate for a sample size of 2,000 items and the sample size that is needed to reach a target precision of 1.5. Based on these results we conclude that a more complex lifetime test considerably reduces the estimation uncertainty, the needed sample size decreases from 180 units to 34 units in the example. While it seems to be efficient to always use two stress factors, this only increases the accuracy if both stress factors really accelerate the failure process and do not interact. The risk of interaction increases if more stress factors are used and when this is not taken into account during the design, the true estimation accuracy is much larger than the expected accuracy.

Situation	Test Configurations					Test Results		
	Low 1	Low 2	Middle 1	Middle 2	High	t_p	SE(t_p)	SS
One stress factor	98W 53%	- -	108W 20%	- -	120W 27%	9564	594	**180**
Two stress factors	117W, 50°C 25%	79W,200°C 28%	120W,107°C 5%	97W,200°C 15%	120W,200°C 27%	9605	260	**34**

Table 7: Comparison One Stress Factor Test vs. Two Stress Factors Test

6.3 One Failure Mode, One Stress Factor, Two or More Product Design Variants

This section focuses on accelerated lifetime tests for products for which one or more product design variants are available. As described in section 3.4 the design variants differ on some specific points and the remaining properties are the same. In the scientific literature, no methods to develop test plans for this scenario under stress conditions are available.

We represent the differences in design with use of design parameters; this can be discrete or continuous variables. For the design of a test plan, the continuous variables have the advantage that they have much in common with a stress factor, so methods for the two stress factors scenario can be adapted to this situation. On the other hand, the discrete design parameters are easier to handle, the test levels are fixed (yes/high(1) and no/low(0)), and only the allocation of the items over the test levels must be optimized. We focus in this section on products with one design parameter, but the methods we use can also be applied to situations with more design parameters. We split the two product design variants scenario in four subscenarios based on the type of design parameter, discrete or continuous and the expected difference in failure probability between the value of the design parameter, equal or unequal. With equal we mean that we assume an equal effect or beforehand it is unknown if there is a difference in lifetime between the variants and it is unknown which design variant is the best. With unequal we mean that we assume that the design parameter influences the reliability and we have an idea about the size of the influence. Figure 15 shows this classification. Per subscenario the decision variables are indicated. The remaining part of this section uses this classification; we describe per subsection a method to design a test plan. The expression for percentile points depending on a stress factor and a design parameter based on the generalized log linear relationship, see equation 4 and equation 18, equals:

$$t_p(\boldsymbol{x}) = exp(\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \Phi_{SEV}^{-1}(p) * \sigma) \qquad (38)$$

x_1: $transformed\ value\ of\ stress\ factor$

x_2: $design\ parameter, can\ be\ distreet\ \ 0/1\ \ or\ continuous$

		Discrete	Continuous
Failure probabilities	Unequal	* Test level 0 and 1 * Allocation of items is decision variable	* Test levels are decision variables * Allocation of items is decision variable
	Equal	* Test level 0 and 1 * Equal allocation of items	* Test levels are decision variables * Equal allocation of items

Design parameter

Figure 15: Decision Variables Design Parameter Scenario per Subscenario

6.3.1 Type of test plan

4a. Discrete Design Parameter- Equal Failure Probabilities

If the influence of the design parameter is unknown and the design parameter is discrete, we allocate an equal fraction of the items to the specific values of the design parameter, zero and one. The test levels of the stress factor and the allocation of the items over the stress levels can be determined based on the methods for the one failure mode, one stress factor, one product design scenario as described in Chapter 5. Based on the stress levels and the allocation for this scenario, we split the allocation fraction per stress level into two equal parts, one for each design parameter value. The sample size needed to reach certain accuracy must be based on the standard error of the best test plan for both product designs, and not the standard error caused by one of the product designs.

To determine the efficiency of this combined stress and design test, we compare a situation for which two separate tests are done, one per product design variant, and the combined test. We measure the efficiency in terms of needed sample size to reach a precision factor of 1.5. Based on the case we use in this thesis, we show the results of this comparison.

Two different lamp designs are tested; the only difference is the color of the light they produce. We use a design variable to represent this color difference. The value of this variable is 1 if the color is white and 0 if the color is yellowish white. We compare two different situations. In the first situation the two tests are execute separately, and in the second situation the tests are combined. For both situations we measure the standard error of the 10th percentile estimate and the sample size necessary to estimate the 10th percentile with a precision factor of 1.5. We generate a two level statistical optimum plan. This means for situation 1 that per design variant the product is tested at two different stress levels, and for situation 2 that four different stress combinations are used.

Table 8 shows the test settings and the needed sample size for both situations to estimate the 10th percentile point. If the design test is combined with the stress test, the sample size can be reduced with 50%, because the parameters of the lifetime model for both designs are the same. If in reality there is a difference between the two designs, the sample size reduction is somewhat smaller. Suppose that in reality the failure probability for the white design variant at use stress equals 0.3% instead of 0.2% as for the yellowish white design variant, this difference is not known during the design of the test, only during the performance comparison. Table 9 summarizes the test settings and the test results per situation.

Based on this, we conclude that a combined lifetime-design test is more efficient than two separate tests when the failure probabilities are beforehand assumed to be equal. In this example the sample size can be reduced with 50% by combining the design test with the stress test if the failure probabilities are equal and with 45% if there is a small difference. The reason for this sample size reduction is that a large part of the lifetime model is the same for both designs, see equation 38.

Situation		**Test Configurations**				**Test Results**		
		Low 1	Low 2	High 1	High 2	t_p	SE(t_p)	Sample Size
One-factor tests	Test 1	98W, 1 70%	- -	120W, 1 30%		X	564	162
	Test 2	98W, 0 70%	- -	120W, 0 30%		X	565	163
	Total							**325**
Combined test		98W, 1 35%	98W,0 35%	120W,1 15%	120W,0 15%	X	0: 557 1: 560	**158**

Table 8: Differences in Sample Size One-Factor Test vs. Combined Stress-Design Test with Equal Failure Probabilities

Situation		**Test Configurations**				**Test Results**		
		Low 1	Low 2	High 1	High 2	t_p	SE(t_p)	Sample Size
One-factor tests	Test 1	98W, 1 70%	- -	120W, 1 30%		x	425	115
	Test 2	98W, 0 70%	- -	120W, 0 30%		x	565	163
	Total							**278**
Combined test		98W, 1 35%	98W,0 35%	120W,1 15%	120W,0 15%	x	0: 557 1: 476	**152**

Table 9: Differences in Sample Size One-Factor Test vs. Combined Stress-Design Test with Unequal Failure Probabilities

4b. Discrete Design Parameter- Unequal Failure Probabilities

Due to the assumed unequal failure probabilities for the different product designs, a more accurate estimation can be obtained if the allocation of the items over the design variants is not equal. This leads to an extra decision variable in the optimization problem. Not only the allocation of the items over the stress factors is a decision variable, but also the allocation of the items over the design variants is. Similarities with the two stress factor scenario, section 6.2, can be found. The only difference is that only the upper and the lower border of the experimental region can be used if we transform the design parameter to a stress factor, because it can only take a value of 0 or 1, and not a value in between. We use the method for the two stress factors scenario with this extra restriction. We assume that the expected probability of failure at use stress, with the design parameter equal to 0 to be 0.2% and equal to 1 to be 0.8%. Table 10 shows the difference in the sample size needed to obtain a target precision factor of 1.5 for the standard error of the 10th percentile estimate. Comparing these results with the results of the section above, equal failure probabilities, we conclude that the efficiency in sample size increases if the difference in performance between the two designs increases. The sample size needed for the combined test in the example is 66% lower than that of the one-factor experiment.

Situation		Test Configurations			Test Results		
		Low 1	Low 2	High	T_{10}	SE(t_{10})	Sample Size
One-factor Tests	Test 1	96W,1 70%	- -	120W,1 30%	x	210	50
	Test 2	98W, 0 70%	- -	120W, 0 30%	x	565	163
	Total						**213**
Combined test		105W, 0 27%	95W,1 43%	120W,1 30%	X	391 257	**76**

Table 10: Differences in Sample Size One-Factor Test vs. Combined Stress-Design Test with Unequal Failure Probabilities

4c. Continuous Design Parameter- Equal Failure Probabilities

In this subscenario we divide the items equal over the design variants per stress level, because of the equal failure probabilities. We use this equal allocation per variant, because no differences in the amount of failures are expected, but we want to test if this assumption holds. The values of the design parameter are not fixed, because they can take each value of a finite interval. But because we expect the failure probability to be the same for all design parameter values, we only use the lowest and the highest stress value of this interval. We solve the optimization problem as the one failure mode, one stress factor, and one product design variant problem as described in Chapter 5.

4d. Continuous Design Parameter- Unequal Failure Probabilities

This scenario can be transformed directly to the two stress factors problem. We transform the continuous design parameter to a stress factor, with a linear life-stress relationship. Based on the assumed failure probabilities under normal use condition (value of design parameter you are interested in) and high stress (highest value design parameter) we obtain the test combinations. Dependent on the situation, a selection of one of the different test plan types appropriate to this scenario can be selected based on Figure 14 and the method described in section 6.2 can be followed.

6.4 **Combined Scenarios**

So far we described different methods to design test plans for different scenarios and build a simulation model to execute these methods and to generate test plan configurations. This section describes how the methods for the scenarios described before can be combined to be able to generate test plans for the more complex scenarios. We describe which part of which methods can be combined, but do not describe the generation of the test plans in detail, because this causes a lot of repetition. When a part of a method deviates from the methods described so far, we describe this new part in more detail.

6.4.1 Two or More Failure Modes, Two or More Stress Factors, One Product Design

A test plan for this scenario can be developed based on a combination of the method to design a test plan for the two failure mode, one stress factor scenario and the one failure mode, two stress factors scenario. An assumption about the lifetime distribution and the parameters must be made for the two failure modes individual. Based on this, the one failure mode, two stress factor problems can be solved for each of the failure modes. This means that the output of the simulation study, the variance-covariance matrix, must be used to determine the variance, or standard error based on simulation for the combined failure modes, as described at the two failure mode scenario. Instead of three parameters for which values must be simulated based on multivariate normal draws, a fourth parameter is included for the second stress factor. Based on the estimates and the variances of the parameters per failure mode, lifetime data can be generated for the failure modes together and the standard error of the total product can be estimated based on these data. The interpretation of the results is the same as for the two failure modes, one stress factor scenario.

6.4.2 Two or More Failure Modes, One Stress Factor, Two or More Product Design Variants

The generation of a test plan for this situation can be done in the same way as for the scenario described before, the two or more failure modes, two or more stress factors, one product design scenario. Only the transformation of a design factor to a stress factor must be taken into account. So instead of using the one failure mode, two or more stress factors, one product design variant scenario as basis scenario per failure mode, the method for the one failure mode, one stress factor, two or more product design variants must be used. After that the same methods as described at section 6.4.1 can be applied.

6.4.3 One Failure Mode, Two or More Stress Factors, Two or More Product Design Variants

We represent product design variants by design parameters. These design parameters can be handled in the same way as stress parameters; section 6.3 described this in more detail. In the earlier sections we described methods to handle with two stress factors via a two-dimensional experimental region, the idea of this method can also be used for this scenario. But now we have at least three stress factors, two for the stresses and one for the design parameter. This means that we have to work with three-dimensional experimental areas or larger. The complexity of finding the best test combinations increases. Instead of equal probability lines, we have to deal with equal probability areas in a three-dimensional case. The test level combinations are based on the intersections of this probability area with the borders of the experimental region. The probability for which these intersections must be searched can be obtained from the combined stress point as described in the one failure mode, two stress factors, one product design scenario, section 6.2. It is possible that there are more intersections with the borders of the experimental region as there are stress factors. Not all these intersections have to be used. We choose the test combinations such that all individual parameters can be estimated. Figure 16 gives an example of a three-dimensional area with test level combinations based on the optimum split plan.

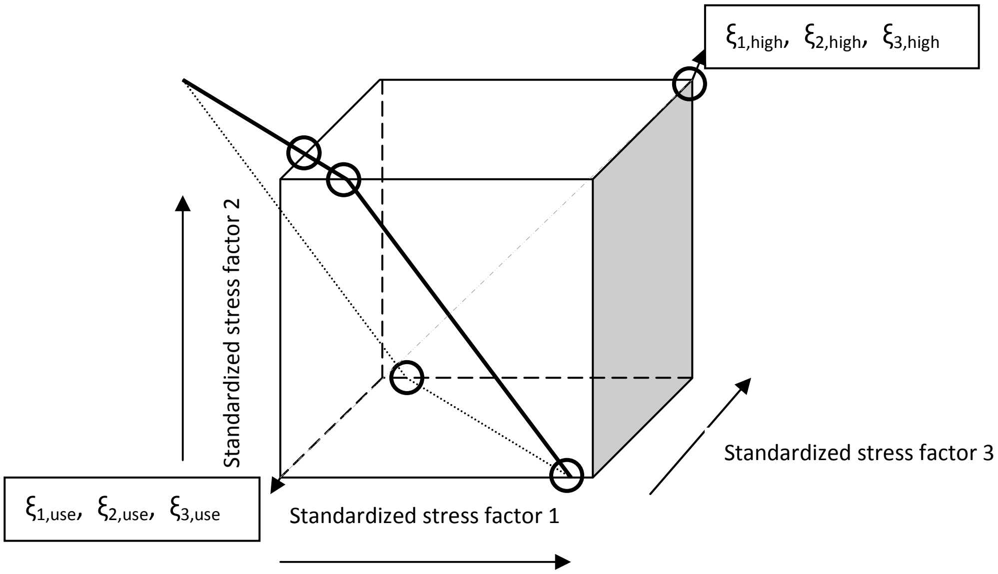


Figure 16: Three-Dimensional Experimental Area with Test Level Combinations with Equal Failure Probabilities on a Surface

6.4.4 Two or More Failure Modes, Two or More Stress Factors, Two or More Product Design Variants

This scenario is the most complex scenario. A general equation for the failure probability based on this situation equals:

$$F(t,\boldsymbol{\xi}) = 1 - \left[1 - \Phi_{SEV}\left(\frac{\ln(t) - \alpha_{1,0} + \sum_{i=1}^{n} \alpha_{1,i}\xi_i}{\sigma_1}\right)\right] * \ldots * \left[1 - \Phi_{SEV}\left(\frac{\ln(t) - \alpha_{M,0} + \sum_{i=1}^{n} \alpha_{M,i}\xi_i}{\sigma_M}\right)\right] \qquad (39)$$

$\boldsymbol{\xi}: vector\ of\ normalized\ (transformed) stress\ factors\ and\ design\ parameters$

$\alpha_{k,i}: model\ parameter\ corresponding\ to\ failure\ mode\ k\ and\ stress\ \ factor/design\ parameter\ i$

$k = 1 \ldots M, i = 1 \ldots n$

We use a combination of the methods described before to develop a method to design a test plan for this scenario. This method is almost the same as the method for the two failure modes, two or more stress factors, and one product design scenario, section 6.4.1. The only difference is the number of stress factors/design parameters. The values for these factors must be chosen based on the method described at the one failure mode, two stress factors, two product design variants scenario, section 6.4.3, so with use of a k-dimensional experimental region.

6.5 Conclusion

Based on the descriptions of the methods to design a test plan per scenario, we can derive a procedure to design test plans appropriate to all the scenarios. We do not describe the individual steps in detail, because they are all based on the methods described before, but we describe the steps that must be taken in order to generate a good test plan.

