Component State Prediction Based on Field Data

Master Thesis in Energy System Engineering

Linnea Johansson
Abstract

Component State Prediction Based on Field Data

Linnea Johansson

This master thesis is part of a big project at Siemens Industrial Turbomachinery (SIT) in Finspång aimed to use the operation experience available at SIT to predict the state of the gas turbines in general and some mechanical components in particular. The objective of the thesis is to continue the development of a prediction model based on experience data for estimations of a components lifetime. In a previous master thesis by Alessandro Olivi statistical analysis of environmental attributes effect on the expected lifetime of components in a gas turbine was performed. Olivi’s thesis constitutes the starting point on which to keep building to create a reliable prediction model.

In this thesis extensive validation tests have been performed in order to further quantify the reliability of the model. Investigations aimed towards finding ways to further develop and improve the prediction model are carried out. The relevant new findings are applied to the model and analysis concerning improvements in the prediction accuracy is carried out. It was revealed that the model is able to make accurate predictions for most of the validation points for each failure mode, but more research is needed to obtain a completely reliable prediction model.
Acknowledgements
The author wish to express a big thanks to the whole service department at Siemens Industrial Turbomachinery (SIT) in Finspång for allowing me to be a part of their team and for taking the time to answer any questions, special thanks is extended to:

- The supervisors at SIT Daniel Dagnelund and Davood Naderi for inspiring, challenging and rewarding guidance throughout the entirety of the master thesis project.

- Erik Ärlebäck for the opportunity to do a master thesis at SIT and for support and motivation during the project.

- Milan Lazic, Erik Lindgren and Simon Strömberg for help navigating the SIT databases and for retrieving additional data when needed.

- Project manager John Ayotte for interesting discussions and for illuminating how this small project is part of a larger project.

In addition the author likes to express great thanks to Bengt Carlsson for acting as the Uppsala University subject reader.
## Glossary

In this section, clarification of the definitions of some words, expressions and abbreviations as used in this report is offered.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Components</strong></td>
<td>Refers to the type of mechanical components analyzed in the thesis. Each gas turbine consists of 50 identical such components; referred to as a set of components.</td>
</tr>
<tr>
<td><strong>Data Set</strong></td>
<td>One data set consists of all the data from one inspection of a gas turbine, and all attributes of the machine e.g. type of filters used and the output. A data set also includes environmental attributes from the machine site e.g. ambient temperature and humidity.</td>
</tr>
<tr>
<td><strong>Dirt</strong></td>
<td>Refers to contaminations detected during inspection of the machine.</td>
</tr>
<tr>
<td><strong>EOC</strong></td>
<td>Equivalent Operation Cycles is a measurement of the number of cycles that takes into consideration the different operation behaviors e.g. number of fast/slow starts and stops (Olivi, 2016).</td>
</tr>
<tr>
<td><strong>EOH</strong></td>
<td>Equivalent Operation Hours is a time measurement that takes into consideration the different operation behaviors e.g. fuel, workload and number of operation cycles (Olivi, 2016).</td>
</tr>
<tr>
<td><strong>Failure Mode</strong></td>
<td>A category in a sorting system for causes of failure, possible failure modes could for example be corrosion or fatigue (Meher-Homji &amp; Gabriles, 1998). For secretes reasons the failure modes analyzed are refer to as failure modes A, B, C and D.</td>
</tr>
<tr>
<td><strong>First Replacement</strong></td>
<td>The first time maintenance is ever performed on a machine.</td>
</tr>
<tr>
<td><strong>g-prior</strong></td>
<td>In statistic this refers to an objective prior for the regression coefficients of a multiple regression i.e. our initial guess for the coefficients (Olivi, 2016).</td>
</tr>
<tr>
<td><strong>Component inspection</strong></td>
<td>Inspection of the components removed at maintenance to determine if they can be repaired or not. The unrepairable components are scrapped and repair of the repairable components are deployed. The components are deemed as good as new and reused after successful repairs (Dagnelund, 2017).</td>
</tr>
</tbody>
</table>
Maintenance Maintenance of a machine entails opening up the gas turbine and removal of all components analyzed in this thesis and replacing them with new ones. The used components are sent to Finspång for inspection and testing (Dagnelund, 2017).

Machine Gas turbine.

Scrap Rate Percentage of components in a gas turbine deemed as unrepairable during component inspection.

Set Size The number of removed components sent to Finspång for inspection and repair.

SIT Siemens Industrial Turbomachinery AB.

Unrepairable Technically impossible or too costly to repair for the SIT workshop.

**weib** An R package developed for internal use at SIT. `weib` implements Bayesian survival analysis based on the Weibull distribution for both censored and uncensored data. Censored data refers to data where the failure time is known only as before or after a point in time i.e. the time of the failure is unknown (Olivi, 2016).

**weibfit** A function within `weib` that constitutes the basis for the prediction model used for prediction of the scrap rate. The function is based on Bayesian survival model using Weibull regression on the scale parameter, leading to the Accelerated Failure Time model (Olivi, 2016).
Sammanfattning

Siemens Industrial Turbomachinery AB (SIT) i Finspång tillverkar gasturbiner som säljs till hela världen. De tillverkas efter olika specifikationer och är ämnade för olika syften. Varje turbin innehåller mekaniska komponenter som arbetar i en för dem fientlig miljö vilket förkortar deras livslängd. På SIT har man tillgång till sensordata från gasturbiner som insamlas under driften samt data insamlad från olika rapporter, t.ex. inspektions- och reparationsrapporter. Denna masteruppsats är del av ett större projekt på SITs serviceavdelning med syfte att bruka de gamla driftdata som finns tillgänglig på Siemens för att utvärdera skicket på gasturbinerna i allmänhet och komponenterna i synnerhet. Under service och komponentinspektionen blir de komponenter som är reparbara reparade och de komponenter som anses vara orepaperbara skrotas. Målet med masterarbetet är att fortsätta utvecklandet av en prediktionsmodell för estimering av antalet skrotad komponenter, vilken sedan kan brukas till estimeringar av komponenternas livstid.


Detta masterarbete består av tre delar som beskrivs nedan:


Den andra delen bestod av att genomföra undersökningar ämnade att förbättra prediktionssäkerheten hos modellen, ett flertal undersökningar genomfördes och de beskrivs kort nedan:

Data från turbiner som stått på samma geografiska plats jämfördes och analyserades. Dessa visade att kännedom om drifttiden och miljön turbinen står i inte är tillräckligt för att prediktera antalet skrotade komponenter. Med bakgrund av detta lades två nya attribut som speglar driften av gasturbinen till: driftstopstid ("downtime") och en indikator på lastnivån ("EHO/OH").
På SIT har det länge funnits en teori att antalet skrotade komponenter är högre första gången en maskin inspekteras än vid senare inspektioner (Barhanko, 2017). Denna teori har i arbetet validerats genom att med olika metoder jämföra data från första inspektioner med data från senare inspektioner. En annan teori var att det finns en korrelation mellan antalet skrotade komponenter och om smuts hittas i inloppet till kompressorn vid inspektionen (Karlsson, 2017). Denna teori undersöktes på ett liknande sätt som den föregående teorin och resultaten indikerade att det finns visst belägg även för denna teori.

Attributanalys genofördes med syfte att hitta de attribut som är korrelerade med antalet skrotade komponenter. Två analysmetoder kallade R²-metoden och b-metoden utvecklades och analyserades. R²-metoden användet kurvanpassning samt beräkning av minstakvadratfelet och b-metoden använder weibfits rangordning av attribut. De två metoderna visade sig brukbara i både teorin och praktiken då de förbättrade modellens noggrannhet.

Den tredje delen bestod i att implementera de relevanta resultaten från genomförda undersökningar i modellen och implementeringen ledde till minskade prediktionsfel. Genom att förändra modellens indata, det vill säga ta bort vissa tidigare brukade attribut samt lägga till nya, kunde modellen göra tillförlitliga prediktioner för en majoritet av valideringspunkterna.

De resultat som tagits fram i detta arbete visar också på ett behov av ytterligare forskning för att ta fram en pålitlig modell.
Executive Summary

For validation of the previous prediction model developed at SIT three validation tests were formulated and validation criteria quantified. The results from the validation tests revealed that the model has the ability to predict systems and make reliable predictions if relevant attributes are provided to the model. Unfortunately, the environmental attributes analyzed by Olivi were found to have limited explanation value, making it hard for the model to make accurate predictions. It was also discovered that the amount of calibration data currently available at SIT yields results with questionable reliability.

It was theorized that the number of unrepairable components is higher the first time the components are inspected than in later component inspections in the same machine. The results reveal that this is a valid statement and that the scrap rate is influenced by whether it is the first replacement of the components in a machine or not. It was also theorized that the components studied in this report are sensitive to dirt in the inlet air. To examine this inspection reports are viewed and mentioning’s of dirt in the inlet to the compressor noted. There are results that indicate that the scrap rate is influenced by whether there is dirt found in a machine or not.

One investigation field was aimed towards finding methods to determine if attributes correlated with the number of scrapped components. Two methods named the R²-method and the b-method, were deployed and analyzed. Both methods were deemed as valid and both proved useful for improving the accuracy of the prediction model.

The results from the investigations were implemented to the model and resulted in improvements in the prediction accuracy. With some changes to the input to the model i.e. adding new attributes and removal of some old, it was discovered that the model is able to make reliable predictions for most of the validation points for each failure mode, but more research is needed to obtain a reliable prediction model.
# Table of Contents

1 Introduction ............................................................................................................. 1  
1.1 Objective ............................................................................................................. 2  

2 Background ............................................................................................................. 3  
2.1 General description of a gas turbine ................................................................. 3  
    2.1.1 Reasons for failure .................................................................................... 3  
2.2 The data .............................................................................................................. 4  

2.3 Useful formulas ................................................................................................... 5  
    2.3.1 Weighted values ....................................................................................... 5  
    2.3.2 Skewness ................................................................................................. 6  

2.4 weib model ......................................................................................................... 7  
    2.4.1 Theory - Parametric analysis .................................................................... 7  
    2.4.2 The model ............................................................................................... 9  
    2.4.3 Assumptions ...........................................................................................11  

3 Methodology ..........................................................................................................12  
3.1 Validation of the weib model ............................................................................12  
    3.1.1 Prediction accuracy of validation data .....................................................12  
    3.1.2 Model stability .......................................................................................13  
    3.1.3 Analysis of the model based on a simulated system ...............................13  

Investigations to further improve the model ............................................................17  
    3.1.4 Same site data analysis ...........................................................................17  
    3.1.5 First replacement of components in a machine; effect on the scrap rate ....17  
    3.1.6 Dirt detected in inlet of machine; effect on the scrap rate .......................19  
    3.1.7 Attribute analysis ...................................................................................19  
    3.1.8 New attributes .......................................................................................21  

3.2 Implementation of investigation findings in the model ....................................21  

4 Results ....................................................................................................................22  
4.1 Validation of the weib model ............................................................................22  
    4.1.1 Prediction accuracy of validation data .....................................................22  
    4.1.2 Model Stability ......................................................................................24  
    4.1.3 Analysis of the model based on a simulated system ...............................26  

4.2 Investigations to further improve the model ....................................................36  
    4.2.1 Same site data analysis ...........................................................................36  

VII
4.2.2 First replacement of components in a machine; effect on the scrap rate .................. 38
4.2.3 Dirt detected in inlet of machine; effect on the scrap rate ................................... 42
4.2.4 Attribute analysis ........................................................................................................ 44
4.3 Implementation of investigation findings in the model .................................................. 47
  4.3.1 Using the $R^2$-method – model one ....................................................................... 47
  4.3.2 Using the b-method – model two .......................................................................... 50
  4.3.3 Outliers ...................................................................................................................... 51
  4.3.4 Discussion - Implementation of investigation findings in the model ................... 52
5 Discussion ............................................................................................................................. 54
  5.1 Limitations and possible error sources ................................................................. 54
  5.2 Suggestions for future work and further analysis .................................................... 54
6 Conclusions .......................................................................................................................... 56
7 Bibliography ......................................................................................................................... 57
  7.1 Oral References ............................................................................................................ 58
8 Appendix ................................................................................................................................. 59
  8.1 Appendix A – Possible outliers ..................................................................................... 59
  8.2 Appendix B – Zoomed Figure 38 to Figure 40 ......................................................... 60
    8.2.1 Figure 38 ................................................................................................................... 60
    8.2.2 Figure 39 ................................................................................................................... 61
    8.2.3 Figure 40 ................................................................................................................... 63
  8.3 Appendix C – Original and best results comparison ............................................ 65
1 Introduction
Prediction models created to help manage the uncertainty of the future are a part of our everyday life e.g. weather forecasts and predictions of the possible effects of carbon dioxide emissions. Manufacturing business strive to predict the future demands of their product to plan the production and the need for storage, and to minimize the waste, i.e. to optimize their production and delivery chain. This optimizing result in financial and environmental benefits for the company e.g. reduced costs and emissions if less superfluous products are simulated and delivered.

Siemens Industrial Turbomachinery AB (SIT) located in Finspång manufactures gas turbines for different specifications, purposes and locations around the world. Each turbine contains mechanical components that are operating in a hostile environment, decreasing their lifetime. During maintenance and inspection some of these components are repaired for reuse in a turbine, and some are classified as unrepairable and scrapped. Manufacturing components for a gas turbine is a complex, time and energy consuming process, and it would be most useful to be able to reuse as many components as possible as this would save money, energy and resources. Accurate predictions of the lifetime of key components in a machine will allow the machine to operate for longer stretches without maintenance i.e. replacement of the key components. Finding the optimal maintenance plan which maximizes the operation time and minimizes the number of unrepairable components would benefit Siemens, their customers and the environment; as it would yield a higher efficiency in the turbines.

