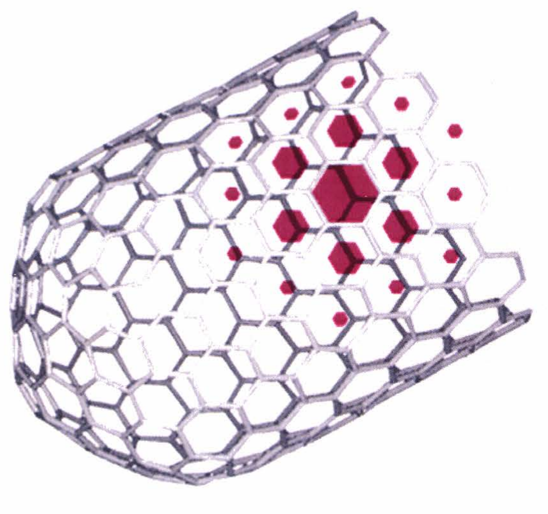
MASTER

Modeling reliability growth of a sample loader

Schroemges, R.P.M.

Award date:
2007

Modeling reliability growth of a Sample loader



FEI COMPANY™
TOOLS FOR NANOTECH

TU/e

Modeling reliability growth of an Sample loader

René Schroemges
Student ID: 499518

Eindhoven University of Technology
Department of Technology Management
Quality and Reliability Engineering

Supervisors
Dr. Ir. P.J.M. Sonnemans
Dr. Ir. J.L. Rouvroye

FEI Company
Electron Optics Product Division

Supervisors
Ir. J.J. ten Thije
Dr. Ir. M.J. Verheijen

Abstract

This report presents the results of a graduation project being carried out at FEI Company in Acht. The project encompasses the modeling and prediction of the reliability of a module during its development. Existing models are investigated, and the relevant models are selected. These models are applied on the data available from the development of the Sample loader; in specific on three of the lifecycle tests. Based on the results the predicted reliability of the Sample loader is presented as well as a recommendation.

Management summary

This report is the result of a graduation project executed at FEI Company in Acht. FEI Company is a large organization developing, building and servicing complex high-end products. The company is situated in the nanotechnology industry. In this industry there are trends towards:

- a growing market,
- higher product utilization,
- increasing product functionality,
- higher levels of software automation,
- large pressure on the time to market,
- an increasing pressure on cost of ownership.

All these factors lead to a larger emphasis on product reliability.

This is a project within the scope of the research fields of the Quality and Reliability Engineering (QRE) department of the Technology Management department at the Technical University of Eindhoven. The goal of this research group is to create methods to predict the reliability of a product during its development and the early stages of the product introduction. FEI wants to give more accurate predictions of reliability numbers during the development phases of its products.
The goal of this project was to predict the reliability of an Sample loader. FEI wants to present scientifically based numbers to customers. The Technical University wants to know the applicability and accuracy of the models used.

To answer these research goals on a system level a reliability growth study is performed. From the available literature the potential models for predicting reliability are investigated. Then the data available from life cycle tests of an Sample loader is used. With a reliability growth model the improvement of the reliability can be modeled. These models are based on the assumption, that the failures found during testing will be solved by changing the design. By elimination the occurrence of failures is decreased and the overall reliability is improved. This kind of strategy is referred to as Test Analyze And Fix (TAAF). In case of a failure, the failure data, including modes of failure, time to failure, and any other relevant information, are collected and analyzed by engineers to discover the cause of failure. The cumulative numbers of failures are modeled against the cumulative number of cycles.

The investigated module (the Sample loader) is a complex repairable system. In case of a failure the Sample loader is restored so that it will perform the intended function without replacing the entire system. Several models are available, based on different mathematical grounds. The most applicable models are those based on stochastical point processes.
Before these models are applied, a general procedure for analyzing data is followed to select the proper type of model. This procedure is based on the results of the Laplace trend test and the kind of repair actions. If the Laplace trend test indicates there is no trend in the failure pattern, a Homogeneous Poisson Process (HPP) or a Renewal Process (RP) is recommended. The choice between these models depends on the distribution of the failure times. If the times between failures are exponentially distributed, the Homogeneous Poisson Process should be used; otherwise a Renewal Process should be used. Since all failure data used in this thesis show trend, both types of models (HPP and RP) are outside the scope of this research.
If the Laplace test does show a trend in the failure pattern, Non Homogeneous Poisson Process can be used. In this research three models were investigated; the Power Law Process, the Log Linear Process and the Log Power Process. In literature the Log Power Process is referred to as a model which has to be applied for software development. If an error in software would be solved it is gone forever, in contrary to hardware. However goal of the reliability improvement during the development phase is to eliminate errors for ever. Therefore this model was investigated as well.

Two methods to fit the before mentioned models on the data are applied; the Least Square method and the Maximum Likelihood method. The Least Square method fits a model on the dataset, by minimizing the sum of the squared deviations between the model and each data point (fit of the cumulative number of failures). The Maximum Likelihood estimation is based on fitting the failure intensity of the underlying Poisson Process (fit of the failure rate).

The discrepancy between the data and the fit is measured by calculating the goodness-of-fit. This goodness of fit describes the ability of the chosen model to fit the current data. The ability of the models to predict the future performance is determined by calculating the Mean Error of Prediction (MEOP).

The Sample loader tests
The reliability of the Sample loader was tested by running three types of cycle test: Cassette test, Cartridge test and Mapping test. Prior to applying the suggested models, first a validation step is required. This is done by using CEI/IEC 1164 [CEI95]. During the validation it showed the dataset from the Cartridge test and the Mapping test had to be split in two test phases (A and B)
From the analysis it can be concluded, applying the Least Square Method on the Power Law Process and the Log Linear Process provides the best fit.

Cassette test
The Cassette test data showed some initial problems. After solving these issues, a quite reliable process was obtained. Such behavior can be modeled best by using a Log Linear Process. The cassette loading is now so robust; it has no significant influence on the reliability of the Sample loader ($MTBF_{i,\ Cassette} \approx$ 8175 cycles, CI-range at a 90% confidence level ≈ [5093, 18476]).

Cartridge test
The Cartridge test revealed many issues. Both processes (Power Law Process and Log Linear Process) fit well to the data. Closer look at the fit shows that the Log Linear Process provides a better fit towards the end of the data set. This leads to a more reliable prediction of the MTBF ($MTBF_{i,\ Cartridge} \approx$ 1038 cycles, CI-range at a 90% confidence level ≈ [906, 1607]). By extrapolating the fitted Log Linear Process to 27.000 cycles (twice the amount of cycles during test phase B) a prediction of the MTBF can be presented ($MTBF_{27.000} \approx$ 801 cycles). This prediction is based on the assumption the current development effort is maintained and no new features are added or removed.

Mapping test
The mapping test shows a large improvement comparing test phase A an B. Both processes (Power Law Process and Log Linear Process) give similar results. The Power Law provides a better fit towards the end of the data set. This leads to a more reliable prediction of the MTBF ($MTBF_{i,\ Mapping} \approx$ 939 cycles, CI-range at a 90% confidence level ≈ [727, $8.47 \cdot 10^6$]). By extrapolating the fitted Log Linear Process to 324.000 cycles (twice the amount of cycles during test phase B) a prediction of the MTBF can be presented ($MTBF_{324.000} \approx$ 1414 cycles). Also based on the assumption the current development effort is maintained and no new features are added or removed.

MTBF of the Sample loader
The overall MTBF of the Sample loader can be calculated by summing the weighted failure rates of each subsequent step. This assumption holds since the three processes are serial processes and mutually independent. These weight factors are dependent on the use case of the customer. As an example the performance of a customer loading each time twelve cartridges with one cassette, mapping each cartridge, is presented. The MTBF of the Sample loader then can be calculated by summing the weighted values of the failure rate. The MTBF of the Sample loader is approximately 120 cycles. By continuing the current development effort and adding or removing no features, the MTBF will be approximately 500 cycles after another test phase of the same length as test phase B.

Recommendations

The research led to a preferred method for estimating the best fit, the Least Squares Method. Also it showed the Power Law Process and the Log Linear Process lead to the best results. Though for each test is has to be evaluated which test fits the data best and thus is able to present a reliable prediction. It is not possible to base this conclusion entirely on the goodness-of-fit (R^2). Especially the 'quality' of the fit at the end of the dataset will be leading. An engineering analysis is required to select the best model.

The data collection can be improved by predefining a template before the actual testing is started. Currently the data is collected by manual entry in an Excel-sheet, this requires a large effort to create a usable data set. An even larger improvement would be creating automatically generated log-files. The time of failure would be unambiguous, but also the diagnosis time would be improved. System parameters can give more insight in the system behavior over time, by measuring for example driver currents in time, or pump down times, the real root causes can be revealed. Also the changes of the configuration (added or removed features) are not captured very well.

The size of the confidence intervals of the MTBF can be decreased by increasing the number of test cycles. This could be done by testing on multiple Sample loaders, or continuously cycle testing rather than only overnight testing.

Contents

2 Project description

2.1 Research goals

The Sample loader will be the sample manipulation system used by FEI, giving customers the opportunity to seamlessly transfer their samples from one system (system type) to another. Due to its pivotal position in the product portfolio a high reliability will be required. Data will be obtained by the Alfa system at R&D. The main research goal of the project is:

Create a model to predict the reliability of the Sample loader.

This research goal has to answer the three main research questions below:

- Which methods are available to predict the reliability of the Sample loader during its development?
- What is the reliability of the Sample loader?
- How accurate are the presented reliability numbers?

The main goal can be divided into University and FEI goals. Below the specific questions regarding these goals are summarized.

University goals

The research under the heading of the Quality and Reliability Engineering deals with modeling and analysis of product quality and reliability. The goals of this project fit within the RAMS I research[1] performed at QRE (Quality and Reliability Engineering). The main topic of the research in this sub-program is to predict the reliability during the early stages of product development. Specific research questions to be answered within this project are:

- Which Reliability Growth models are available in the literature,
  and what type of problem does each model fit best?
- What are the differences between the models?
  Which boundary conditions are of interest?
- How accurate/reliable are these models?

FEI Goals

- Creating a Reliability Growth model that shall have the capability to enable a decision either to:
  - continue the current tests;
  - increase the test capacity (thus increase the speed of reliability growth);
  - stop testing and initiate a major redesign;
  - predict the reliability in time.
- Visualize the progress of the development project with respect to the reliability of the product/module.
- Visualize the largest contributor to reliability issues.
  In this way development can be targeted towards the biggest impact issue and one is able to predict the effects of an improvement (setting target /goals for subproject (SMART[2])).
- Present (scientifically founded) reliability numbers to customers.
- Present possibility to estimate the MTBF[3] for a specific customer with a specific use case.
- Define models which can be used within FEI.
- The models have to have an underlying physical relevance (more than just a mathematical form)

[1] RAMS = Reliability, Availability, Maintainability & Safety
[2] SMART = Specific, Measurable, Achievable, Relevant & Time framed
[3] MTBF = Mean Time Between Failure

2.2 Research approach

In figure 2.1 the structure of this thesis is presented in a graphical way. Each step is shown in a distinct box. The relation of these steps and the chapters in this report are presented by the dotted lines surrounding the boxes.

In the **first step** the background of where this research is performed is presented. To do this a brief profile of FEI Company is sketched and the main product portfolio is described. In the **second step** the actual graduation project is described by defining the actual research goals and questions. In doing this the boundary conditions of this thesis are defined.
In the **third step** the literature on which this thesis is founded is discussed. After defining reliability, a broad scope analysis of reliability models is performed. From these the type of models best fit to the FEI specific case are described in detail. The **fourth step** zooms in on the actual research object; the Sample loader. In this step the functions of this module is explained and the data collection means are presented. The **fifth step** then is the analysis of the data collected with the models proposed during the third step. Based on these analyses in the **sixth step** the conclusions and recommendation are made.

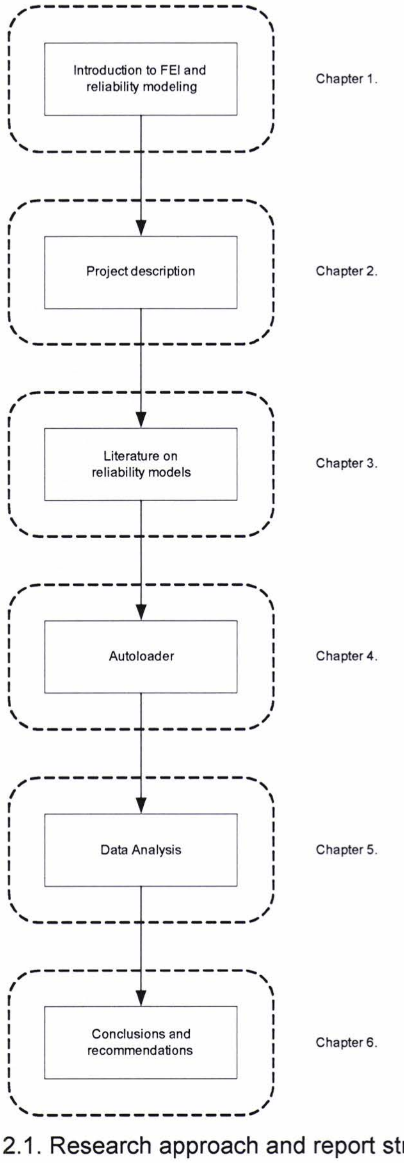


Figure 2.1. Research approach and report structure

3 Literature on reliability models

In the first section the definition 'reliability' is explained. First the Weibull distribution is presented, then the influence of the repair process is explained. In the second section three different approaches how to model the reliability of a repairable system is presented. This overview is used to select a method which will be further explained. In section 3.3 an overview of Reliability Growth Models is presented. This overview is used to select the models suited for the data which shall be analyzed. Three of these models are discussed more in detail in section 3.6. For each model the intensity function, the expected number of failures and the overall properties of the models are presented. In section 3.6.4 two methods to determine the parameters for the chosen models are described. Their applicability is demonstrated on one of the models. Afterwards the parameter estimates are presented in section 3.6.5. Finally in this chapter two methods are presented to evaluate the accuracy of fit of the models.

3.1 Reliability

Before getting into reliability growth first a few basic definitions are presented, starting with the probably most frequent used word in this report: "Reliability". The IEEE (The Institute of Electrical and Electronics Engineers) defines it as "... the ability of a system or component to perform its required functions under stated conditions for a specified period of time." Mathematically this can be described as:

$$R(t) = P(T > t) = \left\{ \begin{array}{l} \text{The probability a system performs its required} \\ \text{function for a specified period of time t} \end{array} \right. \qquad (3.1)$$

$$R(t) = 1 - F(t) = 1 - \int_0^t f(t)dt = \int_t^\infty f(t)dt \qquad (3.2)$$

Where:

R(t) = the reliability of the item in time
F(t) = the probability the item has failed prior to time t (this is only applicable for non-repairable systems, or the first time of failure)
f(t) = the probability density function
$f(t)\Delta t = P(t<T<t+\Delta t)$the probability the item will fail during the interval $<t,t+\Delta t>$)
t = time
T = time of failure

According to the *Reliability Toolkit* [REL07], the failure rate can be estimated by dividing the total number of failures within an item population, by the total time expended by that population, during a particular measurement interval under stated conditions.
By calculating the failure rate for smaller and smaller intervals of Δt, the interval becomes infinitely small, resulting in the failure rate function. The hazard function is then defined as the conditional probability of a failure in $<t, t+\Delta t>$ given the item did not fail up to t:

$$h(t) = P(t \le T \le t + \Delta t \mid T > t) \qquad (3.3)$$

The hazard function now can be defined as:

$$h(t) = \frac{f(t)}{R(t)} \qquad (3.6)$$

From this the expected number of failures can be derived as:

$$M(t) = \int_0^t h(t)dt \qquad (3.7)$$

The Weibull distribution is by far the world's most popular statistical model for life data. This popularity can be explained by its flexibility. It can mimic the behavior of other distributions such as the normal and the exponential distribution. The associated probability density function can be described as:

$$f(t;\beta;\eta)=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(t/\eta\right)^{\beta}} \qquad (3.8)$$

Where:
β = the shape factor (also referred to as skewness)
η = scale factor (or Characteristic Life)

The shape factor is an indication for the behavior of the failure rate in time.
In case of a constant failure rate, $\beta = 1$;
if $\beta < 1$ then the failure rate decreases over time;
if $\beta > 1$ then the failure rate increases over time.

The behavior of the failure rate in time provides insight to the cause of the failures:

- A decreasing failure rate suggests "infant mortality"; defective items fail early thus the failure rate decreases over time as they fall out of the population.
- A constant failure rate suggests that items are failing due to random events.
- An increasing failure rate suggests "wear out"; parts are more likely to fail as time goes on.

These three levels can be presented in the so-called bathtub-curve (see figure 3.1). The bathtub curve is generated by mapping the rate of early "infant mortality" failures when first introduced, the rate of random failures with constant failure rate during its "useful life", and finally the rate of "wear out" failures as the product exceeds its design lifetime. During the accelerated lifetime testing all three types of errors can occur.

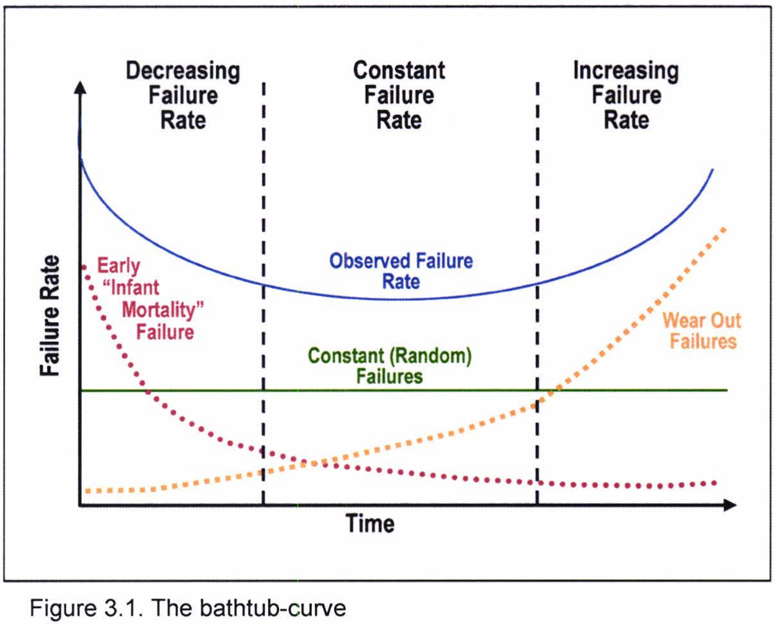


Figure 3.1. The bathtub-curve

A mathematical simplification of the bathtub curve can be presented as:

$$h(t)=\begin{cases} c_0 - c_1 t + \lambda_1 & \textit{for } t=\left[0, {c_0}/{c_1}\right] \quad \text{Early failures} \quad (3.9) \\ \lambda_1 & \textit{for } t=\left\langle {c_0}/{c_1}, t_0 \right\rangle \quad \text{Random failures} \quad (3.10) \\ c_2(t-t_0)+\lambda_1 & \textit{for } t \geq t_0 \quad \text{Wear out failures} \quad (3.11) \end{cases}$$

Where:

c_0, c_1, c_2 = constants
λ_1 = the failure rate [failure/time]

Or graphically:

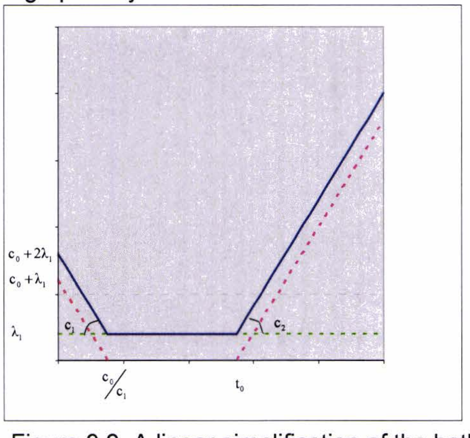


Figure 3.2. A linear simplification of the bathtub curve

3.2 Complex repairable systems

The bathtub curve refers to the first time of failure of a system, this is only in case of a non-repairable system. A complex system (like an Sample loader; the subject of the graduation thesis) will be repaired after a failure. These kind of systems are also referred to as repairable systems. After a repair action the failure distribution is (in general) different from that of a new system. This depends on the time on which the failure occurs (t_f) and the type of repair. Four different kind of repair actions can be defined [BLIS00]:

1. ***Repaired 'As good as new'***
   After the repair the system is as good as a new system. The failure rate of the repaired system is the same as a new system.
   In practice this is only applicable for systems where one specific item determines the failure rate.