1. Obtain prior information per failure mode and initialize the lifetime model.
2. Choose type of test plan based on Figure 9 (one stress factor/design) or Figure 14 (multiple stress factors/design variants).
3. Simulate lifetime data and estimate the model parameters and their variances.
4. If more than one failure mode occurs, simulate the standard error of the p^{th} percentile estimate based on multivariate normal draws for the model parameters.
5. Select the best test plan configurations based on the standard error of the p^{th} percentile estimate.
6. If more than one stress factor is used or more than one product design variants are tested, split the combined values to feasible values on the borders of the experimental region.
7. Test the robustness of the best test plan configurations by changing some prior information values and choose the most robust one.

The methods we use to generate test plans are all based on the methods of the one failure mode, one stress factor and one product design scenario. The two stress factor problem can be transformed to this problem by merging the two stress factors to one stress factor, search the best value for this combined stress factor and split this value afterwards into individual stress levels on the borders of the experimental region. Using two independent stress factors instead of one stress factor increase the accuracy enormously. The two product design variants scenario can also be transformed to this one failure mode, one stress factor, one product design scenario in the same way. It is important to include an extra restriction if the design parameter is a discrete variable, because not all borders of the experimental region can be used in this case. A combined stress-design tests reduces the sample size needed to obtain certain accuracy with approximately 50%. The two failure modes scenario cannot transformed directly to the one failure mode, one stress factor, one product design scenario. The two failure modes must be treated separately and with use of an extra simulation step, p^{th} percentile estimates for the total product can be obtained via the estimates and variances of the individual model parameters. Based on the standard error of the combined failure processes, the best test settings can be selected. We implemented the methods described till and until section 6.3, the remaining test plans can be generated based on these methods. Chapter 0 applies the general procedure to a practical case, which includes design variables and stress factors.

7. Case: Test Plan Design to Determine the Reliability of an Aquarium Lamp

This chapter describes a case to illustrate how the developed procedure can be applied during a practical situation. Company X develops a new type of UV-lamp to use in aquaria for water purification. The high intensity light kills the free-floating micro organisms and cleans the water. The aquarium lamp has to be replaced if the lumen output drops below 70%, this intensity is too low to perform the intended function. The low lumen output, an output below 70%, causes the lamp to fail, so can be seen as failure mode. The normal use level of the lamp is 10°C, but to accelerate the failure process, the temperature can be increased until 80°C (Schuld, 2010).

The company developed different design variants of the lamp and executed a test to measure the effect of the different designs on the product reliability.

The different product design variants can be described with use of four different discrete design parameters, namely:

$$d_1\!: Emitter = \begin{cases} 1 \ if \ type \ A \ emitter \ is \ used \\ 0 \ if \ type \ B \ emitter \ is \ used \end{cases}$$

$$d_2\!: Filling \ pressure = \begin{cases} 1 \ if \ pressure \ is \ high \\ 0 \ if \ pressure \ is \ low \end{cases}$$

$$d_3\!: Neon = \begin{cases} 1 \ if \ amount \ of \ neon \ is \ high \\ 0 \ if \ amount \ of \ neon \ is \ low \end{cases}$$

$$d_4\!: Type \ of \ spiral = \begin{cases} 1 \ if \ type \ A \ spiral \ is \ used \\ 0 \ if \ type \ B \ spiral \ is \ used \end{cases}$$

Design parameter 2, filling pressure, and design parameter 3, neon, can be seen as continuous parameters. The engineers who develop the UV-lamp want to compare two specific levels of these parameters in the test, thus we describe them as discrete variables.

Section 7.1 describes the design of a test plan to test the reliability and to test the influence of the design parameters on the lifetime of this new type of lamp according to the steps and methods described in Chapter 6. Section 7.2 describes the analyses of the test results that can be obtained from the developed test plan. Because there is not enough time to wait on the real test results, we derived the lifetimes used in this section from the two earlier executed experiments; the lifetimes of the 80 items (different designs) under high stress and the stress test for one design variant. We derive an acceleration factor from the latter test and based on this factor the lifetimes of the other test, design variants under high stress, are extrapolated to lifetimes that would be observed at the stress factors that must be used according to the test plan. Section 7.3 compares the performance of the proposed test plan with the worst test plan evaluated during the simulation.

7.1 **Design of a Test Plan: UV-lamp**

We design a test plan for the UV-lamp according to the procedure of Chapter 6. We have to deal with a one failure mode, one stress factor, two or more product design variants scenario with discrete design parameters and equal failure probability as described in section 6.4.3.

1. Obtain prior information per failure mode and initialize the lifetime model.
   We obtained the prior information from the engineers; **Fout! Verwijzingsbron niet gevonden.** summarizes this information.

   The lifetime model we want to obtain after the lifetime test including all design parameters and the interactions equals:
   $\ln(t_p) = \alpha_0 + \alpha_1\xi + \alpha_2 d_1 + \alpha_3 d_2 + \alpha_4 d_3 + \alpha_5 d_4 + \alpha_6 d_1 d_3 + \alpha_7 d_2 d_3 + \Phi_{SEV}^{-1}(p) * \sigma$
   $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7 : model\ parameters\ to\ be\ estimated$
   $\xi : normalized\ transformed\ stress\ level$
   $d_i : value\ of\ design\ parameter\ i, 0\ or\ 1 \qquad i = 1..4$

   The goal of the lifetime test is to measure the effect of the design parameters on the lifetime of the lamp, estimations about these effects are difficult to obtain. Therefore we assume that the different product design variants have equal failure probabilities and leave model parameters $\alpha_2 - \alpha_7$ out our model during the generation of the test plan configurations. We only have to estimate the model parameters $\alpha_0\ and\ \alpha_1$.

   The resulting initial lifetime model becomes:
   $\ln(t_p) = 5.991 + 3.912\xi + \Phi_{SEV}^{-1}(p) * 0.4$

2. Choose type of test plan based on Figure 9, section 5.4.
   We obtained the prior information from engineers based on experiences of similar products, but they are not certain about some of the model assumptions. As described before we assume Weibull distributed lifetimes, but based on the results of some other type of UV-lamps, it can be the case that the Normal distribution fit best. We also assume during the design of the experiment that the design parameters do not influence the lifetime, but in practice some of the parameters will have a significant effect on the lifetimes. For these reasons, we want to design a robust test plan. Based on the guidelines of Figure 9 we choose to use the best compromise plan as basic test plan.

3. Simulate lifetime data and estimate the model parameters and their variances.
   With use of the simulation program and the lifetime model established in step 1, we generate lifetime data for different test plan configurations.

Equipment to increase the temperature is available, but only temperatures with differences of 5°C can be generated (Schuld, 2010). According to the rules described in section 5.2.3, the upper bound of the lowest stress level must be chosen such that the probability of failure is min(50%*2,100%)=100%. So the upper bound of the lowest stress level is the lowest temperature for which the failure probability at the end of the test reaches 100%. The lower bound of the lowest stress level must be chosen such that the probability of failure is 50%. Based on these bounds, the lowest stress level interval becomes [30°C, 45°C], and based on the equipment constraint we use a step size of 5°C. Based on the generation rules of section 5.2.3, the allocation fraction to the lowest stress level varies from 0.33 to 0.63, and we choose to use a step size of 0.1. This results in 16 different test plan configurations to evaluate.

4. Select the best test plan configurations based on the standard error of the p^{th} percentile estimate.
   Table 11 presents the best test configurations, the configurations that perform almost equal and the configurations that perform worst. The performance of test plan 7 and 8 are evaluated during the robustness test of step 6, to choose the final test plan. Test plan 13 is included to compare the performance of the worst and the best test plan.

Test plan	Low stress value	Middle stress value	High stress value	Fraction low	Fraction middle	Fraction high	t_p	$SE(t_p)$
7	35	55	80	0.53	0.2	0.27	x	492
8	35	55	80	0.63	0.2	0.17	x	493
13	45	60	80	0.33	0.2	0.47	X	664

Table 11: Test Plan Configurations Best Compromise Plan

5. Split the combined values to feasible values on the borders of the experimental region.
   Thus far we focused on the stress factor; the design variants are not included yet. We choose to use this order because of the discrete design parameters with equal failure probabilities. Section 6.3.1 described that in this situation, both variants of a design parameter must be tested on each stress level. But because we assume that there can be some interaction between the design parameters, we choose the design parameter combination such that after the lifetime test, significant interactions can be measured and included in the lifetime model. For these combinations, we use the Design of Experiments (DoE) theory. Because the engineers are interested in the interactions between d1 and d3 and d2 and d3 and no other interactions, we can use a half factorial design. This means that with 4 design factors, $½ * 2^4 = 8$ different test runs exist.
   Table 12 displays the settings per run for an 8 run DoE. We want to test an equal number of items per design variant, because we do not know if there are any differences. So, we divide the total sample size equal over these 8 test runs, this means 10 items per test run. To make the estimation for both the stress factor as the design parameters as good estimation as possible, we execute every test run at least one time at every stress level. Table 15 summarizes the resulting test plan including these design parameters and the division of the items over the different stress levels.

	d1	d2	d3	d4
1	0	0	0	0
2	1	0	0	1
3	0	1	0	1
4	1	1	0	0
5	0	0	1	1
6	1	0	1	0
7	0	1	1	0
8	1	1	1	1

0: Low value for design parameter
1: High value for design parameter

Table 12: Four Factor 8 Runs DoE (Montgomery, 2005)

6. Test the robustness of the best test plan configurations by changing some prior information values and choose the most robust one.
   There is some uncertainty about some of the prior information assumptions. Therefore, we compare the performance of the best test plans of step 4 if we change these prior information values. During the first test the Weibull distribution is replaced by the Normal distribution and during the second test the acceleration factor is increased from 50 to 60. Table 18 and Table 19 show the resulting SE(t_{50}) per test plan per robustness test.

Plan	t_{50}	SE(t_{50})	Median
7	x	213	212
8	x	223	223

Table 13: Test 1- Model Departure

Plan	t_{50}	SE(t_{50})	Median
7	x	471	470
8	x	469	466

Table 14: Test 2- Acceleration Increase

Based on the robustness test results we conclude that both test plans perform well. We choose to use test plan 7 as the best test plan, because it scores best on the model departure test and scores almost equal to plan 8 on the acceleration test. Table 15 describes the resulting test plan, including the design parameters. A lifetime test is executed based on the suggested test configurations. The next section describes the test results, the effect of the design parameters and the estimation of the lifetime of the UV-lamp.

Stress level	Design	Number of replications
35	0000	6
35	1001	6
35	0101	5
35	1100	5
35	0011	5
35	1010	5
35	0110	5
35	1111	5
Total		42

Stress level	Design	Number of replications
55	0000	2
55	1001	2
55	0101	2
55	1100	2
55	0011	2
55	1010	2
55	0110	2
55	1111	2
		16

Stress level	Design	Number of replications
80	0000	2
80	1001	2
80	0101	3
80	1100	3
80	0011	3
80	1010	3
80	0110	3
80	1111	3
		22

Table 15: Final Test Plan UV-case

7.2 Results Designed Test Plan UV-lamp

There is not enough time to execute the proposed test plan, but based on the lifetime data of the two test that are executed earlier for this type of lamp, we generate lifetimes for the 80 UV-lamps tested based on the configurations of Table 15 for 4,000 hours. Appendix N presents the resulting lifetimes. These lifetimes are used to estimate the lifetime model parameters; we do this with use of Maximum Likelihood Estimation in STATA. Figure 17 summarizes the resulting parameter values, standard errors of the parameter values and the significance of the parameters. If the P>|z| value is smaller or equal to 0.05 the parameter is significant. We can see that only the model parameter corresponding to x1, d13 and d23 are significant. This means that d4 has no significant influence of the lifetime of the UV-lamp. New parameter values have to be derived without this parameter, and Figure 18 summarizes the results.

Figure 17: STATA Output to Determine Significant Model Parameters for the UV-case Lifetime Model

---

Figure 18: Model Parameter Values of the UV-case Lifetime Model

The model for the lifetime of the UV-lamp, dependent on stress and design parameters becomes:

$$t_p = \exp\left(a + b\xi + cd_1 + ed_2 + fd_3 - gd_3 - hd_2d_3 + \Phi_{SEV}^{-1}(p) * i\right)$$

Optimizing this lifetime function gives the optimal settings for the discrete design parameters, these optimal settings are:

d1=1	Emitter:	Type A
d2=1	Filling pressure:	High
d3=0	Amount of Neon:	Low
d4=0/1	Type of spiral:	No influence on the product reliability

The expected lifetime of the optimal design is expected to be x hours. The standard error of the 50th percentile estimate is 2,595 hours, which leads to a confidence interval of [x hours; x hours]. The estimation is less accurate as expected during the design of the test plan. Reasons for this are the high acceleration factor, the interaction between the design parameters and the unequal failure probabilities of the design variables. To improve the accuracy of the lifetime estimation for such a situation a next time, a pre-test can be used to obtain more information about the acceleration factor and the effects of the design beforehand. During this pre-test a part of the sample is used to improve the prior information assumptions.

The test time for this lifetime test was 4,000 hours. We are interested in the accuracy decrease if a shorter test time is used and the accuracy increase if a longer test time is used. Based on these results, we can conclude if 4,000 hours are really necessary or that the test time can be decreased. We generate a test plan for different test times and compare the standard error of the 50th percentile estimates. Figure 19 shows the results for different test times.

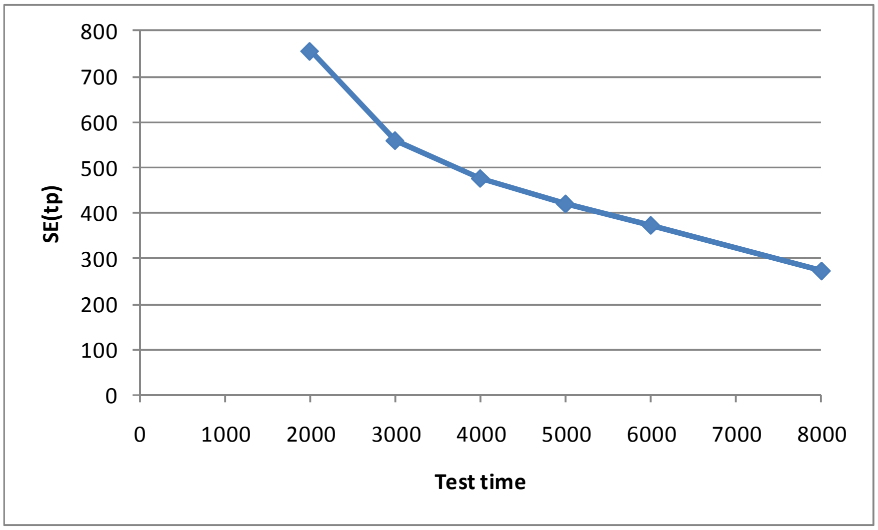


Figure 19: Estimation Accuracy for Different Test Times

From the graph we conclude that the estimation accuracy decreases enormously if the test time decreases from 4,000 to 2,000 hours, and the accuracy increases fewer if the test time is increased from 4,000 to 6,000 hours. In the case the test time and the available sample size are fixed, optimizing the test time can be used to check if the preferred accuracy can be reached within a shorter test time. If the test time is not fixed, it can be optimized based on the target accuracy and the available sample size.

7.3 Comparison

We cannot compare our results with the results of the lifetime test executed by company X, because they measured the influence of the design parameters only under high stress conditions. They make the transformation to normal use conditions based on a small experiment with one design variant and no standard error of the lifetime estimation under normal use condition is available. To give an idea about the performance of the proposed test plan, we compare this plan with our worst test plan, test plan 13 of Table 11. Appendix O summarizes the test plan settings and the lifetime data obtained in this situation. The estimation of the 50th percentile point becomes y hours with a 95% confidence interval of [y hours; y hours]. This percentile estimate is 1,500 hours higher than that of the best test plan and the confidence interval of this worst test plan is larger than that of the best test plan. The precision factor (see section 4.2) of the best test plan is 1.34 and that of the worst test plan is 1.47.