As a step towards the optimal maintenance plan development of a prediction model to predict the number of unrepairable components at the time of maintenance has been deployed at Siemens, this model can also be used for estimations of the components lifetime. The unrepairability of the components can be caused by one or a combination of four independent sources, called failure modes. For secretes reasons the failure modes analyzed are refer to as failure modes A, B, C and D.

All types of predictions are based on experience and/or data from previous events. SIT has access to sensor data collected during the operation of each gas turbine and data from various reports e.g. inspection and repair reports. The reports includes information about the turbines operating time, the number of scrapped components and the failure mode that caused the component to be deemed as unrepairable, but it is unknown when the damage to the component occurred.

This master thesis is part of a big project at the service department at SIT aimed to use the operation experience available at SIT to predict the state of the gas turbines in general and the components in particular, and determine the optimal time for maintenance. In the spring of 2016 Alessandro Olivi and Aleksandra Neupokoeva separately carried out their master thesis at SIT as part of this project. Olivi’s thesis (Olivi, 2016) aimed to estimate the lifetime of components from gas turbines located at different sites, and evaluate the optimal replacement time. Olivi analyzed the influence of some environmental attributes and the failure modes effect on the
components expected lifetime. The aim of Neupokoeva’s thesis (Neupokoeva, 2016) was to find a way to determine the nominal load of the gas turbine based on sensor data. Neupokoeva used different data mining methods to find patterns which could indicate a harmful influence on the lifetime of components in the gas turbine. These two theses constitute the starting point on which to keep building to create a reliable prediction. In this report the focus is on Olivi’s thesis; validate and evolving the method aimed to obtaining a reliable prediction model for prediction of the number of components in need of scrapping i.e. the lifetime of components in a gas turbine.

1.1 Objective
The aim of this master thesis is to continue SIT’s work on describing and interpreting reports and sensor data, and to further aid decision support for maintenance by continuing the development of a prediction model to estimate the number of components in need of scrapping at the time t, where the time T represents the operation time at the time of maintenance in a gas turbine. The model is limited to the predictions of the state of one type of components in one type of gas turbine, and this is the scope for the thesis. The objective of the thesis can be separated into three parts.

1. Validate the existing prediction model which is based on environmental influences on the lifetime of components. Select appropriate validation tests and quantify validation criteria to reflect the purpose of the prediction model.
2. Conduct investigation aimed towards finding methods to further improve the accuracy of the predictions made by the model e.g. investigates possible influence on the lifetime of components and find methods to implement the findings in the model.
3. Implement the new findings from the investigation in the model and analyze the possible improvements in the accuracy of the prediction model.
2 Background
In the background section some necessary background information for the further understanding of the methodology and the results presented in this report is offered. The section includes a general description of a gas turbine and the data used, some useful formulas, and a summary of the theory and methodology used in Olivis thesis.

2.1 General description of a gas turbine
A gas turbine is a heat engine and converts one type of energy to another i.e. it converts heat to mechanical labor. The heat energy is commonly obtained by combusting fuel combined with oxygen taken from the surrounding air. The three main components in a gas turbine are:

- Compressor
- Combustion chamber
- Turbine

Of the air that leaves the compressor only around 25% is used for combustion, the rest is used for cooling purposes to insure that the gas mixture that leaves the combustion chamber has a temperature below the endures limit of the rest of the components (Larsson, 2011).

A figure depicting a general gas turbine can be viewed in Figure 1.

![Figure 1: Picture of a general gas turbine with the main components marked. In the parts pre combustor the air is cold and in the post combustor the air is hot (Siemens Industrial Turbomachinery AB, 2015).](image)

2.1.1 Reasons for failure
The mechanical components used for the analysis in the thesis is part of the hot section of the turbine, see Figure 1. Below some common reasons for failure in hot components in gas turbines are described. The anonymous failure modes have similarities to the reasons for failure expressed below.
2.1.1.1 Fatigue
A significant number of failures of gas turbines and its components are caused by different types of fatigue. In general fatigue is caused by repeated application of fluctuating stresses. Resonant fatigue occurs if the absorption of the periodic input energy is not sufficient in the components; the input energy amplifies and the stress grows causing a fatigue crack. Low cycle fatigues are a result of turbines start/stop cycles; minute flaws grow into cracks that rapture as they reach a critical size. Thermomechanical fatigue is caused by thermal stress e.g. differential expansions of hot section components (Meher-Homji & Gabriles, 1998).

2.1.1.2 Corrosion
Components in the hot section of a gas turbine are often submitted to a combined oxidation-sulphidation phenomena called corrosion. Hot corrosion occurs in two types; high temperature corrosion occurs at temperatures between 825-950°C and low temperature corrosion occurs at temperatures between 700-800°C. A third type of corrosion is standby corrosion; caused by moisture in the ambient air and occurs during shutdown. Fatigue strength is negatively affected by a corrosive environment (Meher-Homji & Gabriles, 1998).

2.1.1.3 Creep
Creep is a type of deformation that solid materials substance over time. One example is the lengthening over time of a metal staff that’s submitted to constant pulling force and temperature (Lundh, 2000). Creep affects hot section components and occurs when components operate over time under high temperature and stress (Meher-Homji & Gabriles, 1998).

2.1.1.4 Erosion
There are two types of erosion that occurs in gas turbines, one milder and one more severe. The milder form is called particle erosion and is caused by particles larger than 5-10 microns in the axial flow. The more severe type of erosion is called hot gas erosion: The components in the hot section of the turbine operate at temperatures hundreds of degrees lower than the gas path temperatures. The components are protected by a natural boundary; a thin layer of cooling air. If the thin film of air brakes or if the cooling effectiveness drops; the roughness on the surface of the component is subjected to high thermal stress cycles. For each cycle the roughness (erosion) on the surface increases, worsening the problem (Meher-Homji & Gabriles, 1998).

2.2 The data
This master thesis analysis field data collected by SIT, SIT manufactures and delivers gas turbines to all parts of the world and preform maintenance on these. This thesis only analyses data concerning one type of mechanical component in one type of gas turbine. Each gas turbine contains a set of 50 such identical components. During maintenance the components are removed from the machine for inspection and replaced with new components, during the inspection the used components deemed as repairable are repaired and those deemed as unrepairable are scrapped. Note that no failure of components has accrued. The components un-repararability is caused by one of or a combination of four failure modes, A, B, C and D, which are assumed to be independent to each other. It is unknown at what time the component
became unrepairable or at what time it would have become unrepairable if operation without maintenance would have continued i.e. the data used is always either right or left censored (Dagnelund, 2017).

The calibration data consist of 41 data sets, each one with approximately 50 components, from 19 different sites. The set size is the number of removed components sent to inspection, for various reasons all 50 of the removed components are not always sent. All data sets have a set size of 50 components except from three of them which have set sizes of 38, 47 and 48 respectively.

In order to calculate the hours of operation equivalent to base load continuous duty operation, SIT has uses a concept of Equivalent Operation Hours (EOH). This time measurement takes into consideration the different operation behaviors e.g. fuel, workload and number of operation cycles. In a similar procedure an equivalent measurement for the number of equivalent operation cycles (EOC) has been developed to reflect e.g. number of fast/slow starts and stops (Olivi, 2016).

An example on the structure of the data is presented in Table 1 (the data is fictional). The total number of scrapped components can be a lower number than the sum of the scrapped components for each failure mode since the un-reparability can be caused by more than one failure mode simultaneously. In addition to the data represented in Table 1 the geographical location for each machine is known, therefore it is easy to find values for environmental attributes that describes the environmental operation conditions for each data set.

Table 1: An example on how the data used in the analysis was structured. Columns A, B, C and D represent the number of scrapped components caused by each failure mode; A, B, C and D. Scrap rate is the fraction of scrapped components in each machine. NOTE: this is an example to display the data structure and all numbers and names are fictional.

<table>
<thead>
<tr>
<th>Machine Number</th>
<th>Site Name</th>
<th>Operation Time (EOH)</th>
<th>Set Size</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total number of Scrapped components</th>
<th>Scrap Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>000001</td>
<td>Polacksbacken</td>
<td>21 569</td>
<td>50</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0,120</td>
</tr>
<tr>
<td>000002</td>
<td>Ångströmlaboratoriet</td>
<td>23 789</td>
<td>50</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0,100</td>
</tr>
<tr>
<td>000003</td>
<td>Ekonomikum</td>
<td>18 623</td>
<td>49</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>0,102</td>
</tr>
<tr>
<td>000004</td>
<td>Blåsenhus</td>
<td>20 147</td>
<td>48</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0,042</td>
</tr>
<tr>
<td>000005</td>
<td>Geoentrum</td>
<td>17 532</td>
<td>50</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0,040</td>
</tr>
</tbody>
</table>

2.3 Useful formulas

2.3.1 Weighted values

Since the set size varies across the data sets, values used to compare data sets require weighting to yield an accurate comparison. The formula for calculation of the weighted mean is expressed in equation (1).

$$\bar{x}^* = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}$$  

(1)
where \( w_i \) are the weights and \( x_i \) are the observations. The formula for calculation of the weighted standard deviation is expressed in equation (2).

\[
\sigma^* = \sqrt{\frac{\sum_{i=1}^{N} w_i (x_i - \bar{x}^*)^2}{(M - 1) \sum_{i=1}^{N} w_i}}
\]

(2)

where \( N \) is the number of observations, \( M \) is the number of nonzero weights, \( w_i \) are the weights, \( x_i \) are the observations and \( \bar{x}^* \) is the weighted mean (National Institute of Standards and Technology, 1996). The weight is the set size for each data set.

2.3.2 Skewness

Skewness is a common way of describing a distribution and it has previously been used at SIT to describe the load distribution in gas turbines. The formula for calculating the skewness factor is expressed in equation (3), normal distribution has skewness factor zero (Indrasukhsri, 2017).

\[
S = \frac{N \sqrt{N - 1}}{N - 2} \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{(\sum_{i=1}^{n} (x_i - \bar{x})^2)^{3/2}}
\]

(3)

where \( N \) is the number of samples, \( x_i \) is the individual values and \( \bar{x} \) is the average of all the values.

An example of positive and negative skewness is depicted in Figure 2.

![Figure 2: Examples on how positively (right-skewed) and negatively (left-skewed) skewed distributions can look (University of California Santa Cruz, 2014).](Figure 2: Examples on how positively (right-skewed) and negatively (left-skewed) skewed distributions can look (University of California Santa Cruz, 2014).)
2.4 weib model

The Weib model estimates the lifetime of components from gas turbines located at different geographical sites by analyzing the influence of various environmental attributes and failure modes effect on the component lifetime (Olivi, 2016).

Olivi applied statistical methods to predict the reliability of components in gas turbines located in different sites and environments, and estimate the lifetime of the components. The environments that hold gas turbines are described with different environmental attribute e.g. latitude, mean relative humidity and mean temperature. Olivi attempted to quantify the influence of the environmental attributes on the lifetime. Olivi also analyzed the optimal replacement time and the effect different failure modes has on the lifetime (Olivi, 2016).

The methods used by Olivi in his master thesis are described below. Olivi created an R package called weib to run all the methods used.

The parametric analysis is the most relevant part of Olivi’s statistical analysis as it constitutes the basis for the whole prediction model, the theory is described below.

2.4.1 Theory - Parametric analysis

In parametric reliability analysis it is assumed that time-to-event data follows a predefined distribution based on a set of parameters, and the most commonly used distribution is the Weibull distribution (Weibull, 1961).

In 1937 Waloddi Weibull invented the Weibull distribution and in 1951 he published a paper on the subject where he claimed that data could both select the distribution and fit the parameters. It has been proven by the mathematician E.J Gumbel (1891-1966) that if a part has multiple failure modes the time to first failure is best estimated using the Weibull distribution. A large number of engineering problems can be solved using Weibull analysis, Weibull analyses includes failure forecasting, maintenance planning and cost effective replacement strategies, and evaluating corrective action plans. The main advantages of Weibull analysis is the ability to provide realistically accurate failure analysis with particularly small samples. The Weibull distribution is defined by two positive parameters; the shape parameter \( \beta \) and the scale parameter \( \eta \). The shape parameter, \( \beta \), is able to provide a clue to the physics of the failure (Abernethy, 2006).

- \( \beta<1 \) indicates infant mortality and the failure rate will decline with time.
- \( \beta=1 \) signifies random failure and the failure rate is constant over time.
- \( \beta>1 \) indicates wear out failures and the failure rate will increase with time.

The Weibull density function \( f(t) \) is displayed in equation (4) (Abernethy, 2006). Examples of the effects of the two parameters on the Weibull distribution can be viewed in Figure 3.