2. ***Minimal repair 'As good as old'***
   After the repair the failure rate of the system is equal to that immediately before it failed. The failure rate is therefore unchanged. This model is well applicable to complex systems. Since only the failing part is repaired, all other parts are in the same state as before, and thus the system is returned to its state just before the failure occurred.

3. ***Imperfect repair I***
   In case of an imperfect repair there are two possibilities, the failure rate becomes worse than just before the failure (b). In case of readjustments, cleaning, greasing, etc. after a repair the failure rate would be lower than just before the failure (a)

   a. ***Repaired items different from new (a)***
      In this case the repair action leads to an improved item (but worse than new).

   b. ***Repaired items different from new (b)***
      In this case the repair action leads to an inferior item, resulting in a less reliable system.

4. ***Imperfect repair II***
   It is also possible the reliability of an item is dependent on the number of times an item is repaired. This is modeled by assuming the average failure rate increases after each repair.

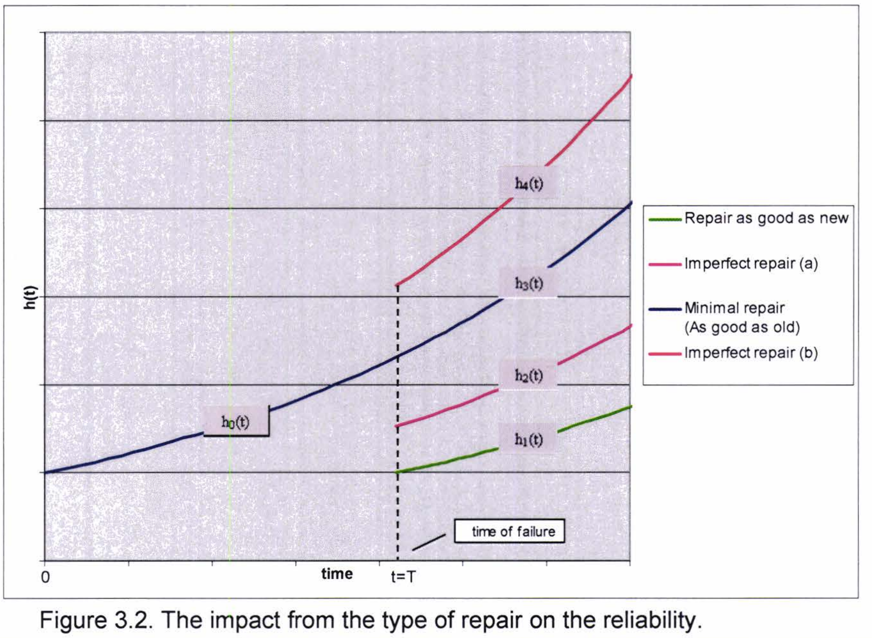


Figure 3.2. The impact from the type of repair on the reliability.

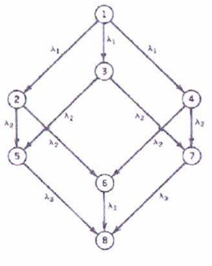
In an ideal case all parameters for each item in a complex system are known, then the failure rate of such system could be described as a function of its age. Often this is quite complex, interactions are hard to model and the results of a repair is often uncertain [CRO93]. Also according to *Roos* [ROO04] it is almost impossible to capture the behavior of a complex repairable system in differential equations. In practice this is solved by using less complicated models which still present viable results.

3.3 Modeling reliability

In this section three different approaches how to model the reliability of a repairable system are presented [GUO01].

3.3.1 State transition diagrams

The background of these kind of models is to calculate the system reliability from the reliability of its components. In a transition state diagram (figure 3) [LEW96] the possible states and the probability of that state in which a particular system can be, is presented. With this kind of model many component failure interactions, as well as systems with independent failures can be modeled using Markov processes. A Markov process is a series of states of a system that has the Markov property. This property means that in a Markov chain the previous states are irrelevant for predicting the probability of subsequent states, given the knowledge of the current state. In this way a Markov chain is "memoryless".



At each time the system may have changed from the state it was in the moment before, or it may have stayed in the same state. The changes of state are called transitions.

This model can only be used if the failure and repair rates can be approximated as time-independent. *Guo et al* [GUO01] though state, these models don't use enough engineering information to reflect the actual physics and are thus not applicable in practice.

Figure 3.3. An example of a state transition diagram for a three-component parallel system

3.3.2 Stochastic/statistical analysis

These models are based on analyzing a given set of data. In this there are two approaches parametric and nonparametric. In case of a nonparametric analysis no assumption regarding the underlying failure distribution is made. Few literature sources could be found discussing issues with these analysis. In an article from *Ascher and Feingold* [ASC84], they mention there is a too large emphasis on developing new models rather than investigating the application of current models. Also *Scarf* [SCA97] underlines this, pointing out too little attention is paid to data fitting and validation and a too large emphasis is made on inventing new models. *Pham and Wang* [PHA96] though state that these models are valuable, provided that the boundary conditions are justifiable with the application of real data.

3.3.3 Discussion

Based on this analysis, it can be concluded that for the graduation thesis transition models can not be used due to their lack of time dependency. Since in the goal statement from FEI the underlying models have to have a physical basis, non-parametrical models are also not applicable. The stochastic models will be described in detail in the next section, also the choice for the parametric models will be explained.

3.4 Reliability Growth Models

During the test phase the prototype is extensively tested. This prototype will contain many design, production and engineering errors. The test program has to be thoroughly designed to find and solve these errors. 'Reliability growth' is defined as the structured process of finding reliability problems, solving these issues and monitoring the improvement [ROO04].

This process usually begins with testing the system under 'accelerated' conditions (for example cycle-testing). In case of a failure all relevant data (time to failure, failure mode and any other relevant data) is collected and analyzed to discover the root causes of the failures. During the failure analysis, the cycle testing is temporary on hold. To reduce the number of failures the appropriate corrective actions are taken. This cycle repeats itself until the required reliability is achieved.
This kind of test cycle is usually referred to as a Test-Analyze-And-Fix program (TAAF). During testing a failure can only be tagged as being "random" or "non relevant" if it is proven that such failure will not occur during normal operation. If multiple prototypes are used during development, corrective actions must be taken as soon as possible on all units. Although this might delay the testing, if faults are not corrected reliability growth will be delayed, since potential failure modes of the next weakest link will not be discovered.

The reliability of a system can be plotted as a function of time or numbers of cycles to identify a possible trend. Since it is assumed the reliability will increase during a project, one speaks of reliability growth. Though it is possible an implemented solution creates new problems. A reliability growth model can project the expected reliability based on the available test data. Since the first publication of J.T. Duane a number of reliability growth models have been developed. In literature [LAG03] these have been divided into two classes:

- Discrete RGM's (Success probability);
  An item in this class has to perform at a certain (discrete) timestamp (also referred to as 'one-shot'). Whether the item performs at a different time is irrelevant. This class is also applied if an item has to perform at multiple discrete timestamps (also referred to as 'repeated cycle'). The key performance indicator for this kind of items is 'Success probability'.
  A discrete RGM is based on test data. If an error occurs during a test, it is solved based on a TAAF strategy.
- Continuous RGMs (MTBF).
  An item in this class has to perform during a certain (continuous) timeframe. This can be a finite time span. The key performance indicator for this kind of items is Mean Time Between Failure (MTBF).
  The MTBF is defined as the mean productive time between failures [SEM01]. Only productive time is included in this calculation. Failures that occur when an attempt is made to change from any state to a productive state are included in this calculation. For a description of the Equipment states and the definition of productive time see Appendix E. Mathematically this can be described as:

$$MTBF = \frac{\text{productive time}}{\#\text{ of failures that occur during productive time}} \qquad (3.12)$$

  A continuous RGM uses two different kind of MTBF, namely:
  - Cumulative MTBF ($MTBF_c$)
    The $MTBF_c$ gives a prediction on t of the average time until the next error, based on the average failure rate ($\bar{\lambda}$) during the period [0, t]
  - Instantaneous MTBF ($MTBF_i$)
    The $MTBF_i$ is an estimate on t of the average time until the next error based on the failure rate (λ_i) at t.

These two classes, discrete and continuous, can be further divided into parametric and non-parametric models.

- Parametric models are models based on an expected distribution of the failure data. This expected distribution furnishes the definition of variables such as failure rate $\lambda(t)$ or MTBF.
- Non-parametric models use curve fitting techniques such as regression analysis, without trying to fit the data to a certain distribution.

Typical differences between these type of models are [MEE98]:

- Nonparametric models require an entire curve, parametric models can be described with a few parameters;
- parametric models have the ability to predict the reliability in time;
- parametric models provide smoother estimates of failure time distributions.

The categorization of models in this section is presented graphically in figure 3.4 [LAG03].

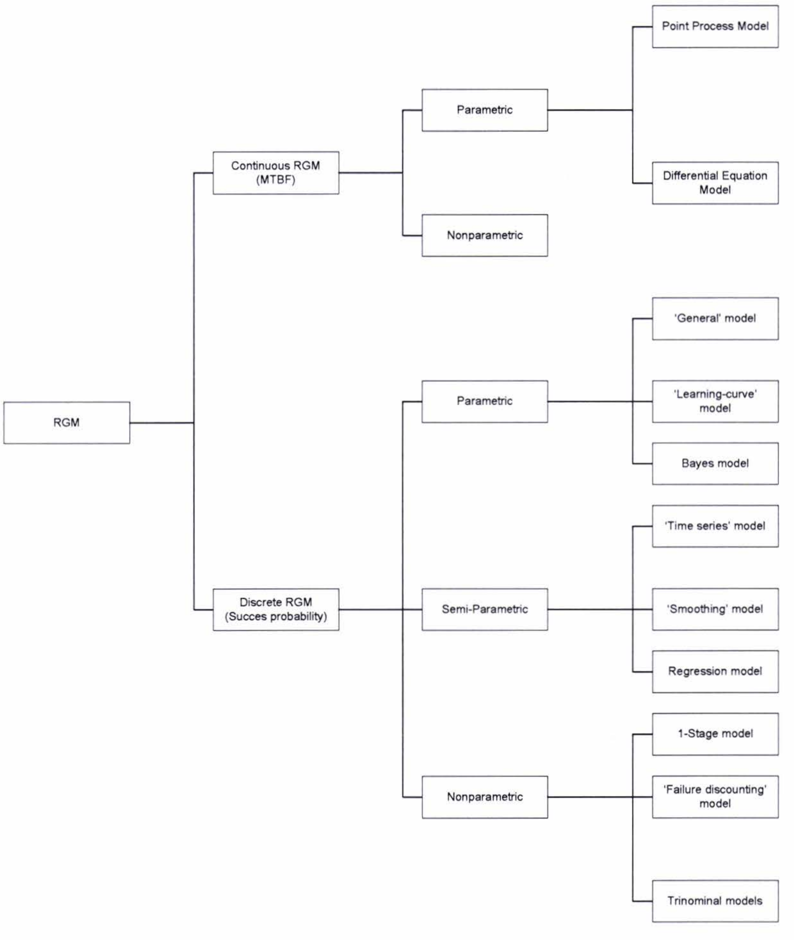


Figure 3.4. A classification of the majority of the reliability growth models

3.5 *Selecting a growth model*

In section 3.1 it is explained what the impact is of minimal and imperfect repair on the reliability. To choose either of them, some assumptions have to be made. The subject of investigation in the graduation, the Sample loader is a complex repairable system. An Sample loader consist of thousands of parts. If one of these parts fails usually a larger module is replaced by a module of the same type (or since the module is still in development by a more reliable module). Even for repair action in the field almost a hundred spares are defined. After the repair action only a small percentage of the specific Sample loader is replaced, thus the probability of a failure after a repair action is assumed to be equal to that just before the failure (thus Minimal repair (see section 3.2)). In case one of these parts fails it will be assumed the actual downtime is negligible with respect to the operational time. Due to these two assumptions it is possible to use the minimal repair model. According to *Rausand and Hoyland* [RAU04] these can be modeled by using Homogeneous Poisson Process (HPP) or Non Homogeneous Poisson Process (NHPP) type of models. By testing the data for a trend, one of both processes can be excluded. The method used to test for a trend is described in Appendix G.

Looking at the tests that will be performed on the Sample loader combined with the properties mentioned above the number of models that have to be investigated can be narrowed down. First the test is run in a continuous operation; cycles are run until a failure occurs. If a failure occurs the test is terminated. This corresponds with the properties of a continuous growth model. The choice between parametric and non-parametric models is based on the arguments presented in the before mentioned article by *Meeker and Escobar* [MEE98]:

- characterization of the data by just a few parameters
- the ability to use the models for prediction

From this type of models both the Power Law and the Log-Linear process will be applied. This is merely a pragmatic choice due the large availability of literature on these models.

Multiple articles [MUS80] and [ASC84] propose to use models originally intended to model software reliability. This might seem odd because if a bug in software is removed it will never reoccur. Though in an article from *Littlewood* [LIT81] it is stated that this also holds for a system in development. For practically each failure which occurs during the Alpha and Beta phase there will be a redesign, and the same error thus will not reoccur. In the articles [MUS80] and [ASC84] two models are described quite thoroughly; the Log-Power model and the Goel-Okumoto model. Both are small modifications of other NHPP models. The Log-Power model is a modification of the Power Law model (it grows much slower). The Goel-Okumoto model is in fact the same as the Log Linear Model, this is proven in Appendix F.

In the next section the actual selected growth models will be discussed in detail.

3.6 Description of selected reliability growth models

In this section the models used for analyzing the data are presented. Successively for each model the expected number of failures, the intensity function and the parameter estimators will be described. Since this thesis is based on a single system, only the parameter estimators for a single system will be presented.

Three basic different mathematical kind of models are used:

1. Power Law Process

$$\text{Cum. Number of failures } M(t) = c_1 \cdot t^{c_2} \tag{3.13}$$

2. Log Linear Process / Goel Okumoto

$$\text{Cum. Number of failures } M(t) = \frac{c_1}{c_2} \cdot \left(e^{c_2 t} - 1\right) \tag{3.14}$$

3. Log Power Process

$$\text{Cum. Number of failures } M(t) = c_1 \ln\left[c_2 (t+1)\right] \tag{3.15}$$

The basic shape of these equations is presented in Figure 3.6. In the consecutive section 3.6.1 to 3.6.3 a more elaborate description of each model is presented.

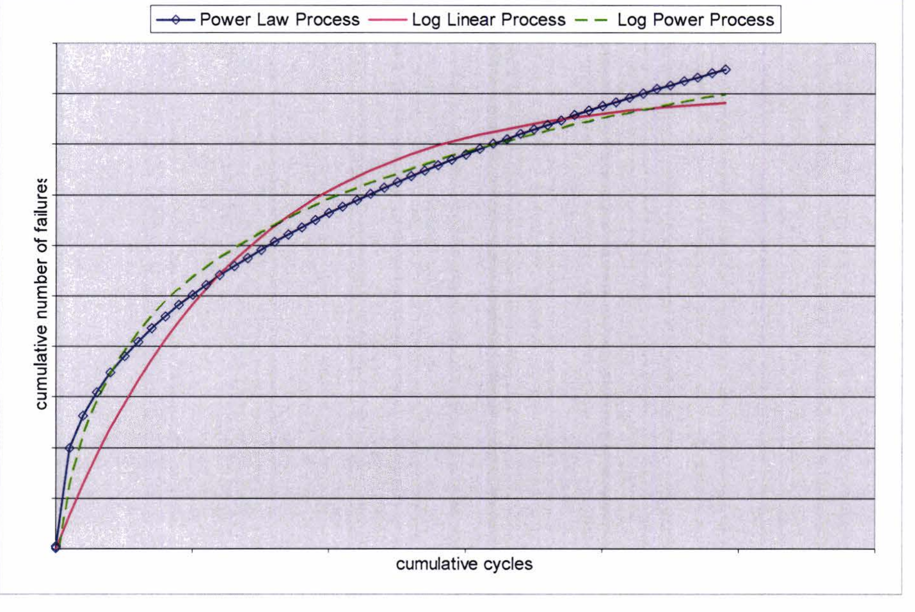


Figure 3.6. Examples of 3 types of growth models

All three models have the following advantages:

- They are reasonably simple to evaluate
- They are easy to use for planning and testing of improvement effectiveness (if employed under equal conditions)

3.6.1 Power Law Process

An improvement process where the actual rate of improvements in time decreases can be modeled by the Power Law process. This means in time, although the time between successive failures keeps growing, the rate of improvement decreases. Implying the time between failures is not identically distributed. The intensity function for the Power Law can be mathematically presented as:

$$h(t) = \alpha_{PL}\beta_{PL}\, t^{\beta_{PL}-1},\ with\ \alpha_{PL} > 0,\ \beta_{PL} =< 0,1 >, t > 0 \tag{3.16}$$

Where:

α_{PL} = scale parameter
β_{PL} = shape factor
t = time

The expected number of failures is:

$$M(t) = \int_0^t \alpha_{PL}\beta_{PL}\, t^{\beta_{PL}-1} \mathrm{d}t = \alpha_{PL} t^{\beta_{pl}} \tag{3.17}$$

Generic properties of the Power model are:

- Both limits ($t \to 0$ and $t \to \infty$) give unrealistic predictions of growth,
  If $t \to 0$ then the failure rate approaches infinity and
  if $t \to \infty$ then the failure rate approaches zero.
  Though in practical use these are rarely limitations
- The model is insensitive to corrective modification, and thus gives a pessimistic estimate of the final MTBF.

3.6.2 Log-Linear Process

The intensity function of the log linear process describes monotonic trends in the failure data. The intensity function for the Log Linear process can be mathematically presented as [LOA92]:

$$h(t) = e^{\alpha_{LL}+\beta_{LL}t} \quad , with\ \alpha_{LL} \in \Re,\ \beta_{LL} < 0 \tag{3.18}$$

Where: α_{LL} and β_{LL} are the parameters for the Log-Linear Process

The expected number of failures is:

$$M(t) = \frac{e^{\alpha_{LL}}}{\beta_{LL}}\left(e^{\beta_{LL}t} - 1\right) \tag{3.19}$$

Properties of this model are:

- In case of $t \to 0$ this model gives a more realistic prediction of growth,

$$= \lim_{t\to 0}(h(t)) = \lim_{t\to 0}\left(e^{\alpha_{LL}+\beta_{LL}t}\right) = \lim_{t\to 0}\left(e^{\alpha_{LL}} e^{\beta_{LL}t}\right) = e^{\alpha_{LL}} \tag{3.20}$$

- if $t \to \infty$ then the failure rate approaches zero, causing the model to present rather optimistic predictions.

3.6.3 Log-Power Process

The Log-Power Process is derived from the Power Law Process. This model was proposed since the Power Law usually overestimates the total number of failures. The intensity function for the Log-Power process [ZHA92] can be mathematically presented as:

FEI Electron Optics B.V.
Building AAE,
P.O. Box 80066
5600 KA Eindhoven

$$h(t)=\frac{\alpha_{LP}}{1+t} \tag{3.21}$$

The expected number of failures is:

$$M(t)=\alpha_{LP}\ln[\beta_{LP}(1+t)],\ with\ \alpha_{LP}>0,\ \beta_{LP}>0, t\geq 0 \tag{3.22}$$

Where: α_{LP} = the expected number of failures which eventually occur
β_{LP} = the failure occurrence rate

Properties of this model are:

- In case of t → 0 this model gives a more realistic prediction of growth,

$$=\lim_{t\to 0}(h(t))=\lim_{t\to 0}\left(\frac{\alpha_{LP}}{1+t}\right)=\alpha_{LP} \tag{3.23}$$

- If t → ∞ then the failure rate approaches zero.

3.6.4 Parameter estimation

In this section two possible methods to estimate the parameters of the fit are presented; Least Squares Regression and Maximum Likelihood Estimation. In literature no underlying theory could be found that points to a preferred solution. In the next two sections the Power Law is used to illustrate both methods, the results for the three models are presented in section 3.6.5.