Based on this comparison we conclude that optimization of the test plan configurations is necessary to generate accurate test plans. Not only the type of the test plan and the rules to create a range of test plan settings lead to good test plans, but also the selection of the most accurate test plan of this range of settings is important.

8. Implementation

It is in the importance of CQM that the results of this research can be used during their projects related to reliability for their customers. This implies that the procedure and the simulation program developed into this research have to be formulated in such a matter that it can be absorbed by CQM. To do this, we identify four steps; each of the sections below describes one step.

8.1 Understand and maintain the developed procedure and the simulation program

To be able to use the developed procedure, it must be understood by its users. We have to transfer the knowledge obtained by this research to the responsible CQM employees. This transfer consists of three parts.

Instructions

The first step of the knowledge transfer is a session in which we explain the procedure and the functionality of the simulation program globally. This session is interactive, which makes it possible to answer questions and focus on some parts found difficult by the CQM employees. We have transferred the procedures and the simulation program during this kind of session.

User Manual

Test plans can be generated with use of the simulation program. This program is based on the procedure described and developed in this research. A user manual to describe the functionality of the simulation program and the link with the develop procedure in detail has to be made. The user manual must look like a recipe that can be followed by individuals with and without knowledge of the procedures to generate a test plan. We developed a user manual to enable CQM to use the simulation program.

Software Administrator

To guarantee efficient and effective maintenance of the software, one CQM employee has to be responsible for the simulation program. This means that this individual has to manage the usage and the maintenance of the program. Roel Wijgers is the CQM employee responsible for the simulation program.

8.2 Test and improve the procedure/ simulation program

If the procedure is transferred (meaning that it is understood by CQM and initial questions are answered), it has to be tested during a pilot project in order to make it part of the existing CQM routines. So far, we only tested the procedure based on case studies. The lifetime data resulting from the test plan we used, it simulated and not actual observed. To guarantee the performance of the developed procedure it must be evaluated during a pilot project.

Pilot Project

The goal of a pilot project is to validate the procedure and make improvements. We suggest searching a project to design a test plan for, with a situation in accordance with the situations we assumed in this research and a test time of around 1,500 hours. The performance of the test plan can be measured after the results become available; this is also the moment to evaluate the performance of the procedure make improvements. To reduce the time needed for the validation and optimization of the procedure, we suggest a period of approximately two months (± 1500 hours) and advise to involve the engineers with this test. Their opinion and knowledge about tests can be used to improve the practical usage of the procedure.

8.3 Apply the procedure during projects

After testing the procedure in practice, it can be integrated in the working procedures with respect to reliability. When using the developed procedure, these two aspects are important:

Emphasize the importance of smart lifetime testing

Many companies know the importance of reliability, but do not know how to test this efficient and/or accurate. CQM must emphasize this importance and support these companies with the insights of this research to reduce costs of warrantee periods and product recalls.

Use a random sample during the lifetime experiments

We assumed that the items placed on life tests are randomly sampled from the population of interest; otherwise the results might be biased. In practice, it is difficult to guarantee that the sample is randomly selected. A control group with known performance can be used as baseline to recognize unexpected problems.

8.4 Extend and adapt the procedure based on the experiences from practice

When using the procedure in practice, new situations can be discovered or some assumptions needed to use the developed procedure do not hold. In these cases, the developed procedure and the simulation program has to be extended or adapted. We describe some possible extensions.

Possible extensions

Extensions we recognized during this research and the case studies are:

- Include a step to optimize the test time based on the sample size and the target accuracy
  We researched the influence of the test time on the accuracy in section 7.2; test time influences the estimation accuracy. Optimizing the needed test time can be included in the test plan development procedure, by specifying the target accuracy and the available sample size.
- Improve and extend the methods with a stress constrained (section 6.2.1).
  We developed a method to design a test plan for the situation in which the stress constrained is a maximum stress constrained and all stress combinations above this maximum are also not feasible. Methods for other stress constraints that occur in practice must be researched and included.
- Include a method that can deal with interactions between stress factors and design parameters in the design of test plans (section 7.1).
  In this research we assume that there is no interaction between the stress and/or design factors. In practice interaction often occurs. If an assumption about the interaction can be made beforehand, this knowledge can be included in the prior information and the situation the test plan is optimized on.

9. Conclusions and Recommendations

So far we treated all research questions as described in section 2.2 Based on their answers, section 9.1 concludes with a complete procedure to design test plans for lifetime testing for different scenarios. Section 9.2 provides recommendations with respect to further research.

9.1 Conclusions

The goal of this research is to develop a procedure for the design of efficient test plans to make an accurate estimation of the reliability of a product and to be able to determine the influence of design parameters on this, by accelerating the occurrence of applicable failure modes. To fulfill this goal, section 2.2 states several sub questions; the conclusions based on the answers to those questions are described below.

The Generalized Log Linear relation is the most appropriate model to describe the lifetime of a product.

This relation describes the product lifetime based on a statistical distribution, stress relations, and/or design parameters. Several existing stress relations, such as the Arrhenius relationship or the Inverse Power Law can be included (section 3.3.4). With use of the Generalized Log Linear relation, also an expression for the failure probability dependent on stress, design and multiple failure modes can be obtained (see equation 39 and section 6.4.4).

Maximum Likelihood Estimation can be used to estimate the unknown parameters of the lifetime model. Based on the variances of these parameter estimations the accuracy of the lifetime estimation can be obtained.

The variance of the parameter estimations can be derived from the inverse Fisher Information Matrix. Based on these variances and the normal approximation, the confidence interval of the percentile estimates can be computed. This determines the accuracy of the estimation; section 4.1 describes this in more detail. To make the percentile estimation as accurate as possible the expected variance, or standard error, of the percentile function must be minimized during the design of a test plan.

The scientific literature provides theories behind different procedures to design test plans for different situations. This means that it does not describe how to generate these test plans in practice and how to include tests to measure the influence of design parameters on reliability.

For the one failure mode, one stress factor, one product design scenario the theory of Meeker and Escobar (1995) and Nelson (1990) can be used to generate test plans for different situations (section 5.2.2). For the one failure mode, two or more stress factors, one product design scenario Meeker and Escobar (1995) developed a method to transform this problem to a one-stress factor problem (section 6.2). For the remaining scenarios, we had to develop our own methods.

We developed a procedure based on the seven steps described below to generate a test plan for all described scenarios. To execute this procedure we developed a program in Microsoft Office Excel to execute the steps.

1. Obtain prior information per failure mode and initialize the lifetime model.
2. Choose type of test plan based on the uncertainty about the prior information assumptions and the available time to optimize the test plan with use of Figure 9 or Figure 14.
3. Simulate lifetime data and estimate the model parameters and their variances.
4. If more than one failure mode occurs, simulate the standard error of the p^{th} percentile estimate based on multivariate normal draws for the model parameters.
5. Select the best test plan configurations based on the standard error of the p^{th} percentile estimate.
6. If more than one stress factor is used or more than one product design variants must be tested, split the combined values to feasible values on the borders of the experimental region.
7. Test the robustness of the best test plan configurations by changing some prior information values and choose the most robust one.

Section 5.2 and section 6.5 describe this procedure in more detail.

The developed procedure is appropriate to well known situations.

With use of the developed procedure and the simulation program we developed test plans for different situations; based on comparison results we conclude that:

- The two-level Statistical Optimum Plan seems to lead to the most accurate lifetime estimations, but the three-level Best Compromise Plan is more robust to model deviations. The three-level New Proposed Plan results in the best or second best test plan type (section 5.4).
- Using two independent stress factors instead of one stress factor reduces the needed sample size to obtain certain accuracy up to 80% (section 6.2).
- A combined stress-design tests reduces the needed sample size to obtain certain accuracy with approximately 40% (section 6.3).

The developed procedure is appropriate to a practical case, the development of aquarium lamps.

When applying the developed procedure to design a test plan to test the reliability of aquarium lamps, we conclude for our developed procedure that:

- The developed procedure is flexible enough to be adjusted to practical scenarios. This became clear from the fact that although our research only assumed two independent design parameters, it can also deal with four interacting design parameters. Small changes in the procedure made it possible to generate a test plan for this “new” scenario (section 7.1).
- Accurate prior information is very important. The difference between the expected accuracy of the lifetime estimation based on the test plan configurations and the accuracy of the lifetime estimation based on the test results is approximately 8%. This difference can be reduced if assumptions about the influence of the design parameters on the reliability are available (section 7.1).

9.2 Recommendations

This research focuses on the design of test plans to estimate product reliability. Based on the assumptions made during this research and the knowledge obtained about this topic, we give some suggestions for further research.

Include a qualitative testing phase, the phase in which the prior information is obtained, in the procedure to design efficient test plans.

To improve the accuracy of the lifetime estimation accurate prior information, such as applicable failure modes or assumed failure percentages, is important. To acquire more accurate information, more research about qualitative testing methods to obtain this information is needed, see section 2.3 .

Research methods appropriate to situations in which the constant shape parameter assumption does not hold.

During this research we assumed that the shape parameter of the Weibull distribution remains constant over stress. This is a reasonable assumption, but in practice this is not always the case. Test plans for situations with a non-constant shape parameter have to be developed and the difference in performance with the constant shape parameter test plans must be examined. Seo, Jung and Kim (2009) describe a numerical method for the one-stress factor scenario, but for the other scenarios extensive research is needed.

Research methods to design test plans for situations based on other assumptions or more complex failure processes.

Chapter 8 describes the implementation of the developed procedure and the simulation program and describes possible extensions of the current method. Further research to incorporate these extensions in the developed procedure is needed.

Research cumulative-exposure models (i.e. models needed for step-stress testing) and the differences in the design of a test plan appropriate for that test method.

Section 3.3.6 describes the basics of cumulative-exposure models. Xu and Fei (2007) published an article about planning step-stress accelerated life tests based on the theories of constant accelerated life tests. The theories and ideas of this article can be used to develop the design of a test plan procedure such that it can also deal with step-stress.

Improve the method to solve the nonlinear optimization problem that leads to the best test plan settings.

We developed a simulation program to solve the nonlinear optimization problem. Improvements to this program can be made to increase the speed of the optimization or make it possible to increase the size of the simulation. We solve the optimization problem with use of simulation in Microsoft Office Excel. Microsoft Office Excel can use a limited number of rows and this leads to a maximum number of replications. Other software packages can deal with more data, such as Microsoft Office Access, or are faster, such as Delphi.

10. Bibliography

Bhote, K., & Bhote, A. (2004). *World Class Reliability.* New York: Amacom.

Breyfogle, F. (2003). *Implementing Six Sigma.* Hoboken, New Jersey: John Wiley & Sons, Inc. .

Brombacher, A. (1992). *Reliability by Design.* John Wiley & Sons: Chichester.

Clark, J., Garganese, U., & Swarz, R. (1997). An approach to design accelerated life test experiments. *Reliability and Maintainability Symposium* (pp. 242-248). IEEE Explore.

Condra, L. (2001). *Reliability Improvement with Design of Experiments.* New York: Marcel Dekker, Inc.

Dale, C. (1985). Application of the Proportional Hazards Model in the Reliability Field. *Reliability Engineering 10* , 1-14.

Deshpande, J., & Purohit, S. (2005). *Life Time Data: Statistical Models and Methods.* Singapore: World Scientific Publishing Co. Pte. Ltd.

Elsayed, E. (2003). Accelerated Life Testing. In H. Pham, *Handbook of Reliability Engineering* (pp. 415-428). London: Springer.

Elsayed, E., & Zhang, H. (2007). Design of PH based accelerated life testing plans under multiple-stress-type. *Reliability Engineering and System Safety* , 286-292.

Escobar, L., & Meeker, W. (2006). A Review of Accelerated Test Models. *Statistical Science* , 552-577.

Escobar, L., & Meeker, W. (1995). Planning Accelerated Life Tests With Two or More Experimental Factors. *Technometrics* , 411-427.

Escobar, L., & Meeker, W. (2004). Reliability: The other dimension of quality. *Quality Technology and Quantitative Management 1* , 1-25.

Gould, W., & Sribney, W. (1999). *Maximum Likelihood Estiamtion with STATA.* USA, Texas: Stata press.

Guo, H., & Pan, R. (2007). D-Optimal Reliability Desugn for Two-Stress Accelerated Life Tests. *IEEE International Conference on Industrial Engineering and Engineering Management*, (pp. 1236-1240). Singapore.

Islam, A., & Ahmad, N. (1994). Optimal Design of Accelerated Life Tests for the Weibull Distribution Under Periodic Inspection and Type I Censoring. *Microelectronics Reliability Vol. 34, No. 3* , 1459-1468.

Kuo, W., & Wan, R. (2007). Recent Advances in Optimal Reliability Allocation. *IEEE Transactions Vol. 37, No. 2* , 143-156.

Larsen, R., & Marx, M. (2006). *An Introduction to mathematical Statistics and Its Applications.* New Jersey: Pearson Eduction Inc. .

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

Liao, C., & Tseng, S. (2006). Optimal Design for Step-Stress Accelerated Degradation Tests. *IEEE Transaction on Reliability* , 59-66.

Lydersen, S., & Rausand, M. (1987). A Systematic Approach to Accelerated Life Testing. *Reliability Engineering 18* , 285-293.

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

Meeker, W. (1984). A Comparison of Accelerated Life Test Plans for Weibull and Lognormal Distributions and Type I censoring. *Technometrics* , 157-171.

Meeker, W., & Hahn, G. (1977). Asymptotically Optimum Over-Stress Tests to Estimate the Survival Probability at a Condition with a Low Expected Failure Probability. *Technometrics* , 381-399.

Meeker, W., & Hahn, G. (1985). *How To Plan an Accelerated Life Test.* Milwaukee, Wisconsin: American Society for Quality Control.

Montgomery, D. (2005). *Design and Analysis of Experiments.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Nelson, W. (1990). *Accelerated testing.* New York: John Wiley & Sons, Inc.

News from CPSC. (2010, March 17). Retrieved March 22, 2010, from U.S. Consumer Product Safety Commission : http://www.cpsc.gov/cpscpub/prerel/prhtml10/10169.html

Pal, S. (2005). Order statistics for some common hazard rate functions with an application. *Interational Journal of Quality & Reliability Management* , 201-210.

Pascual, F. (2008). Accelerated Life Test Planning With Independent Weibull Competing Risks. *IEEE Transactions on Reliability* , 435-444.

Pascual, F. (2007). Accelerated Life Test Planning With Independent Weibull Competing Risks With Known Shape Parameter. *IEEE Transactions on Reliability* , 85-93.

Philips: LED Lighting Systems. (n.d.). Retrieved March 2, 2010, from Philips: http://www.lighting.philips.com

ReliaSoft. (2007). *Accelerated life testing.* Retrieved February 22, 2010, from Weibull.com: http://www.weibull.com/AccelTestWeb/acceltestweb.htm

ReliaSoft. (2007). Design for Reliability: Overview of the Process and Applicable Techniques. *Reliability Edge 8, no 2* , pp. 1-6.

Schuld, M. (2010, July 14). Reliability UV-lamp. (M. Pronk, Interviewer)

Seo, J., Jung, M., & Kim, C. (2009). Design of Accelerated Life Test Sampling Plans with a Nonconstant Shape Parameter. *European Journal of Operational Research* , 659-666.

Tang, L., & Liu, X. (2010). Planning and Inference for a Sequential Accelerated Life Test. *Technometrics* , 103-118.

Tang, L., & Yang, G. (2002). Planning Multiple Levels Constant Stress Accelerated Life Tests. *Reliability and Maintainability Symposium* (pp. 338-342). Singapore: IEEE Reliability.

Tseng, S. (1994). Planning accelereated life tests for selecting the most reliable product. *Journal of Statistical Planning and Inference 41* , 215-230.