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

(4)
The Weibull cumulative distribution function, $F(T)$ (equation (5)), corresponds to the probability that an event occurs before the point $T$ and will later be referred to as the failure function. The probability that an event has not occurred at the point $T$ is described by the survivor function, $S(T)$ (equation (6)) (Abernethy, 2006).

$$F(t) = P(T < t) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta} \tag{5}$$

$$S(t) = 1 - F(t) = P(T > t) = e^{-\left(\frac{t}{\eta}\right)^\beta} \tag{6}$$

The hazard function, $h(t)$ (equation (7)), describes the instantaneous failure rate at any point in time (Abernethy, 2006).

$$h(t) = \frac{f(t)}{F(t) - 1} = \frac{\beta \left(\frac{t}{\eta}\right)^{\beta-1}}{} \tag{7}$$

### 2.4.1.1 Bayesian Weibull AFT model

To include and quantify the effects of different attributes on the lifetime of the components the Weibull Accelerated Failure Time (AFT) model is used. A basic assumption of ATF models is that the effects of attributes are multiplicative with respect to the survival time. The AFT model describes the acceleration of survival time as a function of attributes i.e. AFT estimates the multiplicative or acceleration factor, which is constant over time. In the Weibull AFT model the parameter $\eta$ is reparametrized as equation (8) (Olivi, 2016).
\[ \eta = e^{(b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p)} \]  

where \( b = b_0, \ldots, b_p \) is the set of coefficients that measures the impact of the \( p \) attributes \( x_i \) on the expected lifetime of the components (Olivi, 2016).

In Weibull AFT the parameter \( \beta \) does not vary with the attributes, but for inference purpose it is reparametrized as equation (9) (Olivi, 2016).

\[ \beta = e^{a_0} \]  

For estimation of the distributions defining parameters \( \beta \) and \( \eta \) a Bayesian approach is used, this method combines our prior belief with the observed data. Our prior beliefs are represented by a g-prior which is an objective prior for the regression coefficients of a multiple regression i.e. a guess of the coefficients distributions. The Bayesian approach results in the full posterior distribution of the coefficients, which enable us to calculate credible intervals on the parameters which can be implemented in the failure function (Olivi, 2016).

A normal distributed g-prior on \( a_0 \) and a multivariate normal g-prior on \( b \) are selected as equation (10) and (11) respectively (Olivi, 2016).

\[ p(a_0) = \mathcal{N}(0,2) \]  

\[ p(b) = \mathcal{N}_{p+1}(0, g(X^TX)^{-1}) \]

where \( g \) is a positive scalar set to be equal to \( n \), and \( X \) is the matrix of explanatory variables (intercept included) of dimension \( n \times (p + 1) \) (Olivi, 2016).

Through the AFT model the full posterior distribution of the coefficients has been estimated based on the priors in equation (10) and (11), and on the observed data. Using Monte Carlo Markov Chain (MCMC) methods the parameters (\( \eta \) and \( \beta \)) values are drawn from the posterior distribution of the coefficients (Olivi, 2016).

All posterior distributions from sampled coefficients are bell-shaped. The mean value of the coefficients i.e. best point estimate, and a 95% confidence interval is derived (Olivi, 2016).

### 2.4.2 The model

Within the R-package `weib` there is a function called `weibfit` that constitutes the basis for the prediction model used for prediction of the scrap rate. `weibfit` uses the described theory in section 2.4.1 to predict the values for \( \beta \) as expressed in equation (9), and \( b_0 \) and \( b_1 - b_p \) from the parametrization of \( \eta \) expressed in equation (8). These predicted values can be used for calculations of the predicted value of the failure function, equation (5) (Olivi, 2016).

Olivi has singled out eight environmental attributes believed to influence the lifetime of the components, see Table 2. The values of these attributes are part of the input data to the model. All variables are scaled to mean zero and standard deviation one before applying the statistical
Each coefficient value predicted by `weibfit` is returned as a mean value and with a confidence interval of 95%. If both the high and low value in the confidence interval has the same sign i.e. its statistically significant that the attribute has a negative or positive influence on the estimated lifetime of the components, the coefficient \( b \) is marked with an asterisk (Olivi, 2016).

As the failure modes are assumed to be independent to each other the effect the attributes have on the lifetime of the components varies thus the model returns different output values for each failure mode. Note that according to Olivi erosion factor is not used as an input attribute for failure mode B the as the erosion factor do not affect failure mode B (Olivi, 2016). See Figure 4 for a schematic description of the model.

**Table 2:** The Environmental attributes believed to have effect on the estimated lifetime of the components and their input units as chosen by Olivi (Olivi, 2016).

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Temperature January</td>
<td>°C</td>
</tr>
<tr>
<td>Mean Temperature July</td>
<td>°C</td>
</tr>
<tr>
<td>Mean Relative Humidity January</td>
<td>%</td>
</tr>
<tr>
<td>Mean Relative Humidity July</td>
<td>%</td>
</tr>
<tr>
<td>Distance From Sea</td>
<td>Km</td>
</tr>
<tr>
<td>Altitude</td>
<td>m</td>
</tr>
<tr>
<td>Latitude</td>
<td>°</td>
</tr>
<tr>
<td>Erosion Factor</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Note that according to Olivi erosion factor is not used as an input attribute for failure mode B the as the erosion factor do not affect failure mode B (Olivi, 2016). See Figure 4 for a schematic description of the model.

**Figure 4:** Schematic description of the needed input signals and the received output from the R-function `weibfit`. All output values are given as a mean value and with a confidence interval of 95%.
2.4.3 Assumptions
The following assumptions are made in weib (Olivi, 2016):

- The density function for the number of scrapped components is Weibull distributed.
- The shape parameter $\beta$ for a failure mode is the same for all data sets.
- The failure modes are independent.
3 Methodology
In this section all methods used are presented in three subsections; 3.1 Validation of the \textit{weib} model, 3.2 Investigations to further improve the model, and 3.3 Implementation of investigation findings in the model.

3.1 Validation of the \textit{weib} model
The difficulty in creating models for describing systems is to make the model accurate and reliable. For a model to be useful one must be able to rely on its predictions. To gain confidence and reliability in the model it is validated. The first step in the validation process is to determine the purpose of the model and define validation tests to control if the model is fit to be used for this purpose (Ljung & Glad, 2004).

In this case the purpose of the model is to predict the number of unrepairable components in a gas turbine at the time of maintenance. To determine the usability of the prediction model created by Olive it is validated using the three different validation methods described below.

The validation data consist of 13 data sets, each one with 50 identical components. The validation data has the same structure and known parameters as the calibration data.

After visual inspection of the data and in consensus with the supervisors at SIT it is decided not to include failure mode A in the validation since rarity of the failure mode combined with the low number of data sets makes the basis for the prediction to slim.

3.1.1 Prediction accuracy of validation data
The first validation method aims to quantify the accuracy of the predictions made by the prediction model. An R-function to carry out the following was constructed;

1. Insert calibration data into model to receive predicted values of $\beta$ and the coefficients $b$ for each failure mode.
2. Load the validation data and performed the same pretreatment of that data as for the calibration data e.g. remove the average value.
3. Use the predicted mean values for the coefficients $b$ and the known values $x$ for each attribute to calculate $\eta$ for each validation set and failure mode according to equation (8).
4. Use the calculated $\eta$ and the predicted $\beta$ to calculate the failure function as expressed in equation (5) for each validation set.
5. The predicted number of scrapped components in each of the validation sets is calculated by multiplying failure fraction received from the failure function and the set size.
6. The predicted mean value of number of scrapped components is compared with the real number of scrapped components and an error is calculated as the difference between the true value and the predicted value.

The configuration of the error calculation results in the following interpretation of the error; a negative error reveals that the models prediction of the number of unrepairable components is
higher than the actual number of unrepairable components, and a positive error reveals the opposite.

3.1.1.1 Validation criteria
To pass this validation test the model must for each failure mode display a maximum absolute value of the errors of five and the error must from visual inspection be deemed as centered round zero. The limit of five is based on SIT´s current acceptance of error between production and demand. It is desirable that the errors are centered round zero as this indicates that there is no constant error in the model causing it to always make over or under estimations.

3.1.2 Model stability
To test the stability of the models predictions the model is recalibrated using all available data i.e. calibration and validation data combined, then a comparison is made between the output parameter $\beta$ and the coefficients $b$ when only calibration data was used and the new output. A desired outcome is to find that the mean values of the parameters have remained the same and the confident interval has become smaller i.e. that the model had the correct value for the coefficients $b$ originally and has become more confident in its value. It is also desirable for the mean value of the output parameter remains the same sign as this display the models confidence in the nature of the effect the attribute has on the lifetime of the components.

3.1.2.1 Validation criteria
Only the mean values and not the confidence intervals of the coefficients are included in the criteria as it is the mean value that is used for calculation of the value of the failure function (equation (5)). To pass this validation test the model must for each failure mode display predicted mean values that does not switch sign between the two data sets, and the mean value for the larges data set (calibration and validation data combined) lies within the confidence interval of the smaller data set (only calibration data).

3.1.3 Analysis of the model based on a simulated system
To determine the models ability to accurately identify a Weibull distributed system and predict the output values from that system the model need to be tested on a known system. To be able to make variations in the system structure a simulated system is created in Microsoft Excel.

3.1.3.1 The system
The simulated system consists of a function that determines how the output signal $y$ is depending on the input signals, see equation (12) (compare with the failure function, equation (5)).

$$y = 1 - e^{\left(\frac{T}{\eta}\right)^\beta}$$

(12)

where the input signal $T$ represents the operation time at the time of maintenance and is as a random number between 15000 and 25000, the input signal $\beta$ is chosen as two as a $\beta$ larger than one indicates wear out failures (Abernethy, 2006), and the input signal $\eta$ is created according to equation (13) (compare with the parametrization of $\eta$, equation (8)).
\[ \eta = \frac{e^{(A_1 + A_2 + \cdots + A_7)}}{Z} \]  

where \( A_1 - A_7 \) represents the attributes that influence the lifetime of the components and are random values between different intervals in the three parts described below, and \( Z \) is a size parameter that ensures that the output \( y \), which represent the scrap rate fraction i.e. the failure rate, are spread between zero and one. The influence the attributes have on the output value \( y \) vary depending on how the intervals for \( A_1 - A_7 \) are chosen e.g. if \( A_1 \) and \( A_2 \) both are chosen as random numbers between 0-1 they both have the same influence on the output value (even if they have different values) and if \( A_2 \) instead is chosen as a random number between 0-10, \( A_2 \) has a ten times bigger influence on the output value than \( A_1 \). Note that this is only is accurate for the simulated system i.e. in the true system if e.g. one temperature driven attribute is expressed in °K it will not automatically have a greater influence on the number of scrapped components than another temperature driven attribute expressed in °C in spite of probably having a larger absolute value.

Using the simulated systems \texttt{weibfit}'s ability to both predict the number of scrapped, i.e. unrepairable, components and identify a given Weibull distribution will be analyzed. The error of the predicted number of scrapped components will be calculated as previously described in the first validation method. The error of the estimated \( \eta \) and \( \beta \) will be calculated as 

\[ \text{Error}_\eta = \frac{\eta_{\text{real}}}{\eta_{\text{pred}}} \]

and 

\[ \text{Error}_\beta = \frac{\beta_{\text{real}}}{\beta_{\text{pred}}} \]

The tests performed on \texttt{weibfit} with the simulated systems can be divided into part 1-3 described below.

### 3.1.3.2 Part 1

The aim of this part is to examine if \texttt{weibfit} can predict a system when all the data that influence the output is known, and to see how \texttt{weibfit} performs with different amounts of calibration data for a varied number of influential attributes.

In this part three and seven attributes are used, and all of them have been chosen to have a random value between 0-1 i.e. they all have the same influence on the output value. The size parameters, \( Z \), are 0,0001 and 0,001 respectively. The number of data sets in the calibration data originates as 500 and is reduced to 50 for both numbers of used attributes. The estimated distribution of the number of scarped components for three and seven attributes can be viewed in Figure 5.
15

Figure 5: The estimated distribution of the number of scarped components for three (to the left) and seven (to the right) attributes. The estimations are made using 500 data sets.

3.1.3.3 Part 2

The aim of part 2 is to examine if weibfit can handle attributes with varied influence on the output, and see how it performs with different amounts of calibration data.

In this part two different styles of distribution is used; the first one is similar to the distribution used in Part 1 and the second one has a large number of extreme values i.e. scrap rate fraction values ($y$) of one and zero. The estimated distribution of the number of scarped components for the two styles can be viewed in Figure 6.

In this part seven attributes are used and they are varied in size i.e. their influence on the output value varies. The combination used for the first distribution style is displayed in Table 3. The combination used for the second distribution style is displayed in Table 4. The input in weibfit is the time $T$ and all the attributes that affect the output value and the number of calibration data sets starts as 500 and are then reduced to 50.

Table 3 Displays the variation in value on the used attributes in part 2 for the first distribution style. All attribute values are random numbers within the interval displayed in the table.

<table>
<thead>
<tr>
<th>Value A1</th>
<th>Value A2</th>
<th>Value A3</th>
<th>Value A4</th>
<th>Value A5</th>
<th>Value A6</th>
<th>Value A7</th>
<th>Size Parameter, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.5]</td>
<td>[0,0.5]</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 4: Displays the variation in value on the used attributes in part 2 for the second distribution style i.e. the distribution with a large number of extreme values. All attribute values are random numbers within the interval displayed in the table.

<table>
<thead>
<tr>
<th>Value A1</th>
<th>Value A2</th>
<th>Value A3</th>
<th>Value A4</th>
<th>Value A5</th>
<th>Value A6</th>
<th>Value A7</th>
<th>Size Parameter, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.5]</td>
<td>[0.5]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>1500</td>
</tr>
</tbody>
</table>
Figure 6: Estimations of the two styles of distributions used in Part 2 and 3. The left one is similar to the distributions used in Part 1 and the right one has a large number of extreme values. The estimations are made using 500 data sets.