3.6.4.1 Least Squares Regression (Power Law Process)

The parameter for this model can be estimated for a given data set using curve-fitting methods, to do this the model has to be linearized, using equation 3.16:

$$\ln M(t)=\ln\alpha_{PL}+\beta_{PL}\ln t \tag{3.24}$$

Assuming data available in the form (t_1,M_1), (t_2,M_2), …., (t_N,M_N). The Least Squares principle minimizes the vertical distance between the data points and straight line fitted through the data. The best fitting line then is:

$$\min\sum_{i=1}^{n}(\ln\alpha_{PL}+\beta_{PL}\ln t_i-\ln M(t_i))^2 \tag{3.25}$$

To solve this 3.24 is differentiated to α_{PL} and β_{PL}, and then set these equations to zero, this leads to [REL07]:

$$\hat{\alpha}_{PL}=e^{\frac{1}{n}\left(\sum_{i=1}^{n}\ln(M_i)-\beta_{PL}\sum_{i=1}^{n}\ln t_i\right)} \tag{3.26}$$

$$\hat{\beta}_{PL}=\frac{\sum_{i=1}^{n}\ln(m_i)\ln(t_i)-\sum_{i=1}^{n}\ln(m_i)\sum_{j=1}^{n}\ln(t_j)}{\sum_{i=1}^{n}(\ln(t_i))^2-\sum_{i=1}^{n}\ln(t_i)\sum_{j=1}^{n}\ln(t_j)} \tag{3.27}$$

In literature, when only a single system is considered and the least squares estimation method is used, the power-law process is known as the Duane model [REL07].
Crow, in his article [CRO74], comments that the Duane model could be stochastically represented as a Weibull process, allowing for statistical procedures to be used in the application of this model in reliability growth. This statistical extension became what is known as the Crow-AMSAA (NHPP) model. The Crow-AMSAA model provides a complete Maximum Likelihood Estimation (MLE) solution to the Power Law Process.

3.6.4.2 Maximum Likelihood Estimation (Power Law Process)

The idea behind the Maximum Likelihood Estimation (MEE98) is to determine the parameters that maximize the probability (likelihood) of the sampled data. From a statistical point of view, the method of maximum likelihood is considered to be more robust and yields estimators with good statistical properties [CRO93]. A disadvantage is that the methodology for maximum likelihood estimation is rather complex.

The probability density function (assuming a Weibull distribution) is substituted in the likelihood function. This likelihood function is linearized by taking the natural log on both sides and differentiating respectively to α and β. Then these two equations are set to zero and solved for both parameters α_{PL} and β_{PL}. The maximum likelihood estimates (MLE) of parameters α_{PL} and β_{PL} then are [REL07]:

$$\hat{\alpha}_{PL} = \frac{n}{t_n^{\beta_{pl}}} \tag{3.28}$$

$$\hat{\beta}_{PL} = \frac{n}{\sum_{i=1}^{n} \ln\left(\frac{t_n}{t_i}\right)} \tag{3.29}$$

Where: n = the number of successive failures

3.6.4.3 Parameter estimators

In the analysis phase of this thesis three models will be used. In this section a table is presented with the estimate of the parameters of these models which could be found in literature.

Table 3.1 Overview of the estimates of the parameters

Model	Least Squares Regression		Maximum Likelihood Estimates (MLE)	
Power Law Process	$\hat{\alpha}_{PL} = e^{\frac{1}{n}\left(\sum_{i=1}^{n}\ln(M_i)-\beta_{PL}\sum_{i=1}^{n}\ln t_i\right)}$ $\hat{\beta}_{PL} = \frac{\sum_{i=1}^{n}\ln(m_i)\ln(t_i)-\sum_{i=1}^{n}\ln(m_i)\sum_{j=1}^{n}\ln(t_j)}{\sum_{i=1}^{n}(\ln(t_i))^2-\sum_{i=1}^{n}\ln(t_i)\sum_{j=1}^{n}\ln(t_j)}$	[REL07]	$\hat{\alpha}_{PL} = \frac{n}{t_n^{\hat{\beta}_{pl}}}$ $\hat{\beta}_{PL} = \frac{n}{\sum_{i=1}^{n}\ln\left(\frac{t_n}{t_i}\right)}$	[REL07]
Log-Linear Process	Parameters can only be determined numerically		$\hat{\alpha}_{LL} = \ln\left(\frac{n\hat{\beta}_{LL}}{e^{\hat{\beta}_{LL}t_n}-1}\right)$ $\sum_{i=1}^{n}t_i + \frac{n}{\hat{\beta}_{LL}} - \frac{nt_n}{1-e^{-\hat{\beta}_{PL}t_n}} = 0$	[WEN05]
Log-Power Process	Parameters can only be determined numerically		$\hat{\alpha}_{LP} = \frac{n}{\ln\hat{\beta}_{LP}(1+t_n)}$ $\hat{\beta}_{LP} = \frac{n}{n\ln\ln(1+t_n)-\sum_{i=1}^{n}\ln\ln(1+t_i)}$	[ZHA92]

3.6.5 Confidence bounds

3.6.5.1 Confidence bounds for the Least squares Method

The confidence bounds for the Least Squares method can be obtained from [WEI07]:

Confidence bounds of the cumulative MTBF:

$$MTBF_c = \hat{MTBF_c}(t)\cdot e^{\pm Z_\alpha\sqrt{\frac{\sum_{i=1}^{n}\left[\ln\hat{MTBF_c}(t_i)-MTBF_c(t)\right]^2}{(n-2)}}} \tag{3.30}$$

Then the confidence bounds of the instantaneous MTBF are:

$$[MTBF_i]_L = \frac{MTFB_cL}{\hat{\beta}} \tag{3.31}$$

$$[MTBF_i]_U = \frac{MTFB_cU}{\hat{\beta}} \tag{3.32}$$

The confidence bounds for the instantaneous failure rate are

$$[h_i]_L = \frac{1}{[MTBF_i]_U} \tag{3.33}$$

$$[h_i]_U = \frac{1}{[MTBF_i]_L} \tag{3.34}$$

The confidence bounds for the Cumulative numbers of failures for the Power Law Process are:

$$M_L(t) = \frac{t}{\hat{\beta}}[h_i]_L \tag{3.35}$$

$$M_U(t) = \frac{t}{\hat{\beta}}[h_i]_U \tag{3.36}$$

The confidence bounds for the Cumulative numbers of failures for the Log Linear Process and the Log Power could not be found in literature.

3.6.5.2 Confidence bounds for the Maximum Likelihood Method

The confidence bounds for the Maximum Likelihood method can be obtained from [WEI07]:

$$[MTBF_i]_L = MTFB_i \cdot p_1 \tag{3.37}$$

$$[MTBF_i]_U = MTFB_i \cdot p_2 \tag{3.38}$$

Where the P-Values can be obtained by solving the equation 3.41, $G\left(\frac{n^2}{p}\middle| n\right) = \xi$ for ξ=0.05 and ξ=0.95.

$$G\left(\frac{n^2}{p}\middle| n\right) = \int_0^\infty \frac{e^{-x}x^{n-2}}{(n-2)!}\sum_{i=1}^{n-1}\frac{1}{i!}\left(\frac{n^2}{px}\right)^i e^{-\frac{n^2}{px}}dx \tag{3.39}$$

with $MTBF_i = \left(\hat{\alpha}_{PL} \cdot \hat{\beta}_{PL} \cdot t^{\hat{\beta}_{PL}-1}\right)^{-1}$ for the Power Law Process, (3.40)

$MTBF_i = \left(e^{\hat{\alpha}_{LL}+\hat{\beta}_{LL}\cdot t}\right)^{-1}$ for the Log Linear Process, (3.41)

and $MTBF_i = \left(\frac{\hat{\alpha}_{LP} \cdot \hat{\beta}_{LP}\ln\left[\hat{\beta}_{LP}(1+t)\right]}{(1+t)}\right)^{-1}$ for the Log Power Process (3.42)

Then the confidence bounds for the failure rate can be calculated from:

$$h(t) = \frac{1}{MTBF}$$

and the Cumulative number of failures can be calculated from Equations 3.35 and 3.36 for the Power Law Process,

$$M_L(t) = \frac{\hat{\beta}}{1-e^{-\hat{\beta}t}}[h_i]_L \tag{3.43}$$

for the Log Linear Process and

$$M_U(t) = \frac{\hat{\beta}}{1-e^{-\hat{\beta}t}}[h_i]_U \tag{3.44}$$

$$M_L(t) = \frac{\hat{\beta}}{1+t}[h_i]_L \tag{3.45}$$

for Log Power Process.

$$M_U(t) = \frac{\hat{\beta}}{1+t}[h_i]_U \tag{3.46}$$

3.7 *Model Performance*

To make a sound judgment which model can be applied best two tests are proposed. The first, the goodness of fit tests how good the model fits on the actual data used to determine the parameters of the model. The second one, the Mean Error Of Prediction gives a measure on how well the model is able to predict new data.

3.7.1 Goodness of fit

The Goodness of fit means how well a statistical model fits a set of observations. Measures of goodness-of-fit typically summarize the discrepancy between observed values and the values expected under the model in question. Such measures can be used to test whether outcomes follow a specified distribution.
The method proposed is for calculating the r-square from nonlinear regression. For linear models the value of r^2 quantifies goodness of fit. It is a fraction between 0.0 and 1.0, and has no units. Higher values indicate that the model fits the data better. The R^2-values from nonlinear regressions can be interpreted in a similar way as the r^2 from linear regression. Though it does not have its usual meaning, but nevertheless it can be used to compare different models on the same set of data [SHA07].

$$R^2 = 1 - \frac{SS_E}{SS_T} \tag{3.47}$$

$$SS_E = (\text{Actual value} - \text{Fit value})^2$$

$$SS_T = (\text{Actual value} - \text{Average value})^2$$

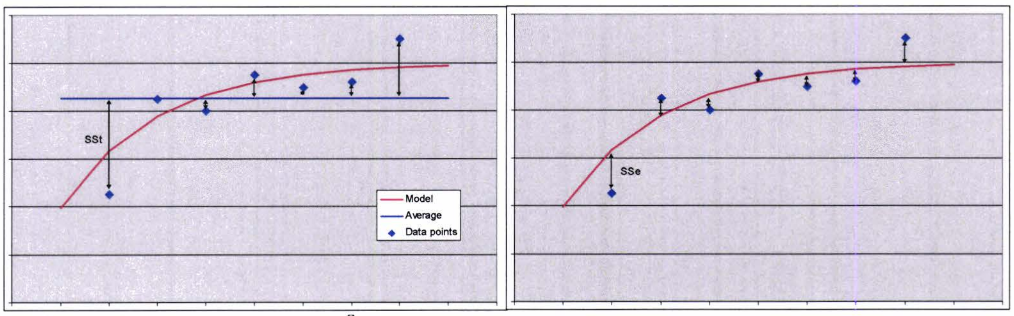


Figure 3.7. Schematic of the R^2 calculation

3.7.2 Mean Error Of Prediction

Besides looking at the accuracy of the fit of a model on the current data, also the accuracy of prediction will be examined (with respect to new measurement data). This Mean Error Of Prediction is defined as:

$$MEOP = \frac{\sum_{i=k}^{n} |n_i - m_i|}{n-k+1} \tag{3.48}$$

Where: m_i = the actual cumulative number of failures at time s_i
n_i = the predicted cumulative number of failures at time s_i
$i = k, k+1, \ldots, n$
k = time from where the prediction starts

The prediction performance is better as the Mean Error Of Prediction approaches zero.

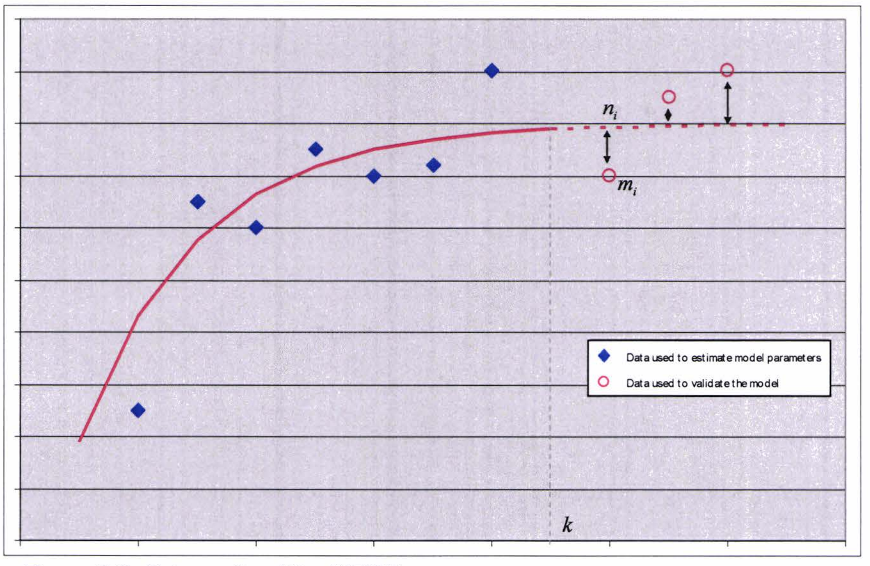


Figure 3.8. Schematic of the MEOP

4 Sample loader

The models described in chapter three will be used to predict the reliability of the Sample loader. In this chapter the goal of the Sample loader is described. In section 4.2 a brief explanation of its functionality is presented. Finally in section 4.3 the way the data is collected is presented.

4.1 Goal of the Sample loader

The goal of the Sample loader project is to provide a standard interface for loading a sample into a TEM or a SmallDualBeam. This encompasses three major success factors:

- Robust (fully automated) loading
- High throughput
- Contamination free (preserve sample at liquid nitrogen or helium temperature (only for LifeScience Sample loader))

Although a standard solution was required, two types of Sample loader were developed:

- Life Science Sample loader;
- Material Science Sample loader.

During this graduation project the Material Science Sample loader project was put on hold, so this thesis focused on the Life Science Sample loader. Consequences of this are:

- It is not possible to validate whether there is variation between Sample loaders
- Less test data will be available

In the first section of this chapter the load steps of the Sample loader are described. In the second section it is explained how the data is collected

4.2 Description of an Sample loader

In this section each step during the loading procedure of a Cartridge will be presented. These steps are aggregated to 'functional steps'. In the drawing below the Sample loader is drawn schematically. The capsule is drawn in red, containing a cassette with specimens in a vertical array. The Airlock-Valve sealing the Airlock (C1) from the environment is yellow. There are two translating arms, a Grid-arm (M1) and a Cassette arm (M2). The Docker is not drawn, but located in the center of the Airlock (C1). At the tip of these arms are V-shaped Grippers (respectively the Cassette-Gripper and the Cartridge-Gripper.

During the loading the (TEM) stage moves to three different positions. This is to prevent the Cartridge arm from colliding into the stage.

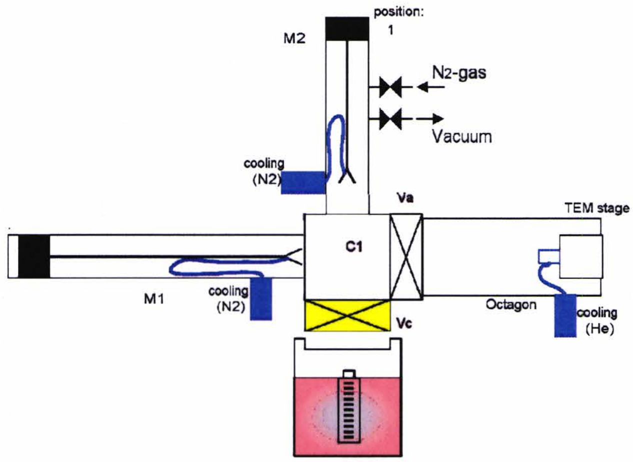


Figure 4.1. Schematic of the Sample loader

1. **Capsule lift**

2. **Cassette loading**
   - A. Vent the Sample loader
   - B. Open the Docker
   - C. Open the Capsule Valve (Vc)
   - D. Open the Cassette Gripper
   - E. Move to Cassette arm to the Capsule
   - F. Close the Cassette Gripper
   - G. Move Cassette arm to the Docker to align the Cassette with respect to the Capsule arm
   - H. Close the Docker
   - I. Close the Capsule Valve (Vc)
   - J. Pump the Sample loader
   - K. Move Cassette arm to Park position (like drawn in the schematic)

3. **Mapping**
   - A. Move Cassette to Cartridge position to align a Cartridge with respect to the Cartridge arm
   - B. Open the Cartridge Gripper
   - C. Move the Cartridge Arm to the Docker
   - D. Close the Cartridge Gripper
   - E. Move the Cartridge Arm to the Park Position (like drawn in the schematic)
   - F. Detect the Cartridge
     If a Cartridge is detected then place back
     If no Cartridge is detected then proceed with the next slot in the Cassette (repeat 3A.)

3. **Load Cartridge (the first steps of the Load cycle are equal to the "Mapping")**
   - A. Move Cassette to Cartridge position
   - B. Open the Cartridge Gripper
   - C. Move the Cartridge Arm to the Docker
   - D. Close the Cartridge Gripper
   - E. Move the Cartridge Arm to the Park Position
   - F. Detect the Cartridge

4. **Load Cartridge**
   - A. Open the Valve to the Octagon[4]/Chamber[5] Chamber
   - B. Move the stage to the 1st position
   - C. Move the Cassette Arm to the Cassette
   - D. Open the Docker
   - E. Move the Cassette to the park position
   - F. Move the Cartridge Arm to the Stage
   - G. Move the stage to the 2nd position
   - H. Open the Cartridge Gripper / Close the Stage Cartridge Gripper
   - I. Move the stage to the 3rd position
   - J. Move Cartridge Arm near the transfer position
   - K. Close the Cartridge Gripper
   - L. Move the Cartridge Arm to the park position
   - M. Close the Valve to the Octagon/Specimen Chamber
   - N. Move the Cassette from the park to the Cassette position
   - O. Close the Docker
   - P. Move the Cassette Arm to the park position
   - Q. Detect Cartridge
   - R. Move the Cartridge Arm to the park position

The sample on the grid can now be analyzed in the microscope. Meanwhile the Sample loader is in the standby position waiting for the operator to 'unload' the Grid. The unloading of the Grid is an exact copy of the loading sequence, but in a reversed order.

[4] An Octagon is the sample area of a TEM
[5] A Chamber is the sample area of a SDB

Beside these 4 functional steps also Software has been defined as a functional error (a glance at the data shows there is a group of errors caused by SW/Firmware error which cannot be assigned to a specific actions of the Sample loader).

4.3 Data collection

Five different kind of lifecycle tests were defined; from complete loads of cartridges to short cycles of sub steps. In the next five subsections these tests will be described in detail. During this testing one central Excel sheet is maintained to collect the measurement data. In appendix H all parameters collected in this sheet will be presented (a summary is presented in section 4.3.6). To validate the data all these parameters were required, the actual modeling only requires the number of cycles ran between two failures and the cumulative number of failures.

4.3.1 Capsule Lock/Unlock

Two different kinds of cycle tests were performed. Cycling a Capsule with a Cassette and cycling with an empty Capsule. During this test step 1 is tested (see section 4.2). During the development phase of the Sample loader no real cycle tests were performed, but any other test requires this step to be performed. No issues were discovered; therefore the results of the two cycle tests will not be used during the analysis phase of this thesis.

4.3.2 Load/Unload Cartridges (Cartridge test)

During this test the Cartridge will be loaded from the Cassette on to the Compustage. After the processing of the sample is finished, this process will take place in reverse order. During this test step 3A to step 4R are tested (see section 4.2). This test is a grouping of 27 different kind of tests. The variations between these tests are which slots were filled and the order of loading. One cycle (including pump down) requires approximately 150 seconds.

4.3.3 Mapping

This test consists of the 3rd step of the cycle presented in section 4.2. During the test phase actually 23 different tests were performed. For the analysis phase these are aggregated into two different tests. The first group is marked with 'Map empty'. For these tests the maps were tested without a Cartridge in the Cassette. The other group is marked with 'Map'. During these tests one or more Cartridge is present in the Cassette. Within these groups there are small differences between the tests. The differences are the slot number with is mapped and the order in which the mapping takes place. One cycle requires approximately 30 seconds.

4.3.4 Cassette Load/Unload

During this test the Cassette will be loaded from the Capsule into the Sample loader. After the processing of the sample is finished, this process will take place in reverse order. The Cassette Load/Unload tests consist of step 2A to step 2K of the entire load cycle (see section 4.2).

4.3.5 Load/Unload All

The entire load cycle is tested during this test. A batch of Cartridge is loaded from the Capsule on to the Compustage and finally unloaded. All steps, from 1 to step 4 are tested (see section 4.2). During the development phase of the Sample loader this was hardly tested, therefore these results will not be analyzed during the analysis phase of this thesis.

4.3.6 Parameters

The logged parameters can be divided in nine different categories. For reliability growth analysis only a few parameters are necessary. These parameters are presented in the table below. All logged parameters are presented in Appendix H.

Table 4.1 Parameters required performing the modeling

M	The cumulative number of failures
This run	The number of cycles during this test run

5 Data analysis

In this section the collected data as presented in the previous chapter, is analyzed with the models described in section 3.6.

5.1 *Validation of the data*

Before the actual modeling starts, the measurement data is validated using the procedure presented in Appendix H. The analysis is performed per dataset. The original data also contained erroneous input such as upgrades and demo's. Some assumptions are made to enable the modeling, but of course create a simplification of the reality.