Werner, P. (2009). Statistical Methods for the Estimation of operational Conditions in Mixed Failure Data of Production Equipment. *International Apllied Reliability Symposium.* Spain: ReliaSoft Corp.

Xu, H., & Fei, H. (2007). Planning Step-Stress Accelerated Life Tests With Two Experimental Variables. *IEEE Transaction on Reliability* , 569-579.

Zhang, Y., & Meeker, W. (2005). Bayesian life test planning for the Weibull distribution with given shape parameter. *Metrika 61* , 237-249.

Appendix A. Definitions

Accelerated stress testing
Stress testing method whereby items are tested under more severe stresses than those encountered during normal use. The goal is to quickly obtain data, which yield desired information on product life under normal use conditions.

Acceleration factor
Ratio of the lifetime under normal use conditions and the lifetime at a higher stress level.

Allocation fraction
Fraction of the items allocated to a specific test setting, i.e. stress level or design setting

Bound ratio
Upper bound of the confidence interval divided by the lower bound

Censoring
Not all the products are failed when the test terminates, so not all items can be used in the same way to determine the lifetime of the product.

Confidence Interval
Range around a given statistical estimate within the true value is said to be located with some special degree of confidence.

Cumulative Exposure (CE) model
A life-stress relationship corresponding to a step-stress test. The cumulative effect of the applied stresses is taken into account. So it relates the life distribution of the units at one stress level to the distribution at the next stress level.

Design parameter
Physical or functional component or product characteristics represented by a discrete or continuous variable

Design of Experiments (DoE)
Structured approach to consider the effects of several independent variables simultaneously in one experiment without evaluating all possible combinations of variable levels.

High Accelerated Life Testing (HALT)
Stress testing method to identify design weaknesses. Aggressive testing conditions are used to identify these weaknesses.

Hazard rate function
Probability of failure in the small interval [t, t+Δ], given survival up to time t, also called failure rate function.

Failure

A product performs no longer its intended function satisfactorily.

Failure mode

Causes of a failure, for example melted or burned components

Failure rate

Probability of failure in the small interval [t,t+Δ], given survival up to time t, also called hazard rate function.

Fisher Information Matrix

The expectation of the symmetric matrix of negative second partial derivatives. It is a way of measuring the amount of information that an observable random variable X carries about an unknown parameter θ upon which the likelihood function of θ, $L(\theta) = f(X;\theta)$, depends.

Life-stress relationship

A method to relate the lifetime under stress to the lifetime under normal conditions.

Lifetime

The period in which a product performs its intended function.

Lifetime distribution

The probability distribution function of the lifetime of a population of items.

Likelihood function

A function which contains the unknown parameters and expresses the "likelihood" of the data, given the values of the parameters.

Location-scale distribution

Univariate probability distribution parameterized by a location parameter and a nonnegative scale parameter.

Log lifetime

The natural logarithm of the lifetime. This transformation is used for computational reasons.

Maximum Likelihood Estimation (MLE)

Method to estimate unknown parameter values with use of a likelihood function. This likelihood function is maximized by changing the parameter values, and results in the parameter values that are most likely to be the true values, the parameter estimates.

Multiple Environment Over Stress Testing (MEOST)

Stress testing method to identify failure modes. Multiple stresses are applied to products in a short period of time and these stresses are increased until the product fails.

Model departures
The “true” family of distributions or life-stress relation deviates from the assumed one.

Monte Carlo simulation
Computational algorithm that relies on repeated random sampling to compute their results.

Normal use conditions
Environmental or physical conditions under which the product is used during normal circumstances

Normalized (transformed) stress level
The (transformed) stress level projected on a one to zero scale. This means that the normal use stress level is one and the highest stress level is zero. All other stress levels take a value between these two values.

Percentile (p^{th})
Point in time below which p% of the items are failed.

Prior information
Assumptions necessary to derive the assumed lifetime model with the corresponding parameter values. This prior information consist of assumption about the family of distribution, assumption about the scale parameter, failure probability estimates per stress factor, design parameter or failure mode and a life-stress relation per stress factor.

Proportional Hazard (PH) model
The proportional hazard regression model can be used to isolate effects of explanatory variables. This model assumes that the explanatory variables have a multiplicative effect on the hazard rate.

Reliability
The probability that a product will perform its intended function until a specified point in time under encountered use conditions

Reliability function
Probability of an items surviving up to time t, also called the survival function.

Robustness
Measure of how sensitive a method is to violations of his assumptions, such as model departures and parameter deviations.

Sample size
The amount of products used during the lifetime test.

Scale Accelerated Failure Time (SAFT) model
A parametric model that assumes that the shape parameter of the distributions is constant for all stress levels and that the scale parameter differs per stress levels. The lifetime of a product is multiplied by

some constant based on the explanatory variables. Based on this, a linear model can be obtained for the log lifetime dependent on the explanatory variables.

Step-stress test
Stress test whereby the items are tested at a pre-specified stress level for a specified period of time. The items that are not failed are tested again, subject to a higher stress level for another period of time. This process continues until the termination time is reached.

Stress factor
A condition in the environment which has an influence on the reliability of the product. In general an increase of the stress factor increases the failure probability of a product. Examples of stress factors are temperature, humidity or power.

Stress level
The value of a stress factor expressed in units. Examples are temperature in degrees Celsius or power in Watt.

Survival function
Probability of an items surviving up to time t, also called the reliability function.

Target precision factor
The fractional deviation between the estimation and one of the confidence bounds.

Test configurations
Settings for a lifetime test. These configurations consist of the sample size, the stress factors, the stress levels and the allocation of the items over the stress levels and design variants.

Test plan
Plan containing the test time, test configurations with as goal to perform a test which result in an as accurate as possible estimation of the percentile of interest.

Test plan type
Kind of test plan appropriate based on situation specific ideas or including specific rules or constraints.

Transformed stress level
Transformation dependent on the stress relationship applied to the original stress level to be able to generate a linear log lifetime-stress relationship.

Variance-Covariance matrix
Matrix of variances of the elements of a factor and the covariances between these elements. The matrix can be generated based on the inverse of the Fisher Information Matrix.

Appendix B. Examples of Product Failures

Example 1: Recall during the hidden 0-hour phase

This is an example of a product recall during the hidden 0-hour phase from the U.S. Consumer Product Safety Commission (News from CPSC, 2010). The failure occurs during the first time the power pack is charged. More insight in the reliability of the product by experiments during the development phase could prevent these kinds of problems.

Mobile Power Packs Recalled By Tumi Due to Fire Hazard

WASHINGTON, D.C. - The U.S. Consumer Product Safety Commission, in cooperation with the firm named below, today announced a voluntary recall of the following consumer product. Consumers should stop using recalled products immediately unless otherwise instructed.

Name of Product: Mobile Power Packs

Units: About 5,000

Manufacturer: Tumi, of South Plainfield, N.J.

Hazard: The lithium-ion cells used in the Mobile Power Pack can ignite or explode while charging, posing a fire hazard. This hazard is only present for units that have not been charged.

Incidents/Injuries: There were two reports of consumers experiencing small fires during their initial charge. No injuries were reported.

Description: The recalled Mobile Power Pack is a mobile device that receives an AC charge in a compact battery pack that will then give five DC charges to small electronic devices including mobile phones, MP-3 players, Blackberries, and PDAs. The power pack is black and silver and is rectangular in shape. The front of it has a small circular control panel. The front also displays the word "Tumi" engraved on a silver button located towards the bottom of the device. Style number 14362 is printed on the power pack packaging.

Sold by: Tumi retail stores, department and specialty stores nationwide and www.Tumi.com from August 2007 through March 2008 for $135.

Manufactured in: China

Remedy: If the unit has not been used and has never been charged, please do not try to charge. Contact Tumi customer care for instructions on how to return the power pack and receive a free replacement power pack. If you have charged the unit previously without incident, you can continue to use the product.

Example 2: Recall during the early wear-out phase

The notification below is an example of failures during the second phase of the life cycle, the early wear-out phase. If the percentage of products failed during this phase is very low, and also the consequences of the failure have no risk, there is no problem. In the situation described below, the failure can cause electrical shocks or incineration. During the development phase these problems must be recognized, improvements of the product in that phase are much cheaper than product recalls and the corresponding loss of goodwill.

DCO5 MOTORHEAD – VEILIGHEIDSBERICHT

Dit veiligheidsbericht is enkel van toepassing op de **Dyson DCO5 Motorhead**, verkocht tussen 2000-2005. Geen enkel ander DCO5 product is getroffen.

Heeft u een Dyson DCO5 Motorhead, neem dan alstublieft contact op met Dyson op het onderstaande nummer.

In zeldzame gevallen breekt het klepje op het DCO5 Motorhead handvat ①. Dit kan ervoor zorgen dat de electrische bedrading bloot komt te liggen ②.

Wanneer u deze bedrading aanraakt als het toestel in het stopcontact zit, bestaat het risico op een electrische schok of verbranding.

We vragen daarom aan alle eigenaars van een DCO5 Motorhead om de volgende maatregelen te treffen:

- **Haal de stekker uit het stopcontact.**
- In het geval dat de slang los zit, raak de electrische bedrading niet aan en gebruik het toestel niet.

Neem meteen contact op met Dyson.

Wanneer de slang goed vast zit, is er niet meteen een risico. Het blijft veilig om uw stofzuiger te gebruiken. Toch willen we, als voorzorgsmaatregel, een verstevigend opzetstuk leveren aan **alle DCO5 Motorhead eigenaars** om te voorkomen dat de slang lost. Neem alstublieft contact op met Dyson op het onderstaande nummer of op de website www.dyson.nl/dc05motorhead voor uw gratis opzetstuk.

Neem alstublieft contact op met Dyson op onderstaand telefoonnummer.

☎ 020 521 9890 www.dyson.nl/dc05motorhead

This notification is retrieved 22 March 2010 from www.productwaarschuwing.nl.

Example 3: Recall during the systematic wear-out phase

The PRO draglink problem is an example of a product where the systematic wear-out starts too early. Material fatigue is the cause of the failure. Accelerated tests during the development phase can be used to recognize these problems in an early stage.

Persbericht

PRO

BELANGRIJKE VEILIGHEIDSINFORMATIE
PRO "Hi-Comp" fietsstuurpennen

PRO heeft een mogelijk veiligheidsrisico ontdekt betreffende PRO Hi-Comp stuurpennen voor racefietsen, geproduceerd tussen november 2003 en juni 2006. Als gevolg van materiaalmoeheid kunnen er scheurtjes in deze stuurpennen ontstaan waardoor ze uiteindelijk kunnen breken. Wanneer dit gebeurt tijdens het fietsen, kan de fietser de controle over zijn rijwiel verliezen wat kan leiden tot ernstige verwondingen.

Omdat we willen dat alle PRO producten voldoen aan de hoogste industrie-eisen, heeft PRO besloten tot een vrijwillige terugroepactie van alle getroffen stuurpennen, als een voorzorgsmaatregel in het belang van de veiligheid van onze klanten.

De getroffen modellen zijn alle PRO stuurpennen voorzien van het "PRO Hi-Comp" logo zoals te zien op de foto.

Wanneer u een PRO stuurpen in bezit heeft voorzien van het "PRO Hi-Comp" logo, ongeacht de lengte of diameter, dient u onmiddellijk te stoppen met het gebruik ervan. Brengt u deze stuurpen terug naar de dealer waar deze is gekocht voor een volledige teruggave van de kostprijs.

Wanneer u hier vragen over heeft kunt u contact opnemen met ons of met uw lokale PRO dealer.

Wij danken u voor uw medewerking en bieden u onze verontschuldigingen aan voor dit ongemak.

SHIMANO EUROPE B.V.
INDUSTRIEWEG 24
8071 CT NUNSPEET
THE NETHERLANDS
+31 (0) 341 272222
WWW.PRO-BIKEGEAR.COM

This notification is retrieved 22 March 2010 from www.productwaarschuwing.nl.

Appendix C. Censoring

When methods are used to accelerate the lifetime of products, censoring occurs in most cases. Two types of censoring are often used in practice.

Type-I censoring involves running each test unit for a predetermined amount of time. In this case, the censoring time is fixed, while the number of failures is random, not all items have to be failed when the test terminates.

Type-II censoring involves simultaneous testing of the units until a predetermined number of them fail. As opposed to type-I censoring, in this case the censoring time is random and the number of failures is fixed (Tseng, 1994).

Within these two types of censoring, different variants exists; left, right and interval censoring. Left censoring means that only the upper bound of the failure time is known, it occurs for example if a unit fails before it first inspection. At right censoring the lower bound of the failure time is known, this occurs for example at type I censoring, all unfailed items are right censored. At interval censoring, both the lower and upper bound of the failure time is known, but the exact failure time is not known. These censored observations influence the accuracy of the parameter estimations. This effect of censoring must be taken into account when designing a test plan; the likelihood function can be adapted to deal with this. In this research we focus on type-I exact censored data, so an observation can be censored because the item is still alive when the test terminates.

Appendix D.Normal and Weibull Distribution

Probability density function

The probability density function describes the relative likelihood that an item fails at time t.

Normal: $f(t) = (2\pi\sigma^2)^{-\frac{1}{2}}\exp\left[-(t-\mu)^2/(2\sigma^2)\right] \qquad -\infty < t < \infty$

Weibull: $f(t) = \frac{\beta}{\eta^\beta}(t)^{\beta-1}\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \qquad \eta, \beta, t > 0$

Reliability Function

The reliability function represents the probability of an items surviving up to time t.

Normal: $R(t) = 1 - \Phi\left[\frac{t-\mu}{\sigma}\right] \qquad -\infty < t < \infty$

Weibull: $R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \qquad \eta, \beta, t > 0$

Hazard Function

The hazard function describes the probability of failure in the small interval [t, t+Δt), given survival up to time t.

Normal: $h(t) = \frac{(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left[-(t-\mu)^2/(2\sigma^2)\right]}{1-\Phi\left[\frac{t-\mu}{\sigma}\right]}$

Weibull: $h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \qquad \eta, \beta, t > 0$

Percentile

The moment by which p*100% of the population will be failed is called the p$^{\text{th}}$ percentile.

Normal: $t_p = \mu + z_p\sigma$

Weibull: $t_p = \eta[-\ln(1-p)]^{\frac{1}{\beta}}$

The scale parameter η is also called the characteristic life. This is always the time at which 63.2 % of the population failed regardless the value of β, because $-\ln(1-0{,}6321..) = 1$. This is also called the 63.2% percentile (Lawless, 2003).

Mean Time to Failure

The mean time to failure (MTTF) of a product is the average lifetime of a product.

Normal: $E[T] = \mu = t_{0,5}$

Weibull: $E[T] = \eta * \Gamma\left(\frac{1}{\beta}+1\right)$

$\Gamma(n) = \int_0^\infty e^{-x}x^{n-1}dx$ (gamma function)

Standard Deviation Lifetime

Normal: $\sigma_t = \sigma$

Weibull: $\sigma_t = \eta * \sqrt{\Gamma\left(\frac{2}{\beta}+1\right) - \Gamma\left(\frac{1}{\beta}+1\right)^2}$

Appendix E. Life-Stress Relations Combined with Location-Scale Distribution

$\mu(x) = \gamma_0 + \gamma_1 x$

Inverse power law $T(U) = \frac{1}{K * U^n}$

$\mu(x) = \ln(\eta) = \ln\left(\frac{1}{K * U^n}\right)$

$= -\ln(K * U^n) = -ln(K) - nln(U)$

$x = \ln(U)$
$\gamma_0 = -\ln(K)$
$\gamma_1 = -n$

Arrhenius $T(V) = C * e^{\frac{B}{V}}$

$\mu(x) = \ln(\eta) = \ln\left(C * e^{\frac{B}{V}}\right)$

$= \ln(C) + B * 1/V$

$x = 1/V$
$\gamma_0 = lnC$
$\gamma_1 = B$

Appendix F. GLL vs. PH

This appendix shows the hazard rate function and the survival rate function for both the Generalized Log Linear (GLL) model and the Proportional Hazard (PH) model. We can write the equations of both models such that for the **standard extreme value distribution** both expressions can be transformed to each other.