3.1.3.4 Part 3

The aim of part three is to examine if weibfit can handle the inclusion of non-influence attributes and the exclusion of influential attributes in the input i.e. can the model identify non-influential attributes and can it compensate for the lack of knowledge of influential attributes to make accurate prediction.

In part 3 up to twenty attributes of varied size are used. The input to weibfit is the time $T$ and the attributes that affect the output value plus 13 non-influential attributes or six of the seven influential attributes, see Table 5. Both distribution styles in Part 2 are represented; the first is presented in test 1 and 2, and the second is presented in test 3 and 4, in Table 5. The number of calibration data sets starts as 500 and are then reduced to 50.

Table 5: Displays the variation in input and in value on the used attributes in part 3. All attribute values are random numbers within the interval displayed in the table.

<table>
<thead>
<tr>
<th>Test</th>
<th>Value A1</th>
<th>Value A2</th>
<th>Value A3</th>
<th>Value A4</th>
<th>Value A5</th>
<th>Value A6</th>
<th>Value A7</th>
<th>Value A8-A20</th>
<th>Input attributes in weibfit</th>
<th>Size parameter, $Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.5]</td>
<td>[0,0.5]</td>
<td>[0,1]</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>A1-A20</td>
<td>0.0001</td>
</tr>
<tr>
<td>2</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>[0,0.5]</td>
<td>[0,0.5]</td>
<td>[0,1]</td>
<td>[0,0.1]</td>
<td>[0,0.1]</td>
<td>A1-A6</td>
<td>0.0001</td>
</tr>
<tr>
<td>3</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.5]</td>
<td>[0.5]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>A1-A20</td>
<td>1500</td>
</tr>
<tr>
<td>4</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.1]</td>
<td>[0.5]</td>
<td>[0.5]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>A1-A6</td>
<td>1500</td>
</tr>
</tbody>
</table>

3.1.3.5 Validation criteria

To pass the third validation test the model must display the ability to accurately predict some systems i.e. maximum absolute value of the prediction errors of five and the error must from visual inspection be deemed as centered round zero, furthermore the errors for $\beta$ and $\eta$ must be smaller than 20%.
Investigations to further improve the model
In this subsection all the methods for the investigations aimed to improve the prediction accuracy of the weib-model is presented.

3.1.4 Same site data analysis
The number of unrepairable components in a set has previously been modelled for prediction based on the assumption that environmental attributes effect the expected lifetime of the components (Olivi, 2016). If exclusively knowledge about environmental attributes is sufficient to predict the scrap rate one would expect to find similar scrap rates in data sets originating from the same geographical site. By looking at data originating from the same location one can hope to find if there is a need to evolve the model beyond environmental attributes.

The components included in the analysis originate from the 15 sites with multiple data points. All data sets from one site share all the same environmental attributes, and also have the same article number on the components. To investigate if the scrap rates vary within the same site the mean scrap rate and its standard deviation is used.

To analyze if the eventual variations in the scrap rates in data originating from the same site are explained by variations in EOH and EOC at the time of maintenance the correlation between the weighted standard deviation in the scrap rates, and EOH and EOC is investigated. The difference may also be a result of the variation in in actual operating hours and operating cycles, thus the correlation between the weighted standard deviation in the scrap rates, and operation hours and operation cycles is investigated. Both analyses are done by creating bubble plots with the weighted standard deviation of the scrap rate represented in the size of the bubbles and the weighted standard deviation of the hours and cycles are on the axis. If the cause of the variated scrap rate at the sites is the difference in hours and cycles, one would find that the size of the bubbles grew bigger as the weighted standard deviation for hours and cycles grew.

3.1.5 First replacement of components in a machine; effect on the scrap rate
A theory at SIT is that the first time you replace components in a machine the scrap rate will be higher than in later replacements. This might be explained by the fact that a new customer has a learning curve and runs the machine in a suboptimal way during the first 20 000 operation hours. Another reason could be that a new machine has a wear in time, during the wear in time particles are released and may clog the cooling air channels in the components causing them to get hotter and therefore more likely to obtain damages beyond reparability. The cause might also be a combination of the two explanations (Barhanko, 2017).

The field experience data described in the Background section is examined and analyzed in an attempt to validate this theory. The methods for the first replacement analysis are described below.

3.1.5.1 Pretreatment of data
Visual inspection of the data reveal two possible outlier sites; Site 6 with one data set (not-first replacement, 50 components) and Site 7 with two data sets (first replacement, 38 components and not-first replacement, 50 components). See Appendix A for more information. Calculations and analysis are performed with and without the two possible outlier sites.
The data is sorted into first replacement data and not-first replacement data, with and without outliers.

3.1.5.2 Scrap rate
To get an initial view of the two data groups (first replacement data and not-first replacement data) scrap rates and to detect possible deviations, the weighted mean of the scrap rate and its weighted standard deviation is calculated for each group. This is performed with and without outliers.

For the result to be statistical significant there should be no overlap between the two intervals e.g. if the scrap rate for first replacement data is 2±1 the interval for not-first replacement data cannot have a larger maximum value than 1 or a smaller minimum value than 3.

3.1.5.3 Distribution
To investigate how the scrap rate varies within the two groups each group is represented in histograms displaying the frequency of the scrap rate in 10% intervals. This is performed with and without outliers.

To compare the two groups there distributions are estimated based on the available data and plotted, and the skewness factor is calculated according to equation (3).

3.1.5.4 Variations in failure mode
Analyzing the cause for the possible increased scrap rate for first replacements requires analysis of the failure modes for the different data groups. If the machines operations are different in the first 20 000 operation hours than during the remaining operating hours the causes of the un-reparability in components (failure mode A, B, C and D) might be differently distributed between the two groups. Pie charts displaying the fraction of repaired and scrapped components, and which failure mode that caused the components to be scrapped will be presented. This is performed with and without outliers.

3.1.5.5 Same site data
The lifetime of the components in a gas turbine is effected by a variety of factors. When validating a factor as an influence on the scrap rate it is therefore desirable to compare data sets that are similar in every way except in that factor e.g. if it is the first replacement or not. Comparing data sets that originates from the same site is a way to come close to accomplice that.

Out of the original 19 sites 5 sites had both first and not-first replacement data. The data sets that originate from the same site have the same input value for the environmental attributes and article number on the components but there are some variations in operation time between the sets.

For each of the five sites the weighted mean of the scrap rate for both the first and not-first replacement data is calculated. For first and not-first replacement data the scrap rate is calculated as weighted mean and weighted standard deviation for the scrap rate for each group. This is performed with and without outliers.
3.1.6 Dirt detected in inlet of machine; effect on the scrap rate
The components examined in this thesis are believed to be sensitive to dirt in the air i.e. if dirt manages to enter the machine it will affect the lifetime of the components negatively (Karlsson, 2017). The environmental attribute Erosion Factor which combines the outside air quality with the quality of the filter, has been developed by SIT partly for this reason. The problem with the usage of the Erosion Factor for this purpose is that the quality of the filter installation and the frequency of filter replacement are not included (Tavast, 2010).

During the inspections of the gas turbines the inlet to the compressor is inspected and if any dirt is detected there is a note of that in the inspection report (Dagnelund & Naderi, 2017).

The field experience data is examined and analyzed with the aim of determining if the mentioning of dirt in the inlet to the compressor correlates with a higher scrap rate. The methods for the dirt analysis are described below.

3.1.6.1 Pretreatment of data
The data is sorted into data that had a mentioning of dirt in the inlet to the compressor in the inspection report and data with no mentioning of dirt in the inlet.

3.1.6.2 Scrap rate
To get an initial view of the two groups (dirt data and not-dirt data) scrap rates and to see if they vary, the weighted mean of the scrap rate and its weighted standard deviation is calculated for each group. The same rule for statistical significance as for the first replacement analysis applies.

3.1.6.3 Distribution
To investigate how the scrap rate varies within the two groups each data group is represented in histograms displaying the frequency of the scrap rate in 10% intervals.

To compare the two groups there distributions are estimated based on the available data and plotted, and the skewness factor is calculated according to equation (3).

3.1.7 Attribute analysis
Since the number of available data sets is quite small (41 sets) the parameters to predict should be kept to a minimum (Ljung & Glad, 2004). It is therefore necessary to find methods to analyze which attributes to be used as input to the model.

In this subsection the two methods of analyzing an attributes influence on the expected lifetime of the components i.e. if it is an influential or a non-influential attribute, are described. The first method, the R2-method, is based on the least square method of determining the error of a prediction. The second method, the b-method, is based on the R-function weibfit's ranking of the coefficients b corresponding to different attributes. The methods for the two approaches are described below.
3.1.7.1 The $R^2$-method

One simple and commonly used method for finding a correlation between two factors is curve fitting. The methods of curve fitting are based on the fitted curve being assigned a measurement that describes the distance between the fitting points and the fitted curve. The assumption for these methods is that a smaller distance measurement equals a better fitted curve. The most commonly used distance measurement is the least square method; the method is based on the squared distances between the line and the fitting points (Jonsson & Norell, 2007).

To determine if curve fitting and the least square can be used to analyze the importance of attributes the simulated systems developed in the validation will be used, as there the influence of the attributes on the output is known. In order to not only consider the value of the attributes effect on the number of scrapped components in machine, but also the time a component was exposed to the attribute, the attributes value will be multiplied with the operation time, $T$. This combined value will be plotted against the number of scrapped components, the linear trend line that best describes the pattern will be drawn and the mean of the least square distance between the line and the fitting points ($R^2$) will be calculated. All calculations will be performed in Microsoft Excel 2010.

The simulated systems will use different numbers of attributes with varied influence as input to the system. The test will be performed for the following combinations:

1. Three attributes of equal size i.e. same influence on the output.
2. Seven attributes of unequal size i.e. varied influence on the output, and a distribution of the number of scrapped components similar to the first combination used in the validation; see Table 3 and Figure 6.
3. Seven attributes of unequal size i.e. varied influence on the output, and a distribution of the number of scrapped components with a large number of extreme values; see Table 4 and Figure 6.

The data for this analysis consists of 500 data sets.

3.1.7.2 The $b$-method

Since all input parameters i.e. attributes are normalized, the size of the corresponding output parameters i.e. coefficients $b$ (see equation (8)) should be an indicator of how influential an attribute is.

This will be analyzed by using the simulated systems and the same combinations as in the $R^2$-method, and in addition seven attributes of unequal size with 13 extra non-influential attributes added as input (Table 5). A desired outcome is that the model returns coefficients $b$ of the same size and correct sign if the attributes are of equal size and if the attribute vary in size coefficients $b$ of different size corresponding to the difference in influence between the attributes.

The data for this analysis consists of 500 data sets.
3.1.8 New attributes
In general there are two categories of attributes that effect the lifetime of the components in a gas turbine; first everything that effects the metal temperature e.g. flame temperature and mixture of cooling air, and secondly everything that effects the tensions in the metal e.g. differences in temperature within the same component (Karlsson, 2017).

For each data set the attributes are added as a scalar and the attributes that are not scalar has been represented as scalars e.g. ambient temperature has been represented by average values for a summer and a winter month (Olivi, 2016). To make useful scalar representations of sensor data it is necessary to preform additional analysis which lies beyond the scope of the thesis.

Two scalar attributes that are aimed towards trying to reflect how the machine has been operated will be added to the model; downtime of the machine and EOH divided with operating hours. Information concerning the two new attributes is presented below.

3.1.8.1 Downtime
The effect the downtime of the machine has on the number of scrapped components is relevant to analyze since the components are more vulnerable to some environmental attributes e.g. humidity, when the machine is still and cold (Karlsson, 2017).

The downtime of the machine is calculated by first calculating the total number of hours the machine has been operational i.e. hours between the installation or last maintenance and the current maintenance, and then subtracting the number of operating hours.

3.1.8.2 EOH/OH
The multidimensional time measurement EOH was developed by SIT for the purpose of having a time measurement that reflects the operation of the machine e.g. starts and stops which are harmful to the machine will result in a higher EOH value compared to the value of the true operating hours (OH) (Olivi, 2016). Therefore if EOH is divided by OH one receives a factor that indicates the load profile of the machine.

3.2 Implementation of investigation findings in the model
The relevant new results from the investigation preformed in this thesis with the aim of finding ways to evolve the model to make more accurate predictions will be implemented in the model. Depending on the nature of the findings the implementation to the model may vary and different combinations of findings may need testing.

For each combination validations test as the first test described in the validation section is required to determine if the accuracy of the predictions made by the model has increased or decreased. This will be determined by comparing the size of the errors between the most accurate current model and the new model.

The two scalar attributes aimed towards trying to reflect how the machine has been operated; downtime of the machine and EOH/OH will be added to the model.
4 Results
In this section all results are presented in the same three subsections as the methodology section; 4.1 Validation of the weib model, 4.2 Investigations to further improve the model, and 4.3 Implementation of investigation findings in the model. After each subsection there is a section where the results presented and the implications of those results are discussed.

4.1 Validation of the weib model
In this subsection all results from the three validation methods are presented.

4.1.1 Prediction accuracy of validation data
The results from the first validation test with different failure modes as described in the methodology section are presented below in Figure 7 to Figure 9.