- The Sample loader under consideration is assumed to be representative for those that will be used in the field. This is not entirely true; the module is still a beta. All changes have to be evaluated with respect to the influence of the reliability of the module.
- In the course of the development process no features will be removed or added. With respect to the goal of the Sample loader this is true, though some features were added (e.g. alignments) to improve usability and performance. These features can have an impact on the reliability of the system.
- The repairs can be considered as minimal repair. This is a valid assumption; a repair on a single item of the Sample loader will have a minor influence on its entire performance.
- After a failure is solved, it will not reoccur. In the data though are some instances, where at that time not the proper corrective action was taken (this caused the issue to reoccur).
- During the development cycle in some cases very long repair times were required, due to redesigns. This was mitigated by modeling the elapsed time by the cumulative number of cycles, rather than calendar days.
- The usage by the engineer is constant in time. In general this is true; only one engineer was responsible for the cycle test. Though there will be a learning effect with respect aligning the Sample loader.

5.2 *Visual examination of the data*

In figure 5.1 to figure 5.3 the failure pattern of the tests described in section 4.3 are plotted. The charts show the cumulative number of failures versus the total number of cycles ran on that particular test. For all tests it can be said they trend concave. The time between each failure tends to become larger, indicating the reliability of the module is improving for each test. Looking more in detail to each chart, specific comments can be made.

The Cartridge test (Figure 5.1) seems to consist of three distinct trends. The first ends after 12.500 cycles. After these 12.500 cycles, TAD (Test And Diagnostic software) was implemented, which caused some initial software errors. Also a hardware design was implemented which caused errors, merely due to improper alignments. The second one after 19.600 cycles, in this period some tests with a different force of the spring-break were performed. Despite these distinct steps all data will be modeled as if it were data from an unchanged Sample loader.

Also the Mapping test shows some increased number of errors during the implementation of the TAD software (after 12.500 cycles). Though this had less impact then on the Cartridge test, also this data set will be modeled as if it were data from an unchanged Sample loader.

The Cassette test contains little data. A large improvement was made by using a coated cassette pin and a new coated cassette. To start a cartridge or a mapping test first the Capsule and the Cassette have to be loaded. These actions were not logged, but they did not reveal any new issues. The tail of the chart also points to a relative long MTBF.

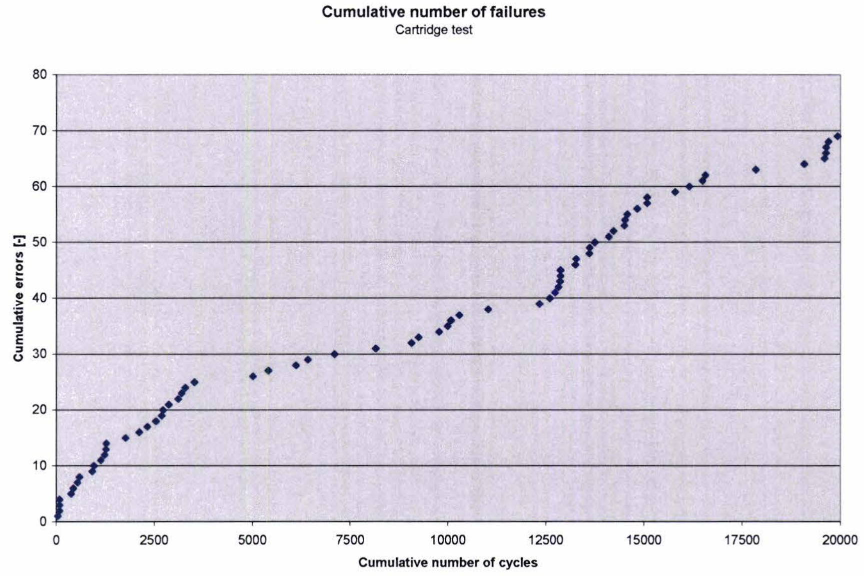


Figure 5.1. Visual presentation of the failure pattern of the Cartridge test.

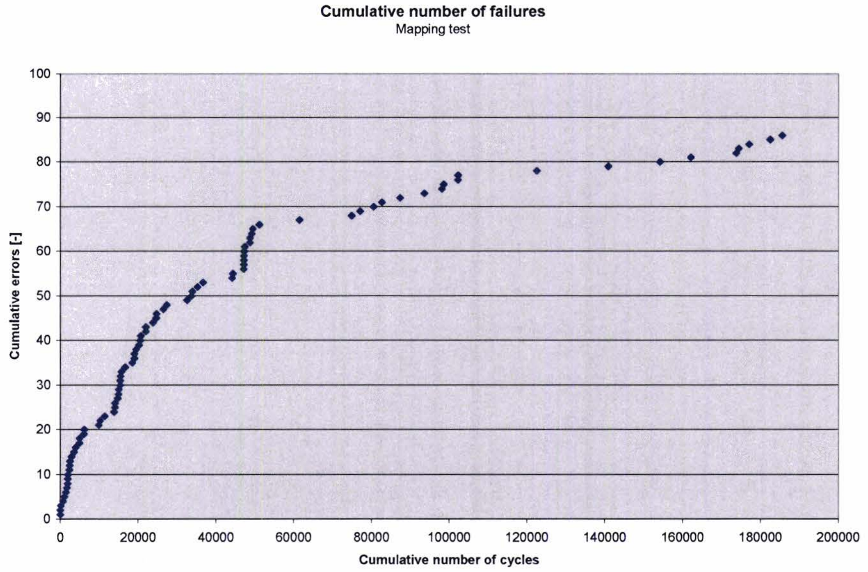


Figure 5.2. Visual presentation of the failure pattern of the Mapping test.

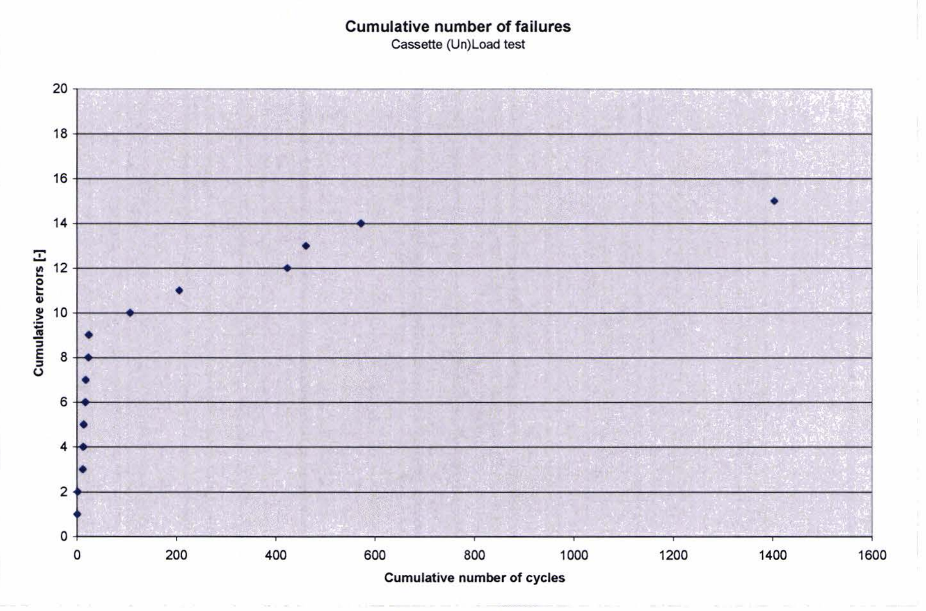


Figure 5.3. Visual presentation of the failure pattern of the Cartridge test.

5.3 Analyzing for trend in the data

Prior to applying a model is has to be examined whether the data actually contains a trend. For this the Laplace-test is used (the complete procedure is described in Appendix G.

The null-hypothesis of a Homogeneous Poisson Process will be rejected at a confidence level α if L = $<\leftarrow, -z_{\alpha/2}> \cap < z_{\alpha/2}, \rightarrow>$. In this a confidence level of α=0.1 will be used. Consequently the null-hypothesis will be rejected if L = $<\leftarrow, -1.645> \cap < 1.645, \rightarrow>$.The results are presented in the table 5.1.

Since the configuration of the Sample loader changed in time and this has influence on the behavior of the cycle performance (see section 5.2), it was decided to divide the Cartridge en the Mapping data into two distinct datasets. At the point in time where TAD was added a split is made. For the Cartridge test this after the 159156th cycle and for the Mapping test after the 865231st cycle. All the Cassette test data is from before the addition of TAD. For each dataset was calculated if growth (positive or negative) can be proven.

Table 5.1 Results of the Laplace test for cumulative number of errors (M)

	Cartridge test		Mapping test		Cassette test
	before TAD (A)	after TAD (B)	before TAD (A)	after TAD (B)	
$\sum_{i=1}^{N-1} T_i$	159156	61952	865231	1502353	3297
T_N	12333	7263	185536	30	1404
N	39	34	86	138438	15
Lp	-3.425	-6.470	-8.312	-2.347	-4.307
P (Z<....)	0%	0%	0 %	0%	0%
Significant trend?	Yes (growth)	Yes (growth)	Yes (growth)	Yes (growth)	Yes (growth)

From row in this table containing Lp it can be concluded that for all three tests the null-hypothesis can be rejected and thus the growth models can be applied.

5.4 Modeling the test data

5.4.1 Presentation of the results

The results of the modeling will be presented per test. Both the results for the Least Square method will and the Maximum Likelihood method will be presented. The accompanying charts will be presented in Appendix J.
The models will be fit on the number of failures (this parameter shows a smoother trend than the failure rate) and the parameters α and β are determined. The parameter can then be used to calculate the instantaneous failure rate and thus the $MTBF_i$. The goodness of fit (R^2) will be calculated by applying equation 3.49. The confidence interval for the expected number of failures of last cycle is presented in the chart (Appendix J). These intervals are calculated with the equations in section 3.6.5. For the Least Square method it wasn't possible to calculate the proper intervals for the Log Linear and the Log Power Process.

For these calculations the test data from January 2006 until April 2007 is used.

5.4.2 Cartridge test

In table 5.3 the results of the Cartridge test are presented. This test ran in total nearly 20.000 cycles. The test is divided in two distinct data sets, from 0 to 12333 cycles (before TAD) and from 12334 to 19596 cycles (after TAD). The first part is referred to as A the second part as B.

Table 5.3 Cartridge test

	Least Squares			Maximum Likelihood		
Parameter	α	β	R^2	α	β	R^2
Power Law Process A	0.388	0.491	0.9841	0.214	0.553	0.9648
Power Law Process B	0.133	0.606	0.9462	0.061	0.681	0.8424
Log Linear Process A	-4.537	$2.931 \cdot 10^{-4}$	0.9779	-4.855	$-1.772 \cdot 10^{-4}$	0.9553
Log Linear Process B	-4.637	$-3.350 \cdot 10^{-4}$	0.9875	-4.717	$-3.068 \cdot 10^{-4}$	0.9785
Log Power Process A	7.089	$8.298 \cdot 10^{-3}$	0.8838	3.591	4.228	-0.0715
Log Power Process B	8.214	$3.09 \cdot 10^{-3}$	0.9729	2.462	5.313	0.9986

In the Appendix J.1 each fit is drawn in a separate chart.

Observations

- When comparing the Least Squares to the Maximum Likelihood method it can be said both methods give comparable results. Overall the Least Squares method performs better (the underlying mechanism of the Least Square method is trying to maximize R^2).Only applying the combination of the Log Power Process and the Maximum Likelihood Method give a poor result for the first part of the dataset (A).
- Comparing the models, it can be said the Log Linear Process creates the best fit, while the Log Power Process performs worst.

Table 5.4 Comparison of the confidence intervals for the Power Law Process for the Cartridge test

Method	h	90% conf interval	MTBF	90% conf interval
Least Squares	$1.78 \cdot 10^{-3}$	$[1.64 \cdot 10^{-3}, 2.17 \cdot 10^{-3}]$	561	[461, 609]
Maximum Likelihood	$2.44 \cdot 10^{-3}$	$[1.57 \cdot 10^{-3}, 2.79 \cdot 10^{-3}]$	410	[358, 635]

From table 5.4 can be concluded that Maximum Likelihood method result in a slightly lower MTBF with a larger confidence interval.

5.4.3 Mapping test

In table 5.5 the results of the Mapping test are presented. This test ran in total nearly 190.000 cycles. The basis for this was not that this was the worst performing part or it required the highest reliability, but mainly due to a long lead time redesign for another part of the Sample loader. The test is divided in two distinct data sets, from 0 to 47098 cycles (before TAD) and from 47099 to 185536 cycles (after TAD). The first part is referred to as A the second B.

Table 5.5 Mapping test

	Least Squares			Maximum Likelihood		
Parameter	α	β	R^2	α	β	R^2
Power Law Process A	0.166	0.5812	0.9756	0.127	0.566	0.9581
Power Law Process B	0.38	0.3609	0.9655	0.235	0.4095	0.9396
Log Linear Process A	-5.819	$-4.817 \cdot 10^{-5}$	0.9652	-5.819	$-4.731 \cdot 10^{-5}$	0.9693
Log Linear Process B	-7.357	$-2.261 \cdot 10^{-5}$	0.8609	-7.690	$-1.258 \cdot 10^{-6}$	0.9082
Log Power Process A	8.909	$3.042 \cdot 10^{-3}$	0.7598	4.993	1.578	-0.7233
Log Power Process B	3.084	$1.262 \cdot 10^{-2}$	0.8371	2.289	3.549	-0.2794

In the Appendix J.2 each fit is drawn in a separate chart.

Observations

- When comparing the Least Squares to the Maximum Likelihood method it can be said both methods give similar (good) results except for the Log Power Process.
- Comparing the models, it can be said both the Power Law Process and the Log Linear Process give good results, the Log Power Process performs worst.

Table 5.6 Comparison of the confidence intervals for the Power Law Process for the Mapping test

Method	h	90% conf interval	MTBF	90% conf interval
Least Squares	$8.38 \cdot 10^{-5}$	$[5.88 \cdot 10^{-5}, 1.56 \cdot 10^{-4}]$	11926	[6423, 16998]
Maximum Likelihood	$8.87 \cdot 10^{-5}$	$[4.93 \cdot 10^{-5}, 1.11 \cdot 10^{-4}]$	11269	[9015, 20284]

From table 5.6 can be concluded that both methods result in similar confidence intervals.

5.4.4 Cassette test

In table 5.7 the results of the Cassette test are presented. This test ran in total nearly 1500 cycles. After some issues at the start this test ran rather successful. This dataset was split since no obvious change in the configuration could be derived from the data (TAD was added after the last Cassette test was performed)

Table 5.7 Cassette test

	Least Squares			Maximum Likelihood		
Parameter	α	β	R^2	α	β	R^2
Power Law Process	2.981	0.236	0.8928	2.614	0.214	0.8519
Log Linear Process	-0.721	$-3.854 \cdot 10^{-2}$	0.9125	-2.694	$-4.498 \cdot 10^{-2}$	0.2339
Log Power Process	2.181	0.834	0.9397	2.007	1.256	0.9319

In the Appendix J.3 each fit is drawn in a separate chart including the confidence intervals for each fit.

Observations

- When comparing the Least Squares to the Maximum Likelihood method is can be said both methods give similar (good) results, expect for the Log Linear Process.
- Comparing the models for the LS method it can be said the Log Power Process creates the best fit.

Table 5.8 Comparison of the confidence intervals for the Power Law Process for the Cassette test

Method	h	90% conf interval	MTBF	90% conf interval
Least Squares	$3.14 \cdot 10^{-3}$	$[1.41 \cdot 10^{-3}, 5.44 \cdot 10^{-3}]$	319	[184, 712]
Maximum Likelihood	$2.58 \cdot 10^{-3}$	$[1.14 \cdot 10^{-3}, 4.13 \cdot 10^{-3}]$	388	[242, 877]

From table 5.8 can be concluded that both methods result in similar confidence intervals.

5.4.5 Mean error of Prediction

In the analysis in section 5.4.2, 5.4.3 and 5.4.4, the data until April 2007 was used. Since this time, only the mapping has been tested substantially to perform a reasonable MEOP. In these 71883 cycles 22 errors occurred. The data is printed italic in table I.1 (Appendix I) and starts with the 87th failure. This leads to a MEOP presented in table 5.9.

Table 5.9 MEOP for the Mapping test

	Least Squares	Maximum Likelihood
Power Law Process	7.147	4.180
Log Linear Process	17.136	17.273
Log Power Process	19.266	23.268

Observations

- The MEOP does not result into large differences between both the Least Squares Regression and the Maximum Likelihood the Power Law Process.
- The Power Law Process gives the most accurate prediction.

5.4.6 Interpretation of fits

In the visual observation it was already mentioned the data consists of multiple distinct datasets. To create a more accurate analysis these dataset would have to be fitted independently. With this approach the performance of both estimation methods are evaluated. By comparing the R^2 the performance of the models and the method can be evaluated. The models are categorized three-ways; good, reasonable and poor. The levels chosen for these categories are:

- Good $R^2 > 0.95$
- Reasonable $R^2 = <0.90,0.95]$
- Poor $R^2 \leq 0.90$

First a comparison of the methods is presented in table 5.10.

Table 5.10 Comparison of the fit performance of the Least Squares and the Maximum Likelihood Method

Fit was R^2	Good >0.95	Reasonable <0.90,0.95]	Poor <0.90
Least Squares Method	6	4	5
Maximum Likelihood Method	6	3	6

From this table can be concluded that both methods give equal results.

Zooming in on the methods the performance of the models can be evaluated. This comparison is presented in table 5.11.

Table 5.11 Comparison between the applied models for the Least Squares Method

Fit was R^2	Good >0.95	Reasonable <0.90,0.95]	Poor <←,0.90]
Power Law Process	2	2	1
Log Linear Process	3	1	1
Log Power Process	1	1	3

Table 5.12 Comparison between the applied models for the Least Squares Method

Fit was R^2	Good >0.95	Reasonable <0.90,0.95]	Poor <←,0.90]
Power Law Process	2	1	2
Log Linear Process	3	1	1
Log Power Process	1	1	3

From these tables can be concluded that using both the Power Law Process and the Log Linear Process will most frequently lead to an accurate fit of the data. Due to the overall low performance of the Log Power Process it will not be in the next section for the calculation of the MTBF. This low performance is caused because of the basic nature of the data. During the Sample loader project many small improvements were made, fitting well on the moderate slope of the Power Law Process, while the Log Power Process creates a too sharp point of inflection.

Based on the MEOP it can be concluded the Power Law Process will provide the most accurate predictions.

5.5 *Calculating the Mean Time between Failures*

In this section the Mean time between failures (MTBF) for each test is presented. From these numbers the overall MTBF for the Sample loader will be calculated. This is done by using the parameter estimates from the Power Law Process and the Log Linear Process using the Maximum Likelihood method (since for this method the confidence intervals for all the processes could be calculated). For the Mapping and the Cartridge test the parameters of the fit from the most recent data is used (the "B" part of the data). Due to its poor fit the Log Power Process is not used in the further calculations.

5.5.1 MTBF of the Mapping test

In table 5.13 and table 5.14 the instantaneous failure rate of the mapping test and the $MTBF_i$ is presented after 190.000 cycles (June 2007; the end of the data set). The instantaneous Mean Time Between Failure is calculated from:

$$\mathrm{MTBF}_{t=190.000\text{ cyclces}} = \left[\frac{1}{h(t)}\right]_{t=190.000\text{ cycles}} = \left[\left(\frac{\mathrm{dM}}{\mathrm{dt}}\right)^{-1}\right]_{t=190.000\text{ cycles}} \tag{5.1}$$

A visual presentation of the confidence interval (CB) of the instantaneous failure rate and $MTBF_i$ are presented in Figure 5.4. The lower and upper limits are drawn with a broad line, the average with a smaller cross.

In figure 5.5, it can be observed that while the MTBF has a large confidence interval, the average is very close to the lower confidence number.

Table 5.11 Failure rate & MTBF of the Mapping test using the Power Law Process

	Least Squares Method		Maximum Likelihood Method	
	Average	CI-range	Average	CI-range
$M_{t=190.000\text{ cycles}}$	83	[76, 93]	86	[73, 94]
$h_{t=190.000\text{ cycles}}$	$5.88 \cdot 10^{-5}$	$[8.38 \cdot 10^{-5}, 1.56 \cdot 10^{-4}]$	$7.69 \cdot 10^{-5}$	$[4.27 \cdot 10^{-5}, 1.11 \cdot 10^{-4}]$
$MTBF_{i,t=190.000\text{ cycles}}$	11926	[6423, 16998]	11269	[9015, 22441]

Table 5.12 Failure rate & MTBF of the Mapping test using the Log Linear Process

	Least Squares Method		Maximum Likelihood Method	
	Average	CI-range	Average	CI-range
$M_{t=190.000\text{ cycles}}$	83	-	86	[73, 94]
$h_{t=190.000\text{ cycles}}$	$2.52 \cdot 10^{-5}$	-	$8.02 \cdot 10^{-5}$	$[4.46 \cdot 10^{-5}, 1.00 \cdot 10^{-4}]$
$MTBF_{i,t=190.000\text{ cycles}}$	39635	-	12467	[9974, 22441]

Both processes give similar results. The confidence intervals of the Power Law Process using the Least Square method are slightly larger.