GLL, smallest extreme value distribution

$$\mu(\boldsymbol{x}) = \ln\big(\eta(x)\big) = \ln(\exp(\alpha_0 + \textstyle\sum_{i=1}^{m} \alpha_i\, x_i)) = \alpha_0 + \sum_{i=1}^{m} \alpha_i\, x_i$$

$$h(t;\boldsymbol{x},\sigma) = \frac{1}{\sigma \exp(\alpha_0 + \sum_{i=1}^{m} \alpha_i x_i)} \left[\frac{t}{\exp(\alpha_0 + \sum_{i=1}^{m} \alpha_i x_i)}\right]^{\frac{1}{\sigma}-1}$$

$$R(t;\boldsymbol{x},\sigma) = \exp\left(-exp\left[\frac{\ln(t) - \alpha_0 + \sum_{i=1}^{m} \alpha_i\, x_i}{\sigma}\right]\right) = \exp\left(-\exp\left(\frac{\ln(t)}{\sigma}\right) * \exp\left(\frac{\alpha_0 + \sum_{i=1}^{m} \alpha_i\, x_i}{-\sigma}\right)\right)$$

$$= \exp\,(-t^{\frac{1}{\sigma}} * \exp\left(\frac{\alpha_0 + \sum_{i=1}^{m} \alpha_i\, x_i}{-\sigma}\right))$$

PH, baseline hazard rate via smallest extreme value distribution

$$h(t,\boldsymbol{x}) = h_0(t) * \exp\left(\sum_{i=1}^{m} \alpha_i x_i\right) = \frac{1}{\sigma\mu} * (\frac{t}{\mu})^{\frac{1}{\sigma}-1} * \exp\left(\sum_{i=1}^{m} \alpha_i x_i\right) = \frac{1}{\sigma} * t^{\frac{1}{\sigma}-1} * \exp\left(\alpha_0 + \sum_{i=1}^{m} \alpha_i x_i\right)$$

$$R(t;\boldsymbol{x}) = e^{-\int_0^t h(u)du} = \exp\,(-t^{\frac{1}{\sigma}} * \exp\left(\alpha_0 + \sum_{i=1}^{m} \alpha_i x_i\right))$$

Based on the expression for the reliability function of the generalized log linear model and the proportional hazard model, we can see that both functions can be transformed to each other.

The transformation we use is: $\alpha_i, GLL = \alpha_i, PH * -\sigma$

Appendix G.Likelihood Functions

Weibull distribution without censoring

$$f(t) = \frac{\beta}{\eta^{\beta}}(t)^{\beta-1}\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$

MLE

$$\mathcal{L}(\beta,\eta;\boldsymbol{t}) = \sum_{i=1}^{n} ln\left(\frac{\beta}{\eta^{\beta}}(T_i)^{\beta-1}\exp\left[-\left(\frac{T_i}{\eta}\right)^{\beta}\right]\right)$$

$$= nln(\beta) - n\beta\ln(\eta) + (\beta-1)\sum_{i=1}^{n}\ln(T_i) - \sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}$$

$$\frac{\partial\mathcal{L}}{\partial\beta} = \frac{n}{\beta} - nln(\eta) + \sum_{i=1}^{n}\ln(T_i) - \sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}\ln\left(\frac{T_i}{\eta}\right)$$

$$\frac{\partial\mathcal{L}}{\partial\eta} = -\frac{n\beta}{\eta} + \frac{\beta}{\eta}\sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}$$

$$\frac{\partial^2\mathcal{L}}{\partial\beta^2} = \frac{-n}{\beta^2} - \sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}\left(\ln\left(\frac{T_i}{\eta}\right)\right)^2$$

$$\frac{\partial^2\mathcal{L}}{\partial\eta^2} = \frac{n\beta}{\eta^2} - \frac{\beta^2+\beta}{\eta^2}\sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}$$

$$\frac{\partial^2\mathcal{L}}{\partial\beta\,\partial\eta} = \frac{-n}{\eta} + \frac{1}{\eta}\sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta} + \frac{\beta}{\eta}\sum_{i=1}^{n}\left(\frac{T_i}{\eta}\right)^{\beta}\ln\left(\frac{T_i}{\eta}\right)$$

Weibull-Arrhenius model without censoring

$$f(t) = \frac{\beta}{\left(C*\exp\left(\frac{B}{V}\right)\right)^{\beta}}(t)^{\beta-1}\exp\left[-\left(\frac{t}{C*\exp\left(\frac{B}{V}\right)}\right)^{\beta}\right]$$

$$\mathcal{L}(\beta,B,C;\boldsymbol{t}) = \sum_{i=1}^{n} ln\left[\frac{\beta}{\left(C*\exp\left(\frac{B}{V_i}\right)\right)^{\beta}}(T_i)^{\beta-1}\exp\left[-\left(\frac{T_i}{C*\exp\left(\frac{B}{V_i}\right)}\right)^{\beta}\right]\right]$$

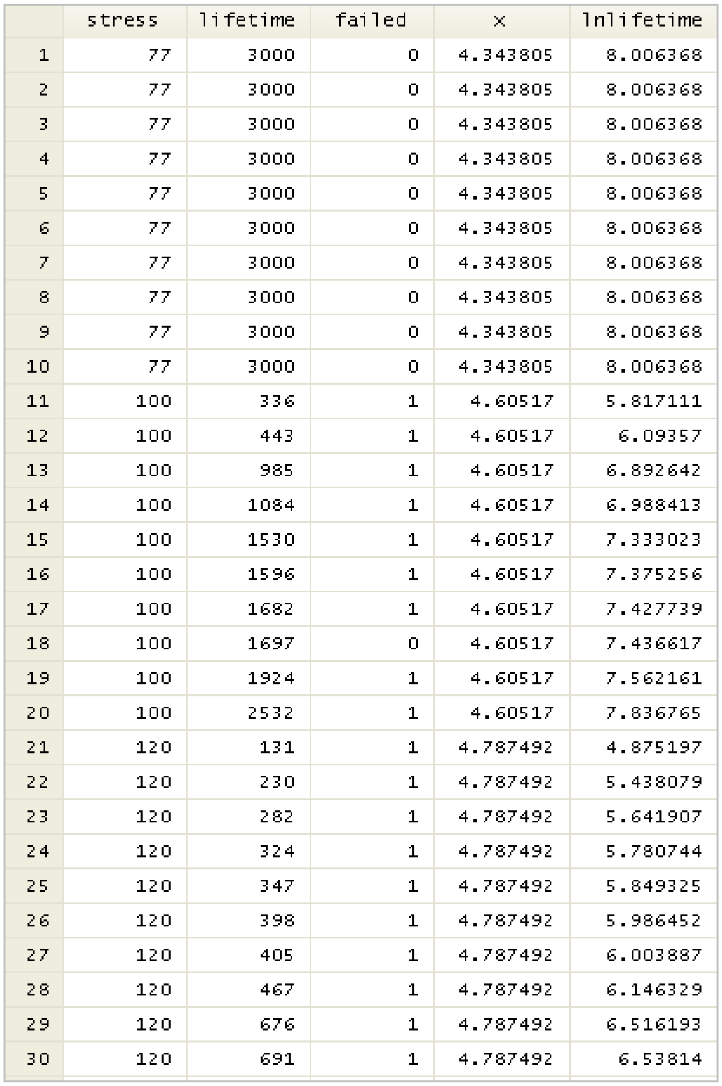
$$= nln(\beta) - n\beta\ln\left(C*\exp\left(\frac{B}{V_i}\right)\right) + (\beta-1)\sum_{i=1}^{n}\ln(T_i) - \sum_{i=1}^{n}\left(\frac{T_i}{C*\exp\left(\frac{B}{V_i}\right)}\right)^{\beta}$$

Appendix H.MLE with STATA

Stata code:

```
program weibull
        args lnf theta1 theta2

        tempvar R
        quietly gen double `R'= ($ML_y1-`theta1')/`theta2'
        quietly replace `lnf'= -ln(`theta2')+`R'-exp(`R') if $ML_y2==1
        quietly replace `lnf'= ln(1-(1-exp(-exp(`R')))) if $ML_y2==0
end
. ml model lf weibull (mu: lnlifetime failed= x) /sigma
. ml maximize
```

Data sheet:



	stress	lifetime	failed	x	lnlifetime
1	77	3000	0	4.343805	8.006368
2	77	3000	0	4.343805	8.006368
3	77	3000	0	4.343805	8.006368
4	77	3000	0	4.343805	8.006368
5	77	3000	0	4.343805	8.006368
6	77	3000	0	4.343805	8.006368
7	77	3000	0	4.343805	8.006368
8	77	3000	0	4.343805	8.006368
9	77	3000	0	4.343805	8.006368
10	77	3000	0	4.343805	8.006368
11	100	336	1	4.60517	5.817111
12	100	443	1	4.60517	6.09357
13	100	985	1	4.60517	6.892642
14	100	1084	1	4.60517	6.988413
15	100	1530	1	4.60517	7.333023
16	100	1596	1	4.60517	7.375256
17	100	1682	1	4.60517	7.427739
18	100	1697	0	4.60517	7.436617
19	100	1924	1	4.60517	7.562161
20	100	2532	1	4.60517	7.836765
21	120	131	1	4.787492	4.875197
22	120	230	1	4.787492	5.438079
23	120	282	1	4.787492	5.641907
24	120	324	1	4.787492	5.780744
25	120	347	1	4.787492	5.849325
26	120	398	1	4.787492	5.986452
27	120	405	1	4.787492	6.003887
28	120	467	1	4.787492	6.146329
29	120	676	1	4.787492	6.516193
30	120	691	1	4.787492	6.53814

Appendix I. Description of Additional Methods to Design a Test Plan

Not all the test plan types are used during this research; some are not taken into account because they can only be used for specific situations. To give a complete overview of the existing methods, we give of a short summery of these methods.

One failure mode, one stress factor, one product design scenario

Two Level Statistical Optimum Plan with Interval Censoring

This method uses the same way of reasoning as the statistical optimum plan presented above, but is also able to use interval censored observations. Interval censoring means that the products are inspected on interval basis; not the exact failure time, but the interval of time in which a product fails is known. The likelihood function can be adapted to $L(\boldsymbol{\theta};\, x) = \prod_{\mathrm{i=1}}^{\mathrm{n}} \int_{t_{i-1}}^{t_i} f(t;\boldsymbol{\theta})dt = \prod_{i=1}^{n}(F(T_i;\boldsymbol{\theta}) - F(T_{i-1};\boldsymbol{\theta}))$ for this situation. Islam and Ahmad (1994) describe a method to develop a 2 level test plan with use of a two step procedure to minimize the asymptotic variance of the p$^{\mathrm{th}}$ percentile. The first step is to optimize with respect to the allocation percentage of the low stress level. The second step is to optimize with respect to the low stress level, by grid search. The asymptotic variance is evaluated on the grid $x_L = d, 2d, 3d, \ldots$ where d is the grid size. The optimal plan is that plan for which the asymptotic variance is minimal among all the grid points considered. They also carried out a sensitivity analysis to see the influence of estimation errors and conclude that 5 or more inspections give the same asymptotically results as when k goes to infinity. This implies that more than 5 inspections are unnecessary.

Three Level Best Linear-Quadratic Discriminator Plan

The Best Linear-Quadratic Discriminator Plan is a three level plan and minimizes the variances of γ_2 in the quadratic model: $\mu(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2$. This variance is a measure of the power to detect quadratic departures from the linear model $\mu(x) = \gamma_0 + \gamma_1 x$. This method can be used if a curvilinear relationship between the stress and the lifetime is expected (Meeker W. , 1984).

Other ideas and guidelines

Tang and Yang (2002) use contour plots to determine the sample allocation and the stress levels. Three different approaches are described. The first approach follows the strategy that the average of the low and middle stress levels, weighted by their respective allocation in a near optimal situation equal to the optimal low stress level in the statistical optimum plan (plan 1a.). The second approach is based on the assumption that the purpose of the middle stress level is validating the stress-life relationship. In this case, minimum allocation to this level is preferred, such that there are sufficient failures to detect nonlinearity. The third approach is a hybrid of both. Because contour plots are difficult to compute, these methods are not used directly, but the idea behind these methods can be used in developing new tests plan methods.

Meeker and Hahn (1985) give some guidelines for planning accelerated life test which can be used in developing new procedures. The guidelines are based on a compromise between extrapolation in time and extrapolation in stress. Extrapolation in time is necessary if less than p100% is failed if one wants to estimate the pth percentile. Extrapolation in stress is necessary when the lowest test stress level is higher than the normal use condition. If three stress levels are used, they must be chosen such that the failure probability at the highest stress level is at least 2p(100)%, at the middle stress level at least p(100)% and at the lowest stress level p/3(100)%. The minimum of failures per stress level equals 5, to generate significant results.

One failure mode, two or more stress factors, one product design scenario

Factorial plan

The factorial plan is based on the idea of Design of Experiments. Traditional DoE uses for each factor a number of levels on which it is tested and the test runs are designed such that all interactions can be determined based on the results. These plans are called full factorial plans (Condra, 2001). The disadvantage of such plans is that for an experiment with n different factors tested at k different levels, k^ndifferent test runs are necessary. Standard the factors have 2 settings, for example low and high, or without and with. No optimization criterion is used to choose these levels. Half or quarter factorial test plans can be developed to overcome the large number of runs problems, but in this way it is not possible to determine all interaction effects.

D-Optimal Reliability Test Design

The test plans that can be obtained with this method are based on the DoE theory. If no censoring takes place, the boundaries of the feasible regions are used as test values. For the D-optimal plan, the determinant of the Fisher Information Matrix is maximized, and this occurs in the no censoring case when the allocation to each level is the same. When censoring is included, the Information Matrix is also dependent on the value of the other parameters, and no exact analysis can be done. No suggestions for planning values in this situation are given. (Guo & Pan, 2007)

Proportional Hazard plan

This method is based on the Proportional Hazard models, with a quadratic baseline hazard rate function and an exponential relative risk ratio. The optimization criterion for this kind of test plans is to minimize the variance of the estimated hazard rate function. Per stress factor, two test levels are used; and both the low and the high stress levels are chosen during the optimization. A restriction to the minimum number of failures for each stress level is given. Based on initial estimates of all the model parameters and numerical methods for nonlinear optimization, the problem is solved. The output is the four test levels, and their allocation proportions (Elsayed & Zhang, 2007).

One failure mode, one stress factor, two or more product design variants scenario

Selection of Most Reliable Product Plan

Tseng (1994) describes a method to design a test plan for selecting the most reliable design. The method is based on type-II censoring; the test terminates after a predetermined number of failures per stress level. The method uses a probability of correct selection and based on this probability and initial parameter estimates, the test plan is edited. A two level test plan with a given low and high stress level are assumed, only the allocation over the levels and the number of failures per level before the test terminates are computed. The minimization criterion is the estimated standard error of the characteristic life. The situation we use during this thesis differs from this situation, but some ideas can be used to develop a method to design a test plan in our case.