Figure 7 displays the results of the validation test for failure mode B; one can observe that the error has its center round zero and that the absolute value of the error never exceeds 4.

Figure 8 displays the results of the validation test for failure mode C; one can observe that the absolute value of the error never is below 34 and frequently is around the maximum possible error of 50. For all validation points the model has made predictions larger than the real value and the errors are not centered round zero.

Figure 9 displays the results of the validation test for failure mode D; one can observe that the absolute value of the error variates between 3 and 50, and for 6 points it is around the maximum possible error of 50. For 12 of the 13 validation points the model has made predictions larger than the real value and the errors are not centered round zero.

According to the validation criteria established in in the methodology section the model passed the validation test for one out of three failure mode; failure mode B.
Figure 7: The results of the validation tests of the model's ability to predict the number of scrapped components caused by failure mode B. The table in the top left corner displays the output from `weibfit`. The table in the bottom left corner displays the real number of scrapped components, the mean and confidence interval of 95% of the predicted number of scrapped components. The table also contains the prediction error which is also visualized in the scatter plot on the right.

Figure 8: The results of the validation tests of the model's ability to predict the number of scrapped components caused by failure mode C. The table in the top left corner displays the output from `weibfit`. The table in the bottom left corner displays the real number of scrapped components, the mean and confidence interval of 95% of the predicted number of scrapped components. The table also contains the prediction error which is also visualized in the scatter plot on the right.
Figure 9: The results of the validation tests of the model's ability to predict the number of scrapped components caused by failure mode D. The table in the top left corner displays the output from *weibfit*. The table in the bottom left corner displays the real number of scrapped components, the mean and confidence interval of 95% of the predicted number of scrapped components. The table also contains the prediction error which is also visualized in the scatter plot on the right.

4.1.1 Discussion - Prediction accuracy of validation data

The results from the first validation test reveal that the validation criteria only is met for failure mode B and that the worst predictions are made for failure mode C, see Figure 7, Figure 8, and Figure 9. The reason for those results might be found in how frequent each failure mode is. The last pie chart in Figure 34 displays the distribution of failure modes for the components that were deemed as unrepairable. There one can deduce that failure mode B is twice as common as Failure mode C and in-between the two lays failure mode D. The distribution of the results from the first validation test could be described in a similar way; results are best for failure mode B, worst for failure mode C and in-between the two for failure mode D. If the lack of calibration data for failure mode C and D is the cause of the test failure the results will improve as the available data increases.

Another possible explanation for the results is that the attributes that cause failure mode C and D are missing to a greater extent as input to the model than for failure mode B i.e. failure mode B is more dependent on environmental attributes than failure modes C and D. If this is the cause of the failed test new attributes that better correlates with failure mode C and D must be found and implemented as input to the model.

4.1.2 Model Stability

The results from the stability validation tests are displayed in Figure 10 to Figure 12.

Figure 10 displays the results for failure mode B, one can observe that when the validation data was added to the calibration data; four mean values of parameters has switched sign and that two mean values have moved so far that they no longer lies within the original parameters confidence interval. The corresponding numbers for failure mode C displayed in Figure 11 is
According to the validation criteria established in the methodology section the model did not pass the validation test for any failure mode.

**Figure 10:** Displaying the value of all output parameters for failure mode B when calibrations are made with first only the calibration data and second with calibration and validation data combined. The mean value is represented with a dot and the confidence interval of 95% with interval markers.

**Figure 11:** Displaying the value of all output parameters for failure mode C when calibrations are made with first only the calibration data and second with calibration and validation data combined. The mean value is represented with a dot and the confidence interval of 95% with interval markers.
Figure 12: Displaying the value of all output parameters for failure mode D when calibrations are made with first only the calibration data and second with calibration and validation data combined. The mean value is represented with a dot and the confidence interval of 95% with interval markers.

4.1.2.1 Discussion - Model stability
The results from the second validation test reveal that none of the two validation criteria were meet for the three failure modes, see Figure 10 to Figure 12. All failure modes had coefficients that changed sign and parameters that moved beyond the confidence interval for the first data set. This is an indication that the model needs a larger amount of input data to accurately be able to determine the values of the shape parameter $\beta$ and the coefficients $b$.

As discussed in the previous section the cause of the bad predictions for failure mode C and D might be caused by the lack of calibration data for those failure modes. If that is the cause one can view the failed second validation test for failure mode C and D as a positive result since the values for the coefficients $b$ clearly were incorrect. But for the same reason it would have been desirable to find a more stable predictions for failure mode B, this is not the case.

4.1.3 Analysis of the model based on a simulated system
The results from the validation tests using manufactured systems are presented below.

4.1.3.1 Part 1
The results from part 1 where all attributes were of equal size and all values that influence the output value were included as input to weibfit are displayed in Figure 13 to Figure 16 and in Table 6.

In the results one can observe that when 500 data sets are used weibfit can without any errors predict the number of components that needs to be scrapped when three and seven attributes are used, see Figure 13 and Figure 15. weibfit can when 500 data sets are used describe the Weibull distribution, see Figure 14, Figure 16 and Table 6. When 50 data sets are used weibfit can with only small errors predict the number of components that needs to be
scrapped, see Figure 13 and Figure 15, and describe the Weibull distribution with small deviations, see Figure 14, Figure 16 and Table 6. The predictions of $\eta$ are larger than the true value and the prediction of $\beta$ is smaller than the true value.

Figure 13: Figure displaying the error in the number of components in need of scrapping when three attributes of equal size were used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 14: Figure displaying the error in the prediction of the scale parameter $\eta$ when three attributes of equal size were used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.
Figure 15: Figure displaying the error in the number of components in need of scrapping when seven attributes of equal size were used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 16: Figure displaying the error in the prediction of the scale parameter η when seven attributes of equal size were used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Table 6: Table displaying the error in the prediction of the shape parameter β for different numbers of calibration data sets and number of attributes used; all attributes have the same size.

<table>
<thead>
<tr>
<th>Error_β</th>
<th>500 data sets</th>
<th>50 data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 attributes</td>
<td>1.02</td>
<td>1.23</td>
</tr>
<tr>
<td>7 attributes</td>
<td>1.02</td>
<td>1.26</td>
</tr>
</tbody>
</table>

4.1.3.2 Part 2

The results from part 2 where all attributes were of unequal size (view Table 3), all parameters that influence the output value were included as input to `weibfit` and a distribution style similar to the one were all attributes have the same influence (view Figure 6) is used are displayed in Figure 17, Figure 18 and Table 7.

In the results one can detect that when the influence of the attributes vary `weibfit` can with without errors accurately predict the number of components to be scrapped when 500 data sets are used and with maximum absolute value of the errors of two if 50 data sets are used, see Figure 17. `weibfit` can also quite accurately identify and describe the Weibull distribution, see
Figure 18 and Table 7. The predictions of $\eta$ are larger than the true value and the prediction of $\beta$ is smaller than the true value.

Figure 17: Figure displaying the error in the number of components in need of scrapping when seven attributes of unequal size were used, all values that influence the output value were included in weibfit and a distribution style similar to the one were all attributes have the same values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 18: Figure displaying the error in the prediction of the scale parameter $\eta$ when seven attributes of unequal size were used, all values that influence the output value were included in weibfit and a distribution style similar to the one were all attributes have the same values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Table 7: Table displaying the error in the prediction of the shape parameter $\beta$ for different numbers of calibration data when seven attributes of unequal size was used, all values that influence the output value were included in weibfit and a distribution style similar to the one were all attributes have the same values is used.

<table>
<thead>
<tr>
<th>$\text{Error}_{\beta}$</th>
<th>500 data sets</th>
<th>50 data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.04</td>
<td>1.34</td>
</tr>
</tbody>
</table>

The results from part 2 where all attributes were of unequal size (view Table 4), all parameters that influence the output value were included as input to weibfit and a distribution style with a large number of extreme values (view Figure 6) is used are displayed in Figure 19, Figure 20 and Table 8.
In the results one can detect that when the influence of the attributes vary *weibfit* can with only a few small errors accurately predict the number of components to be scrapped regardless of the number of data sets used, see Figure 19. However *weibfit* cannot accurately identify and describe the Weibull distribution, see Figure 20 and Table 8. The predictions of $\eta$ are smaller than the true value and the prediction of $\beta$ is larger than the true value.

![Figure 19: Figure displaying the error in the number of components in need of scrapping when seven attributes of un-equal size were used, all values that influence the output value were included in *weibfit* and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.](image)

![Figure 20: Figure displaying the error in the prediction of the scale parameter $\eta$ when seven attributes of un-equal size were used, all values that influence the output value were included in *weibfit* and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data. Note the difference in scale on the y-axis and that in the left plot one value of 43.3 is missing.](image)

![Table 8: Table displaying the error in the prediction of the shape parameter $\beta$ for different numbers of calibration data when seven attributes of un-equal size was used, all values that influence the output value were included in *weibfit* and a distribution style with a large number of extreme values is used.](image)

<table>
<thead>
<tr>
<th>Error $\beta$</th>
<th>500 data sets</th>
<th>50 data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.70</td>
<td>0.30</td>
<td></td>
</tr>
</tbody>
</table>

### 4.1.3.3 Part 3

The results from part 3 where the attributes were of unequal size (view Table 5), non-influential attributes were included and influential attributes were excluded as input in *weibfit*, and a
distribution style similar to the one were all attributes have the same influence (view Figure 6) is used are displayed in Figure 21 to Figure 24 and Table 9.

From the results one can deduct that when an influential attribute is removed as input to weibfit, weibfit cannot accurately predict the number of components in need of scrapping or the Weibull distribution, see Figure 21, Figure 22 and Table 9.

When non-influential attributes were added weibfit can predict the number of components in need of scrapping and the Weibull distribution if 500 data sets are used, see Figure 23, Figure 24 and Table 9. When 50 data sets are used weibfit can predict the number of components in need of scrapping with absolute error values no larger than 4, see Figure 23. But the distribution cannot accurately be identified when 50 data sets is used, see Figure 24 and Table 9.

Figure 21: Figure displaying the error in the number of components in need of scrapping when seven attributes of un-equal size were used, one influential attribute was removed as input in weibfit and a distribution style similar to the one were all attributes have the same influence is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 22: Figure displaying the error in the prediction of the scale parameter η when seven attributes of un-equal size were used, one influential attribute was removed as input in weibfit and a distribution style similar to the one were all attributes have the same influence is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.
Figure 23: Figure displaying the error in the number of components in need of scrapping when seven attributes of un-equal size were used, 13 non-influential attribute were added as input in \textit{weibfit} and a distribution style similar to the one were all attributes have the same influence is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 24: Figure displaying the error in the prediction of the scale parameter $\eta$ when seven attributes of un-equal size were used, 13 non-influential attribute were added as input in \textit{weibfit} and a distribution style similar to the one were all attributes have the same influence is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Table 9: Table displaying the error in the prediction of the shape parameter $\beta$ for different numbers of calibration data sets and number of attributes used as input in \textit{weibfit}. A distribution style similar to the one were all attributes have the same influence is used.

<table>
<thead>
<tr>
<th>$\text{Error}_{\beta}$</th>
<th>500 data sets</th>
<th>50 data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input: A1-A6</td>
<td>1.11</td>
<td>1.82</td>
</tr>
<tr>
<td>Input: A1-A20</td>
<td>1.12</td>
<td>1.66</td>
</tr>
</tbody>
</table>

The results from part 3 where the attributes were of unequal size (view Table 5), non-influential attributes were included and influential attributes were excluded as input in \textit{weibfit}, and a distribution style with a large number of extreme values (view Figure 6) is used are displayed in Figure 25 to Figure 28 and Table 10.

From the results one can deduct that when an influential attribute is removed as input to \textit{weibfit}, \textit{weibfit} cannot accurately predict the number of components in need of scrapping.
or the Weibull distribution, see Figure 25, Figure 26 and Table 10. This is the same result as for the other distribution style used.

When non-influential attributes were added weibfit can with few larger errors (one for 500 data sets and six for 50 data sets) predict the number of attributes in need of scrapping, see Figure 27. However weibfit cannot accurately identify and describe the Weibull distribution, see Figure 28 and Table 10.

![Figure 25](image1)

**Figure 25:** Figure displaying the error in the number of components in need of scrapping when seven attributes of un-equal size were used, one influential attribute was removed as input in weibfit and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

![Figure 26](image2)

**Figure 26:** Figure displaying the error in the prediction of the scale parameter $\eta$ when seven attributes of un-equal size were used, one influential attribute was removed as input in weibfit and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.
Figure 27: Figure displaying the error in the number of components in need of scrapping when seven attributes of un-equal size were used. 13 non-influential attribute were added as input in weibfit and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Figure 28: Figure displaying the error in the prediction of the scale parameter $\eta$ when seven attributes of un-equal size were used, 13 non-influential attribute were added as input in weibfit and a distribution style with a large number of extreme values is used. The plot on the left is when 500 data sets were used as calibration data and the right one when 50 data sets were used as calibration data.

Table 10: Table displaying the error in the prediction of the shape parameter $\beta$ for different numbers of calibration data sets and number of attributes used as input in weibfit. A distribution style with a large number of extreme values is used.