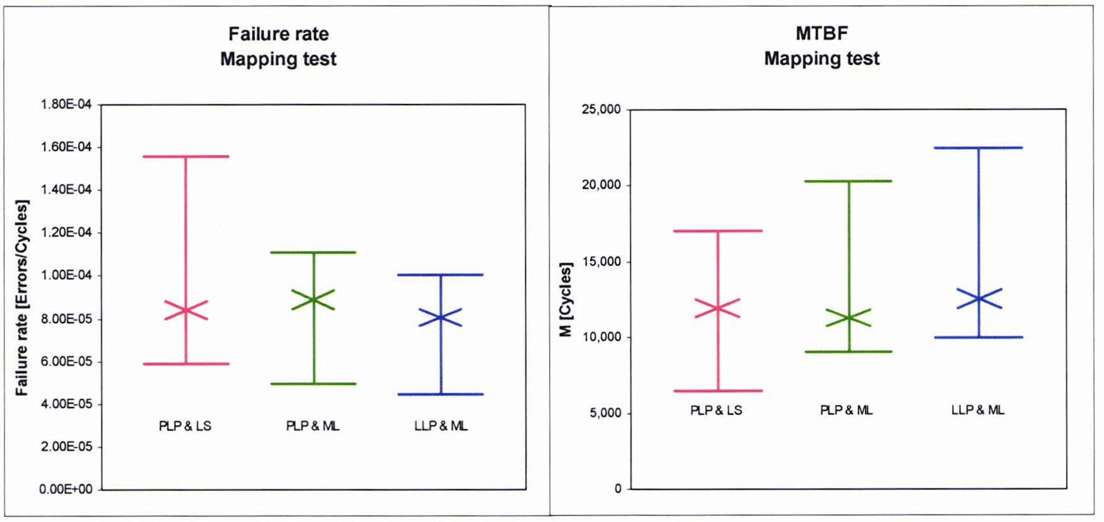


Figure 5.4. Visual presentation of the confidence intervals of the Mapping test.

5.5.2 MTBF of the Cartridge test

In table 5.13 and table 5.14 the instantaneous failure rate of the Cartridge test and the $MTBF_i$ is presented after 20.000 cycles (June 2007; the end of the data set). The instantaneous failure rate is calculated from:

$$\text{MTBF}_{\text{t=20.000 cyclces}} = \left[\frac{1}{h(t)}\right]_{\text{t=20.000 cycles}} = \left[\left(\frac{dM}{dt}\right)^{-1}\right]_{\text{t=20.000 cycles}} \qquad (5.2)$$

A visual presentation of the confidence interval of the instantaneous failure rate and $MTBF_i$ are presented in Figure 5.5. The lower and upper limits are drawn with a broad line, the average with a smaller cross.

Table 5.13 Failure rate & MTBF of the Cartridge test using the Power Law Process

	Average	CI-range
$M_{t=20.000\ cycles}$	65	[55, 69]
$\lambda_{t=20.000\ cycles}$	$2.44 \cdot 10^{-3}$	[$1.57 \cdot 10^{-3}$, $2.79 \cdot 10^{-3}$]
$MTBF_{i,t=20.000\ cycles}$	410	[538, 635]

Table 5.14 Failure rate & MTBF of the Cartridge test using the Log Linear Process

	Average	CI-range
$M_{t=20.000\ cycles}$	65	[56, 69]
$\lambda_{t=20.000\ cycles}$	$1.90 \cdot 10^{-2}$	[$1.23 \cdot 10^{-2}$, $2.18 \cdot 10^{-2}$]
$MTBF_{i,t=20.000\ cycles}$	53	[81, 46]

From the chart can be concluded that the resulting MTBF for each model differs quite a lot.

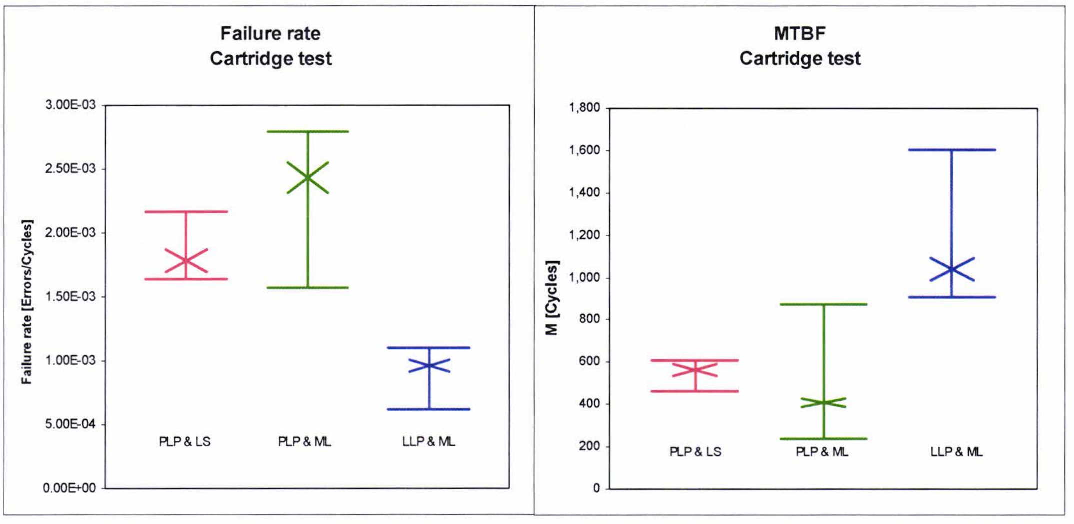


Figure 5.5. Visual presentation of the confidence intervals of the Cartridge tests.

Looking at the table 5.13 and 5.14 it can be said while the Cumulative number of errors are much alike, both the Failure rate and the MTBF differ significantly using the Power Law Process or the Log Linear Process. The Least Squares Method results in smaller confidence intervals with respect to the Maximum Likelihood Method. Applying the Power Law Process leads to a confidence band for the MTBF around 410 cycles, while the Log Linear Process leads to an average of 1100 cycles with a small confidence band. These are large difference despite the relative equal confidence intervals for the Cumulative number of failures. The $MTBF_i$ is derived from the slope of the fit of the M-function. For both models some distinct differences can be observed. The Power Law Process limits to a constant failure rate, while the Log Linear Process trends to zero failure rate. Looking at the data the Log Linear Process provides the best fit on the end of the data set. This could be validated by means of the MEOP, though no test data is available from beyond June 2007. Assuming the Log Linear provides the best fit it is most likely the current MTBF will be close to 1100 cycles.

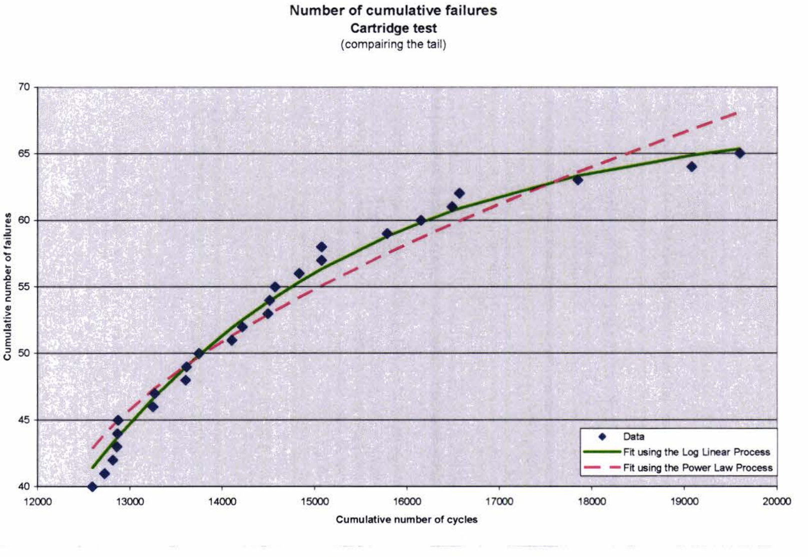


Figure 5.6. Comparison of the M-Fit on the Cartridge test of the Power Law and the Log Linear Process after the addition of TAD

5.5.3 MTBF of the Cassette test

Since the initial issues with the cassette loading were solved, the performance increased with a large step. Since that time hardly any tests were performed (Besides that the Cassette has to be loaded before any of the other two tests can be started. With these 'pre-steps' no errors were observed). The MTBF of the Cassette is very large with respect to the other tests and can therefore be discarded in the calculation of the overall MTBF.
(M_{avg} = 15 errors; $h_{avg} = 1.22 \cdot 10^{-4}$ errors/cycle; $MTBF_{avg}$ = 8175 cycles/error).

5.5.4 MTBF of the Sample loader

The overall MTBF of the Sample loader can be calculated by summing the failure rates of each subsequent step. This assumption holds since the three processes are serial processes and mutually independent. The actual ratio of the failure rate is though dependent on the use case (for example one customer might load three cartridges with one cassette, while the other one only loads cassettes containing 12 cartridges). The equations then becomes:

$$MTBF_{AL} = \frac{1}{h_{AL}} \tag{5.3}$$

$$h_{AL} = A \cdot h_{Cartridge} + B \cdot h_{Mapping} + C \cdot h_{Cassette} \tag{5.4}$$

Where:

A, B = Variables based on the customer use case
- $A = [1,12] \cap A \in N$ [Cartridge loads / Sample loader Cycle]
- $B = [0,12] \cap B \in N$ [Mappings / Sample loader Cycle]
- $C = 1$ [Cassette loads / Sample loader Cycle]

As an example a use case is presented of a customer who will load twelve Cartridges with one Cassette performing a map for each Cartridge. This would be the typical use case for a production test of the Sample loader. Using equation 5.3 and 5.4 and the $MTBF_i$:

$$\mathrm{h}_{AL} = 12 \cdot \mathrm{h}_{\text{Cartridge, Log Linear Process}} + 12 \cdot \mathrm{h}_{\text{Mapping, Power Law Process}} + 1 \cdot \mathrm{h}_{\text{Cassette, Log Linear Process}}$$
$$= 12 \cdot 6.22 \cdot 10^{-4} + 12 \cdot 4.93 \cdot 10^{-5} + 1 \cdot 5.41 \cdot 10^{-5} = 8.11 \cdot 10^{-3}$$

$$\mathrm{MTBF}_{AL} = \frac{1}{8.11 \cdot 10^{-3}} = 123 \text{ loads of a capsule filled with 12 cartridges}$$

The confidence bounds are calculated similarly by summing the outer bounds; resulting in a 90% confidence interval for $MTBF_{AL}$ of [78,77519].

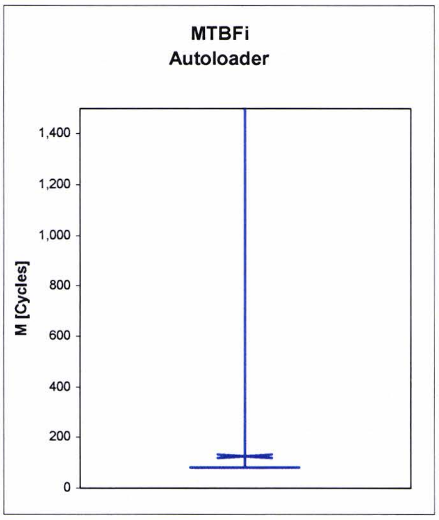


Figure 5.7. Visual presentation of the MTBF confidence intervals of the Sample loader

5.6 *Prediction of the MTBF of the Sample loader*

A prediction of the MTBF of the Sample loader is calculated by extrapolating the fits of the second part of the data set (after the addition of TAD). Some boundary conditions have be taken into account before doing this calculation:

- The addition or removal of features during the extrapolation period is not taken into account.
- All issues found during the extrapolation period will have to be solved.
- When extrapolating the confidence interval it is assumed not only the development process will continue, but also that the characteristics of the new measurement data is equal to that of the current (same variance)

As an example in this section the MTBF after twice the amount of cycles performed (and the accompanying development work) is performed. The total number of cycles will then be the number of cycles during A plus twice the number of cycles during B.

In table 5.15 a prediction of the average failure rate and the average MTBF of the Mapping and the Cartridge test are presented. The Cassette test is so robust the failure rate does not influence the overall failure rate of the Sample loader.

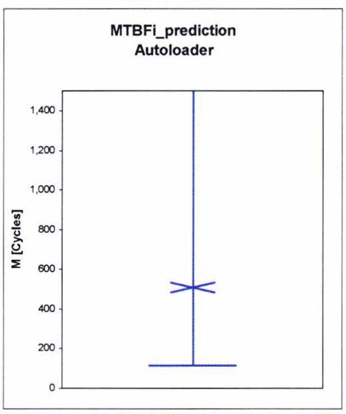
Table 5.15 Prediction of the Failure rate & MTBF

$t = 2t_b - t_a$	Mapping test t = 324.000 cycles		Cartridge test t = 27.000 cycles	
	Power Law Process		Log Linear Process	
M_t	96	[76, 106]	68	[58, 72]
h_t	$5.893 \cdot 10^{-5}$	$[3.274 \cdot 10^{-5}, 7.367 \cdot 10^{-5}]$	$1.04 \cdot 10^{-4}$	$[6.71 \cdot 10^{-4}, 1.19 \cdot 10^{-5}]$
$MTBF_t$	16968	[13575, 30543]	9639	[8408, 14921]

The overall MTBF of the Sample loader can be calculated by summing the failure rates of each subsequent step. The actual ratio of the failure rate is though dependent on the use case of the customer. In this section the same example is used as in section 5.5.4 (loading twelve Cartridges with one Cassette performing a map for each Cartridge). The selection of the usage of the Power Law Process for the Mapping test and the Log Linear Process for the Cartridge test are based on the same assumptions as in sections 5.5.1 and 5.5.2. By using equation 5.3 and 5.4, the predicted overall MTBF can be calculated:

$$\mathrm{h}_{\mathrm{AL}} = 12 \cdot h_{\mathrm{Cartridge}} + 12 \cdot h_{\mathrm{Mapping}} + h_{\mathrm{Cassette}} = 12 \cdot 5.89 \cdot 10^{-5} + 12 \cdot 1.04 \cdot 10^{-4} + 1 \cdot 2.21 \cdot 10^{-7} = 1.96 \cdot 10^{-3}$$

$$\mathrm{MTBF}_{\mathrm{AL}} = \frac{1}{1.96 \cdot 10^{-3}} = 511 \text{ loads of a capsule filled with 12 cartridges}$$



The confidence bounds are calculated similarly by summing the outer bounds; resulting in a 90% confidence interval for $MTBF_{AL}$ of [118,14936170].

6

Figure 5.8. Visual presentation of the prediction of the MTBF confidence intervals of the Sample loader

Evaluation & Recommendations

Before finalizing this report with recommendation to FEI Company and TU/e, first the research goals as presented in section 2 will be evaluated.

6.1 Evaluation of the research goals

In section 2.1 the primary goal of this graduation project has been defined as:

Create a model to predict the reliability of the Sample loader.

The literature research showed there are a large number of models available. One of the boundary conditions of this project -to use a model with an underlying physical relevance- led to the use of Non Homogeneous Poisson models. From the available models, three models were evaluated; the Power Law Model, the Log Linear Process and the Log Power Process. The literature research didn't result in a clear selection criterion to decide for a specific model out of these three.
By means of the R-squared Goodness-of-fit test the accuracy of each model to fit the data is tested. The ability to predict future results is tested by applying the Mean Error of Prediction.
The analysis showed by applying the Least Squares method on the both Power Law Process and the Log Linear Process reliable fits were obtainable. The other model (Log Power Process) was able to model the shape of the trend, though the fit was rather poor. For these models just one out of five analyses would have been accepted at an R^2 value of larger than 0.95. Due to the nature of the data (having a large variance in the failure rate) the Maximum Likelihood generated multiple 'misfits' (14 out of 15 analyses).

The underlying research goals can now be answered.

Which methods are available to predict the reliability of the Sample loader during its development?
To properly model the reliability of the Sample loader an applicable model has to be selected. This is only possible by knowing the behavior of the system that has to be modeled. In section 3.3 a categorization of reliability growth models is presented. The selected model can be applied to the available data set.
The data set has to consist of two parameters:

1. The cumulative number of failures, these failures have to be classified as a 'relevant failure' (see Appendix G)
2. The exact time of failure. In the thesis the time of failure is described by the cumulative number of cycles. (see section 4.3.6)

With this data the parameters of the model are determined. The resulting model is compared to the actual data set by means of a goodness of fit test. The larger R^2 is the more accurate the fit is.

How accurate/reliable are these models?
The reliability is dependent on the quality and the shape of the data set. By analyzing the goodness-of-fit in section 5.4.2 to 5.4.3 the quality of the models can be determined. The results where obtained by accepting a goodness-of-fit where $R^2>0.95$.

Table 6.1 Test quality analysis (number of good fits)

	Least Squares Method	Maximum Likelihood method
Power Law Process	2	2
Log Linear Process	3	3
Log Power Process	1	1

The R2 show little differences for the Power Law process and the Log Linear Process), also the confidence intervals are similar. Based on these result there is no preference for a certain model, this decision has to be based on fitting the models on the specific data set.

What type of problem does each model fit best?
The Log Power Process did not fit well on any of the data. For data that trended towards a zero failure rate the Log Linear Process provides the best fit. If the data trends towards a constant failure rate ($\lambda \neq 0$), then the Power Law Process provides the best fit. The Log Linear Process provided the best overall fit.

What is the reliability of the Sample loader for a specific customer?
The reliability for a specified use case and thus for a specific customer can be calculated by summing the failure rate of the individual tests. In section 5.5.4 an example was presented for a customer loading 12 Cartridges with one Cassette (Mapping each Cartridge). The predicted MTBF is calculated based on the assumption the current development speed is continued for the same amount of cycles as the second dataset (see the example in section 5.6). The MTBF will then be:

Table 6.2 Summary of the MTBF

Test	$MTBF_i$ [cycles]		Predicted MTBF [cycles]	
Cartridge	134	92.0%	801	63.8%
Mapping	1690	7.3%	1414	4.6%
Cassette	18476	0.7%	$4.52 \cdot 10^6$	0.1%
Overall	**123**		**511**	

How can a Reliability Growth model help with the decision process during development?
Applying a Reliability Growth model on each test makes it possible to pin point the current issues. By extrapolating the failure rate it is possible to make a decision how to progress with the further development.
In case the $MTBF_i$ is equal to the desired MTBF no further improvements are necessary. The cycle test can be continued to decrease the confidence intervals and provide a more accurate MTBF.
If the $MTBF_i$ is too low several options are possible:

- By extrapolating the current development rate, a prediction of the MTBF can be calculated. If this predicted MTBF satisfies the requirements, the current test rate and solving speed can be continued.
- If the predicted MTBF will reach the required MTBF, but not within the desired time frame, the speed of solving the issues should be increased (to maintain the confidence level it is preferred to increase the number of test cycles as well).
- If the prediction of the MTBF will never reach the requirements, the testing can be stopped. A major engineering step will be required to reach the goal of the project.

6.2 Recommendations

Model selection
The research lead to a preferred method for estimating the best fit, the Least Squares Method. Also it showed the Power Law Process and the Log Linear Process lead to the best results. Though for each test is has to be evaluated which test fits the data best and thus is able to present a reliable prediction. It is not possible to base this conclusion entirely on the goodness-of-fit (R^2). Especially the 'quality' of the fit at the end of the dataset will be leading. An engineering analysis is required to select the best model.

Data collection
Currently the data is collected by manual entry in an Excel-sheet. A large effort was required to create a usable data set. This can be improved by predefining a template before the actual testing is started (preferably containing drop down boxes containing predefined error selections to ensure reproducible data). Creating automatic generated log-files would create even more accurate measurements. The time of failure would be unambiguous, but also the diagnosis time would be improved. System parameters can give more insight in the system behavior over time, by measuring for example driver currents in time, or pump down times, the real root causes can be revealed. Also the changes of the configuration (added or removed features) are not captured very well.
After a repair cycle no real test runs are done, this causes several early failures related to the repair action. These early failures have a large influence on the fitting of the models. An improvement would be to do a minimum test run prior to starting a cycle test.