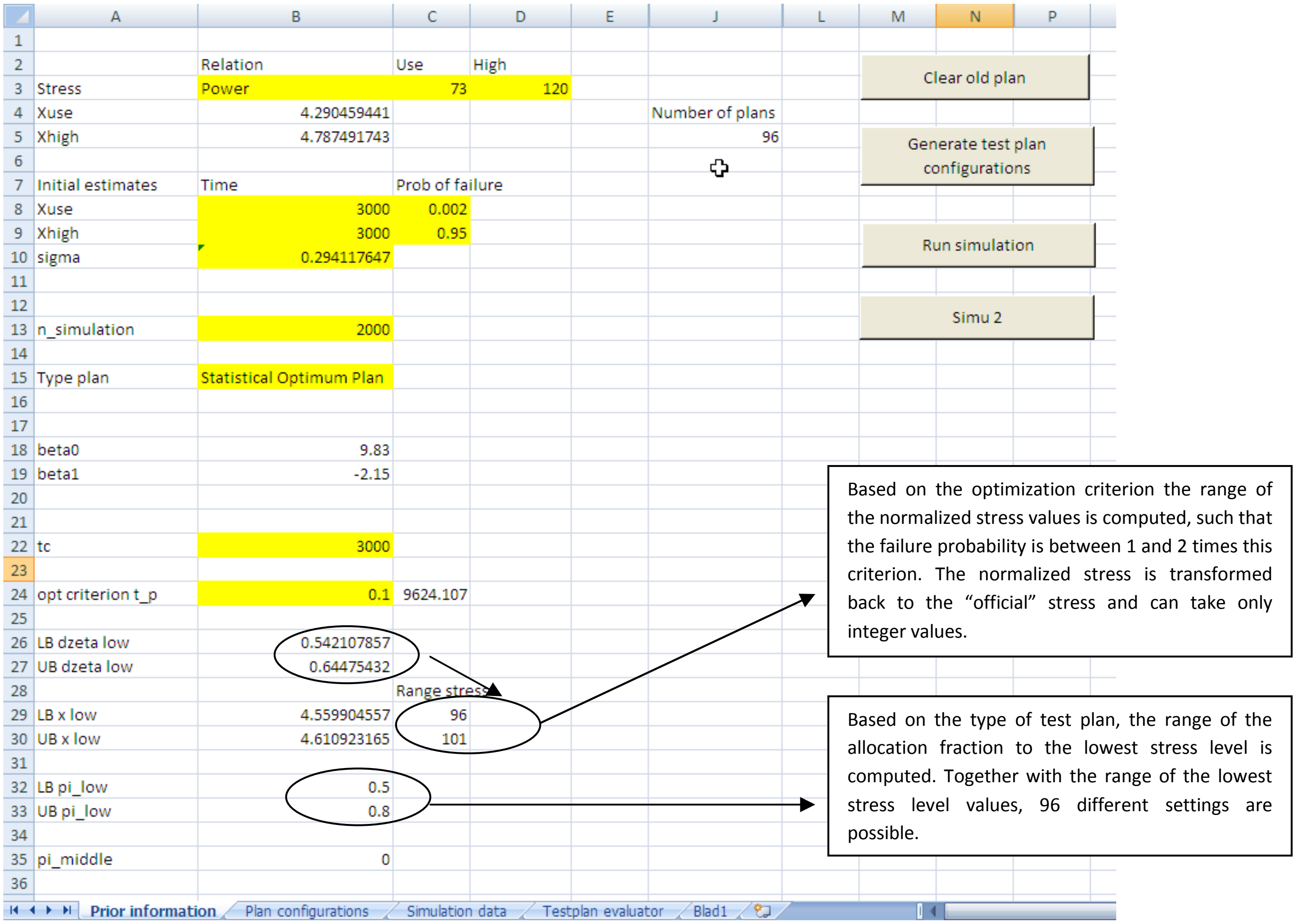
Appendix J. Data Generator

This appendix describes the methods that are used to select the best test plan configuration per step and illustrates how this method are validated. The steps to generate the best test plan are:

1. Prior Information
2. Test Plan Configurations
3. Simulation
4. Maximum Likelihood Estimation
5. Test Plan Evaluation

These steps are described one by one.

Step 1. Prior information



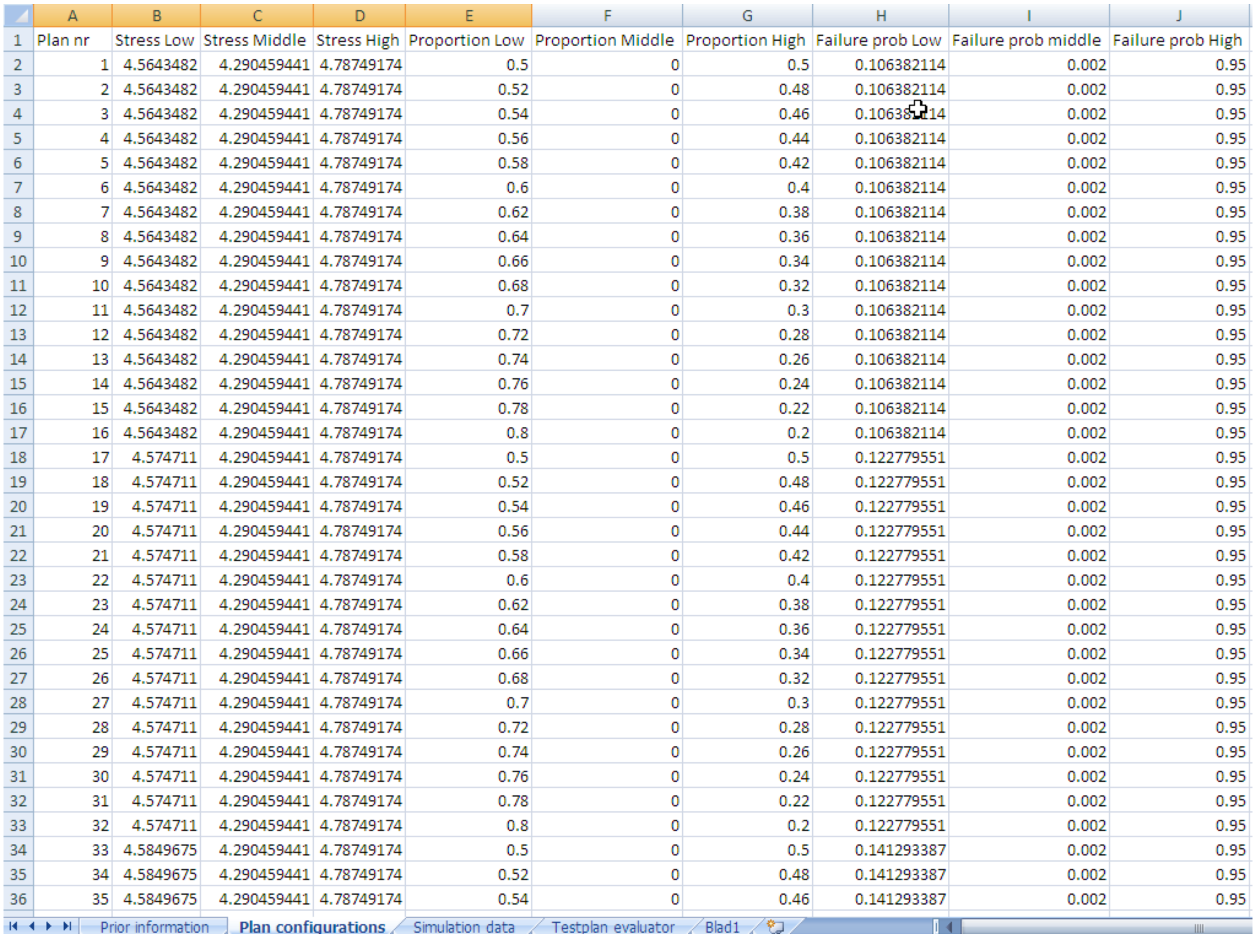
	A	B	C	D	J
1					
2		Relation	Use	High	
3	Stress	Power	73	120	
4	Xuse	4.290459441			Number of plans
5	Xhigh	4.787491743			96
6					
7	Initial estimates	Time	Prob of failure		
8	Xuse	3000	0.002		
9	Xhigh	3000	0.95		
10	sigma	0.294117647			
11					
12					
13	n_simulation	2000			
14					
15	Type plan	Statistical Optimum Plan			
16					
17					
18	beta0	9.83			
19	beta1	-2.15			
20					
21					
22	tc	3000			
23					
24	opt criterion t_p	0.1	9624.107		
25					
26	LB dzeta low	0.542107857			
27	UB dzeta low	0.64475432			
28			Range stress		
29	LB x low	4.559904557	96		
30	UB x low	4.610923165	101		
31					
32	LB pi_low	0.5			
33	UB pi_low	0.8			
34					
35	pi_middle	0			
36					

Based on the optimization criterion the range of the normalized stress values is computed, such that the failure probability is between 1 and 2 times this criterion. The normalized stress is transformed back to the "official" stress and can take only integer values.

Based on the type of test plan, the range of the allocation fraction to the lowest stress level is computed. Together with the range of the lowest stress level values, 96 different settings are possible.

The marked cells are input cells and contain the prior information. Why this information is necessary is explained in section 5.2.1. The values of the other cells are computed based on the input values and the configuration rules.

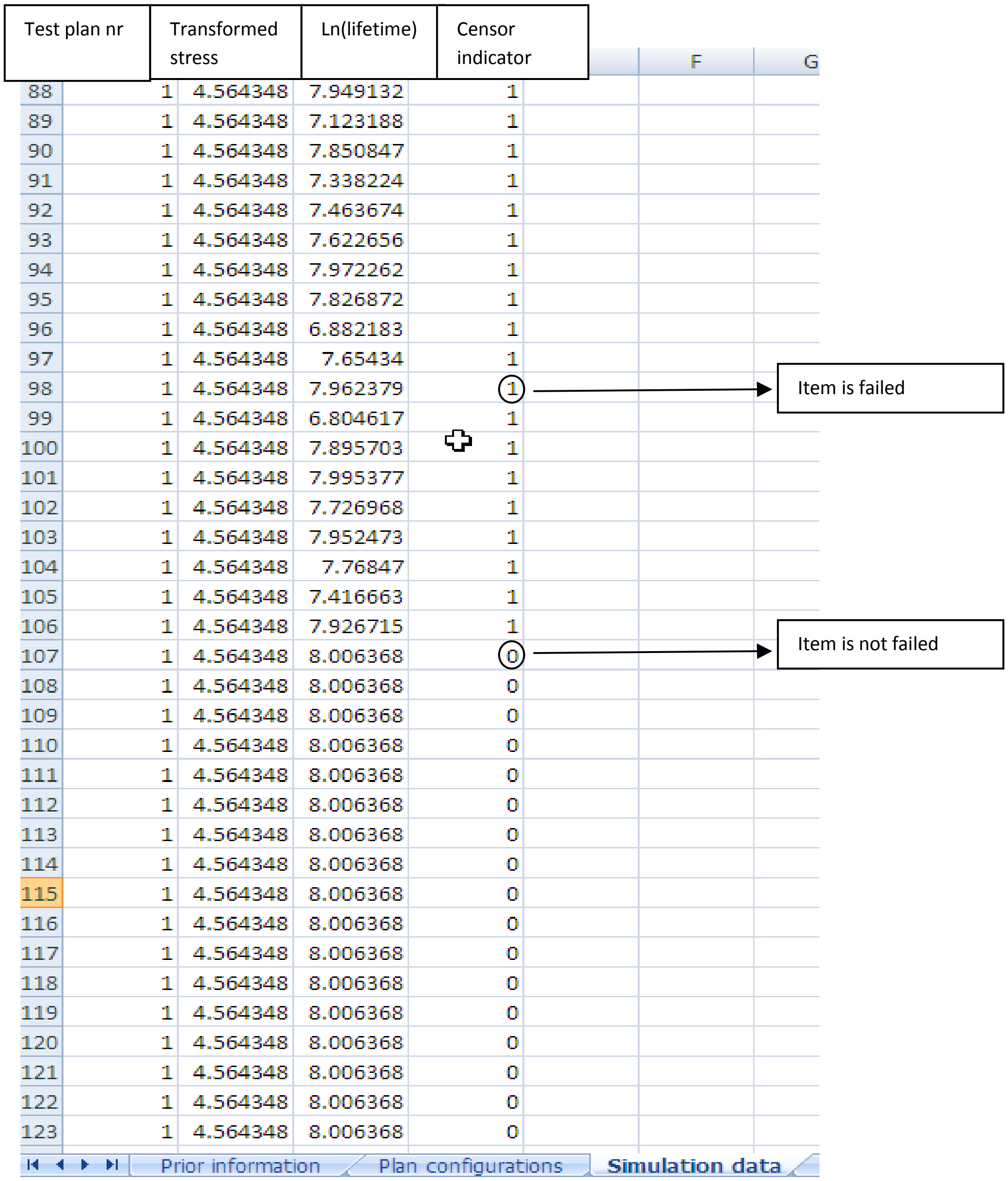
Step 2. Test Plan Configurations



	A	B	C	D	E	F	G	H	I	J
1	Plan nr	Stress Low	Stress Middle	Stress High	Proportion Low	Proportion Middle	Proportion High	Failure prob Low	Failure prob middle	Failure prob High
2	1	4.5643482	4.290459441	4.78749174	0.5	0	0.5	0.106382114	0.002	0.95
3	2	4.5643482	4.290459441	4.78749174	0.52	0	0.48	0.106382114	0.002	0.95
4	3	4.5643482	4.290459441	4.78749174	0.54	0	0.46	0.106382114	0.002	0.95
5	4	4.5643482	4.290459441	4.78749174	0.56	0	0.44	0.106382114	0.002	0.95
6	5	4.5643482	4.290459441	4.78749174	0.58	0	0.42	0.106382114	0.002	0.95
7	6	4.5643482	4.290459441	4.78749174	0.6	0	0.4	0.106382114	0.002	0.95
8	7	4.5643482	4.290459441	4.78749174	0.62	0	0.38	0.106382114	0.002	0.95
9	8	4.5643482	4.290459441	4.78749174	0.64	0	0.36	0.106382114	0.002	0.95
10	9	4.5643482	4.290459441	4.78749174	0.66	0	0.34	0.106382114	0.002	0.95
11	10	4.5643482	4.290459441	4.78749174	0.68	0	0.32	0.106382114	0.002	0.95
12	11	4.5643482	4.290459441	4.78749174	0.7	0	0.3	0.106382114	0.002	0.95
13	12	4.5643482	4.290459441	4.78749174	0.72	0	0.28	0.106382114	0.002	0.95
14	13	4.5643482	4.290459441	4.78749174	0.74	0	0.26	0.106382114	0.002	0.95
15	14	4.5643482	4.290459441	4.78749174	0.76	0	0.24	0.106382114	0.002	0.95
16	15	4.5643482	4.290459441	4.78749174	0.78	0	0.22	0.106382114	0.002	0.95
17	16	4.5643482	4.290459441	4.78749174	0.8	0	0.2	0.106382114	0.002	0.95
18	17	4.574711	4.290459441	4.78749174	0.5	0	0.5	0.122779551	0.002	0.95
19	18	4.574711	4.290459441	4.78749174	0.52	0	0.48	0.122779551	0.002	0.95
20	19	4.574711	4.290459441	4.78749174	0.54	0	0.46	0.122779551	0.002	0.95
21	20	4.574711	4.290459441	4.78749174	0.56	0	0.44	0.122779551	0.002	0.95
22	21	4.574711	4.290459441	4.78749174	0.58	0	0.42	0.122779551	0.002	0.95
23	22	4.574711	4.290459441	4.78749174	0.6	0	0.4	0.122779551	0.002	0.95
24	23	4.574711	4.290459441	4.78749174	0.62	0	0.38	0.122779551	0.002	0.95
25	24	4.574711	4.290459441	4.78749174	0.64	0	0.36	0.122779551	0.002	0.95
26	25	4.574711	4.290459441	4.78749174	0.66	0	0.34	0.122779551	0.002	0.95
27	26	4.574711	4.290459441	4.78749174	0.68	0	0.32	0.122779551	0.002	0.95
28	27	4.574711	4.290459441	4.78749174	0.7	0	0.3	0.122779551	0.002	0.95
29	28	4.574711	4.290459441	4.78749174	0.72	0	0.28	0.122779551	0.002	0.95
30	29	4.574711	4.290459441	4.78749174	0.74	0	0.26	0.122779551	0.002	0.95
31	30	4.574711	4.290459441	4.78749174	0.76	0	0.24	0.122779551	0.002	0.95
32	31	4.574711	4.290459441	4.78749174	0.78	0	0.22	0.122779551	0.002	0.95
33	32	4.574711	4.290459441	4.78749174	0.8	0	0.2	0.122779551	0.002	0.95
34	33	4.5849675	4.290459441	4.78749174	0.5	0	0.5	0.141293387	0.002	0.95
35	34	4.5849675	4.290459441	4.78749174	0.52	0	0.48	0.141293387	0.002	0.95
36	35	4.5849675	4.290459441	4.78749174	0.54	0	0.46	0.141293387	0.002	0.95

This sheet gives an overview of the configurations per test plan. The transformed stress is used instead of the normalized stress, because this is easier for the maximum likelihood calculations and the determination of the standard error of the p^{th} percentile estimate. The failure probabilities are stored to calculate the expected number of failures for the best test plan at the end of the optimization.