<table>
<thead>
<tr>
<th>Error $\beta$</th>
<th>500 data sets</th>
<th>50 data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input: A1-A6</td>
<td>2.77</td>
<td>1.90</td>
</tr>
<tr>
<td>Input: A1-A20</td>
<td>0.81</td>
<td>0.15</td>
</tr>
</tbody>
</table>

4.1.3.4 Discussion - Analysis of the model based on a simulated system

Since the model failed the previous validation criteria it is important to find out if weibfit has the ability to predict a Weibull distributed system. If not the whole model must be reevaluated and new statistical methods must be analyzed and used as basis for the model. Visual inspection of the results from the third validation test reveals that weibfit has the ability to predict some systems based on Weibull distribution if 500 calibration data sets are used and the relevant attributes are included.
The third validation test also has the purpose of evaluating the possibility's and limitations of the model. After a visual inspection of the results the following notes can be made:

1. The sizes of the errors are smaller if all attributes have the same influence on the output than if the influence varies across the attributes.

2. If the influence varies across the attributes and distributions similar to the one were all attributes have the same influence on the output (view Figure 6) is used; weibfit can make accurate predictions, when 500 calibration data sets are used.

3. If the influence varies across the attributes and a distribution with a large number of extreme values is used; weibfit can predict the number of components in need of scrapping but not correctly identify the Weibull distribution, when 500 calibration data sets are used.

4. Weibfit is unable make accurate predictions if an influential attribute is missing, for both distribution styles when 500 calibration data sets are used.

5. If non-influential attributes are added as input to weibfit it can predict the number of components in need of scrapping but not correctly identify the Weibull distribution, for both distribution styles when 500 calibration data sets are used.

As stated previously one incentive of the simulated system was to find out if the model has the ability to accurately predict a system, a second incentive is to analyze how the model performs with different amounts of data. When 500 calibration data sets the model can accurately predict systems but when only 50 calibration sets are used the model struggles. The current number of calibration data sets available at SIT is around 50; therefore the reliability of the model at this stage becomes questionable. It is also important to note that the simulated system is totally without noise which unlikely is the case for the real system.

Depending on the attended use of the prediction model its abilities to predict the number of components in need of scrapping and the Weibull distribution becomes of varied importance. If the model is only to be used as a method of estimating the production and storage needs, then the only error size that matters is that of the predicted number of components verses the true value. If the model is also to be used to predict the state of the gas turbine components, then an accurate prediction of the distribution would be useful e.g. the distribution could help predict the optimal maintenence time and the value of \( \beta \) would help to determine if the failures are infant, random or wear out failures.

If 500 data sets can be used as calibration data the results from the validation with a simulated system reveals that it is more important to find all influential attributes then it is to sort out the non-influential attributes. But as it is unknown what the sufficient number of calibration data set is and it is likely to vary depending on the number of input parameters it is recommended to always try to remove non-influential attributes.

How the number of components in need of scrapping is distributed has an effect on the models ability to make accurate predictions. When a distribution with a large number of extreme values
is used the model cannot make accurate estimations of the true system’s distribution even when 500 calibration data sets are used. But weibfit make more accurate predictions of the number of scrapped components. This might be explained by the fact that an accurate estimation of the Weibull distribution parameters, $\beta$ and $\eta$, is less important for the correct prediction of scrapped components when the distribution contains a large number of extreme values e.g. regardless of how large $\eta$ is the scrap rate can never be lower than zero which frequently is an accurate prediction.

4.2 Investigations to further improve the model
In this subsection all the results from the investigations to further improve the model is presented.

4.2.1 Same site data analysis
The results from the analysis of data sets originating at the same site using weighted mean of the scrap rate and its standard deviation is displayed in Table 11. From the table one can read that the weighted standard deviation of the scrap rate at a site can vary between 1.4 and 25, indicating that even if machines are exposed to the exact same environmental attributes there scrap rate may vary greatly.

Table 11: The biggest difference in amount of scrapped components, the weighted mean of the scrap rate and its standard deviation is displayed for each site with multiple data sets. At the bottom of the table the totals weighted mean of the scrap rate and its standard deviation for all sites included in the table.

<table>
<thead>
<tr>
<th>Site number</th>
<th>Number of data sets</th>
<th>Scrap rate difference [amount of scrapped components]</th>
<th>Scrap rate weighted mean [%]</th>
<th>Scrap rate weighted standard deviation [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>9.9</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>9.9</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>21</td>
<td>33</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>9.9</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>25</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>4</td>
<td>34</td>
<td>5.7</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4.2</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>8</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>18</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>13</td>
<td>29</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>16</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>4.2</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>8.5</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>1</td>
<td>57</td>
<td>1.4</td>
</tr>
</tbody>
</table>

**WEIGHTED MEAN** | **20** | **12** |

The bubble plots that investigate if there is a correlation between the standard deviation of the scrap rate for data sets originating from the same site and the standard deviation of different expressions of hours and cycles are displayed in Figure 29 and Figure 30. If the cause of the variated scrap rate at the sites was the difference in EOH and EOC one should find that the size of the bubbles in Figure 29 grew bigger as the weighted standard deviation for EOH and EOC grew. If the cause of the variated scrap rate at the sites was the difference in the number of operation hours and
cycles one should find that the size of the bubbles in Figure 30 grew bigger as the weighted standard deviation for operation hours and cycles grew. Visual inspection of the figures (Figure 29 and Figure 30) shows no such pattern.

![Correlation Weighted Standard Deviation; EOH, EOC and Scrap Rate for Each Site](image)

**Figure 29:** A plot to investigate if the standard deviation for the scrap rate is correlated to the standard deviation of EOH and EOC. The weighted standard deviation for EOH and EOC are on the axis. The size of the bubble is the weighted standard deviation of the scrap rate for each site. The value is also written in the bubbles.

![Correlation Weighted Standard Deviation; OH, C and Scrap Rate for Each Site](image)

**Figure 30:** A plot to investigate if the standard deviation for the scrap rate is correlated to the standard deviation of operation hours (OH) and operation cycles (C). The weighted standard deviation for OH and C are on the axis. The size of the bubble is the weighted standard deviation of the scrap rate for each site. The value is also written in the bubbles.

### 4.2.1.1 Discussion - Same site data analysis

The result from the analysis of data originating from the same site indicates that knowledge about environmental attributes and the runtime is not enough to determine the lifetime of the
components. It is therefore important to find other categories of attribute e.g. load attributes, to be able to accurately predict the lifetime of the components.

The variations in scrap rate between machines originating from the same site might also be due to random failures. The impact random failures have not been estimated in this thesis.

4.2.2 First replacement of components in a machine; effect on the scrap rate
In this section the results from the different methods used in the first replacement analysis is presented.

The first replacement data consists of 22 set and the not-first replacement data consists of 19 set, when possible outliers are included.

4.2.2.1 Scrap rate
The results of the calculations of weighted scrap rate and its standard deviation for first and not-first replacement data is displayed in Table 12.

Looking at Table 12 one can calculate that for the data that included outliers the weighted mean of the scrap rate is 2.4 times higher for first replacement data than for not-first replacement data. When the weighted standard deviation is included the scrap rates for the two groups overlap i.e. the result that the scrap rate is higher for first replacement data is not statistically significant. For the data not including the possible outliers the weighted mean of the scrap rate is 4.5 times higher for first replacement data than for not-first replacement data. When the weighted standard deviation is included the scrap rates for the two groups do not overlap i.e. the result that the scrap rate is higher for first replacement data is statistically significant.

<table>
<thead>
<tr>
<th></th>
<th>Number of data sets</th>
<th>Scrap rate weighted mean [%]</th>
<th>Scrap rate weighted standard deviation [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Replacement Data</td>
<td>22</td>
<td>29</td>
<td>16</td>
</tr>
<tr>
<td>First Replacement Data (no outliers)</td>
<td>21</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Not-First Replacement Data</td>
<td>19</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>Not-First Replacement Data (no outliers)</td>
<td>17</td>
<td>6.7</td>
<td>7</td>
</tr>
</tbody>
</table>

4.2.2.2 Distribution
The histograms for first and not-first replacement data with and without outliers are displayed in Figure 31 and Figure 32.

The histogram over not-first replacement data with the outliers included show that a majority (84%) of the sets in that data group have a scrap rate between 0-20%, see the left histogram (red) in Figure 31. This trend is strengthened when the outliers are removed, as the number then becomes 94%, see the left histogram (red) in Figure 32.

The histogram over first replacement data with the outliers included show that a majority (59%) of the sets in that data group have a scrap rate between 20-40%, see the right histogram (blue)
in Figure 31. This trend is strengthened when the outliers are removed, as the number then becomes 62%, see right histogram (blue) in Figure 32.

The plot of the estimated distribution for first replacement and not-first replacement data is shown in Figure 33. One can observe that the distribution for the scrap rate in first replacement data is close to normal distribution which is confirmed by the calculated skewness factor of -0.16. The distribution of the scrap rate for not-first replacement data can be observed as having positive skew also confirmed by the calculated skewness factor of 2.2.
Figure 33: Displaying the estimated distribution of the scrap rate; on the left for first replacement data (red) and on the right for not-first replacement data (blue).

4.2.2.3 Variations in failure mode

The pie charts for first and not-first replacement data that display the total fraction unrepairable components and which failure mode that caused the component to become unrepairable with and without outliers is shown in Figure 34 and Figure 35.

After a review of the results from Table 12, Figure 31 and Figure 32 it is not unexpected to find that the repair levels for not-first replacement data is higher than for first replacement data and that this trend is strengthen when the outliers have been removed, see Figure 34 and Figure 35.

Looking at the pie chart for not-first replacement data including outliers one can see that the failure modes for the scrapped percentage are quite evenly distributed between failure mode B, C and D, see the second pie chart in Figure 34. For not-first replacement data not including outliers the scrapped percentage is totally dominated by failure mode B with a small fraction of failure mode C, see the second pie chart in Figure 35. For first replacement data with and without outliers failure mode B and D is more dominating, and failure modes A and C is less dominating, see the first pie charts in Figure 34 and Figure 35.

Figure 34: Pie charts showing the distribution of repaired and scrapped components, and which failure mode that caused the scrapping for; only first replacement data, only not-first replacement data and all the data. The outliers have not been removed.
4.2.2.4 Same site data

The results from the analysis of sites with both first and not-first replacement data is displayed in Table 13. There one can see that for four out of the five sites the scrap rate is higher for the first replacement than for not-first replacement. The site that does not display this pattern (Site 7) is one of the possible outlier sites; see Appendix A for information about the outliers.

Using only the data from these five sites and calculating the same values as in Table 12 one cannot get a statistically significant result that first replacement data displays a higher scrap rate than not-first replacement data, both with and without outliers. See scrap rate in Table 13.

Table 13: Table showing the sites that have both first replacement data and not-first replacement data. For each set the number of data sets, number of components and the weighted mean of the scrap rate is given. For first and not-first replacement data the scrap rate is calculated as weighted mean and weighted standard deviation for the scrap rate for each group.

<table>
<thead>
<tr>
<th>Site name</th>
<th>Number of data sets</th>
<th>Number of components</th>
<th>Scrap rate weighted mean[%]</th>
<th>Number of data sets</th>
<th>Number of components</th>
<th>Scrap rate weighted mean[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>1</td>
<td>50</td>
<td>14</td>
<td>1</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>Site 4</td>
<td>2</td>
<td>50</td>
<td>44</td>
<td>1</td>
<td>50</td>
<td>12</td>
</tr>
<tr>
<td>Site 7</td>
<td>1</td>
<td>38</td>
<td>0</td>
<td>1</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Site 12</td>
<td>1</td>
<td>50</td>
<td>30</td>
<td>2</td>
<td>100</td>
<td>18</td>
</tr>
<tr>
<td>Site 15</td>
<td>3</td>
<td>147</td>
<td>29</td>
<td>1</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>Without Site 7(outlier)</td>
<td></td>
<td></td>
<td>SCRAP RATE: 28±15</td>
<td>SCRAP RATE: 16±19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2.2.1 Discussion - First replacement of components in a packages; effect on the scrap rate

The result from the analysis of the importance of first replacement shows that the statement “the scrap rate is higher for first replacement data than for not-first replacement data” is not statistically significant proven, see Table 12. This result is not surprising when the distribution of the scrap rate is viewed for both data groups and found to be overlapping, see Figure 33. The overlap of the distributions reveile that the same scrap rate value can be obtained for data sets in both data groups.

Even if the possibility for the same scrap rate to be found within the two data groups exist the analysis of the histograms reveals that a lower scrap rate is more common to find within the not-first replacement data and that a higher scrap rate is more common to find within the first replacement data, see Figure 31 and Figure 32. This indicates that the lifetime of the
components is effected by whether it is the first replacement in a machine or not. The results indicate the same with and without possible outliers. Since the scale of the distributions is the same for both data groups the skewness factor can be used as a comparing measurement and confirms the results. The distribution for first replacement data is close to normally distributed with a slightly negative skew i.e. the scrap rates found in this data group varies and higher scrap rates are slightly more common. The distribution for not-first replacement data is far from normally distributed and has a big positive skew i.e. the scrap rates found in this data group does not vary as much and lower scrap rates are more common. This can be an indication that the scrap rate for first replacement sets in a machines might be harder to predict than the scrap rate for later replacements. It might also be an indication that the cause for the increased scrap rate in first replacement data has multiple explanations.

The results for the sites containing both first and not-first replacement data might include the answer as to why data sets originating from the same site sometimes display a big difference in scrap rate (see Table 11); as four out of five sites are found to have a large difference in scrap rate between first and not first replacement data, see Table 13.

Difference in the distribution of the failure modes between the two groups might be an indication that something happens only during the first 20000 operation hours that causes failure modes that do not occur in later in the operation, see Figure 34 and Figure 35. To validate this theory more research is needed.

4.2.3 Dirt detected in inlet of machine; effect on the scrap rate
In this section the results from the different methods used in the dirt analysis is presented.

The dirt data consists of 18 set and the not-dirt data consists of 23 set.

4.2.3.1 Scrap rate
The results of the calculations of weighted scrap rate and its standard deviation for dirt and not-dirt data is displayed in Table 14.

In Table 14 one can see that the weighted mean for the scrap rate is 1,4 times higher for the data were dirt has been detected in the inlet of the compressor than for the data were dirt has not been detected. Looking at the weighted standard deviation for both groups one can conclude that the result is not statistically significant.

Table 14: Table showing the weighted scrap rate and its weighted standard deviation for; Dirt and not-dirt data.

<table>
<thead>
<tr>
<th></th>
<th>Number of data sets</th>
<th>Scrap rate weighted mean [%]</th>
<th>Scrap rate weighted standard deviation [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dirt Data</td>
<td>18</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Not-Dirt Data</td>
<td>23</td>
<td>18</td>
<td>15</td>
</tr>
</tbody>
</table>
4.2.3.2 Distribution

The histograms for dirt and not-dirt data are displayed in Figure 36. There one can observe that the interval with the highest frequency for both groups is the 0-10% interval. No not-dirt data set has a scrap rate above 40%, whereas the dirt data has 28% of its data sets above that limit.

![Histograms for dirt and not-dirt data](image)

**Figure 36:** Histograms (10% intervals) showing the scrap rate for; only dirt data (red) and only not-dirt data (blue).

The plots of the estimated distribution for dirt and not-dirt data are shown in Figure 37. One can observe that the distribution for the scrap rate in both data groups are close to normal distribution which is confirmed by the calculated skewness factor; 0.44 for dirt data and 0.20 for not-dirt data. But note the difference in scale between the two distributions.

![Estimated distributions](image)

**Figure 37:** Displaying the estimated distribution of the scrap rate; on the left for dirt data (red) and on the right for not-dirt data (blue).

4.2.3.1 Discussion - Dirt detected in inlet of machine; effect on the scrap rate

Both the histograms and the distributions developed for the dirt analysis indicates the same thing: dirt in the machine has a negative effect on the lifetime of the components, see Figure 36 and Figure 37. The histograms show that not-dirt data never has a scrap rate higher than 40% and that dirt data has 28% over that limit. The distributions of the two data groups look similar and have a similar value for the skewness factor, but the scales of the distributions vary. Normalization of the distributions would reveal a positive skew for the not-dirt distribution. Even with normalized distribution the two groups would have a big overlap in possible scrap rates.
The dirt analysis is performed with a binary attribute (dirt found: yes/no) it would be superior if a continuous attribute to describe the amount of dirt in the machine could be found and used. Then the possible effects of dirt to the expected lifetime of the components might become clearer.

4.2.4 Attribute analysis

4.2.4.1 The $R^2$-method

The results from the analysis of the $R^2$-method are displayed in Figure 38 to Figure 40. From the results one can observe that the $R^2$-method can identify attributes of equal importance and assign those similar $R^2$-values, see Figure 38. The method can regardless of the style of the distribution (view Figure 6) identify and sorts attributes of un-equal size (view Table 3 and Table 4), but not correctly determine their proportions, see Figure 39 and Figure 40.

Figure 38: Scatter plot that describes the number of scrapped components as a function of attributes A1, A2 and A3 multiplied with the time parameter. A1, A2 and A3 have equal influence on the output. The $R^2$-value for A1 is 0.23, the $R^2$-value for A2 is 0.20 and the $R^2$-value for A3 is 0.22. See Appendix B to view zoomed in versions.

Figure 39: Scatter plot that describes the number of scrapped components as a function of attributes A1, A2 and A3 multiplied with the time parameter. A1, A4 and A6 have un-equal influence on the output, see Table 3, and a distribution style similar to the one were all attributes have the same values is used, see Figure 6. The $R^2$-value for A1 is 0.01, the $R^2$-value for A4 is 0.07 and the $R^2$-value for A7 is 0.23. See Appendix B to view zoomed in versions.
Figure 40: Scatter plot that describes the number of scrapped components as a function of attributes A1, A2 and A3 multiplied with the time parameter. A1, A4 and A6 have un-equal influence on the output, see Table 4 and a distribution style with a large number of extreme values is used, see Figure 6. The $R^2$-value for A1 is 0.00, the $R^2$-value for A4 is 0.07 and the $R^2$-value for A7 is 0.29. See Appendix B to view zoomed in versions.

4.2.4.2 The b-method

In Figure 41 the predicted output parameters for a simulated system were all attributes have equal influence on the output are displayed, there one can see that the coefficients $b$ for all attributes are of similar size and correct sign.

In Figure 42 and Figure 43 the predicted output parameters for a simulated system were the attributes have unequal influence on the output are displayed. There one can see that the coefficients $b$ for all attributes are of a size corresponding to the size of the attributes influence i.e. the coefficient $b$ for attributes A1-A3 are close to fifth of the size compared to the coefficient $b$ for A4-A5, which are close to half the size compared to the coefficient $b$ for attributes A6-A7. All attributes are assigned the correct signs.

In Figure 44 and Figure 45 one can observe that when twenty attributes are added as input to weibfit the model has no problems finding the relationship of influence between the influential attributes, and able to detect which attributes that are non-influential and set there corresponding coefficient $b$ close to zero i.e. the coefficient $b$ for attributes A1-A3 are close to a fifth of the size compared to the coefficient $b$ for A4-A5, which are close to half the size compared to the coefficient $b$ for attributes A6-A7, and A8-A20 are close to zero. All attributes are assigned the correct signs.

Figure 41: The predicted output parameters for a simulated system were all attributes have equal influence on the output. The value for each attribute is a random number within the intervals displayed to the right in the figure.
Figure 42: The predicted output parameters for a simulated system were the attributes have unequal influence on the output and a distribution style similar to the one were all attributes have the same values is used. The value for each attribute is a random number within the intervals displayed to the right in the figure.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>0.025</th>
<th>0.975</th>
<th>A1: [0, 0.1]</th>
<th>A2: [0, 0.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta (intercept)</td>
<td>-1.915</td>
<td>1.702</td>
<td>2.119</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>0.028</td>
<td>0.013</td>
<td>0.045</td>
<td>A3: [0, 0.1]</td>
<td>A4: [0, 0.5]</td>
</tr>
<tr>
<td>A2</td>
<td>0.029</td>
<td>0.013</td>
<td>0.045</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>0.032</td>
<td>0.016</td>
<td>0.049</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>0.148</td>
<td>0.127</td>
<td>0.172</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.153</td>
<td>0.132</td>
<td>0.178</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>0.300</td>
<td>0.266</td>
<td>0.338</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>0.306</td>
<td>0.272</td>
<td>0.348</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Figure 43: The predicted output parameters for a simulated system were the attributes have unequal influence on the output and a distribution style with a large number of extreme values is used. The value for each attribute is a random number within the intervals displayed to the right in the figure.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>0.025</th>
<th>0.975</th>
<th>A1: [0, 0.1]</th>
<th>A2: [0, 0.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta (intercept)</td>
<td>-2.1</td>
<td>-2.13</td>
<td>-2.03</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>0.2</td>
<td>0.18</td>
<td>0.23</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>0.2</td>
<td>0.18</td>
<td>0.22</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>0.2</td>
<td>0.18</td>
<td>0.23</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>1.2</td>
<td>1.13</td>
<td>1.36</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>1.2</td>
<td>1.12</td>
<td>1.35</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>2.2</td>
<td>2.05</td>
<td>2.45</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>2.2</td>
<td>2.03</td>
<td>2.42</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Figure 44: The predicted output parameters for a simulated system were the attributes have unequal influence on the output, non-influential attributes are added as input in weibfit and a distribution style similar to the one were all attributes have the same values is used. The value for each influential attribute is a random number within the intervals displayed to the right in the figure.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>0.025</th>
<th>0.975</th>
<th>A1: [0, 0.1]</th>
<th>A2: [0, 0.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta (intercept)</td>
<td>1.8e-00</td>
<td>1.554</td>
<td>2.011</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>3.3e-02</td>
<td>0.016</td>
<td>0.051</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>3.2e-02</td>
<td>0.016</td>
<td>0.050</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>3.2e-02</td>
<td>0.014</td>
<td>0.051</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>1.6e-01</td>
<td>0.133</td>
<td>0.187</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>1.6e-01</td>
<td>0.138</td>
<td>0.194</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>3.2e-01</td>
<td>0.282</td>
<td>0.374</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>3.2e-01</td>
<td>0.281</td>
<td>0.366</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A8</td>
<td>2.2e-04</td>
<td>-0.017</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A9</td>
<td>-2.1e-03</td>
<td>-0.020</td>
<td>0.015</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A10</td>
<td>5.7e-04</td>
<td>-0.016</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A11</td>
<td>-2.5e-04</td>
<td>-0.017</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A12</td>
<td>-1.3e-03</td>
<td>-0.019</td>
<td>0.016</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A13</td>
<td>1.8e-03</td>
<td>-0.016</td>
<td>0.019</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A14</td>
<td>2.6e-04</td>
<td>-0.017</td>
<td>0.018</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A15</td>
<td>1.9e-03</td>
<td>-0.015</td>
<td>0.019</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td>1.8e-03</td>
<td>-0.014</td>
<td>0.019</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A17</td>
<td>9.5e-04</td>
<td>-0.016</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A18</td>
<td>8.3e-04</td>
<td>-0.016</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A19</td>
<td>-9.1e-05</td>
<td>-0.017</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A20</td>
<td>1.4e-03</td>
<td>-0.015</td>
<td>0.017</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Input in weibfit: A1-A20
4.2.4.1 Discussion - Attribute analysis

Both methods for analyzing the importance and influence of attributes works in theory as can be seen in the result section labeled "Attribute analysis". In the result section "Implementing of investigation findings in the model" it is evident that both methods were useful when trying to improving the accuracy of the predictions made by the model. The R²-method found the attributes that yielded the smallest errors for failure mode B, and the b-method found the attributes that yielded the smallest errors for failure mode C and D.

4.3 Implementation of investigation findings in the model

The results and effects of the implementation of relevant new findings are presented in the subsections below. See Appendix C for a comparison between the original model and the new model with the smallest predicted error.

Attributes will be analyzed and chosen using the R²-method and the b-method. The results from the analysis of the importance of the first replacement in a machine indicates that the scrap rate is influenced by wether it is the first replacement or not, therefor first replacement is added as a binary attribute in the model. The effects dirt in the inlet has on the scrap rate are deemed as needing more research before adding it as an attribute in the model.

4.3.1 Using the R²-method – model one

In section Attribute Analysis it was discovered that creating trend lines and calculating the least square R² for the attributes multiplied with a time unit as a function of the scrap rate might be a

---

**Figure 45**: The predicted output parameters for a simulated system were the attributes have unequal influence on the output, non-influential attributes are added as input in `weibfit` and a distribution style with a large number of extreme values is used. The value for each influential attribute is a random number within the intervals displayed to the right in the figure.
useful way of determine if an attribute influences the scrap rate or not. This method of analyzing the attributes called the $R^2$-method were carried out as describes in the methodology section for all attributes (see Table 15). Based on the results the attributes were chosen as input to the different failure modes, see Table 16. The effect the first replacement in a machine has on the scrap rate is a binary attribute and can therefore not be analyzed using the $R^2$-method, but the investigations indicate that it influences all failure mode types and is therefore also included.

Table 15: The results from the analysis of attributes using the $R^2$-method.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>$R^2$ Failure Mode B</th>
<th>$R^2$ Failure Mode C</th>
<th>$R^2$ Failure Mode D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Temperature January * EOH</td>
<td>0.089</td>
<td>0.0029</td>
<td>0.0051</td>
</tr>
<tr>
<td>Mean Temperature July * EOH</td>
<td>0.12</td>
<td>0.027</td>
<td>0.062</td>
</tr>
<tr>
<td>Mean Relative Humidity January * EOH</td>
<td>0.20</td>
<td>0.016</td>
<td>0.021</td>
</tr>
<tr>
<td>Mean Relative Humidity July * EOH</td>
<td>0.0079</td>
<td>0.023</td>
<td>0.13</td>
</tr>
<tr>
<td>Distance From Sea * EOH</td>
<td>0.0080</td>
<td>0.062</td>
<td>0.024</td>
</tr>
<tr>
<td>Altitude * EOH</td>
<td>0.00040</td>
<td>0.013</td>
<td>0.13</td>
</tr>
<tr>
<td>Latitude * EOH</td>
<td>0.057</td>
<td>0.012</td>
<td>0.0010</td>
</tr>
<tr>
<td>Erosion Factor * EOH</td>
<td>0.0059</td>
<td>0.0028</td>
<td>0.015</td>
</tr>
<tr>
<td>EOH/OH</td>
<td>0.42</td>
<td>0.054</td>
<td>0.082</td>
</tr>
<tr>
<td>Downtime</td>
<td>0.21</td>
<td>0.092</td>
<td>0.068</td>
</tr>
</tbody>
</table>

Table 16: The attributes chosen in the first attempt to increase the accuracy of the predictions made by the model are listed in the table. The attributes have been chosen based on the $R^2$-method.

<table>
<thead>
<tr>
<th>Failure Mode B</th>
<th>Failure Mode C</th>
<th>Failure Mode D</th>
</tr>
</thead>
<tbody>
<tr>
<td>MeanTempJanuary</td>
<td>DistanceFromSea</td>
<td>MeanTempJuly</td>
</tr>
<tr>
<td>MeanTempJuly</td>
<td>FirstRep</td>
<td>MeanRelHumidJuly</td>
</tr>
<tr>
<td>MeanRelHumidJanuary</td>
<td>Downtime.H</td>
<td>Altitude</td>
</tr>
<tr>
<td>FirstRep</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EOH/OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Downtime</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The comparison of the error between the original model (model zero) and the new model (model one) for each failure mode is displayed in Figure 46 to Figure 48.

In Figure 46 one can see that for failure mode B the sizes of the errors are not notably changed. The opposite is true for Failure mode C and D represented in Figure 47 and Figure 48 respectively, there one can observe a massive improvement in prediction accuracy as the size of the prediction errors have become smaller and centered round zero.
Figure 46: A comparison of the size of the prediction errors between model zero (on the left) and model one (on the right) for failure mode B.

Figure 47: A comparison of the size of the prediction errors between model zero (on the left) and model one (on the right) for failure mode C.

Figure 48: A comparison of the size of the prediction errors between model zero (on the left) and model one (on the right) for failure mode D.
4.3.2 Using the b-method – model two

In section Attribute analysis it was discovered that examining the values of the coefficients $b$ for the different attributes within one failure mode type might be a useful way of determine if an attribute influences the scrap rate or not. Therefore a run with all available attributes as input in *weibfit* is done, followed by an analysis of the output to determine which attributes that effect the lifetime of the components. The result from the run and analysis is displayed in Table 17, were the grey marked cells are deemed as non-influential and removed.

Table 17: The predicted mean values for the coefficients $b$ for each attribute and failure mode, when all available attributes are used as input are presented in the table. The grey marked cells are deemed as non-influential and removed.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>$b$ values, failure mode B</th>
<th>$b$ values, failure mode C</th>
<th>$b$ values, failure mode D</th>
</tr>
</thead>
<tbody>
<tr>
<td>MeanTempJanuary</td>
<td>0.037</td>
<td>-1.398</td>
<td>3.54</td>
</tr>
<tr>
<td>MeanTempJuly</td>
<td>-0.025</td>
<td>1.724</td>
<td>3.39</td>
</tr>
<tr>
<td>MeanRelHumidJanuary</td>
<td>0.897</td>
<td>2.568</td>
<td>1.92</td>
</tr>
<tr>
<td>MeanRelHumidJuly</td>
<td>-0.396</td>
<td>0.383</td>
<td>1.94</td>
</tr>
<tr>
<td>DistanceFromSea</td>
<td>-0.084</td>
<td>-0.443</td>
<td>2.49</td>
</tr>
<tr>
<td>Altitude</td>
<td>0.183</td>
<td>1.357</td>
<td>0.20</td>
</tr>
<tr>
<td>Latitude</td>
<td>-0.332</td>
<td>-1.129</td>
<td>3.21</td>
</tr>
<tr>
<td>ErosionFactor</td>
<td>N/A</td>
<td>-0.245</td>
<td>0.58</td>
</tr>
<tr>
<td>FirstRep</td>
<td>-0.201</td>
<td>0.259</td>
<td>-0.38</td>
</tr>
<tr>
<td>EOH/OH</td>
<td>-0.048</td>
<td>-0.205</td>
<td>-0.18</td>
</tr>
<tr>
<td>Downtime.H</td>
<td>-0.201</td>
<td>0.025</td>
<td>0.43</td>
</tr>
</tbody>
</table>

The comparison of the error between model zero and the new model (model two) for failure mode B is displayed in Figure 49; one can observe that accuracy of the predictions have improved slightly.

The comparison of the error between model one and model two for failure mode C and D is displayed in Figure 50 and Figure 51; one can observe that using the attributes displayed in Table 17 has worsened the prediction accuracy compared to usage of the attributes displayed in Table 16.

![Figure 49: A comparison of the size of the prediction errors between model zero (on the left) and model two (on the right) for failure mode B.](image)
4.3.3 Outliers

In the first replacement analysis some possible outliers were discovered and discussed, see Appendix A. In an attempt to further improve the accuracy of the predictions made by the model the best combinations of attributes (see Table 18) i.e. the combination that yielded the smallest errors, will be used and the possible outliers will be removed from the calibration data set. The lists of attributes used in the two models can be found in Table 16 and Table 17 respectively.

Table 18: Table displaying the models and its corresponding failure mode that has been deemed as the best combination of attributes i.e. the combination that yielded the smallest errors so far.

<table>
<thead>
<tr>
<th>Failure Mode</th>
<th>Best Current Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Model two</td>
</tr>
<tr>
<td>C</td>
<td>Model one</td>
</tr>
<tr>
<td>D</td>
<td>Model one</td>
</tr>
</tbody>
</table>

The comparison of the error between the models displayed in Table 18 with and without outliers is displayed in Figure 52 to Figure 54. There one can observe that the removal of possible outliers have no notable effect on the size of the errors i.e. prediction accuracy for failure mode C and a decline in prediction accuracy for failure modes B and D.
4.3.4 Discussion - Implementation of investigation findings in the model

When applying the new results found in the investigations the prediction accuracy was improved, the biggest difference in accuracy can be seen for failure mode C and D, see Appendix C for a comparison between the original model and the new model with the smallest
prediction error. For failure mode C and D the absolute value of the error is often close to zero but occasionally it is still as large as twenty. This indicates that further research and development is necessary to receive an accurate and reliable prediction model for prediction of components in need of scrapping, however as the true system is unknown it is unknown if the models ability to predict the Weibull distribution has improved or not.
5 Discussion
In this section the limitations and the possible error sources are presented and discussed, and there is a discussion about possible future developments of the project.

5.1 Limitations and possible error sources
A big limitation for creating a reliable prediction model is the small number of calibration data sets. As this number becomes larger the model will become more predictable and evaluation the importance of attributes will become easier.

During the simulations R occasionally warns about divergent transitions after warmup. R is constructed warn about all possible errors and then it is up to the user to determine if the warnings are relevant to the reliability of the results (Stan Development Team, 2016). In this thesis no further investigations concerning the warnings have been made.

To be able use continuous attributes as inputs to the model scalar representations has been made e.g. relative humidity has been represented as a mean value. These representations reduce the worth of the attribute and might not be the optimal representation for determining the attributes effect on the lifetime of the components. There might be better suited representations for this purpose. Each geographical site has a scalar value for each environmental attribute e.g. average temperature in a summer month, and these values are not changed based on when the machine was operating at the site. Constant environmental attributes e.g. distance from the sea and altitude is not limited by this, but continuous attributes e.g. temperature and humidity is.

The values for the different attributes are partially gathered from different reports and transferred by hand into the data frame used as input to model, therefore the human error must be considered as a possible source of errors. Possible errors made during the transfer could have a big influence as the data set is small. A preferred approach would be to create data code to extract the desired information from the reports, until then the employees are urged to be meticulous.

The analysis preformed in this report is only for one type of component in one type of gas turbine. A turbine has many more types of components and SIT manufactures other types of turbines, and it would be desirable to be able to predict the state and lifetime of all of them. Since the analysis is restricted to one type of component its unknown how applicable the model would be for other components.

5.2 Suggestions for future work and further analysis
When the simulated system was used it was discovered that the number of calibration data sets currently available at SIT has limitations and yield results with questionable reliability. To find out what the sufficient number of calibration data sets is further analysis is necessary. The analysis should be performed with variation in number of calibration sets and distribution styles, and varied types and amounts of noise added.
For the best model found in the result section “Implementation of investigation findings in the model” only the first validation test (Prediction accuracy of validation data) is preformed, further validation of the best current model i.e. validation test two (Model stability), is necessary to fully evaluate the new model. A sensitivity analysis to quantify the impact the coefficients b has on the predicted scrap rate would also be good future validation test.

The total scrap rate in a machine is not always equal to the sum of the scrapped components for each failure mode as a component un-reparability might be caused by a combination of two or more failure modes. Analysis is necessary to find the optimal approach for finding the total scrap rate and to determine the size of the total error.

For first replacements of components in a machine it is theorized that during the wear in time particles are released and clog the ventilation in the components causing them to get hotter and therefore more likely to break beyond reparability (Barhanko, 2017). To find if this is the case it would be desirable to look at the mean total ventilation flow thru the components in each machine and compare first and not-first replacement data. Another cause for higher scrap rate might be customer behavior (Barhanko, 2017). To analyze the role of the customer and there learning curve one could separate new customers from experienced ones and see if a pattern can be detected.

There are additional attributes believed to effect the lifetime of the components and some of these attributes are continues sensor data e.g. flame temperature. For this data to be used as input in the model it first has to be transformed into a scalar representation, therefore analysis to determine the optimal approach is necessary. Possible conversions include using the highest and lowest value, the difference between the highest and lowest value, the mean value or the number of times the value oscillates in a predetermined interval. Some already included continuous environmental attributes are limited by their scalar representation and analysis to find more optimal representations is required.

When a reliable model has been developed and validated, it would be desirable to test the model on other components within the same turbine type and similar components in other gas turbine types, this to determine how versatile the model is.

In a previous master thesis Neupokoeva used different data mining methods to find patterns which could indicate a harmful effect on the lifetime of components in the gas turbine (Neupokoeva, 2016). Using the found patterns one could combine attributes to make more influential attributes.

Both the surrounding air and the fuel used in the gas turbine are of importance for the expected lifetime of the components. Olivi analyzed the importance of environmental attributes to evaluate the air quality and the effect it has on the expected lifetime. Evaluating the fuel properties is harder since it is the customer that controls the choice of fuel. Evaluating and quantifying fuel properties would be an interesting direction for future work.
6 Conclusions
The aim of the thesis is to continue the development of a model to estimate the number of components in need of scrapping in a gas turbine. Validation of an existing prediction model developed at SIT is deployed. Some investigations to further develop the model are performed and the results are implemented in the model.

The following conclusions can be drawn from the report:

The original model weibfit created at SIT (Olivi, 2016) gave poor predictions when only environmental attributes and measurements of operation time were used as input variables to the model. However, by including other attributes e.g. downtime and first replacement as inputs, and removing some original attributes; the model gave accurate predictions for a majority of the validation data sets. More effort and data is needed to obtain an even better prediction model.

Using simulated systems as a validation tool it was concluded that the number of calibration data sets currently available at SIT has limitations and yield results with questionable reliability. The model called weibfit has the ability to identify some systems based on Weibull distribution and make reliable predictions concerning the number of components in need of scrapping when 500 calibration data sets are used. weibfit's prediction accuracy is affected by the structure of the input attributes.

Results from the investigations reveal that the scrap rate is influenced by whether it is the first replacement of the components in a machine or not. In addition there are results that indicate that the scrap rate is influenced by whether there is dirt found in a machine or not.

Both the R²-method and the b-method can be used for determining if an attribute impact the lifetime of the components and both proved useful methods for improving the accuracy of the predictions made by weibfit.
7 Bibliography


Siemens Industrial Turbomachinery AB. (2015, April 30). SGT-800 Image Inventory. Finspång, Sweden: Siemens Industrial Turbomachinery AB.


### 7.1 Oral References


8 Appendix

8.1 Appendix A – Possible outliers

To get an overview of the available data a bubble plot that displays all data sets and there operation times were created. The size of the bubble represents the scrap rate and the colors represent if it is the first replacement or not, see Figure 55. One can observe that two of the bubbles representing not-first replacement data (blue) are significantly larger than the rest with a scrap rate of 66% and 50% respectively. Those data sets correspond to the two sites Site 6 and Site 7 and these sites are therefore considered possible outliers. The values of the outliers are displayed in Table 19.

![Scrap rate – First replacement analysis](image)

*Figure 55:* All calibrations data and every bubble represent a repair event. The size of the bubble corresponds with the scrap rate (the value is also written in the bubble). The red bubbles are the first replacements data and the blue are not-first replacements data.

**Table 19:** The values of the possible outliers.

<table>
<thead>
<tr>
<th>Site</th>
<th>First replacement</th>
<th>Set size</th>
<th>Failure mode A</th>
<th>Failure mode B</th>
<th>Failure mode C</th>
<th>Failure mode D</th>
<th>Scrapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 6</td>
<td>No</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Site 7</td>
<td>No</td>
<td>50</td>
<td>0</td>
<td>2</td>
<td>23</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Site 7</td>
<td>Yes</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
8.2 Appendix B – Zoomed Figure 38 to Figure 40

8.2.1 Figure 38

Three Attributes of Equal Size; RAND(0,1)

Three Attributes of Equal Size; RAND(0,1)
8.2.2 Figure 39

Three Attributes of Equal Size; RAND(0,1)

Seven Attributes of Unequal Size
8.2.3 Figure 40

Seven Attributes of Unequal Size

- Figure 40a: Scatter plot with the equation $R^2 = 0.0032$
- Figure 40b: Scatter plot with the equation $R^2 = 0.0697$
8.3 Appendix C – Original and best results comparison

Figure 56: Original model on the left and best result on the right.

Figure 57: Original model on the left and best result on the right.

Figure 58: Original model on the left and best result on the right.