Test plan
The results of the MTBF analysis should lead to a decision which tests are most effective to run (subtests or complete cycles). By continuously adapting the test plan –focusing on the major contributors of the failures- a high solving speed can be reached.

Confidence intervals
To create more accurate estimates and predictions of the MTBF the number of test cycles has to be increased. This could be done by testing on multiple Sample loaders, or continuously cycle testing rather than using the Sample loader for demonstrations and software testing during the day. If multiple Sample loaders are used it has to be taken into account that it could be possible the systems will not show identical behavior with respect to their reliability. This will make the analysis phase more complex. It can be expected that the 'solving' speed increases, because this creates the possibility to do mutual comparison and it gives more flexibility in the test plan.

References

[ASC84] Ascher H.E., Feingold H., "Repairable systems reliability: Modeling, interference, misconceptions and their causes", Marcel Dekker Inc., 1984

[ASC98] Ascher H.E., Hansen C.K., "Spurious exponentiality observed when incorrectly fitting a distribution to nonstationary data", IEEE Transactions on Reliability, 1998

[BLIS00] Blischke W.R., Murthy, DNP, "Reliability, Modeling Prediction and Optimization", Wiley, 2000

[CAR00] Carragher N.O., Frame M.C, Fincham V.J.. Wyke J.A.*, "V-SRC'S hold over and cell adhesions", Molecular Cell Biology, 2002

[CRO74] Crow L.H., "Reliability Analysis for Complex Repairable Systems", Relliability and Biometry: Statistical Analysis of Life Length, 1974

[CRO93] Crow L.H., "Confidence intervals on the reliability of repairable systems", Proceedings of the Annual Reliability and Maintainability Symposium, 1993

[GOE80] Goel A.L., Okumoto K., "Time-dependent and error detection rate model for software reliability and other performance measures", 1980

[GUO01] Guo A.L., Ascher H.E, Love E., "Towards practical and synthetical modeling of repairable systems", Economic Quality Control, 2001

[IEC95] IEC/CEI 1164, "Reliability growth – Statistical test and estimation method", 1995

[IEE01] http://ewh.ieee.org/soc/rs/Reliability_Engineering/index.html

[PHA96] Pham H., Wang H., "Imperfect maintenance", European Journal of Operational Research 94, 1996

[KAL98] Kales P., "Reliability for Technology", Engineering and Maintenance, Prentice Hall, 1998

[LAG03] Lagendijk A.J., "Reliability Growth", Gradation report Technische Universiteit Eindhoven, 2003

[LEW96] Lewis, E.E., "Introduction to Reliability Engineering", New York, Wiley, 1996

[LIT81] Littlewood B., "Stochastical reliability growth: A model for fault removal in computer-programs and hardware designs", IEEE Transactions on Reliability, 1981

[LOA92] Loader C.R., "A Log-Linear Model for a Poisson Process Change Point", The Annals of Statistics, Vol. 20, No. 3 (Sep., 1992), pp. 1391-1411

[MEE98] Meeker, W.Q., "Statistical methods for Reliability Data", New York, Wiley Interscience, 1998

[MUS80] Musa, J.D., "The measurement and management of software reliability", Proceedings of the IEEE, 1980RAUS04 Rausand M, Hoyland A., "System Reliability Theory – Models, Statistical Models and Applications", Wiley-Interscience, 2004

[REL07] The Reliability Growth & Repairable System Analysis Reference, ReliaSoft Publishing, 2005

[RAU04] Rausand M, Høyland A., "System reliability theory: models, statistical methods, and applications", Wiley, 2004

[ROO04] Roos C.S., "Reliability groei bij PMS cardiovasculair", Gradation report Technische Universiteit Eindhoven, 2004

[SHA07] http://shazam.econ.ubc.ca

[SCA97] Scarf P.A., "On the application of mathematical models in maintenance", European Journal of Operational Research, Volume 99, Number 3, 1997

[SEM01] Specification for definition and measurement of equipment reliability, availability and maintainability (RAM), SEMI E10-0710:2001

[TAN92] Taneja, V.S., Safie F.M., "An overview of Reliability Growth Models and their potential use for NASA", NASA, 1992

[WEI07] www.weibull.com

[WEN05] Wensing, T, "Predicting and Monitoring the Reliability of Complex Repairable Systems", Gradation report Technische Universiteit Eindhoven, 2005

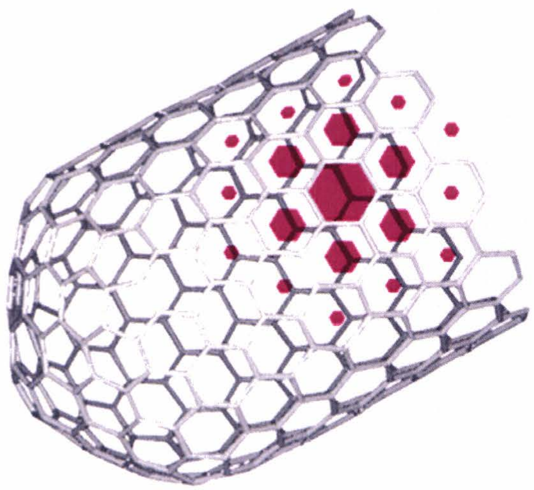
[ZHA92] Zhao M., Xie M., "On the Log-Power NHPP Software reliability model", Proceedings on the Third National Symposium on Software Reliability Engineering, 1992

Abbreviations

AL	Sample loader
AMSAA	US Army Materiel Systems Analysis Activity
BOM	Bill of Materials
CAD	Computer Aided Design
CI	Confidence Interval
COGS	Cost Of Goods Sold
Cryo-EM	Cryo-Electron Microscopy
DB	DualBeam system
FIB	Focused Ion Beam
HPP	Homogeneous Poisson Process
HT3DEM	High-Throughput 3D Electron Microscopy
IEC	International Electro-technical Commission
IEEE	Institute of Electrical and Electronics Engineers
IMS	Integrated Management System
MDR	Module Design Report
MEOP	Mean Error Of Prediction
MLE	Maximum Likelihood Estimate
MRR	Module Release Reports
MTBF	Mean Time Between Failure
NHPP	Non Homogeneous Poisson Process
PIB	Product Improvement Board
PMP	Product Management Process
PRB	Product Review Board
QRE	Quality and Reliability Engineering
R&D	Research and Development
RAMS	Reliability, Availability, Maintainability and Safety
RGM	Reliability Growth Models
RP	Renewal Process
SDB	Small Dual Beam System
SEM	Scanning Electron Microscope
SEMI	Semiconductor Equipment and Materials International
SMART	Specific, Measurable, Achievable and Time framed
SQA	Software Quality Assurance
TAAF	Test Analyze And Fix
TAD	Test And Diagnostics software
TEM	Transmitting Electron Microscope

Modeling reliability growth of a Sample loader

Appendices



FEI COMPANY™
TOOLS FOR NANOTECH

TU/e

Modeling reliability growth of an Sample loader

René Schroemges
Student ID: 499518

Eindhoven University of Technology
Department of Technology Management
Quality and Reliability Engineering

Supervisors
Dr. Ir. P.J.M. Sonnemans
Dr. Ir. J.L. Rouvroye

FEI Company
Electron Optics Product Division

Supervisors
Ir. J.J. ten Thije
Dr. Ir. M.J. Verheijen

Appendix A Life Science Sample loader

The Life Science Sample loader is part of the HT3DEM-project. HT3DEM is an abbreviation of "High-Throughput 3D Electron Microscopy". The goal of this project is to create a reliable, high-throughput transmission electron microscope. The scope of this microscope is to serve the Life-Science community requiring the ultimate spatial 3D resolution with maximum sample integrity.

In recent years, two trends have emerged: efforts to achieve a comprehensive coverage of individual protein structures (so-called structural genomics) and efforts to analyze structures of large complexes. In particular in the latter area cryo-electron microscopy (Cryo-EM) has unique capabilities in providing detailed 3D structural information about those large complexes – the molecular machines - as well as about their localization and dynamics within the 3D architecture of the biological cell. In this way scientists will be able to create exceptional new break-throughs in basic biology and in the understanding and treatment of human diseases.

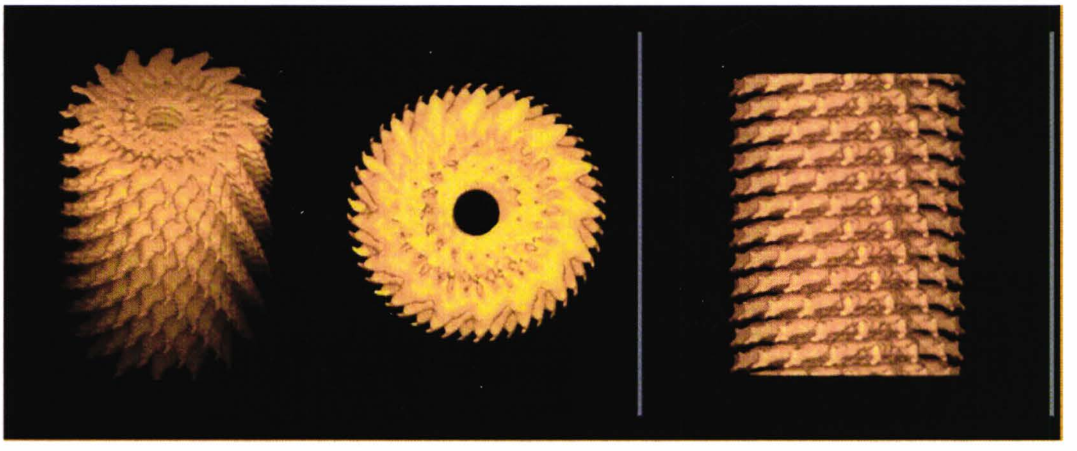


Figure A1. 3D reconstruction of Tobacco Mosaic Virus to approximately 1nm resolution; completely automated from specimen grid to final 3D map. *(Specimen courtesy: Dr. Bridget Carragher – Scripps, San Diego, USA)*

Compared to state-of-the-art electron microscopic imaging available today, the academic world expects a 100-fold increase in data throughput, a 5-fold increase in resolution, and large increase in the level of automation for 3D imaging.
Consequently the long term scientific and commercial success of Cryo-EM in that market segment requires instrumentation that offers high productivity and high quality data, which means that the instrumentation has to be fully automated. Automation meaning:

- robotic sample handling
- minimal or no user intervention
- uptime > 95%

Nowadays it typically takes weeks to months to collect and analyze a dataset in order to reconstruct a map at 10-20Å resolution. In near future this resolution has to increase into the range of 7Å ($7 \cdot 10^{-10}$m). However, that will require an order of magnitude increase in the amount of data collected and analyzed. It is clear that this will only be practical as a mainstream technique if the image acquisition and analysis processes are highly automated and the overall throughput greatly improved. Over the past decade there has been considerable work in automating electron tomography and this has been a key factor in establishing that technique as a powerful tool for studying cellular and macromolecular structures. Recent work by the Scripps group[6] and others[7] has led to an increasing acceptance that automated image acquisition will also be a necessary next step for Molecular Microscopy.

[6] Carragher et al.2001; Carragher et al.2000; Potter et al.1999; Zhu et al.2001
[7] Zhang et al.2001

FEI Electron Optics B.V.
Building AAE,
P.O. Box 80066
5600 KA Eindhoven

The major objective of the HT-3DEM project is to transform TEM 3D structure determination into a rapid, efficient, high throughput process. It is envisioned that with automation, image acquisition and calculation of 3D maps at ~ 7-10 Å resolution for a range of specimen types, will take place in a few days rather than the weeks to months now common.
For the long-term scientific and commercial success of cryo-electron tomography it is important that the operation of the instruments remains not a specialist craft, which requires months or even years of training, but becomes more or less a routine task.
Besides automating the imaging, loading of the sample must be automated and preserved at a cryo environment. This part of the project is covered by the life science sample loader.

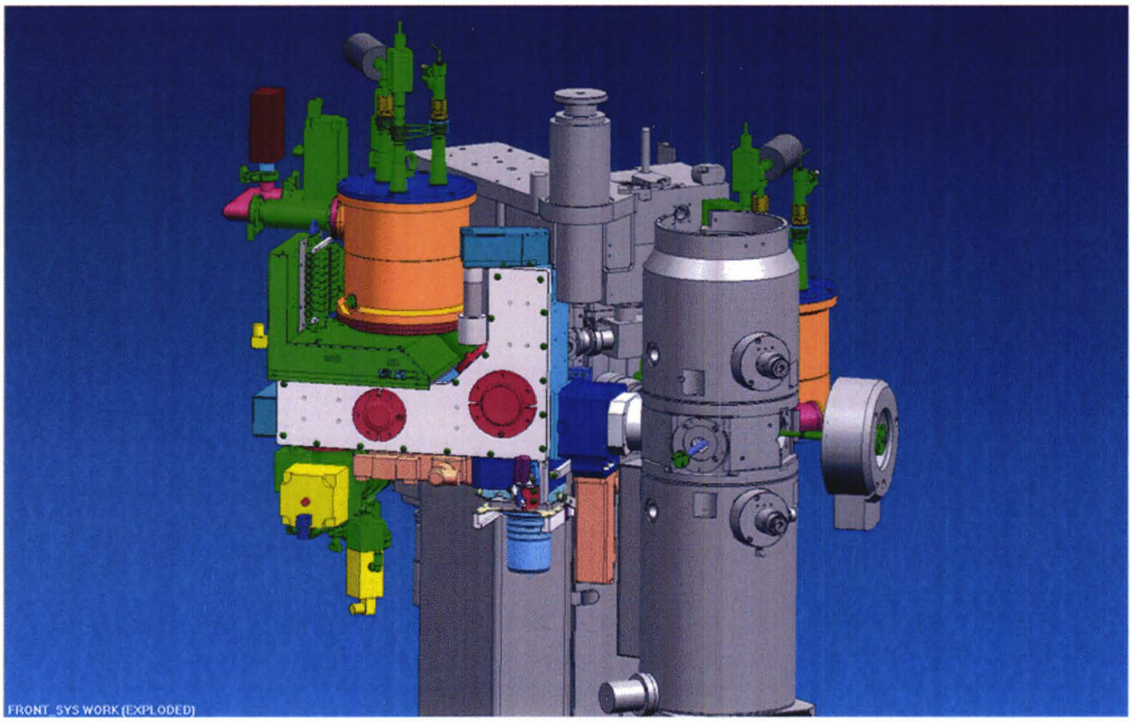


Figure A.2. CAD-model of the Life Science Sample loader on a Transmitting Electron Microscope

Appendix B Material Science Sample loader

The goal of the UltraView project is to create a standard interface for Sample Handling and Data Management. Any system which has to comply with this standard will be an UltraView Process Module.
The Semi conductor market is constantly striving for smaller dimensions. Current developments are aiming for layers smaller than 65nm. By doing this at least 50% of their activities during Research, Development and Production ramp will require process analysis in the (S)TEM space. TEM imaging will be part of the main stream analysis and must therefore be kept affordable. In order to reach this goal improvements across the entire analysis process are required with respect to:

- Sample Creation (more than 4 samples per hour)
- Sample Handling (automated handling (no tweezers)
- Sample/Data Tracking

In the picture below the interfaces between each system are presented..

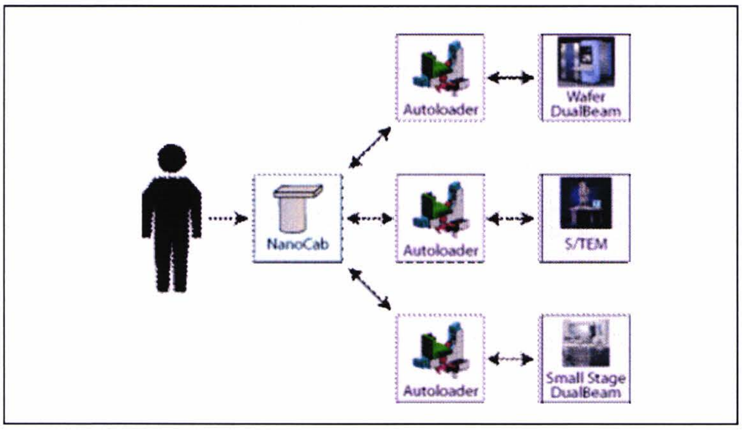
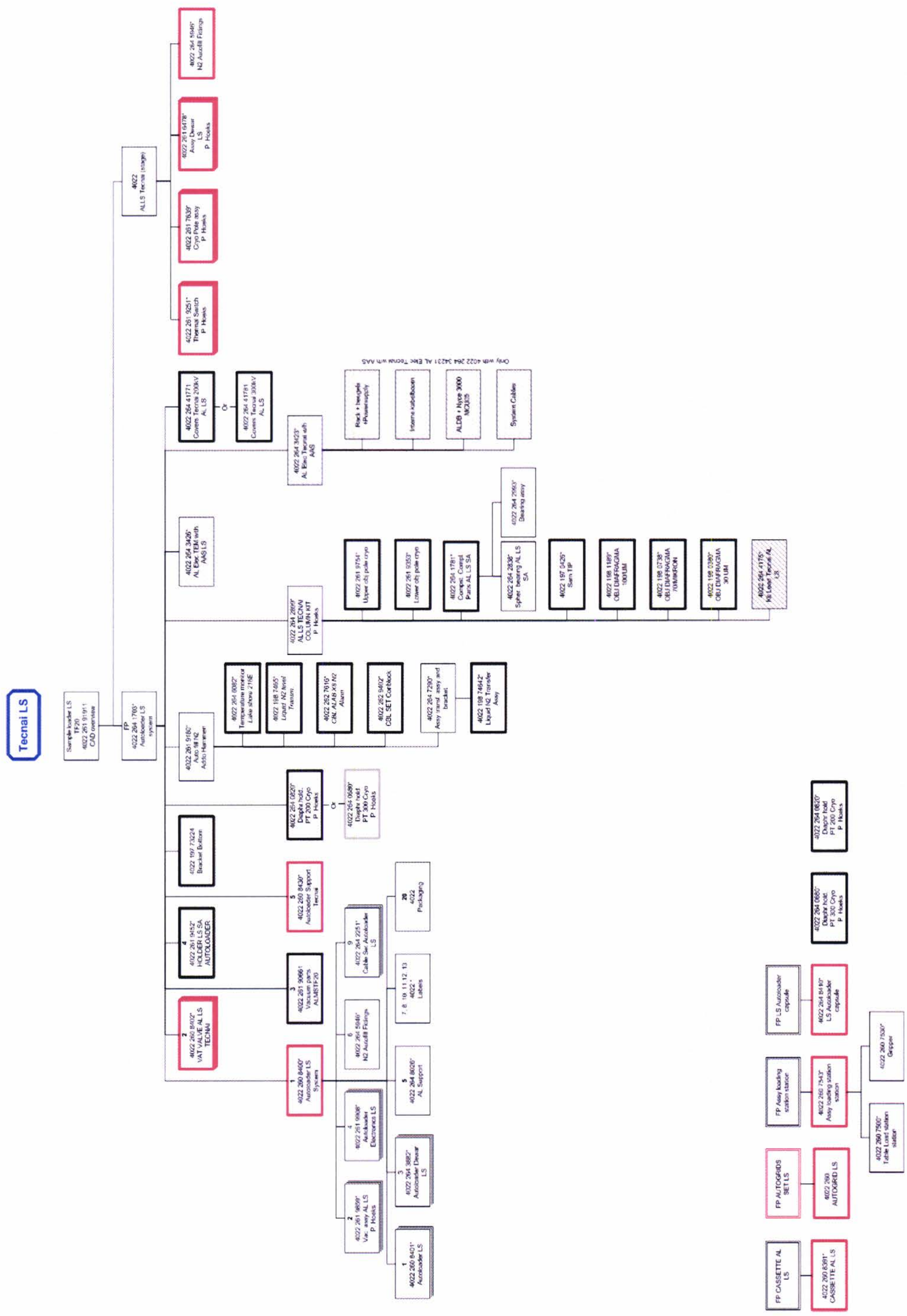


Figure B.1. Process flow

Using a NanoCab to exchange a sample, rather than the sample itself has to following advantages:

- Easy transfer of TEM Samples between systems
- Best throughput, repeatability, and usability.
- Supports transport of samples in an inert environment.
- Low initial drift in TEM.

Appendix C Product Structure of the Sample loader



Appendix D The development process at FEI[8]

The Product Management Process (PMP) is FEI's global approach to develop and introduce new products to the market and support them throughout their lifecycle.
The Product Management Process is based on a series of phases that are linked to project milestones. At the conclusion of each phase, defined criteria are reviewed and must be approved prior to the project's transition to the next phase.

The PMP Deliverables Checklist provides a detailed roadmap for teams to follow in order to develop and release products. The Checklist also supplies the guidelines by which a project's stakeholders can assess the completion of each phase.
At the completion of each phase, each PMP project is reviewed first by its cross-functional stakeholders and finally by the Product Division's Product Review Board (PRB). The stakeholders conduct a detailed review of the phase deliverables and overall project health. The PRB assesses the project's phase deliverables along with a range of other business factors, including market changes, new opportunities, and the Division's overall product portfolio. Based on this array of factors, the PRB decides whether and when projects should proceed through the PMP phases.

In the figure below the seven phases are presented in a graphical way.

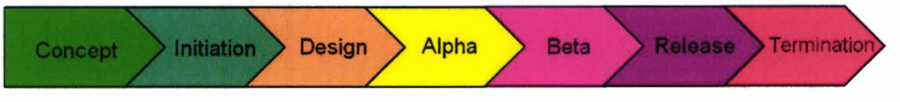


Figure D.1. The PMP Phases

In the next sections each phase of the PMP process is explained briefly.

D.1 The Concept phase

In the Concept-phase it is assessed whether the proposal does fit FEI business both commercially and technically. The goal of the Concept phase is fact finding. During this phase, the Product Manager identifies a new or improved product and establishes the points to be considered if the development of this product should proceed. The Product Manager nominates a study team for the product. This team will advise on the execution of the plan.
The product Manager solicits and analyzes qualified customer input (for example customer reviews), addresses manufacturing and service issues and determines if there are business and strategic justifications for moving forward into the Initiation phase.
Minimal time is spent on the project when it is in Concept phase; no detailed engineering work is done and no commitments to customers are made.

1. Product Manager
   To conclude the research performed during the Concept phase, the Product Manager prepares the initial documentation (Executive Summary, Business Plan, Execution Plan/Risk Analysis). At the Concept gate the Business Plan includes the following:
   a. Financial sketch of the business opportunity
   b. Rough sizing of the project resource requirements
   c. Proposed technology concepts
   d. Preliminary product brochure
   e. Analysis of qualified customer survey input
2. Project Manager
   In the Concept phase the Project manager is primarily responsible for providing input to the Product Manager. This includes initial drafts of the following:
   a. Concept Study Report
   b. Project Execution Plan
   c. Rough sizing of the project resource requirements

[8] IMS (Integrated Management System Manual) Section : 09 Design and Development

   d. Proposed technologies
   e. Resource list
3. Study Team
   At this preliminary stage, the study team provides input for the product's initial business plan the concept study report.

D.2 The Initiation phase

After the initiation phase it should be clear how the new product shall promote FEI business. During the Initiation phase the Product Manager assesses the commercial viability of the proposed product. The project manager will look at the technical viability of the product, form a project team and develop the plan for completing the project.
In the Initiation phase the project team starts the engineering investigation and does some first experiments. Together with all appropriate stakeholders the target specification are discussed and set. Based on the conclusions reached in this phase, the Product Manager updates the Business plan and the Execution plan/Risk analysis. To complete this phase the business plan needs to completed with a detailed sizing of the project resource requirements. The other documents from the Concept phase have to be updated if necessary.
The Project Manager delivers the overall Project Plan. This plan includes:

a. The draft engineering specification, the electronic and software architecture, and system integration plan
b. The Execution plan (including the task sizing, resource allocation, project schedule and technical/schedule risk analysis)
   By assigning the resources to their task a baseline for tracking is created, also this gives insight into possible scheduling conflicts.
c. The Budget (Resource load from Manufacturing and Service, and the preliminary COGS)
d. Input on make or buy decisions (input from Purchasing).

D.3 The Design phase

After the Design phase it should be clear how the product will do its work, the technology to built the product is selected. During the Design phase the majority of the engineering for a new product takes place. The three primary questions that have to be answered during the Design phase are:

- Does the design still meet the market needs with respect to cost, specification and timing?
- What compromises have to be made to meet timing and specification/functionality?
- Are the technical and timing risks and mitigation plans acceptable?

Meanwhile the Project Manager is tasked with determining which technical solutions are available to achieve the desired specification, developing a configuration plan and detailed specifications. These requirements have to include:

- SW, electronic, mechanical, vacuum and electron/ion beam modules and specifications
- Test and Diagnostics (TAD)
- Product engineering specifications
- Options, interlocks and environmental specifications
- Application specifications
- Industrial and ergonomic design, safety, and regulatory requirements/specifications

The Product manager and the Project manager work together to develop budgets and specifications, also they have to refine the Execution plan and create an Alfa plan.

As the Design phase leads to Alpha prototypes, there is an emphasis on finalizing technical specifications in this phase. At the end of this phase the Project team must have completed the design/drafting of a prototype this includes software design, target engineering and system specifications, draft Bill Of Materials (BOM), preliminary built plan and a detailed Alpha plan. The manufacturing and service requirements need to be further detailed and adapted on a system level if new technologies are applied (for example accessibility, remote diagnostics, exchangeability and test depth of modules. The results are given in a Manufacturing/Service Preparation Plan.
In this phase also the specification have to be worked out for the functional modules and supermodules. For each module or supermodule a Module Design Report (MDR) has to be written in such a way that all

stakeholders (development, production and service engineers) can make a proper evaluation. When the MDR is ready and approved the parts for prototyping and the buildup of Alpha systems may be ordered.
Approval of the Design phase documentation by the Product Review Board moves the project into the Alpha phase.

D.4 The Alpha phase

The Alpha phase initiates the prototype process. In this phase the complete system and the engineering specification are proven. Proof-of concept and system level engineering development takes place during this phase. The separate modules should be sufficiently complete for system integration.
The level of prototyping that has occurred in the Alpha phase is dependent on the complexity of the project (smaller projects usually have completed more prototyping in the Design phase, where large tool developments have a larger amount of system-level engineering occurring during the Alpha phase).
The manufacturing members of the project team begin working with the development members to built and test the prototypes. For the manufacturing the draft BOMs developed in the Design phase are used.
Alpha models shall only be sold in Joint Development Programs (the emphasis here should be put on direct customer feedback).
The Product Manager's Alpha phase documents will have to include:

a. An updated business plan
b. An updated technology plan
c. The number of Beta tools and finalized Beta site commitments
d. A preliminary product launch plan

At the gatepass the Project Manager has to report on the progress of the prototype-building. The update of the Project plan will have to include:

- Requirements - target vs achieved specifications, SQA test report, safety and regulatory testing plan, service and testing procedures
- Execution plan – resource allocation and project schedule, technical and schedule risks (and mitigation plans), draft manufacturing procedures and BOMs, draft install and training documents
- Updated budget

During the first phase of the Alpha phase, modules and supermodules are built and tested in parallel (as much as possible). During the development of the individual modules the project team has to focus on system integrity as much as possible.. The team member representing software department is responsible for initiating the feature freeze, the feedback on the integration testing and the test code.

After individual module and supermodules are tested, the result will be published in so called Module Release Reports (MRRs). These are written to give the team members representing the other departments a proper opportunity to review the results. When the MRR is approved by management, part with a long lead-time for the Beta and the initial release systems should be ordered.
If all Alpha phase documents are approved, the project advances to the Beta phase.

D.5 The Beta phase

During the Beta phase it is verified whether the customer needs are met. During the Beta phase, the product is installed at one or more customer. At their sites the field performance and the readiness for release to the entire market is verified. The production planning is ready and the service training is planned. In case of a new system, a product launch will take place in this phase. This can only take place if system configuration, specification and pricing are ready.
Manufacturing builds, tests and ships the Beta systems in cooperation with the project team. This should be done based on the normal production schedule. Product engineers will train manufacturing in using new procedures or techniques. Based on the feedback from building these systems the documentation shall be finalized. Also problems occurring at supplies are solved.
The Product manager has to make sure that customer feedback is integrated into the release version of the system. If all Beta phase documents are approved, the project advances to the Release phase.

D.6 The Release phase

During the Release phase it is verified whether the system is ready to be sold, produced and serviced. At the exit of the Release phase the system can be quoted to customers and at the Application Laboratory demo's are available. Service courses are available and for ach region a key engineer is already trained.

The sustaining engineering team takes over; the system is now owned by the Product Improvement Board (PIB).
During this phase the Project Manager does the final project review. All the system specifications are proven and the safety tests are passed. For software the finale code shall be released. Finally the Product Manager and the Project Manager give a detailed product evaluation, collecting input from all members of the project team regarding lessons learned and opportunities for improvement.

D.7 The Termination phase

During the Termination phase it is decided whether the system still meets FEIs business needs. Based on business or strategic reasons it is decided a system is at the end of its product lifecycle. Although production stops, aftercare and service issue continue in the field. When a system reaches Termination, the Product manager must develop a phase-out plan that shows a clear product migration path. The Product Manager also initiates production ramp down.

Appendix E Equipment states

To clearly measure equipment performance (Reliability, Availability, Maintainability), six basic equipment state are defined into which all equipment conditions and periods of time must fall [SEM01]. In figure E.1 a stack chart of the six basic system states is presented. Key blocks of time associated with the basic states are given in figure E.2.

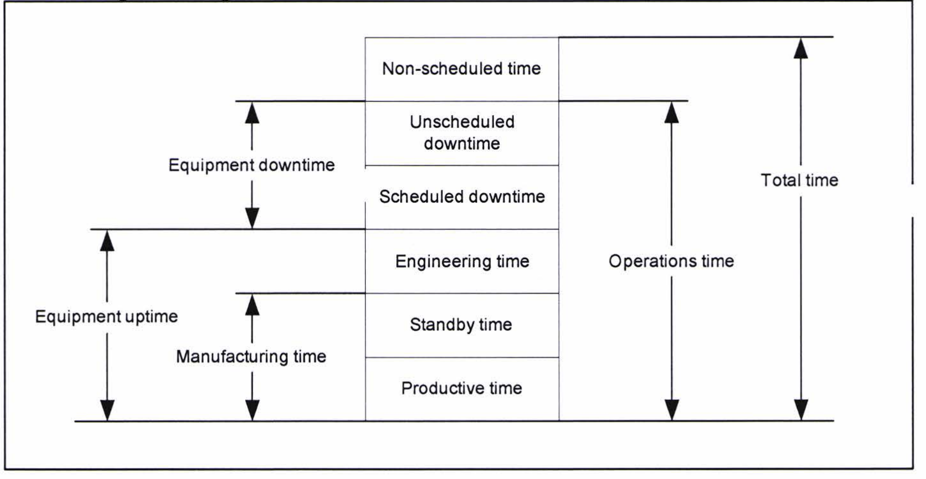


Figure E.1 Equipment states stack chart

The total time consists of the following states:

1. Productive state
   The productive state is defined as the time that the equipment is performing its intended function. The productive state includes:
   - Regular production
   - Work for third parties
   - Rework
2. Standby state
   The standby state is defined as the time that the equipment in a state capable of performing its intended functions, but not operated. Standby state includes:
   - No operator available (including breaks, lunches and meetings)
   - No units available
   - No support tools (e.g. capsules, cassettes)

   No input from external automation system
3. Engineering state
   The engineering state is defined as the time that the equipment is performing its intended function but is operated to conduct engineering experiments. The engineering state includes:
   - Process engineering (e.g. process characterization)
   - Equipment engineering (e.g. equipment evaluation)
   - Software engineering (e.g. software qualification)
4. Scheduled downtime state
   The scheduled downtime state is defined as the time that the equipment is not able to perform its intended function due to planned down events. The scheduled downtime state includes:
   - Maintenance
   - Production test
   - Preventive maintenance
   - Change of consumables/chemicals
   - Setup
   - Facility related

5. Unscheduled downtime state

The unscheduled downtime state is defined as the time that the equipment is not able to perform its intended function due to unplanned down events. The unscheduled downtime state includes:

- Repair
- Out-of-spec input
- Maintenance delay
- Facility related

6. Non-schedule time state

The non-scheduled state is defined as the time that the equipment is not scheduled to be utilized in production. The productive state includes:

- Unworked shifts, weekends and holidays
- Training, installation, modification, rebuilt or upgrade

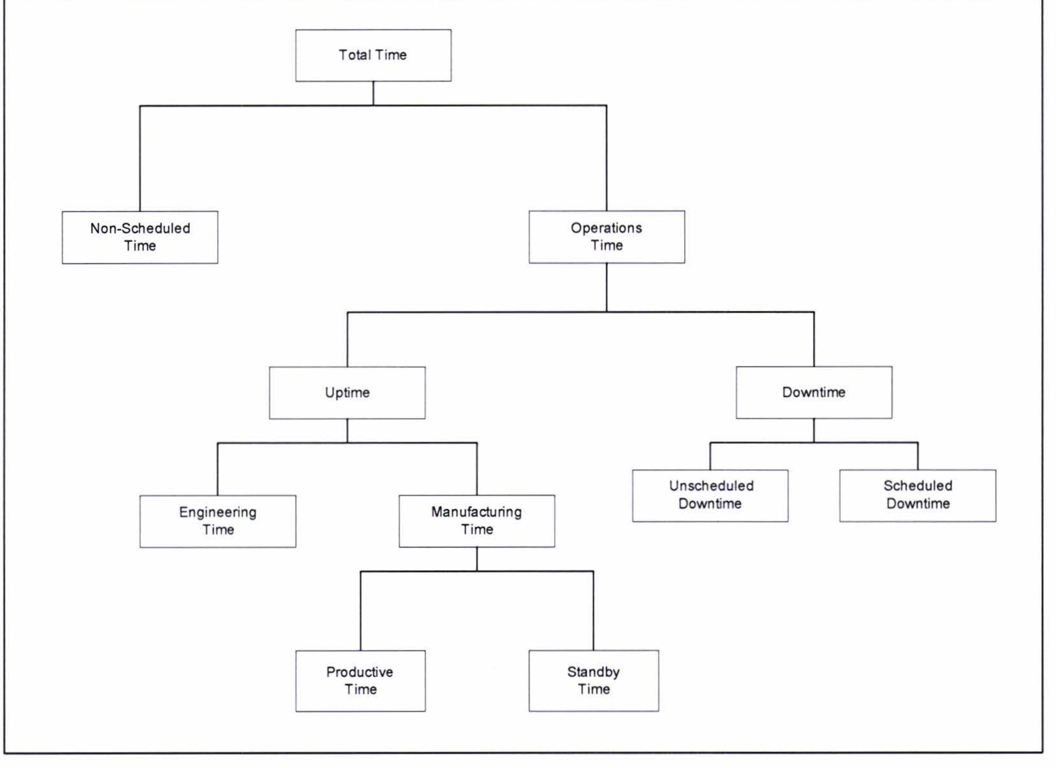


Figure E.2 Summary of time

Appendix F Goel-Okumoto Process

In an article from *Goel and Okumoto* a growth model is proposed [GOE]. This model has the following intensity function:

$$\lambda(t) = \alpha_{GO} \cdot \beta_{GO} e^{-\beta_{GO} t} \quad , with\ \alpha_{GO} > 0,\ \beta_{GO} > 0, t \geq 0 \tag{F.1}$$

Where: α_{GO} = the expected number of failures which eventually occur
β_{GO} = the failure occurrence rate

By integrating this equation this leads to an expected number of failures of:

$$M(t) = \alpha_{GO}\left(1 - e^{-\beta_{GO} t}\right) \tag{F.2}$$

By some mathematical substitutions it can be proven this model is actually a Log-Linear process. The expected number of failures assuming a Log Linear Process can be presented as (Equation 3.19)

$$M(t) = \frac{e^{\alpha_{LL}}}{\beta_{LL}}\left(e^{\beta_{LL} t} - 1\right) \quad \Leftrightarrow \tag{F.3}$$

$$M(t) = -\frac{e^{\alpha_{LL}}}{\beta_{LL}}\left(1 - e^{\beta_{LL} t}\right) \quad \Leftrightarrow \tag{F.4}$$

$$\text{subsitute} \quad \beta_{LL} = -\beta_{GO} \quad and \quad e^{\alpha_{LL}} = \alpha_{GO}\beta_{GO}$$

$$M(t) = -\frac{-\alpha_{GO}\beta_{GO}}{-\beta_{GO}}\left(1 - e^{-\beta_{GO} t}\right) = \alpha_{GO}\left(1 - e^{-\beta_{GO} t}\right) \quad Q.E.D. \tag{F.5}$$

Appendix G Analyzing measurement data

Before a model can be applied, the quality of the dataset which will be analyzed has to be validated. The procedure to analyze the Sample loader-data is discussed briefly. This procedure is derived from CEI/IEC1164 Reliability growth; Statistical test and estimation methods [CIE95].

Step 1
To prevent making wrong conclusions it has to be investigated whether the data contains inconsistencies or errors. IEC1164 proposes the usage of IEC1014; this standard presents a guideline for excluding non relevant failures.

Step 2a
The raw data has to be assembled into a clean dataset with relevant test times at which each failure occurred. There are different possibilities towards presenting the timescale; calendar, operational, online hours or even cumulative number of cycles. Each method could give a different pattern of failure.
After creating a clean dataset some choices have to be made explicit, as different models apply for different types of repair. Also there is a distinct different between models used for single or multiple systems.

Step 2b
Although visual examination of the data is not part of the flowchart in the IEC standard, it can be helpful to plot the data. In an article from *Ascher and Feingold*[ASC84] it is proposed to always present the data in such way, since it gives the ability to analyze the data without assuming an underlying distribution. This prevents possible errors caused by making incorrect assumptions regarding the distribution. Further they state only in a few cases reasoning would lead to the choice of a specific distribution. In case measurement data is presented using a nonparametric technique this always gives an indication on how much the data really implies and how much is being added by a specific model.
For the analysis of multiple systems the Nelson-Aalen estimator is mainly used. If only one system is under consideration this estimator coincides with a plot of the cumulative numbers of failures versus the cumulative elapsed time.

Step3
In step 3 it is determined if there is any trend in the dataset. This test has to show evidence if there is a systematic change in the time between failures, otherwise the interarrival time would be identically distributed. So to test whether there is reliability growth (positive or negative), it is tested if there is no growth. In case of zero growth, the failure times would follow a Homogeneous Poisson Process (HPP). Under this hypothesis the Laplace statistic Lp is distributed as a standard normal variable with mean 0 and standard deviation 1. The Laplace statistic can be used to test if there is evidence of a trend, independent of the reliability growth model. A two sided test for growth at the α significance level has critical values u(1- α/2) and u(1+ α/2). In case there is a trend, the average value deviates from the midpoint of the observation interval. If Lp<0 this implies a negative trend, and if Lp>0 there is a positive trend.

$$Lp = \frac{\sum_{i=1}^{N-1} T_i - (N-1)\frac{T_N}{2}}{T_N\sqrt{\frac{N-1}{12}}} \qquad \text{(G.1)}$$

Where:
- N = the total number of relevant failures;
- T_N = the total accumulated test time for type II test;
- T_i = the accumulated test time a the i^{th} failure.

If there is evidence of positive or negative growth the analysis can be continued with step 4. Otherwise it has to be determined whether the times between failures are exponentially distributed. To test the goodness-of-fit with an exponential distribution *Ascher and Hansen* [ASC98] propose to use the total time on test (TTT) plot. If the time between failures (X_j) are presented in an increasing fashion, the TTT statistics for the times between failures $X_1, X_2, \ldots X_n$, are:

$$R_i \equiv \sum_{j=1}^{i}(n-j+1)(X_j - X_{j-1}) \tag{G.2}$$

$$S_i = \frac{R_i}{R_m} \tag{G.3}$$

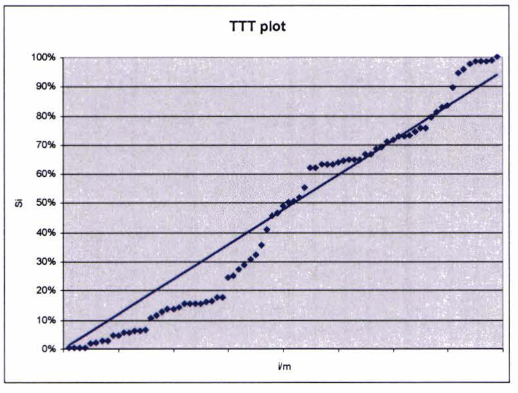


Figure G.1. An example of a TTT-plot

In a TTT plot, Si is plotted against i/m (see Figure G.1). In case the dataset is exponentially distributed the deviation of the points in this plot to the trend line are normally distributed. The null hypothesis for this test, H_0 the data is not exponentially distributed in case the test statistic $W = [-\infty, z_\alpha> \cup <z_\alpha, \infty]$

$$W = \frac{\left(\sum_{i=1}^{m-1} S_i\right) - \frac{m-1}{2}}{\sqrt{\frac{m-1}{12}}} \tag{G.4}$$

If H_0 is rejected, it is assumed the data is exponentially distributed. Then the reliability can be modeled by applying a Homogeneous Poisson Process (HPP), otherwise a Renewal Process can be applied.

Step 4

If in step 3 it was the hypothesis of no growth was rejected, a Non-Stationary model can be selected. However, in literature no clear guidance is given to choose either between minimal and imperfect repair. This choice should be based on a qualitative analysis of the actual repair process.

Appendix H Overview of the logged parameters

In the table below all parameters logged are presented. For the growth analysis only few parameters are necessary. Other parameters could be useful for more in depth analysis. For each parameter a small description is presented.

Date	
Date	For each event only the date of the test/failure is captured.
Test type	
Original description	Since the sheet started before this thesis was started. Some of the cells contain
Remarks	an ambiguous naming. Later in the project the 'Original description' and the
Test type	'Remarks' were merged into one 'Test type'
Test description	
Slot order sequence	Twelve cells are available to capture with slots were tested [1,12]
Random order	Yes or No, to capture whether the test was ran in a random or sequential order
Slots used	The sum of unique slots tested
Environment	
Pressure	The test can be performed at either **Vacuum** or at **Atmosfere**
Temperature	The test can be performed at either room temperature **Environment** or at cryo temperature **Nitrogen**
Comment	
Comment	One cell is available to capture the Comments regarding the Test types
Counters	
This run	The number of cycles during this test run
Capsule lock/unlocks	The number of Capsule locks/unlocks during this test run
Cassette loads/unloads	The number of Cassette loads/unloads during this test run
Docker open/closed	The number of Docker open/closed actions during this test run
Cartridge maps	The number of Cartridge maps during this test run
Cartridge load to/from stage	The number of Cartridge load to/from stage during this test run
Corrected Capsule lock/unlocks	The number of Capsule locks/unlocks during this test run
Corrected Docker open/closed	The number of Docker open/closed actions during this test run
Error analysis	
Error analysis	This is the root cause at the time of the failure according to the test engineer (to prevent different descriptions by different testers the file is filled each time by the same engineer)
Action	
Action to remove cause	This is the containment action to remove the root cause mentioned above.
Machine version	
Machine version	Despite the fact that this thesis models the test strategy as a continuous improvement project. It was

Discipline

Mechanical Electrical Electro mechanical Software	After a failure in these cells is pointed towards the "discipline source" of the failure. To do so only four categories are available (E, M, EM and SW)

Actions

Action taken	The actions have been aggregated to eight categories (Adjust, Change setting, Clean, None, Repair, Replace, Software change, Substitute)
Action remark	This cell contains some extra space for the test engineer to enter comments with respect to the corrective action

Appendix I Measurement data

In the tables below the data used to determine the analysis presented in section 5 are presented. Only the data actually used for the calculation are shown, all other presented in Appendix J are not crucial for the calculation.

Table I.1 Mapping test data

M	Cum cycles	M	Cum cycles	M	Cum cycles	M	Cum cycles
1	24	28	15021	55	44320	82	173790
2	45	29	15061	56	47098	83	174384
3	186	30	15423	57	47134	84	177097
4	772	31	15440	58	47146	85	182425
5	1110	32	15490	59	47183	86	185536
6	1478	33	15732	60	47246	*87*	*209773*
7	1743	34	16723	61	47392	*88*	*235009*
8	1920	35	18538	62	48743	*89*	*238083*
9	1981	36	19092	63	48770	*90*	*238166*
10	2005	37	19101	64	49184	*91*	*238612*
11	2366	38	19485	65	49490	*92*	*241211*
12	2454	39	20315	66	51210	*93*	*241212*
13	2533	40	20563	67	61508	*94*	*244263*
14	2957	41	20695	68	74881	*95*	*244334*
15	3570	42	21994	69	77101	*96*	*244515*
16	3971	43	22053	70	80517	*97*	*244553*
17	4946	44	23798	71	82659	*98*	*244705*
18	4947	45	24704	72	87384	*99*	*244871*
19	6063	46	24719	73	93618	*100*	*244884*
20	6165	47	26521	74	98112	*101*	*245964*
21	9902	48	27320	75	98569	*102*	*249062*
22	10318	49	32599	76	102232	*103*	*254998*
23	11420	50	33667	77	102291	*104*	*255021*
24	13908	51	33917	78	122503	*105*	*255308*
25	14085	52	35232	79	140910	*106*	*255379*
26	14155	53	36660	80	154217	*107*	*255439*
27	14821	54	44105	81	162061	*108*	*257419*

The italic data is gathered from May 2007 until July 2007, this data is used to calculate the predictive performance of the investigated models.

Table I.2 Cassette test data

M	Cum cycles	M	Cum cycles	M	Cum cycles	M	Cum cycles
1	1	5	14	9	24	13	461
2	2	6	17	10	107	14	572
3	12	7	18	11	206	15	1404
4	13	8	23	12	423		

Table I.3 Cartridge test data

M	Cum cycles	M	Cum cycles	M	Cum cycles	M	Cum cycles
1	41	19	2693	37	10285	55	14569
2	79	20	2721	38	11023	56	14829
3	80	21	2869	39	12333	57	15072
4	81	22	3109	40	12596	58	15073
5	372	23	3199	41	12729	59	15788
6	435	24	3291	42	12816	60	16150
7	540	25	3519	43	12860	61	16487
8	588	26	5018	44	12869	62	16565
9	916	27	5413	45	12874	63	17848
10	959	28	6116	46	13249	64	19078
11	1137	29	6417	47	13269	65	19596
12	1236	30	7101	48	13602	66	19630
13	1267	31	8147	49	13613	67	19640
14	1276	32	9065	50	13747	68	19695
15	1767	33	9250	51	14105	69	19925
16	2117	34	9768	52	14219		
17	2320	35	9991	53	14495		
18	2540	36	10077	54	14512		

Table I.4 Transformed Cartridge test data

M*	Cum cycles*	M*	Cum cycles*	M*	Cum cycles*	M*	Cum cycles*
1	263	8	936	15	2179	22	4154
2	396	9	1269	16	2236	23	4232
3	483	10	1280	17	2496	24	5515
4	527	11	1414	18	2739	25	6745
5	536	12	1772	19	2740	26	7236
6	541	13	1886	20	3455		
7	916	14	2162	21	3817		

M = M* + 39
Cum Cycles = Cum Cycles* + 12333

Appendix J Data Analysis

In this appendix the detailed results for the Cartridge, Cassette and the Mapping test are presented. In section J.1 to section J.3 subsequently the charts with the fitted models are plotted (where it was possible to calculate the confidence intervals these are presented by dotted lines).

J.1 Results of the Cartridge test

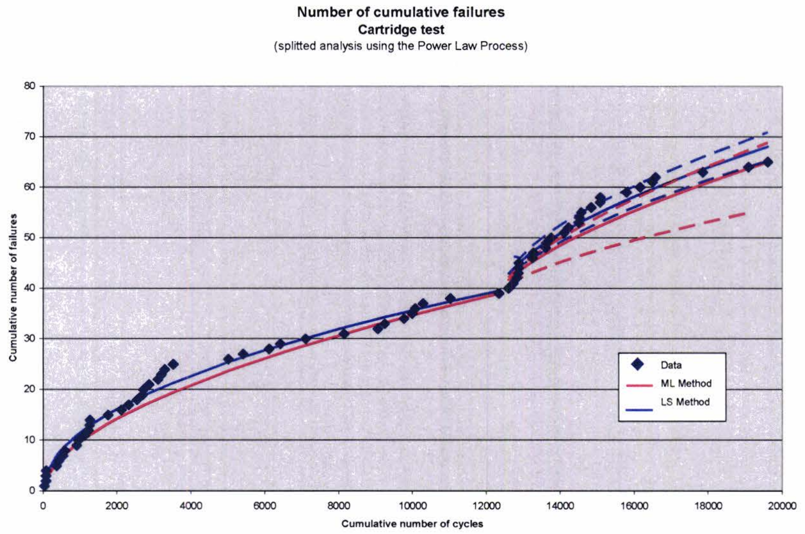


Figure J.1. Fit of M(t) using Power Law Process for the Cartridge test

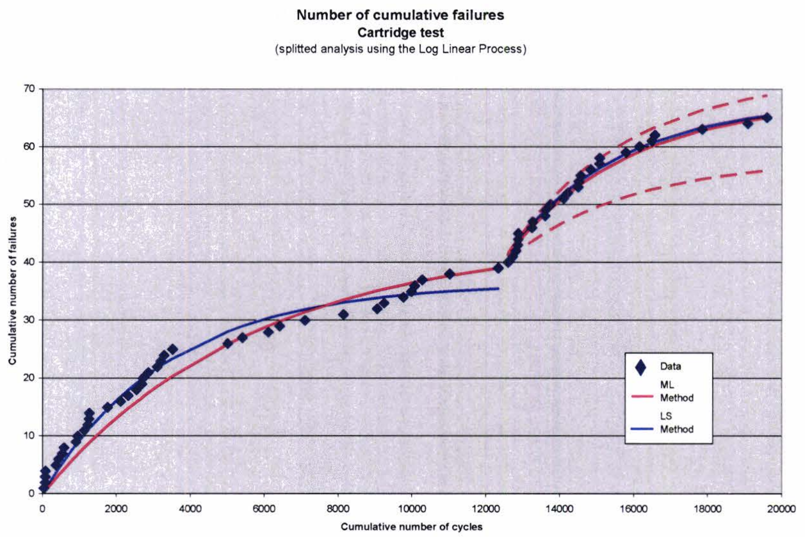


Figure J.2. Fit of M(t) using LL Process for the Cartridge test

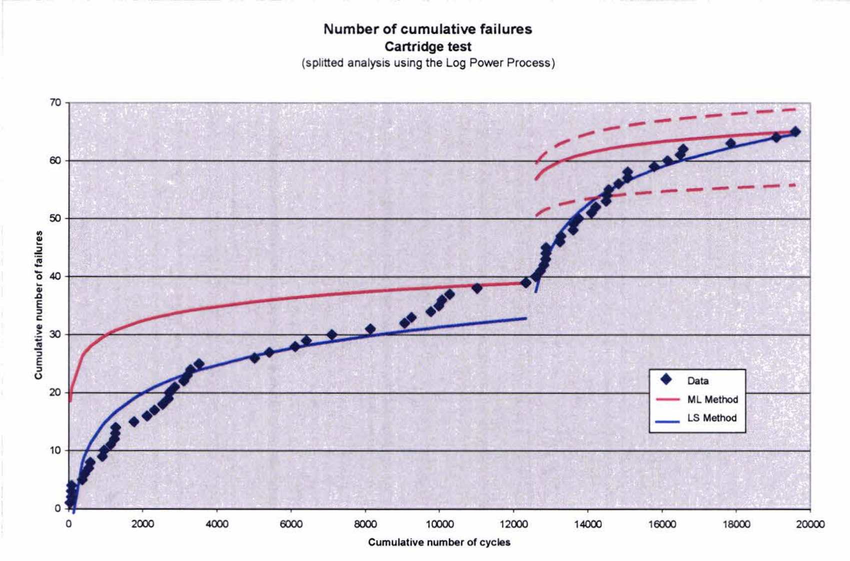


Figure J.3. Fit of M(t) using LP Process for the Cartridge test

J.2 Results of the Mapping test

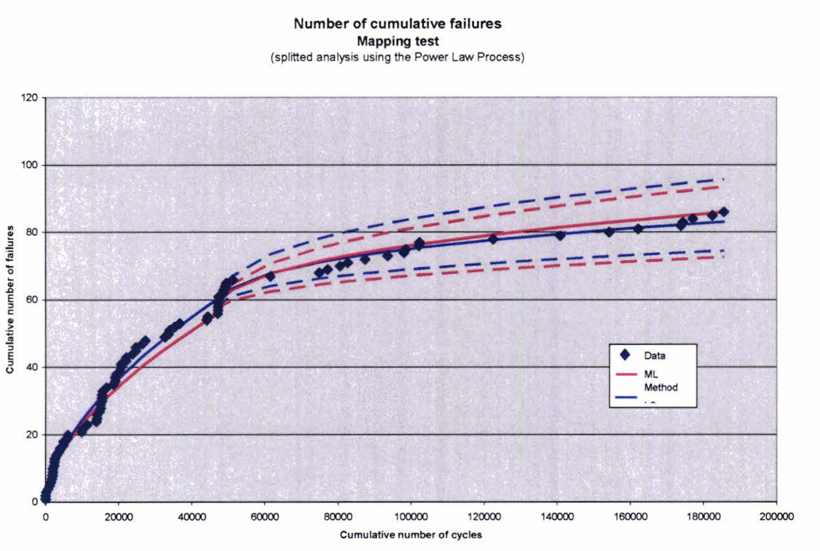


Figure J.4. Failures for the Mapping test fitting a Power Law Process.

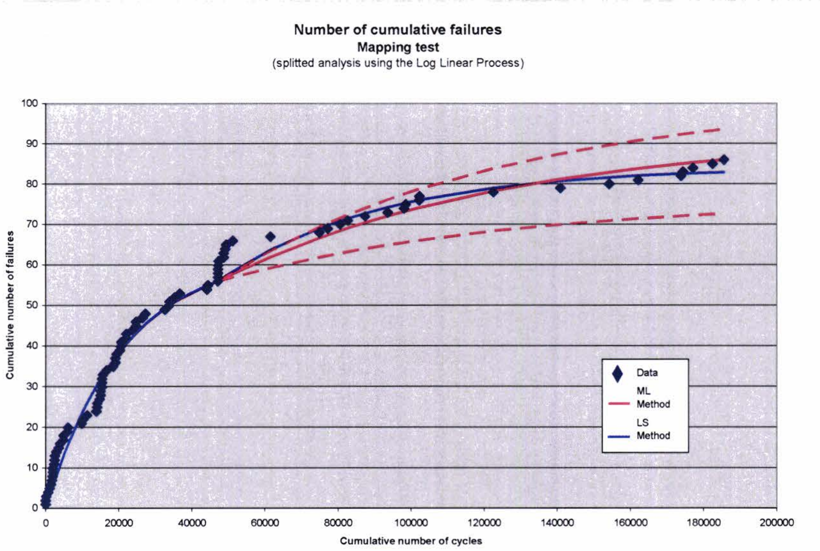


Figure J.5. Failures for the Mapping test fitting a Log Linear Process.

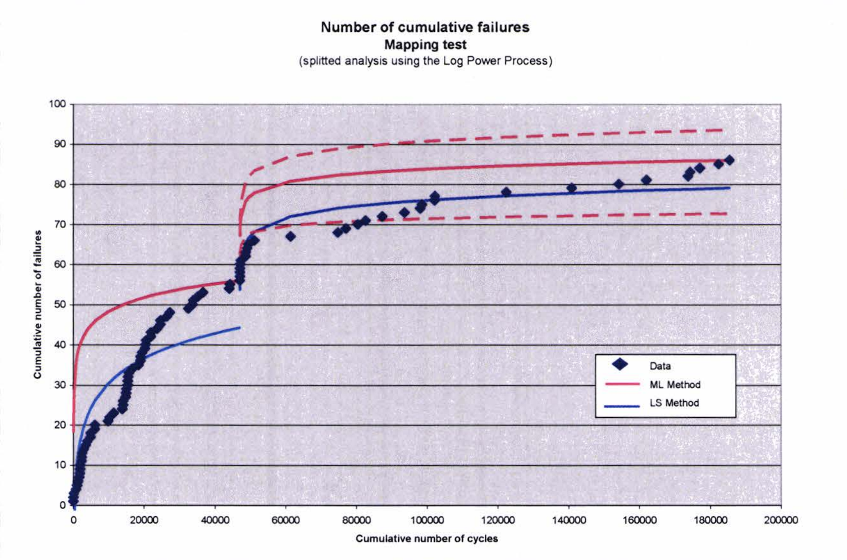


Figure J.6. Failures for the Mapping test fitting a Log Power Process.

J.3 Results of the Cassette test

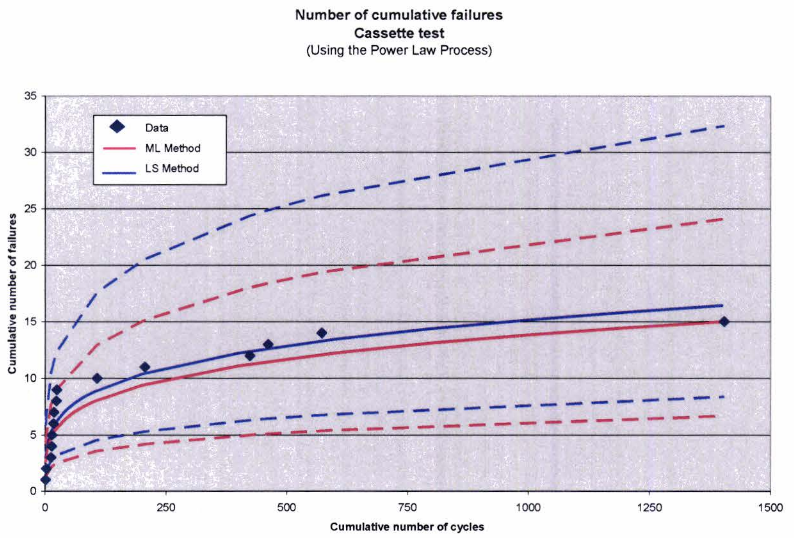


Figure J.7. Failures for the Cassette test fitting a Power Law Process.

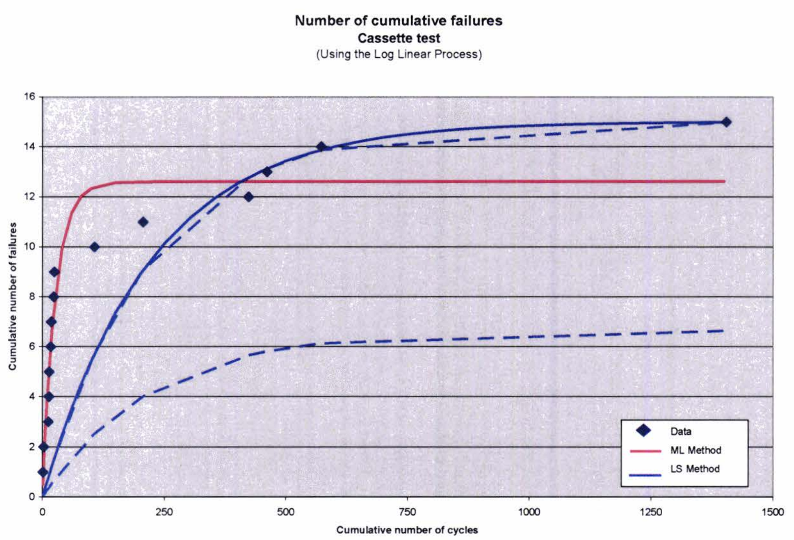


Figure J.8. Failures for the Cassette test fitting a Log Linear Process.

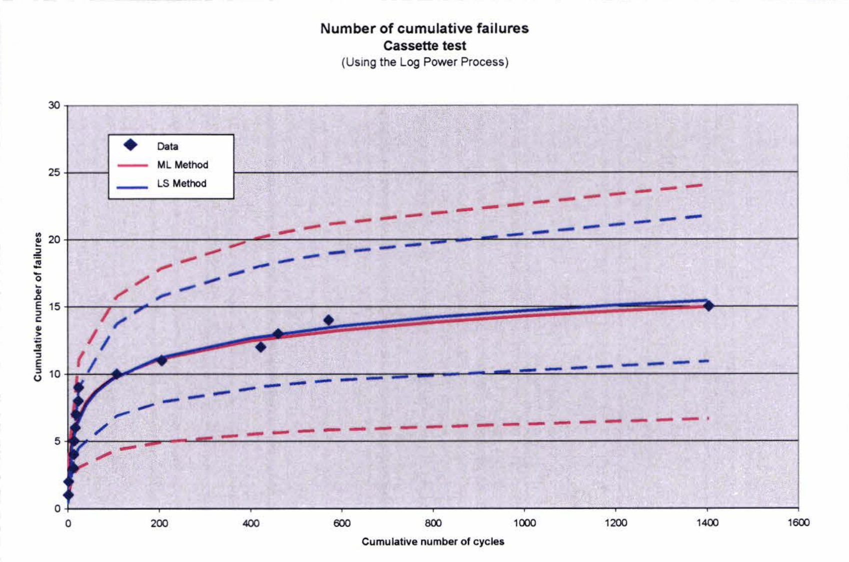


Figure J.9. Failures for the Cassette test fitting a Log Power Process.