Step 3. Simulation



	Test plan nr	Transformed stress	Ln(lifetime)	Censor indicator
88	1	4.564348	7.949132	1
89	1	4.564348	7.123188	1
90	1	4.564348	7.850847	1
91	1	4.564348	7.338224	1
92	1	4.564348	7.463674	1
93	1	4.564348	7.622656	1
94	1	4.564348	7.972262	1
95	1	4.564348	7.826872	1
96	1	4.564348	6.882183	1
97	1	4.564348	7.65434	1
98	1	4.564348	7.962379	①
99	1	4.564348	6.804617	1
100	1	4.564348	7.895703	1
101	1	4.564348	7.995377	1
102	1	4.564348	7.726968	1
103	1	4.564348	7.952473	1
104	1	4.564348	7.76847	1
105	1	4.564348	7.416663	1
106	1	4.564348	7.926715	1
107	1	4.564348	8.006368	⓪
108	1	4.564348	8.006368	0
109	1	4.564348	8.006368	0
110	1	4.564348	8.006368	0
111	1	4.564348	8.006368	0
112	1	4.564348	8.006368	0
113	1	4.564348	8.006368	0
114	1	4.564348	8.006368	0
115	1	4.564348	8.006368	0
116	1	4.564348	8.006368	0
117	1	4.564348	8.006368	0
118	1	4.564348	8.006368	0
119	1	4.564348	8.006368	0
120	1	4.564348	8.006368	0
121	1	4.564348	8.006368	0
122	1	4.564348	8.006368	0
123	1	4.564348	8.006368	0

This sheet contains the simulation data. First of all, per replication a set of n (n=sample size) lifetimes under normal use conditions are generated based on the model derived from the prior information. Per test plan configuration, a fraction π_L of the items is allocated to the lowest stress level, π_M to the middle stress level and π_H to the highest stress level. The lifetimes of these items are transformed to the lifetimes observed under the stress corresponding to the level. If this accelerated lifetime of an item is larger than the test time, the item is censored and the test time becomes the censored lifetime of the product.

An example of the generation of the lifetime data based on the case:

$\ln(tp) = 7.684 + 2.15\xi + \Phi_{SEV}^{-1}(p) * 0.294$

The lnlifetimes are generated under normal use conditions, so $\xi = 1$. p is a random value from a uniform distribution between 0 and 1. Based on this life-stress relation: $\ddot{lnlifetime}(i) = 9.834 + \ln(-\ln(1 - \mathrm{p})) * 0.294 \quad \forall i = 1 \ldots n$, n random lnlifetimes under normal use conditions are generated.

For a set of different test plan configurations these generated lifetimes are accelerated based on the life-stress model and the observed lifetimes per stress can be determined. We show this for one specific test plan configuration. The configurations we use are:
stress$_{\mathrm{Low}}$=96W ($\xi = 0.45$) stress$_{\mathrm{Middle}}$=107W ($\xi = 0.22$) and stress$_{\mathrm{High}}$=120W ($\xi = 0$)
allocation π_{Low}=0.33, π_{Middle}=0.2 and π_{High}=0.47
sample size: 2000 items
test termination time, t_c: 3000 hours

The first 660 items (0.33*2000) are allocated to the lowest stress level and the lnlifetimes for these items become: $lnlifetime(i) = \min\,[\ln\,(tc), (\ddot{lnlifetime}(i) - 2.15 * (1 - 0.45)] \quad \forall i = 1 \ldots 660$
The lnlifetimes for the other stress levels become:
$lnlifetime(i) = \min\,[\ln\,(tc), \left(\ddot{lnlifetime}(i) - 2.15 * (1 - 0.22)\right)] \quad \forall i = 661 \ldots 1060$
$lnlifetime(i) = \min\,[\ln\,(tc), \left(\ddot{lnlifetime}(i) - 2.15 * (1 - 1)\right)] \quad \forall i = 1061 \ldots 2000$

This process is repeated for each replication. So within a replication all test plan configurations use the same lifetime data under normal use conditions, but between replications these lifetime data is different.

This lifetime data per test plan configuration and per replication is used in STATA to estimate the parameter values per test plan and compute the variance and covariance of these estimated.

Step 4. Maximum Likelihood Estimation

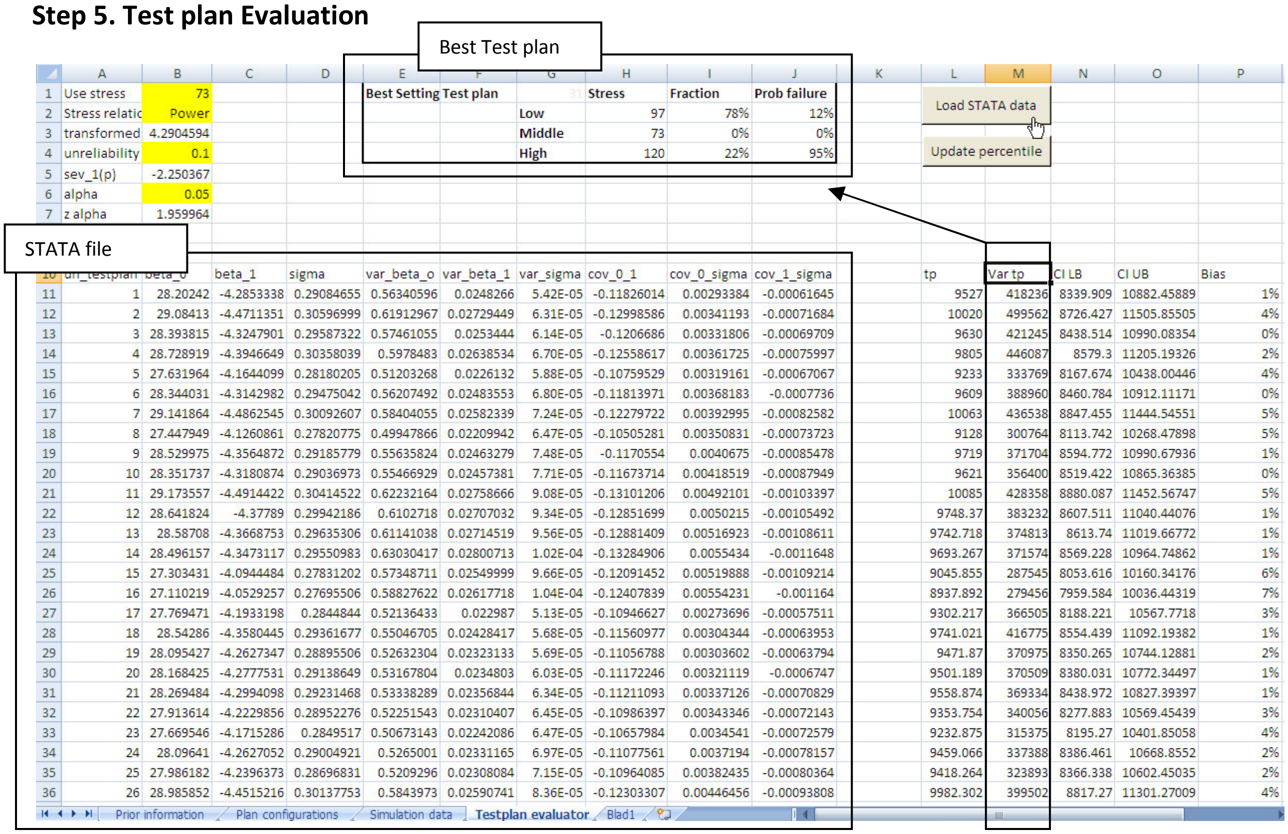
The program STATA uses is:

```
insheet using "C:\TestDesign\Simulation\simulationdata.csv", comma clear
rename v1 un_testplan
rename v2 x
rename v3 lnlifetime
rename v4 failed


file open myfile9 using "C:\TestDesign\Simulation\modelfits.txt", replace write
file write myfile9 "un_testplan;beta_0;beta_1;sigma;var_beta_o;var_beta_1;var_sigma;cov_0_1;cov_0_sigma;cov_1_sigma;" _n

foreach num of numlist 1/480 {
                              ml model lf weibull (lnlifetime failed= x) /s if un_testplan == `num'
                              ml init x=-4.326 eq1:_cons=28.395 s:_cons=0.294
                              ml maximize, nolog
                              matrix b = e(b)
                              matrix v = e(V)
                              file write myfile9 (`num') ";" (el(b,1,2)) ";" (el(b,1,1)) ";" (el(b,1,3)) ";"
                              file write myfile9 (el(v,2,2)) ";" (el(v,1,1)) ";" (el(v,3,3)) ";" (el(v,2,1)) ";" (el(v,3,2)) ";" (el(v,3,1)) ";" _n
                        }

file close myfile9
```

The function ml init is used to speed up the MLE process. The init values are based on the prior information. The use of these true values does not affect the parameter estimations, but speeds up the process, because STATA starts searching for values close to the estimates.

The output is a text file with per test plan, the parameter estimated and the covariances. This is file loaded in Microsoft Office Excel.

Step 5. Test plan Evaluation



	A	B
1	Use stress	73
2	Stress relatic	Power
3	transformed	4.2904594
4	unreliability	0.1
5	sev_1(p)	-2.250367
6	alpha	0.05
7	z alpha	1.959964

Best Setting	Test plan	31	Stress	Fraction	Prob failure
		Low	97	78%	12%
		Middle	73	0%	0%
		High	120	22%	95%

10	un_testplan	beta_o	beta_1	sigma	var_beta_o	var_beta_1	var_sigma	cov_0_1	cov_0_sigma	cov_1_sigma	tp	Var tp	CI LB	CI UB	Bias
11	1	28.20242	-4.2853338	0.29084655	0.56340596	0.0248266	5.42E-05	-0.11826014	0.00293384	-0.00061645	9527	418236	8339.909	10882.45889	1%
12	2	29.08413	-4.4711351	0.30596999	0.61912967	0.02729449	6.31E-05	-0.12998586	0.00341193	-0.00071684	10020	499562	8726.427	11505.85505	4%
13	3	28.393815	-4.3247901	0.29587322	0.57461055	0.0253444	6.14E-05	-0.1206686	0.00331806	-0.00069709	9630	421245	8438.514	10990.08354	0%
14	4	28.728919	-4.3946649	0.30358039	0.5978483	0.02638534	6.70E-05	-0.12558617	0.00361725	-0.00075997	9805	446087	8579.3	11205.19326	2%
15	5	27.631964	-4.1644099	0.28180205	0.51203268	0.0226132	5.88E-05	-0.10759529	0.00319161	-0.00067067	9233	333769	8167.674	10438.00446	4%
16	6	28.344031	-4.3142982	0.29475042	0.56207492	0.02483553	6.80E-05	-0.11813971	0.00368183	-0.0007736	9609	388960	8460.784	10912.11171	0%
17	7	29.141864	-4.4862545	0.30092607	0.58404055	0.02582339	7.24E-05	-0.12279722	0.00392995	-0.00082582	10063	436538	8847.455	11444.54551	5%
18	8	27.447949	-4.1260861	0.27820775	0.49947866	0.02209942	6.47E-05	-0.10505281	0.00350831	-0.00073723	9128	300764	8113.742	10268.47898	5%
19	9	28.529975	-4.3564872	0.29185779	0.55635824	0.02463279	7.48E-05	-0.1170554	0.0040675	-0.00085478	9719	371704	8594.772	10990.67936	1%
20	10	28.351737	-4.3180874	0.29036973	0.55466929	0.02457381	7.71E-05	-0.11673714	0.00418519	-0.00087949	9621	356400	8519.422	10865.36385	0%
21	11	29.173557	-4.4914422	0.30414522	0.62232164	0.02758666	9.08E-05	-0.13101206	0.00492101	-0.00103397	10085	428358	8880.087	11452.56747	5%
22	12	28.641824	-4.37789	0.29942186	0.6102718	0.02707032	9.34E-05	-0.12851699	0.0050215	-0.00105492	9748.37	383232	8607.511	11040.44076	1%
23	13	28.58708	-4.3668753	0.29635306	0.61141038	0.02714519	9.56E-05	-0.12881409	0.00516923	-0.00108611	9742.718	374813	8613.74	11019.66772	1%
24	14	28.496157	-4.3473117	0.29550983	0.63030417	0.02800713	1.02E-04	-0.13284906	0.0055434	-0.0011648	9693.267	371574	8569.228	10964.74862	1%
25	15	27.303431	-4.0944484	0.27831202	0.57348711	0.02549999	9.66E-05	-0.12091452	0.00519888	-0.00109214	9045.855	287545	8053.616	10160.34176	6%
26	16	27.110219	-4.0529257	0.27695506	0.58827622	0.02617718	1.04E-04	-0.12407839	0.00554231	-0.001164	8937.892	279456	7959.584	10036.44319	7%
27	17	27.769471	-4.1933198	0.2844844	0.52136433	0.022987	5.13E-05	-0.10946627	0.00273696	-0.00057511	9302.217	366505	8188.221	10567.7718	3%
28	18	28.54286	-4.3580445	0.29361677	0.55046705	0.02428417	5.68E-05	-0.11560977	0.00304344	-0.00063953	9741.021	416775	8554.439	11092.19382	1%
29	19	28.095427	-4.2627347	0.28895506	0.52632304	0.02323133	5.69E-05	-0.11056788	0.00303602	-0.00063794	9471.87	370975	8350.265	10744.12881	2%
30	20	28.168425	-4.2777531	0.29138649	0.53167804	0.0234803	6.03E-05	-0.11172246	0.00321119	-0.0006747	9501.189	370509	8380.031	10772.34497	1%
31	21	28.269484	-4.2994098	0.29231468	0.53338289	0.02356844	6.34E-05	-0.11211093	0.00337126	-0.00070829	9558.874	369334	8438.972	10827.39397	1%
32	22	27.913614	-4.2229856	0.28952276	0.52251543	0.02310407	6.45E-05	-0.10986397	0.00343346	-0.00072143	9353.754	340056	8277.883	10569.45439	3%
33	23	27.669546	-4.1715286	0.2849517	0.50673143	0.02242086	6.47E-05	-0.10657984	0.0034541	-0.00072579	9232.875	315375	8195.27	10401.85058	4%
34	24	28.09641	-4.2627052	0.29004921	0.5265001	0.02331165	6.97E-05	-0.11077561	0.0037194	-0.00078157	9459.066	337388	8386.461	10668.8552	2%
35	25	27.986182	-4.2396373	0.28696831	0.5209296	0.02308084	7.15E-05	-0.10964085	0.00382435	-0.00080364	9418.264	323893	8366.338	10602.45035	2%
36	26	28.985852	-4.4515216	0.30137753	0.5843973	0.02590741	8.36E-05	-0.12303307	0.00446456	-0.00093808	9982.302	399502	8817.27	11301.27009	4%

The data file of STATA is used in Microsoft Office Excel to select the best test plan configurations based on the minimal standard error of the pth percentile estimate. A functionality is created that can read the data file of STATA and computes an estimation of the percentile of interest and the standard error of this percentile per test plan.

The variance of the pth percentile is calculated based on equation 27:

$$var\left(g(\widehat{\boldsymbol{\theta}})\right) = \sum_{i=1}^{n}\left(\frac{\partial g}{\partial \theta_i}\right)^2 var(\widehat{\theta_\iota}) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(2 * \frac{\partial g}{\partial \theta_i}\frac{\partial g}{\partial \theta_j} cov(\widehat{\theta_\iota}, \widehat{\theta_J}))$$

$$g(\widehat{\boldsymbol{\theta}}) = t_p(\widehat{\gamma_0}, \widehat{\gamma}_1, \hat{\sigma}) = \exp\,(\widehat{\gamma_0} + \widehat{\gamma}_1 x + \Phi_{SEV}^{-1}(p) * \hat{\sigma})$$

$$var(t_p) = {t_p}^2 * var(\gamma_0) + (t_p * x)^2 * var(\gamma_1) + \left(t_p * \Phi_{SEV}^{-1}(p)\right) * var(\sigma) + 2 * \left(t_p * t_p * \Phi_{SEV}^{-1}(p)\right) * cov(\gamma_0, \sigma) + 2 * \left(t_p * t_p * x\right) * cov(\gamma_0, \gamma_1) + 2 * \left(t_p * t_p * \Phi_{SEV}^{-1}(p) * x\right) * cov(\gamma_{1,}\sigma)$$

The average variance over the replications per test plan configuration is calculated, and based on this the average standard error. Microsoft Office Excel selects the test plan with the smallest average standard error. Based on the saved settings per test plan, the configurations of this best test plan are displayed in the block “Best Test Plan”. Also the confidence interval of the percentile of interest is displayed and the sample size needed to obtain a certain bound ratio is computed based on the expected variance and equation 26. The confidence level, optimization criterion and normal use stress level can be changed, to optimize other percentiles or different use levels.

Validation of Optimization Method and Improvements

The different parts of the simulation model are validated with use of Minitab and the software package ALTA, Accelerated Life Testing data Analysis. Minitab is used to validate the MLE and the estimation of the variances. ALTA is used to validate the best test plan settings. Below the functionality of both programs and the validations are described in more detail.

Minitab has a functionality to estimate parameters of life-stress models and a given stress-relation based on MLE. The values that are estimated based on Minitab are the same as based on the program I made to use MLE within STATA. Also the variance and the confidence interval of the p^{th} percentile calculated in Excel are the same as the one calculated by Minitab. For this research, Excel and STATA are used instead of Minitab, because they can be adapted to more generic situations by programming and Minitab cannot. Minitab has some functionality to develop a test plan, but only for simple situations and stress levels cannot be computed, only the allocation over the stress levels. Therefore a more complete procedure is developed in Excel, with use of STATA.

The resulting best settings for a certain type of test plan is compared with the test plan ALTA suggest for the same situation. ALTA is a software package designed for quantitative accelerated life testing data analysis. The package has the functionality to design test plans for one or two stress factors based on some specified stress relations and test plan types. Extensions to more complicated situations are not possible, therefore this package cannot be used for this research and an own method has to be developed. But for the simple scenario's, the results of the Excel simulation and the ALTA results can be compared. In most of the situations the results are the same. Sometimes the solution of the simulation changes if you do the simulation again. We tried to overcome these problems by:

- Use of the same failure times under normal circumstances per simulation run for all test plan configurations, to make a fair comparison and to reduce variance.

- Make use of a sample size that generates stable results, this means that the estimation of a percentile for a test plan configuration for one replication does not deviate that much from the percentile estimation of another replication for the same configurations. The effect of the estimation error on the variance is minimized in this case. Figure 20 shows the differences in estimation error (true t_p – estimated t_p) for different sample sizes. Ten replications are used to show the differences in estimations between the different runs. The values of the percentile estimates are sorted to make the effect of the sample size increase more clear. Based on the results, we choose to use a sample size of 2,000 items per simulation run. We can conclude that the differences in estimations become small for a sample size of 2,000, and the difference in estimation error between a sample size of 1,000 and a sample size of 2,000 is not that large that a further reduction by increasing the sample size can be expected.

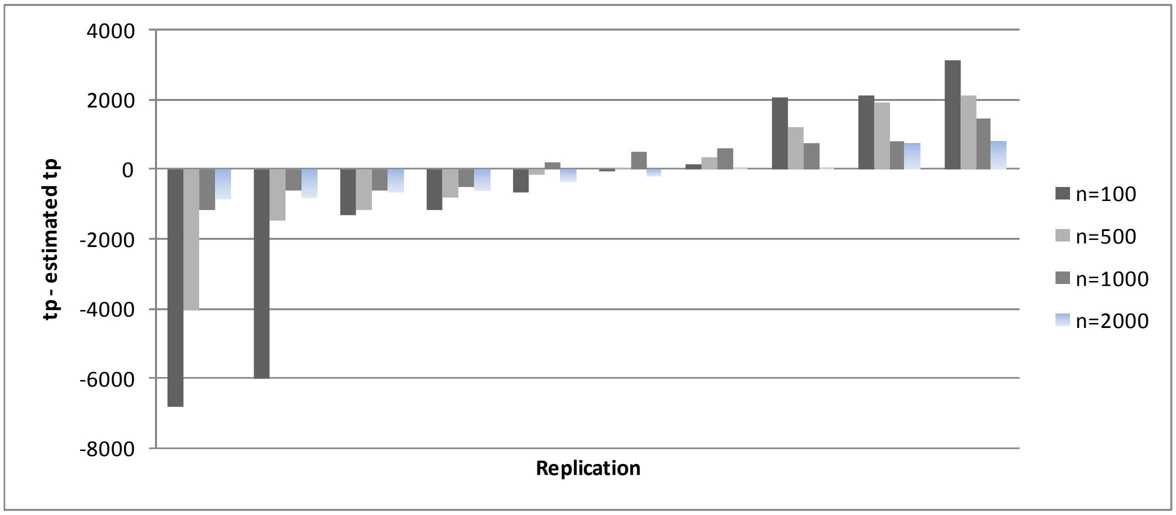


Figure 20: Sorted Estimation Error for Ten Replications per Sample Size

- Make use of enough replications to exclude the dependency of the variance on the simulation values and the parameter estimates. To test the number of replications that are necessary to generate stable results and be able to conclude which test plan settings minimize the var(t_p), the average variances over different numbers of replications are calculated. Figure 21 shows the result of such experiment. Based on the figure, we can conclude that if a small sample size is used (<5) the best variance is very dependent on the simulated values and it is difficult to find the best test plan settings. For a number of replications larger than 38, the variance remains almost constant if the number of replications increases. From the figure can be concluded that at least 10 replications must be used to obtain stable results.

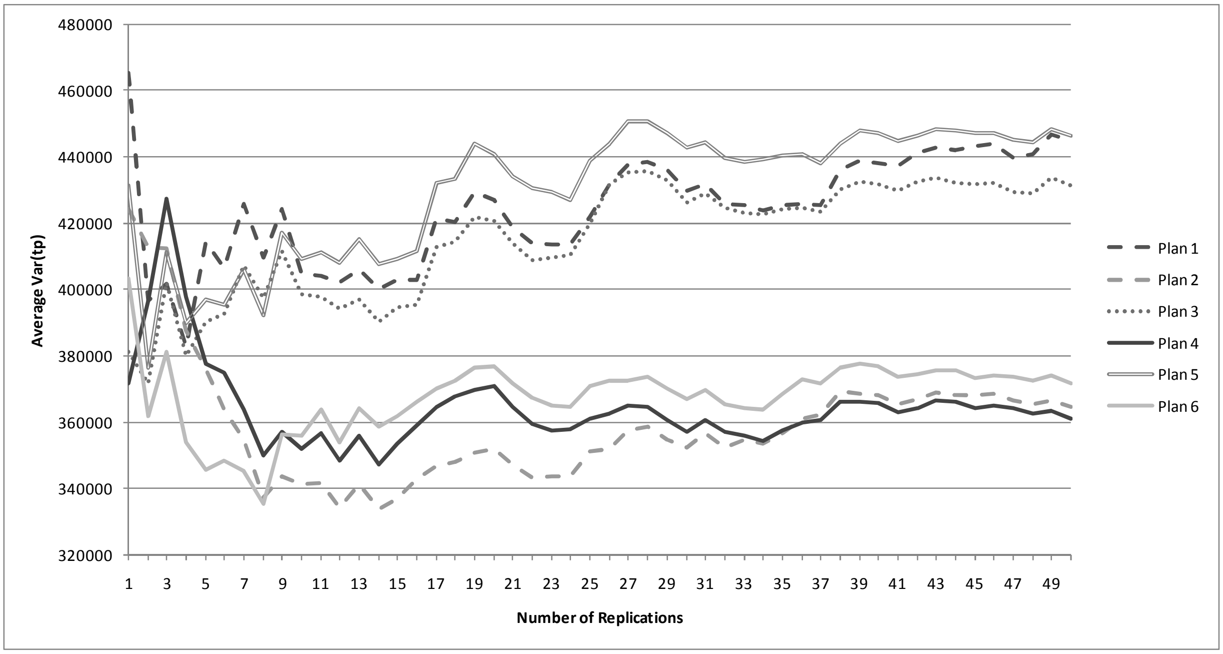


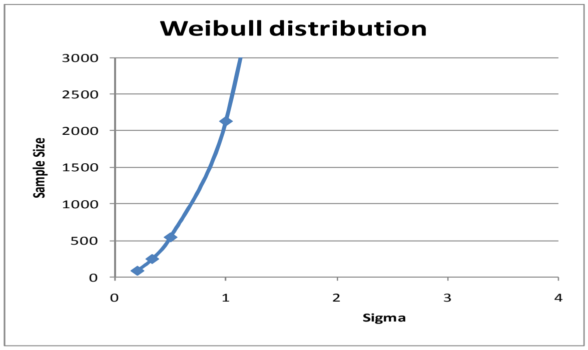
Figure 21: Average Var(t_p) over a Number of Replications

Appendix K. Influence Scale Parameter on Sample Size

The scale parameter of a distribution has no influence on the test plan configurations, because the variance increases or decreases proportional to the parameter change. This means that for a given test plan, the variance increases or decreases if the scale parameter changes, but relative to other test plans, there are no differences. Because the variance is changing, also the sample size needed to obtain certain precision changes. We show this effect for the Weibull, Lognormal and the Normal distribution and a precision factor of 1.5. The example based on the case is used to determine the values. For other situations, the sample size values will change, but the effect of the scale parameter on these values do not change.

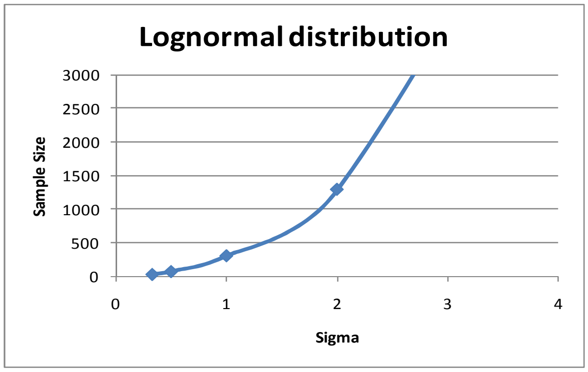
Weibull Distribution

beta	sigma (1/beta)	SS
0.5	2	9600
1	1	2130
2	0.5	540
3	0.333333	242
5	0.2	83



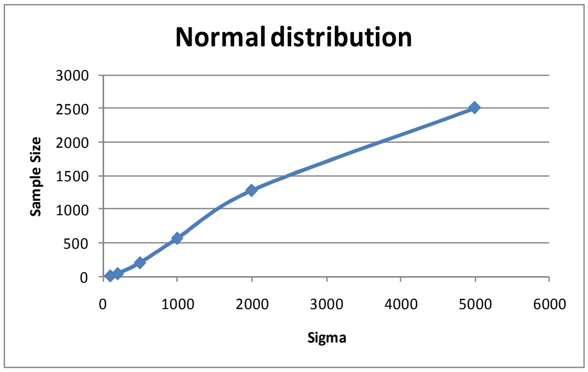
Lognormal Distribution

sigma	SS
5	9300
2	1288
1	310
0.5	77
0.33	34



Normal Distribution

sigma	SS
5000	2514
2000	1284
1000	571
500	208
200	45
100	12



Appendix L. Percentile Expression Two Failure Modes

$$F(t) = p = 1 - \left[1 - 1 - \exp\left(-\exp\left(\frac{\ln(t_p) - \mu_1(x)}{\sigma_1}\right)\right)\right] * \left[1 - 1 - \exp\left(-\exp\left(\frac{\ln(t_p) - \mu_2(x)}{\sigma_2}\right)\right)\right] \Rightarrow$$

$$1 - p = \exp\left(-\exp\left(\frac{\ln(t_p) - \mu_1(x)}{\sigma_1}\right)\right) * \exp\left(-\exp\left(\frac{\ln(t_p) - \mu_2(x)}{\sigma_2}\right)\right)$$

$$= \exp\left(-\exp\left(\frac{\ln(t_p) - \mu_1(x)}{\sigma_1}\right) - \exp\left(\frac{\ln(t_p) - \mu_2(x)}{\sigma_2}\right)\right) \Rightarrow$$

$$\ln(1 - p) = -\exp\left(\frac{\ln(t_p) - \mu_1(x)}{\sigma_1}\right) - \exp\left(\frac{\ln(t_p) - \mu_2(x)}{\sigma_2}\right) \Rightarrow$$

$$\ln(-\ln(1 - p)) = \frac{\ln(t_p) - \mu_1(x)}{\sigma_1} + \frac{\ln(t_p) - \mu_2(x)}{\sigma_2} = \frac{\ln(t_p)}{\sigma_1} - \frac{\mu_1(x)}{\sigma_1} + \frac{\ln(t_p)}{\sigma_2} - \frac{\mu_2(x)}{\sigma_2} \Rightarrow$$

$$\frac{\ln(t_p)}{\sigma_1} + \frac{\ln(t_p)}{\sigma_2} = \ln(-\ln(1 - p)) + \frac{\mu_1(x)}{\sigma_1} + \frac{\mu_2(x)}{\sigma_2}$$

An exact expression for t_p only exists if $\sigma_1 = \sigma_2$.

Appendix M. Example of a Test Plan Design

Two failure modes, one stress factor and one product design variant scenario

---

Appendix N.Lifetime Data UV-case Best Test Plan

Appendix O. UV-case Worst Test Plan

The test settings for the worst test plan become:

Stress level	Design	Number of replications
45	0000	4
45	1001	4
45	0101	3
45	1100	3
45	0011	3
45	1010	3
45	0110	3
45	1111	3
Total		26

Stress level	Design	Number of replications
60	0000	2
60	1001	2
60	0101	2
60	1100	2
60	0011	2
60	1010	2
60	0110	2
60	1111	2
		16

Stress level	Design	Number of replications
80	0000	4
80	1001	4
80	0101	5
80	1100	5
80	0011	5
80	1010	5
80	0110	5
80	1111	5
		38

The lifetimes generated with the worst test plan are:
