**DAS** Departamento de Automação e Sistemas
**CTC** Centro Tecnológico
**UFSC** Universidade Federal de Santa Catarina

# Development of a service-supporting software tool for maintenance advice

*Relatório submetido à Universidade Federal de Santa Catarina*

*como requisito para a aprovação na disciplina*

***DAS 5511: Projeto de Fim de Curso***

***Carolina Ensfeld Lueders***

*Würzburg, July 2014*

# Acknowledgements

First of all I would like to thank my parents for all the support and the amazing chances they have given me over the years. Without you I would not be the person that I am today. An extra thank you to my three little sisters, who had made me learn to share and turned my life much more colorful. I would also like to mention the rest of my family, thanks for your love, support and for believing in me even when I was not around.

I appreciate the trust the Central Department of Quality Management and Methods of Bosch Rexroth in Würzburg relied on me. It was an outstanding opportunity to work on such a challenging project.

It has been my privilege to work with Danny Köllmann, an exceptional engineer, boss and supervisor who I also consider a good friend. His frequent insights and patience with me were always appreciated. I am very proud of what we have achieved together.

I am grateful to Jens Göbel for sharing with me his great knowledge and experience. Nowadays such hard workers willing to help are not easy to find, I hope we get the chance to work together one more time.

I would like to thank my supervisor Professor Ubirajara Franco Moreno for the meetings and feedbacks and Professor Ricardo José Rabelo for the whole bureaucratic support.

I am in debt to two of my best friends who have helped me over the graduation years: Igor and Thiago, the ones I tend to turn to first when I have something to discuss or just need a break.

Lastly I owe much to Ulisses, without his love and understanding the days would be much harder and meaningless. Thank you for always being my "consultant for important matters", but most of all thank you for being my best friend.

## Resumo Estendido

A digitalização de informações abriu portas para novos modelos de negócios inconcebíveis há 20 anos. Empresas como Google, Facebook, Amazon e Apple vem se beneficiando do poder adquirido com a análise de dados em massa propondo soluções inovadoras para necessidades que seus clientes ainda nem sabiam que tinham.

Inserida nesse contexto está a Bosch Rexroth, empresa especializada em tecnologias de acionamento e comando, subsidiária do grupo Bosch. Com um amplo portfolio de soluções sob medida, Rexroth está presente em uma variedade de setores apoiando a construção de máquinas e instalações a nível mundial. Com foco no setor de aplicações industriais, o setor de IA (Industrial Applications) é responsável pelo fornecimento de produtos hidráulicos, mecânicos, acionamentos e comandos elétricos. A sede da empresa em Lohr, no sul da Alemanha, é a principal responsável pela produção e manutenção de acionadores e comandos elétricos na Europa.

Pensando a frente de seu tempo, dados sobre entregas e reparos de produtos foram armazenados pelos últimos 20 anos na planta de Lohr. Seguindo a tendência da digitalização a Bosch Rexroth decidiu utilizar essas informacoes para tornar-se mais competitiva, intensificar o relacionamento com seus clientes e fortalecer sua imagem através de um novo modelo de negócios baseado em aconselhamento de estratégias de manutenção.

O principal objetivo desse projeto foi o desenvolvimento de uma metodologia para decidir a estratégia de manutenção mais adequada para cada cliente (manutencao corretiva ou preventiva) e sua implementação em uma ferramenta de software. A decisão quanto a estratégia de manutenção é tomada com base na análise de Weibull dos dados de tempo de vida dos produtos seguida por uma previsão de falhas para os próximos anos. Com essa previsão é possível calcular os custos futuros e avaliar qual estratégia de manutenção permitirá maior economia para o cliente.

O resultado final alcançado foi uma ferramenta de software implementada em linguagem de programação Python, baseada na metodologia desenvolvida. A ferramenta auxilia no esclarecimento e discussão das estratégias de manutenção com os clientes, elimina a necessidade de conhecimentos especializados sobre a teoria de Weibull, estabelece um método sistemático e facilita a realização das análises.

Este trabalho apresenta a metodologia e a ferramenta de software mencionadas, desenvolvidas pela estudante Carolina Ensfeld Lueders em seu projeto de fim de curso no Departamento de Gerenciamento e Métodos de Qualidade da Bosch Rexroth.

O capítulo 2 apresenta a base teórica para a realização do projeto. Conceitos básicos de confiabilidade sao apresentados assim como os principios por traz da análise de Weibull e previsão de falhas.

A metodologia desenvolvida é descrita no capítulo 3 através de um passo a passo dos procedimentos realizados para obter a melhor estratégia de manutenção a partir dos dados de tempo de vida dos produtos.

No capítulo 4 descreve-se a implementação da metodologia em uma ferramenta de software.

Finaliza-se o trabalho com a verificação e validação da ferramenta no capítulo 5. A discussão dos resultados obtidos e sugestões para trabalhos futuros são apresentados na conclusão.

# Abstract

The digitalization of information has created new business opportunities that would be inconceivable 20 years ago. Companies like Google, Facebook, Amazon and Apple have made great use of mass data analysis to propose innovative solutions to their customers.

Inserted in this context is Bosch Rexroth, a subsidiary company of the Bosch group specialized in drive and control technologies. Thinking ahead of its time Rexroth has stored over the last 20 years a massive amount of detailed information on delivery and maintenance of products. Following the trend of digitalization the company has decided to use this information to become more competitive enhancing its relationship with customers through a new business model based on maintenance strategy advice.

The main goal of this project was to develop a methodology to decide which maintenance strategy is the most appropriate for each customer (corrective or preventive maintenance) and its implementation in a software tool. The decision is made based on Weibull analysis of products lifetime data followed by a failure forecast for the next years. With this prediction is possible to calculate future costs related to failures and evaluate which strategy has the greater earning-saving potential for the customer.

The developed tool assists in the discussion of maintenance strategies with customers, eliminates the need for expert knowledge on the Weibull theory, establishes a systematic method and facilitates the evaluation process.

This paper presents the mentioned methodology and software developed by the student Carolina Ensfeld Lueders during her internship in the Central department of Quality Management and Methods of Bosch Rexroth.

**Keywords:** *Lifetime data analysis, Weibull analysis, Failure forecast, Maintenance strategy.*

# Contents

# List of figures

# List of Tables

# Simbology

CCC: Critical Correlation Coefficient

DC: Drive and Control subsidiary of Bosch (Bosch Rexroth)

ERP: Enterprise Resource Planning

GLP: General Public License

IA: Industrial Applications subdivision of DC

MLE: Maximum Likelihood Estimation

MTTF: Mean Time to Failure

OEM: Original Equipment Manufacturer

RRY: Rank Regression on Y

RRX: Rank Regression on X

TCO: Total Cost of Ownership

VBA: Visual Basic

# Chapter 1: Introduction

The digitalization changed the way companies develop their business and the whole economy environment as fundamentally as the first three industrial revolutions did. Along with the mechanization of production using water and steam power, the introduction of mass production with help of electric power and the use of electronics and information technology to further automate the production, collection and analysis of massive amount of digital data are considered by many as the beginning of a fourth industrial revolution, often referred as "Industrie 4.0" [1].

When everything is connected to everything and everyone with everyone, there are a lot of opportunities to create new business models that can transform entire industries and sometimes even eliminate parts of them [2]. Google, Facebook, Amazon and Apple concentrate their efforts in this new concept, applying tremendous computing power to collect and analyze data in order to generate innovative business ideas. In the current era, information, represented by data, is the treasure; the one who controls it controls the customer. Possessing the right information enables companies to make interesting offers to the customers that have not been considered yet. Big industries in general have not followed this trend and still face difficulties to adapt their business to this new concept [2].

Inserted in this context is Bosch Rexroth, one of the worldwide leading specialists in drive and control technology, known as the Drive and Control (DC) subsidiary of Bosch. With more than 36,500 associates in 80 countries all over the globe, the company supplies tailored solutions for driving, controlling and moving through high quality electrical, hydraulic and mechatronic components and systems to more than 150,000 customers. Rexroth products can be found in a wide range of sectors such as industrial applications, factory automation, mobile applications and use of renewable energies [3] [4] [5].

Focused on plant equipment and automation, the Industrial Applications (IA) subdivision of DC is responsible for supplying products in the fields of hydraulic control and power, linear motion, assembly technology, electric drives and controls.

Their customers range from producers of minuscule microchips to companies dealing with the motion of heavy loads, weighting several tons [4].

Located in southern Germany, the headquarter of Rexroth in the city of Lohr am Main is the main responsible for the production and maintenance of electric drives, controllers and motors of Bosch Rexroth in Europe. Thinking ahead of its time Rexroth has stored over the last 20 years a massive amount of detailed product information on delivery and maintenance of products. This information can now be used to generate new business cases in several different areas and Rexroth can benefit from the digitalization trend.

## 1.1: Problem description

The Service department of Rexroth, responsible for maintenances and repairs of electric products, has recently realized that the company could extend its business when it comes to products repairs and maintenances. In order to gain new service contracts the Service department in Lohr intends to intensify the relationship with their customers and get closer to them by offering maintenance advice.

With that in mind, a new business model is being implemented in the Lohr Service and Sales departments of DC. The so called "Maintenance plan advice" model is an innovative method of customer advising. The goal is not only to find the best maintenance plan for the customer but also to ensure that whether the customer goes for preventive or corrective, the service is done by Bosch Rexroth. However, decide the optimal maintenance plan for the customers is not a simple task for Rexroth and convince the customer of that is also a challenge.

For the realization of the aforementioned business model the need of a reliable methodology arose. The pursuit for a valid solution led the Service department to seek the required knowledge together with the Central department of Quality Management and Methods of Bosch Rexroth, where the present work was realized.

The implementation of this methodology in a software tool supports customer advice regarding maintenance strategy through reliable methods, facilitate the contact with customers and also help them to better understand the benefits of a particular maintenance contract.

## 1.2: Proposed solution

The main goal of this project was to develop a methodology to decide which type of maintenance is the most adequate for each customer and its implementation in a software tool for supporting the Sales department, as seen in Figure 1.
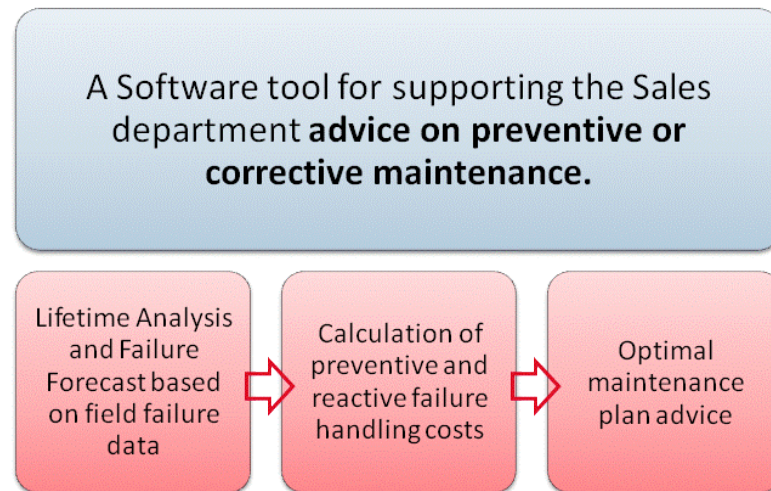


*Figure 1 – Proposed solution*

Another aim of the project is to make this analysis using a well known, experienced and valid statistical method, producing precise and reliable results.

The department of Quality Management and Methods has been using lifetime evaluations such as the Weibull distribution to perform cause analysis and to determine reliability parameters. The Weibull distribution is also one of the most widely used lifetime distributions in reliability engineering because of its versatility in modeling different sets of lifetime data. Therefore to generate a reliable maintenance advice, a Weibull analysis based on historical data (over the last years) was established as standard method, followed by a failure forecast of the customer's products for the next years. With this forecast, the customer's operating data and costs it's then possible to predict whether the costs of preventive services are lower than the costs of reactive maintenance services or not.

The Weibull analysis and failure forecast are done based on product information stored in the past years by Bosch Rexroth, available in the form of data sets with delivery and claim information. The costs are calculated based on customers' specific operating data: failure downtime, downtime costs, costs of

15

overhaul and costs of repair. As the data sets used to perform the analysis are only available to Rexroth, the methodology and the tool are considered a unique selling proposition, since no competitor would be able to achieve the results here presented.

To advice the customer, the developed software tool shows the best possible plan of maintenance and calculates its earning-saving potential for the customer (when comparing to the other option of maintenance).

Another reason for implementing the evaluation as a software tool is to make it simpler and easier to perform. A manual evaluation requires hours of work of a Service department experienced worker. Additionally, the software tool was developed in modular features so that further development could be easily integrated.

## 1.3: Structure of the document

Chapter 2: presents the technical background necessary to the accomplishment of this project. In this chapter a brief review of reliability concepts and its applications in the industry is presented to set the context of the project to the reader. Also the principles behind Weibull analysis and failure forecast are explained to situate the reader that is not familiar to the lifetime analysis theory.

Chapter 3: addresses the developed methodology to go from raw data to the desired maintenance advice. Each step of the procedure is explained in details and a global vision of the methodology is also given.

Chapter 4: presents how the developed methodology was implemented as a software tool. Here the selection of python as the programming language is justified, the software structure is explained and an overview of the development iterations is given.

In Chapter 5: the tool verification and validation results are presented.

Chapter 6: presents the conclusion of the developed work, suggestions for improving the software tool and further possibilities to continue the development of the project.

# Chapter 2: Theoretical fundaments of reliability evaluations for industry services

In this chapter, a review of the techniques and background knowledge for the development of this project is presented. Emphasis is given to the lifetime data evaluation (Weibull analysis of failure data) and forecast of future failures since these are the main theoretical subjects in the evaluation of a maintenance advice.

To give the reader the context of the project this chapter begins with a brief discussion about reliability, the different situations where it may be applied in industry and how Bosch Rexroth uses these concepts. Further the basic theory of life data analysis is explained and the techniques used to evaluate product lifetime data through a Weibull analysis are clarified. Finally the procedure to perform a failure forecast based on the previous results is detailed.

The theory of life data analysis presented in the section 2.3 was mainly adapted from Chapters 2, 3 and 4 from The New Weibull Handbook [7] and Chapters 1, 2, 3, 4 and 8 from Life Data Analysis Reference [8].

## 2.1: Basic principles of reliability

Nowadays, the term reliability is part of our daily vocabulary, particularly regarding the functionality of a product [9]. The technical definition for the term is the following:

"Reliability is the product property to execute a required function in a specified operating range during a specified service time." [10]

Among customers' surveys, reliability is often highlighted as a desired product characteristic. Therefore, the guarantee of adequate product reliability is a key factor for business success. Alongside with the traditionally important experimental tests to demonstrate reliability, new approaches are necessary to achieve the required reliability level in the current times of reduced development cycles and increasingly product requirements. The growing use of design for reliability techniques, combined with arithmetical prediction of reliability and its optimization has shifted the work on

reliability to an earlier phase of product development in order to reduce overall costs [11]. From the development of the product to after sales, reliability analysis can be used in a wide range of applications such as seen in Figure 2.

**Product development**
- Calculation of product lifetime
- Reduction of development time and costs (reduced tests efforts)
- Minimization of failure costs
- Early detection of weaknesses

**Series production**
- Early detection of weaknesses
- Minimization of failure costs
- Improvement of product quality

**After sales**
- Calculation of expected failure rates for products
- Calculation of expected failures for customers (failure costs)
- Development of solutions to reduce maintenance costs

*Figure 2 - Application of reliability analysis in the product life-cycle*

## 2.2: Strategies for reliability at Bosch Rexroth

"It has always been an unbearable thought to me that someone could inspect one of my products and find it inferior in any way."

"I would rather lose money than trust."

Robert Bosch 1861-1942

Since its early times reliability has been an extremely important value for Bosch. The subsidiary DC has also embraced this idea adopting a reliability management strategy through the whole product life cycle.

The reliability of Bosch Rexroth products is determined in the development phase meaning that the design has a big influence over the product reliability. The Bosch Group considers reliability as a product characteristic that must be designed

systematically, understanding the cause and effect relationships. This is achieved through a range of approaches applied in the design phase of the product [11].

On the launch of series production the reliability of products must be ensured. In order to do so, products are subjected to a range of trials during their development to verify that the requirements set have been met and validate the reliability [11].

To complete the reliability management over the product life cycle observations of the product behaviour in the field are required. Field data provide the final proof that the targets defined in the development phase were achieved, signalize possible improvements for the product's next generations and indicate sources of unreliability in the production processes [12].

## 2.3: Theory of lifetime analysis

Reliability lifetime analysis refers to the study and modeling of observed product lives and failures. Lifetime data can be described as the time the product operated successfully or the time the product operated before it failed [8].

If the load of a specific component exceeds its load capacity, or its local stress exceeds its strength, a failure occurs. Conducting a series of experiments with several components under the same load will result in different failure times, despite the same test conditions and (nominally) identical components are used. Failure time is therefore not a deterministic but a statistically distributed quantity. Due to randomly fluctuating material and manufacturing conditions, the failure time of a component cannot be exactly predicted, only statistically estimated [11].

In lifetime analysis, the failure behavior of a large number N of initially identical units can be described either by the probability density function, *pdf* or the cumulative density function, *cdf.* The *pdf* indicates the relative frequency of the random variable failure time as a function of time, $f(t)$ and the *cdf* designates the probability of failure before a time t, $F(t)$ [8] [13]. The *cdf* represents the cumulative values of the *pdf.* That is to say that the value of a point *a* on the *cdf* curve represents the area under the *pdf* curve to the left of that point as seen in Figure 3.
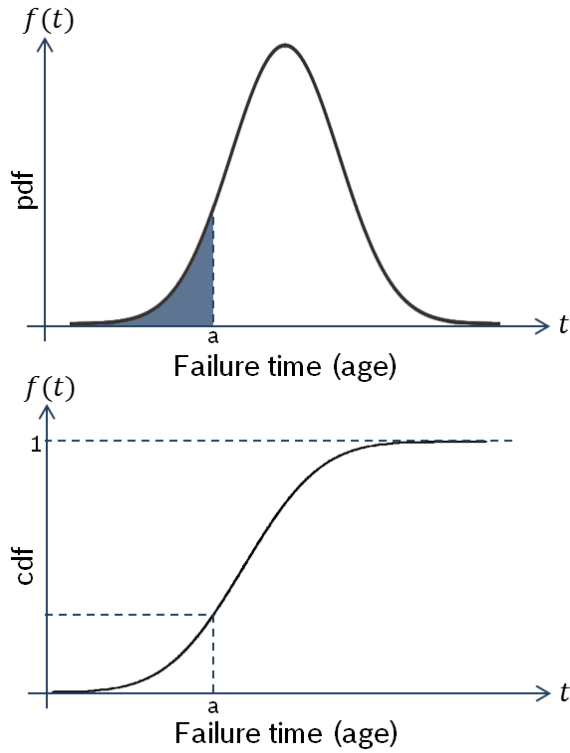
*Figure 3 -* pdf *and* cdf *curves based on [8]*

The mathematical relationship between the *pdf* and the *cdf* is given by:

$$F(t) = \int_0^t f(x)\,dx$$

The unreliability value attributed to products describes their probability of failure before time *t,* and can be defined by the unreliability function $Q(t)$, a cumulative distribution function.

$$Q(t) = \frac{n_f(t)}{N} = \frac{number\ of\ failed\ products\ before\ t}{total\ number\ of\ products} = F(t)$$

$$Q(t) = \int_0^t f(x)\,dx$$

On the other hand, the reliability value of the same products is defined as probability that they work correctly until a specified time *t.* The reliability function $R(t)$ is defined as the complement of the unreliability function:

$$R(t) = 1 - Q(t) = \frac{n(t)}{N} = \frac{number\ of\ correct\ functioning\ products\ until\ t}{total\ number\ of\ products}$$

$$R(t) = 1 - \int_0^t f(x)\,dx = \int_t^\infty f(x)\,dx$$

The failure rate function $\lambda(t)$ enables the determination of the number of failures occurring per unit of time and is defined as:

$$\lambda(t) = \frac{f(t)}{R(t)}$$

The failure rate is often used to describe the failure characteristic of a population of products through its entire life cycle. Some products exhibit failure characteristics that can be described by the bathtub curve of Figure 4 [8]. The bathtub curve consists of three periods: an early failure period with a decreasing failure rate followed by a useful life period of random failures with a relatively constant failure rate and concluding with a wear-out failure period that exhibits an increasing failure rate.
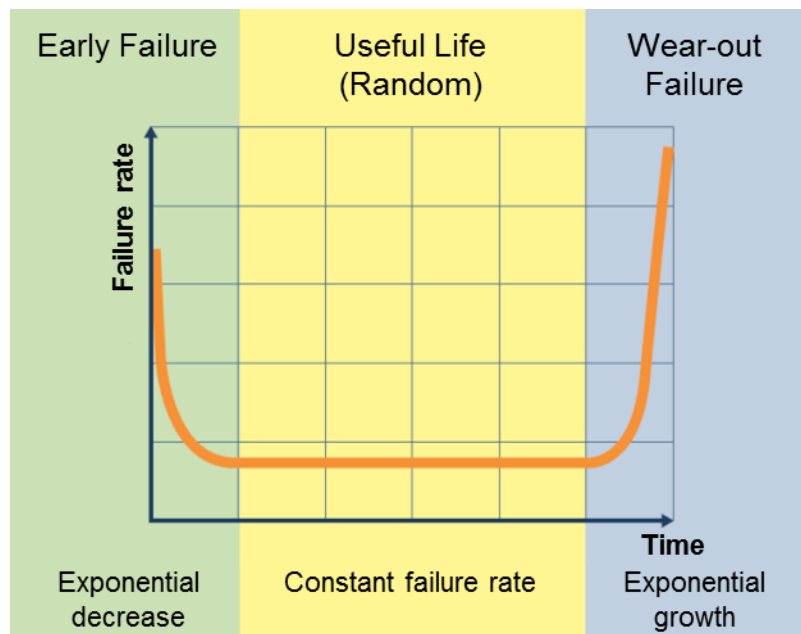


*Figure 4 - Bathtube curve based on [8]*

The goal of a lifetime analysis is to make predictions about the life and failures of a product population by fitting a statistical distribution (model) to lifetime data. Important characteristics of the product can be estimated through the distributions parameters, such as reliability or probability of failure at a specific time. In order to analyze a product population four main steps are necessary:

21

1. Gather lifetime data for the product in the form of survival and failures ages;
2. Select an adequate lifetime distribution that describes the data and model the life of the product;
3. Estimate the parameters that fit the distribution to the data;
4. Generate the desired life characteristics of the product using the distributions parameters.

### 2.3.1: Lifetime Distributions

The term lifetime distributions is used to describe a collection of statistical probability distributions used in reliability engineering and lifetime analysis to mathematically model the lifetime behavior of a product [8].

Different distributions exist, such as the normal, lognormal, exponential, gamma, Weibull, etc. As seen before, each statistical probability distribution can be described by a specific mathematical function, the probability density function (*pdf*).

The Weibull distribution, formulated by Professor Waloddi Weibull, is a popular distribution for analyzing lifetime data. In this project, Weibull distributions are used to describe lifetime behavior of products. The procedure to do so is explained in the following subsection.

### 2.3.2: Weibull Distribution

The Weibull distribution can be used to model a wide range of products life behaviors, being therefore broadly used in reliability and life data analysis [8].

A variety of Weibull models exist to analyze and describe lifetime data, including 1-parameter, 2-parameter, 3-parameter or mixed Weibull. The 2-parameter Weibull distribution is the most widely used distribution to analyze lifetime data [7] and its *pdf* can be described as:

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

$t = statistical\ variable\ that\ represents\ the\ age\ of\ survival\ and\ failed\ products$

$\beta = shape\ parameter\ or\ slope.\ Describes\ the\ shape\ of\ the\ pdf\ function$

$\eta = characteristic\ life\ or\ scale\ parameter.\ Is\ the\ age\ at\ which\ 63.2\%\ of\ the\ units$ $will\ fail\ (ureliability = 63.2\%)\ and\ indicates\ the\ location\ of\ the\ distribution\ along$ $the\ abcissa$

The other mentioned reliability functions of the 2-parametric Weibull distribution are defined as:

$$Reliability\ = R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

$$Unreliability = Q(t) = \ 1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

$$Failure\ rate = \lambda(t) = \frac{f(t)}{R(t)} = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}$$

The mean time to failure (MTTF) function provides a measure of the average time of operation until a unit fails and indicates the reliability of a component. The MTTF of the 2-parameter Weibull is given by:

$$\bar{T} = \ \eta * \Gamma(\frac{1}{\beta} + 1)$$

where the gamma function is defined as:

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

B10 is another reliability measure and indicates the time at which the probability of failure is of 10% of the products. Its value can be determined from the Weibull unreliability equation as follows:

$$0,1 = \int_0^t 1 - e^{-\left(\frac{x}{\eta}\right)^{\beta}}\ dx$$

### 2.3.3: Parameter estimation

Several methods have been devised to estimate the parameters that will fit a lifetime distribution to a particular data set. The following explanations describe four available parameter estimation methods: probability plotting, rank regression on y (RRY), rank regression on x (RRX) and maximum likelihood estimation (MLE). The appropriate analysis method will vary depending on the data set.

*Probability plotting*

23

Probability plotting consists on physically plotting the lifetime data in a probability plot paper. Probability plot papers scales are constructed in a way that its distribution's *cdf* is linearized [7].

The first step of the method consists in linearizing the Weibull unreliability function to the standard form $y = mx + b$ as follows:

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

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

$$\ln\left(1 - Q(t)\right) = -\left(\frac{t}{\eta}\right)^{\beta}$$

$$\ln\left(-\ln\left(1 - Q(t)\right)\right) = \beta * \left(\ln\left(\frac{t}{n}\right)\right)$$

$$\ln\left(\ln\left(\frac{1}{1 - Q(t)}\right)\right) = \beta * \ln(t) - \beta * \ln(\eta)$$

Then, by setting:

$$y = \ln\left(\ln\left(\frac{1}{1 - Q(t)}\right)\right)$$

$$x = \ln(t)$$

the equation can be rewritten as:

$$y = \beta x - \beta \ln(\eta)$$

In a Weibull probability plotting paper the x-axis transformation is logarithmic. The y-axis requires a double logarithmic transformation:

$$y = \ln\left(\ln\left(\frac{1}{1 - Q(t)}\right)\right)$$

The y-axis represents unreliability and the x-axis represents time. Both of these values must be known for each point to be plotted. Such papers can be easily founded in internet sources and related literature, an example is seen in Figure 5.
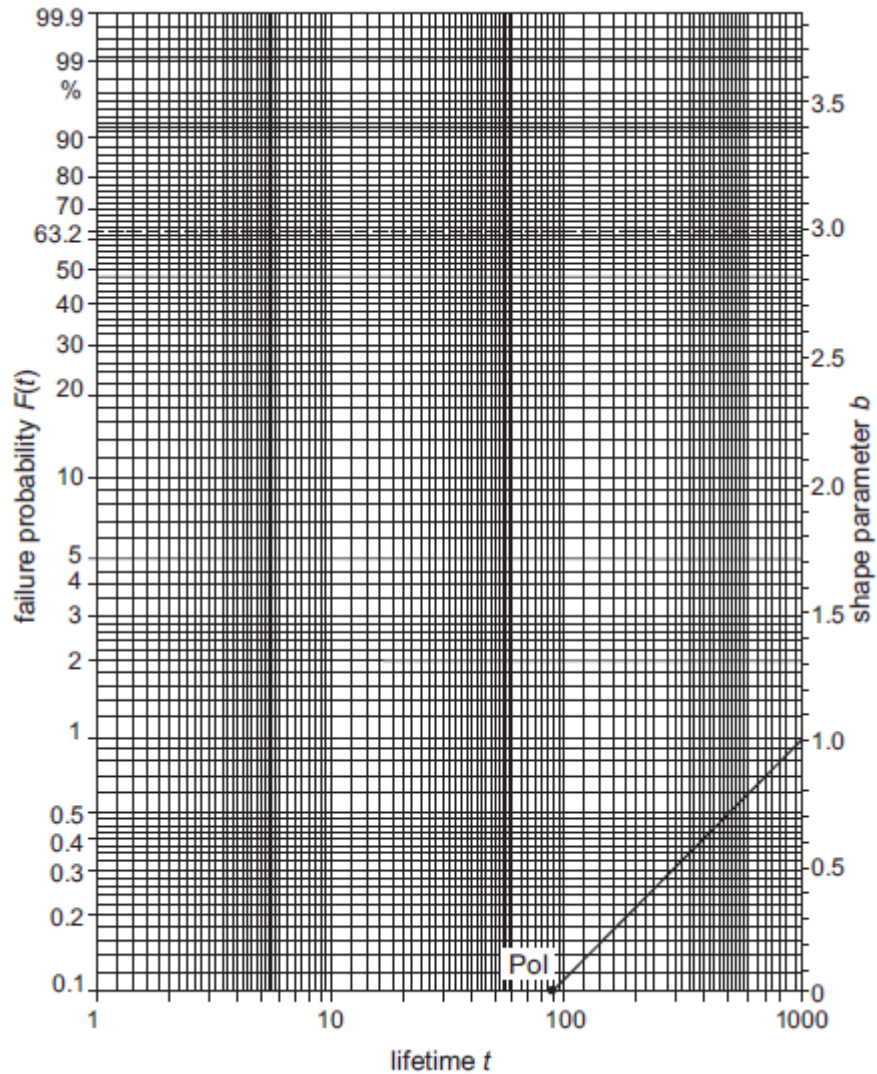
*Figure 5 - Probability plot paper [9]*

The points on the plot represent the failures analyzed. The x values of the points correspond to the ages of failed units in the data set [7]. To determine the y plotting positions it is necessary to find out a value indicating the corresponding unreliability of each failure. In order to do so, first, failed and survival units must be ranked according to its ages, from the younger to the older. Next, the rank of failed units must be adjusted using the equation below:

$$Adjusted\ Rank = \frac{(Reverse\ Rank)x(Previous\ Adjusted\ Rank) + (N + 1)}{(Reverse\ Rank) + 1}$$

$N = total\ number\ of\ units$

As an example the units of Table 1 where ranked and the failed units ranks were adjusted in Table 2.

| Unit | Age |
|---|---|
| Survival | 10 |
| Failure | 20 |
| Failure | 25 |
| Survival | 40 |
| Survival | 50 |
| Survival | 60 |
| Failure | 65 |

*Table 1 – Probability plot example: failed and survival units*

| Rank | Reverse Rank | Unit | Age | Adjusted Rank | Median Rank |
|---|---|---|---|---|---|
| 1 | 7 | Survival | 10 | - | |
| 2 | 6 | Failure | 20 | $\frac{6*0 + (7+1)}{6+1} = 1.143$ | 11.4% |
| 3 | 5 | Failure | 25 | $\frac{5*1.143 + (7+1)}{5+1} = 2.286$ | 26.4% |
| 4 | 4 | Survival | 40 | - | |
| 5 | 3 | Survival | 50 | - | |
| 6 | 2 | Survival | 60 | - | |
| 7 | 1 | Failure | 65 | $\frac{1*2.286 + (7+1)}{1+1} = 5.143$ | 65.4% |

*Table 2 – Probability plot: adjusted ranks and median rank*

After the ranks of the failures were adjusted, the Benard's Approximation for Median Rank is used to obtain an estimate of unreliability for each failed unit as follows:

$$MR = \frac{j - 0.3}{N + 0.4}$$

$N = total\ number\ of\ units$

$j = failure\ adjusted\ rank$

Then, given the y and x value for each failure point, they can easily be pltted. Next, the best possible straight line is drawn through these points. Once the line has been drawn, the parameter β, can be obtained, it corresponds to the line slope. To determine the scale parameter, η, it is necessary to read the time from the x-axis corresponding to $Q(t) = 63,2\%$.

Although this method lacks mathematical precision it is easily implemented by hand, making it helpful for understanding fundamental ideas behind the process of parameter estimation.

*Rank regression*

Essentially, the rank regression is a mathematically based version of the method discussed previously. Following the idea of probability plotting, the rank regression method requires that a straight line be fitted to a set of data points obtained through ranking the units and estimating the failures median ranks. the straight line is fitted mathematically. The fitting is done such that the sum of the squares of the distances of the points to the fitted line is minimized [8]. The minimization can be realized in either vertical or horizontal direction, as seen in Figure 6.
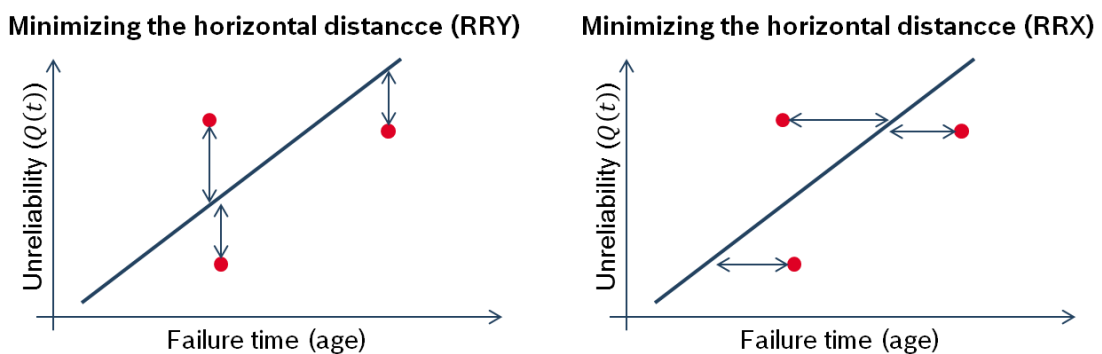


*Figure 6 - Minimization methods of the rank regression based on [8]*

Regressing on Y (RRY) means that the distance of the vertical deviations from the points to the line is minimized. If the regression is on X (RRX) then the line is fitted so that the horizontal deviations from the points to the line are minimized.

RRY:

According to the least squares principle the best straight line to minimize the vertical distance to the obtained data points is the $y = a + bx$ such that:

$$\sum_{i=1}^{N}(\hat{a} + \hat{b}x_i - y_i)^2 = \min \sum_{i=1}^{N}(a + bx_i - y_i)^2$$

where $\hat{a}$ and $\hat{b}$ are the least squares estimates of $a$ and $b$ and $N$ is the number of data points. The distance is minimized by estimating $\hat{a}$ and $\hat{b}$ as follows:

$$\hat{a} = \frac{\sum_{i=1}^{N} y_i}{N} - \hat{b}\frac{\sum_{i=1}^{N} x_i}{N}$$

$$\hat{b} = \frac{\sum_{i=1}^{N} x_i y_i - \frac{\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N}}{\sum_{i=1}^{N} x_i^2 - \frac{(\sum_{i=1}^{N} x_i)^2}{N}}$$

Using the linear transformations:

$$x = \ln(t)$$

$$y = \ln\left(ln\left(\frac{1}{1 - Q(t)}\right)\right)$$

The line equation can be rewritten as:

$$y = \beta x - \beta \ln(\eta)$$

and b*eta* and *eta* can be obtained.

RRX:

Similarly, the best straight line to minimize the horizontal distance to the obtained data points is the $y = a + bx$ such that:

$$\sum_{i=1}^{N}(\hat{a} + \hat{b}y_i - x_i)^2 = \min \sum_{i=1}^{N}(a + by_i - x_i)^2$$

where $\hat{a}$ and $\hat{b}$ are the least squares estimates of $a$ and $b$ and $N$ is the number of data points.

The distance is minimized by estimating $\hat{a}$ and $\hat{b}$ as follow:

$$\hat{a} = \frac{\sum_{i=1}^{N} x_i}{N} - \hat{b}\frac{\sum_{i=1}^{N} y_i}{N}$$

$$\hat{b} = \frac{\sum_{i=1}^{N} x_i y_i - \dfrac{\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N}}{\sum_{i=1}^{N} y_i^2 - \dfrac{\left(\sum_{i=1}^{N} y_i\right)^2}{N}}$$

Again, using the linear transformations:

$$x = \ln(t)$$

$$y = \ln\left(ln\left(\frac{1}{1 - Q(t)}\right)\right)$$

The line equation can be rewritten as:

$$y = \beta x - \beta \ln(\eta)$$

and b*eta* and *eta* can be obtained.

<u>Correlation Coefficient</u>:

The correlation coefficient, ρ, measures how good the fitting of the linear regression model to the analyzed data is and can be defined as:

$$\rho = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 * \sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

A correlation coefficient value of zero indicates that the data have no correlation relation to the linear regression model. On the other hand, the closer the absolute value is to 1, the better the linear fit.

*Maximum Likelihood Estimation (MLE)*

The central idea of MLE method is to obtain the most likely parameters values of a distribution. In order to do so a likelihood function is developed based on the available data. The most likely values of the parameters are the ones that maximize the likelihood function. A method to find the parameters estimates involve taking the partial derivatives of the likelihood functions with respect to the desired parameters, setting the resulting equations to zero and solving them simultaneously to determine

the values of the parameter estimates. Considering the *pdf, $f(t)$,* and the *cdf, $F(t)$,* for the 2-parametric Weibull distribution the likelihood function is as follow:

$$L(\beta, \eta | T_1, \dots, T_R, S_1, \dots, S_M) = \prod_{i=1}^{R} f(T_i; \beta, \eta) * \prod_{j=1}^{M} [1 - F(S_j; \beta, \eta)]$$

$R = number\ of\ failures$

$M = number\ of\ survivals$

$T_i = failures\ ages$

$M_j = survivals\ ages$

### 2.3.4: Failure forecast

A failure forecast is a prediction of the expected number of future failures that may occur in a specific period of time [7]. Such prediction can be calculated using the Weibull parameters together with the age of survival units. The failure forecast for a defined period of time sums the predicted failures for each survival in the period. If a set of *M* products, with lifetime of $M_j$ will accumulate *u* additional time in a determined future period the failure forecast for this period can be calculated as:

$$Failure\ forecast = \sum_{j=1}^{M} \frac{F(M_j + u) - F(M_j)}{1 - F(M_j)}$$

$M = number\ of\ survivals$

$M_j = survivals\ ages$

# Chapter 3: Maintenance advice based on failure forecast and lifetime evaluation of field failures

As seen in the previous chapter, the Weibull analysis is a powerful tool that can be used for different reliability purposes. The forecast of future failures is a particular useful one and it is also the basis of the maintenance advice calculation in the present work. In order to find out which type of maintenance is the most adequate to each individual customer, a Rexroth specific methodology was developed.

The maintenance advice addressed in the methodology can be either corrective or preventive maintenance. The first part of this chapter is dedicated to explain the concepts behind those terms. Following, the required data inputs for the developed methodology are exposed and also the procedure to prepare them for evaluation. The rest of the chapter will address the methodology itself, explaining the process to evaluate field failure data in order to obtain the final maintenance advice. The chapter ends with an overview of the whole evaluation methodology.

## 3.1: Corrective and preventive maintenance

The maintenance strategies offered by the Service department of Rexroth are labeled corrective or preventive maintenance. According to the DIN 13306 norm [14] on maintenance terminology:

"**Preventive maintenance**: maintenance carried out at predetermined intervals or according to prescribed criteria and intended to reduce the probability of failure or the degradation of the functioning of an item.

(…)

**Corrective maintenance:** maintenance carried out after fault recognition and intended to put an item into a state in which it can perform a required function."

Rexroth has high reliability standards for the intended lifetime of its products; therefore, in this period of lower failure occurrence, corrective maintenance is usually advised. As the products become older and enter the wear out phase failures occur much more frequently. Although the products have already exceeded its intended

lifetime they are still in operation. In these cases Rexroth advises preventive maintenance to reduce the product high probability of failure and reduce losses related to operation downtime.

At Rexroth, when a product comes to maintenance the actions performed to recover it depend on its age. For failed products with less than five years only a repair is performed, while older products are repaired and overhauled. Units sent for preventive maintenance are overhauled but not repaired since they have not failed.

## 3.2: Data Input

The methodology main data input consists of information about deliveries and failures of Bosch Rexroth products, referred as field failure data or SAP data, since that is where it can be obtained. SAP data will be addressed in more details in the next four subsections: 3.2.1:, 3.2.2:, 3.2.3: and 3.2.4:.

Additionally, a few advanced evaluation parameters regarding the Weibull analysis will be described in subsection 3.2.5:. Also some parameters concerning operation conditions are required. Those values are associated with failures downtime and maintenance costs and are usually related to the customer analyzed being individually settled for each evaluation. More details on these parameters can be found in the last subsection 3.2.6:.

### 3.2.1: SAP data

The SAP/R3, produced by SAP AG (Walldorf, Germany), is the standard Enterprise Resource Planning (ERP) software currently implemented in most of the Bosch Rexroth facilities worldwide. It is used, among other features, to collect, store and manage product data.

Every time a product of Bosch Rexroth is sold, a record regarding its information is stored (e.g. delivery date, customer, material description, production plant…). Similarly, every time a failed product comes to maintenance the information about the findings and the repair is documented (e.g. complaint date, failure type, failed parts, failure description…). With this approach, Bosch Rexroth has built over the last years a huge data base of product information. Among the benefits of having such information is the possibility to realize a range of reliability analysis based on

lifetime data. It is worth to mention that without this resource the present tool would be inconceivable. Such information is referred to as "SAP data" and can be retrieved as Excel tables through the ERP system graphic user interface. The files obtained are similar to the ones showed in Figure 7.

| Delivery number | Customer | Quantity | Delivery date | Production Plant | Material Description | ... |
|---|---|---|---|---|---|---|
| 177899887 | A | 123 | 21.05.2001 | Lohr, DE | XXX04.2 | ... |
| 177323224 | A | 140 | 26.04.2001 | Lohr, DE | XXX04.4 | ... |
| 177829383 | C | 211 | 18.06.2001 | Lohr, DE | XXX03.1 | ... |
| 187903020 | C | 431 | 21.05.2001 | Lohr, DE | XXX01.5 | ... |
| 940563235 | C | 244 | 21.05.2001 | Atibaia, BR | XXX03.2 | ... |
| 940473791 | B | 176 | 21.05.2001 | Atibaia, BR | XXX04.4 | ... |
| 940629438 | A | 120 | 21.05.2001 | Atibaia, BR | XXX01.5 | ... |
| 940428186 | A | 47 | 21.05.2001 | Lohr, DE | XXX03.2 | ... |
| ... | ... | ... | ... | ... | ... | ... |

| Complaint number | PROD SHORT-DESC | Quantity | Delivery date | Complaint date | K-SN | Final Customer | Failure Type | Failed Parts | ... |
|---|---|---|---|---|---|---|---|---|---|
| 940420107 | XXX04.2 | 1 | 21.05.2001 | 01.06.2009 | 27828-067 | A | 0000 | Part 72 | ... |
| 940160830 | XXX04.4 | 1 | 26.04.2001 | 01.04.2006 | 27828-013 | A | 0000 | Part 83 | ... |
| 940526163 | XXX03.1 | 1 | 18.06.2001 | 01.12.2010 | 27728-078 | C | 0001 | Part 65 | ... |
| 940344058 | XXX01.5 | 1 | 21.05.2001 | 01.01.2013 | 27828-071 | A | 0013 | Part 90 | ... |
| 940563235 | XXX03.2 | 1 | 21.05.2001 | 01.06.2008 | 7828-0763 | B | 0008 | Part 65 | ... |
| 940473791 | XXX04.4 | 1 | 21.05.2001 | 01.06.2011 | 277828-04 | B | 0011 | Part 83 | ... |
| 940629438 | XXX01.5 | 1 | 21.05.2001 | 01.04.2010 | 2828-0766 | B | 0010 | Part 47 | ... |
| 940428186 | XXX03.2 | 1 | 21.05.2001 | 01.04.2012 | 27728-068 | C | 0012 | Part 12 | ... |
| ... | ... | ... | ... | .... | ... | ... | ... | ... | ... |

*Figure 7 - Examples of delivery and claim data files*

The useful data of the delivery and claim files can be categorized in three groups according to their use in the methodology: the evaluation data set, the preprocessing data and the data used for the selection process. One type of data, such as the delivery date, can have more than one use and therefore belong to more than one group, as seen in Figure 8.
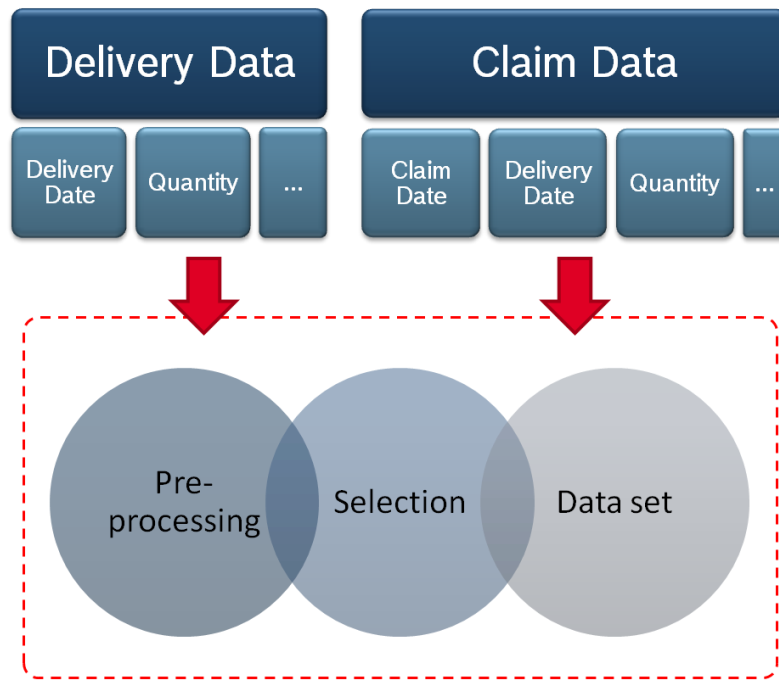
*Figure 8 - Uses of delivery and claim file data*

For the Weibull analysis and failure forecast, it is necessary to know the age of the analyzed products [7]. In the industry field usually no better aging parameter is available than the time between product delivery and claim data, called service time. The delivery date describes the date on which the product was delivered to the customer while the complaint date corresponds to the date on which a failure was reported by the customer. This group of data is called analysis data and is available in two files, the delivery data file and the claim data file. Together they compose the data set of the evaluation as seen in Figure 9.



*Figure 9 - Components of the evaluation data set*

34

In the delivery data file, each line of the table corresponds to a delivery. The basic description of a delivery consists of the quantity delivered and the date the delivery was made. In the claim data file, each line of the table corresponds to a reported complaint. The basic description of a complaint consists of the complaint date, delivery date of the failed products and failed quantity.

A second important group of data is referred as pre-processing data. Ideally, the analysis data should be ready for use in the Weibull evaluation but that is not the case. Some extra columns of the delivery and claim data files are necessary for data verification and manipulation. The problems and inconsistencies found in the input files require the pre-processing of data in order to make it suitable for the Weibull analysis. Inconsistencies and pre-processing are explained in detail in the upcoming 3.2.2: and 3.2.3: sections.

The delivery and claim data files contain several information about a product. In other words it means that information about customers, product generations, failure types, delivery dates, Original Equipment Manufacturer (OEM), among others, are mixed in big files. It might be of interest to perform an analysis with the whole information about one product but usually some specific evaluation is needed such as one specific customer or product generation. Therefore, a variable third group of data can be optionally used to select the desired data through the whole set. This process will be described in the section 3.2.4:.

### 3.2.2: Data inconsistencies

The effort made by DC to keep the SAP/R3 up-to-dated and consistently working is huge. Nevertheless data inconsistencies occur frequently, especially when it comes to standardization and maintainability of fields (meaning that not all records provide the required fields). The main reason of these inconsistencies is that until now the SAP/R3 data was hardly ever used to generate profit.

In the past decades there was a successive changing in the ERP software due to several acquisitions of companies. These constant changes generated inconsistencies in the data base that remains in the newest ERP software (e.g. some data was lost in the process).

Since the SAP/R3 was settled at Lohr in October 2003, it would be reasonable to assume that the records from this period are consistent but this is not entirely true. Three main issues explain most of the problems found in the products information. The first relates to change of standards in the SAP recording, for example changes in the way failures are described and identified through time (new codes and standards are periodically defined). The second issue is due to lack of standardization worldwide regarding the ERP software used. Different SAP machines are used and some locations don't even have it (e.g. Brazilian facilities report their data directly on a database tool. Once a month those data are retrieved to the world-wide database). Though clear guidelines on how to document failures and sales exist, in some countries they're not strictly followed. This can occur due to the different appearance, messages and fields of the SAP systems that make the guidelines hard to follow, or due lack of knowledge in identifying the right codes to be used (e.g. failure types categories are not always clear). The last important cause of inconsistencies is related to customer's names and customer's numbers. Companies are mutable and just like DC, several customers companies have been sold to or bought other companies. Those transactions have not been followed by the SAP data and as result there are some customers with more than one number and/or name.

Summarizing, the delivery data file has standardization problems regarding some information fields, such as customer's names and numbers, and also issues related to maintainability of fields. The claim data file presents the same problems observed in the delivery data file and additionally problems related to different codes usage.

Special attention was drawn to SAP data inconsistencies subject during the development of the methodology. The issues were analyzed and reported to the head of the Service department in Lohr who was sensitized to the importance of the matter. Big changes are needed and the process will be long and slow. To make the project feasible without having to wait for the 'perfect' data a pre-processing method was defined by the Service Department. This process is done not only to deal with data inconsistencies but it is also necessary to select the right types of failures as it will be described in the next section.

### 3.2.3: Pre-processing of data

This phase of the methodology can be divided in two major steps, pre-processing of the delivery data file and pre-processing of the claim data file. The problems found in the delivery data files are not critical and despite the difficulties imposed due to lack of standardization, if the sorting and selecting of data are done properly, no effect is seen in the analysis. On the other hand, the data issues in the claim data files can result in wrong analysis hence bad maintenance advice. Therefore the preprocessing of claim data files is much more complex and hard to automate since expert knowledge is required to evaluate specific fields.

The first step in the pre-processing of the delivery data file is to verify if every delivery record has a delivery date. Without this date the record cannot be used in the analysis and must be discarded. Next the quantity values must be verified. Negative values correspond to product returns and must be discarded.

Each failure recognized in the maintenance of a product has a record in the claim data file. If more than one failure has occurred in the same unit each failure will be described in one record, being one of them the primary failure and the others the secondary failures. Therefore, to begin with the pre-processing of the claim data file, the records corresponding to secondary failures must be deleted. Following, the delivery date and complaint date must be verified. Again, all claim records must have a delivery and complaint date. In case of missing dates, if the record has a serial number it is possible to search for the missing product information in SAP. Nevertheless, if delivery or complaint dates are not found the record cannot be used and must be discarded. For the lifetime analysis only wear out failures should be considered, hence failures due to customer misuse, improper installation and conditions, transport, among others must be discarded. Records that describe products where no failures where found must also be deleted. Finally the last step corresponds to the selection of the first failure timewise occurred for each unit. The same unit can fail several times during its lifetime, however due to product modifications during the maintenance it was established that the unit age should be "resettled" after being repaired. Also it is not known how the reliability of the repaired product is affected, for this reason separated analyses should be performed for first, second, third failures and so on.

After the completion of the mentioned steps the data set is ready for evaluation.

### 3.2.4: Data selection

As mentioned before, much more than delivery and complaint dates are stored in the delivery and claim data files, they can actually contain a lot of information and characteristics of a product, therefore the selection of the information to be used should be performed.

The most frequently used sorting parameters are delivery date, Original Equipment Manufacturer (OEM), final customer and material name (which describes the product generation). However, a range of more than 20 fields is available for both delivery and claim data files, therefore different sorting parameters can be used depending on the goal of the analysis. After sorting by desired parameters it is necessary to select the required values from the tables.

For instance, let's suppose that an analysis for customer A, which purchased $2^{nd}$ generation products, delivered between 2001 and 2004 is being requested. In this case both delivery and claim data files must be sorted by final customer, material name and delivery date. Then the entries with the values: customer = 'A', generation = '2' and delivery date = 2001 - 2004 should be selected.

### 3.2.5: Advanced Weibull input parameters

Chapter 2: mentions different distributions to perform a Weibull analysis. In the current lifetime evaluation methodology two different elements are considered, the distribution type and the parameters estimation technique.

Concerning the distribution type, the data set is usually described by a 2-parametric Weibull distribution. In some cases, when the distribution fit is very poor, a 3-parametric Weibull distribution can be used introducing a third parameter, *gamma,* to describe a failure free time [9]. In the developed methodology a 2-parametric Weibull was used as the standard distribution but the 3-parametric could also be used if necessary and adequate.

Weibull authors differ when it comes to parameters estimation techniques. The following paragraphs show the opinion of two references, [7] and [8] respectively.

According to Abernethy [7], the standard method and best practice for small to moderate size samples is the median rank regression curve. In order to choose the regression method (RRX or RRY), it is recommended to select the scale with the largest variation as the dependent variable. The author also mentions that for lifetime Weibull plots, the X variable (the ages of the units) tends to have much more error than the median ranks; therefore the regression should be X on Y. When it comes to Maximum Likelihood Estimation (MLE) Abernethy is skeptical. He does admit that under some specific conditions MLE can provide more accurate results but argues that difficulties exist in the use of the method such as its complexity, the need of large samples (more than 100 failures) and the lack of a proper graphic display of the data, i.e. if RRX or RRY plotting positions are used the MLE curve will often appear to have a poor fit.

According to ReliaSoft Weibull experts [8], when enough failure records are available Maximum Likelihood Estimation (MLE) should always be used, since it is considered from a statistical point of view the most robust of the parameters estimation techniques. This means that the MLE method is asymptotically consistent[1], unbiased[2] and efficient[3]. The number of failures required (30 to 50 to more than 100, depending on the application) is sometimes prohibitive given that with small samples the method can be badly biased. In those situations it is recommended to use rank regression techniques regressing from the scale with largest variation to the scale with less variation. However it is stressed that MLE can handle suspensions better than rank regression, especially when a heavily censored data set is analyzed.

For DC, as a wide range of analyses is intended, both rank regression and MLE techniques are considered depending on the data input. The approximation method used on the median rank regression is Y on X. This is done because it is

---

[1] For large sample sizes the estimates converge to the right values.

[2] For large sample sizes the estimates are the most precise.

[3] For large sample sizes on average the right values are obtained.

known that many customers do not report every occurred failure (instead they either buy a new product or repair it by themselves). Under those circumstances the Y scale will contain the largest variation (missing failures will rise the ranking uncertainty) consequently the regression Y on X will be used (RRY).

### 3.2.6: Additional data inputs

Alongside with the product data gathered from SAP some other inputs concerning the data set itself and customers' characteristics are used in the evaluation. These inputs are grouped as mandatory and optional parameters and are categorized according to its uses in the methodology as seen in Figure 10.
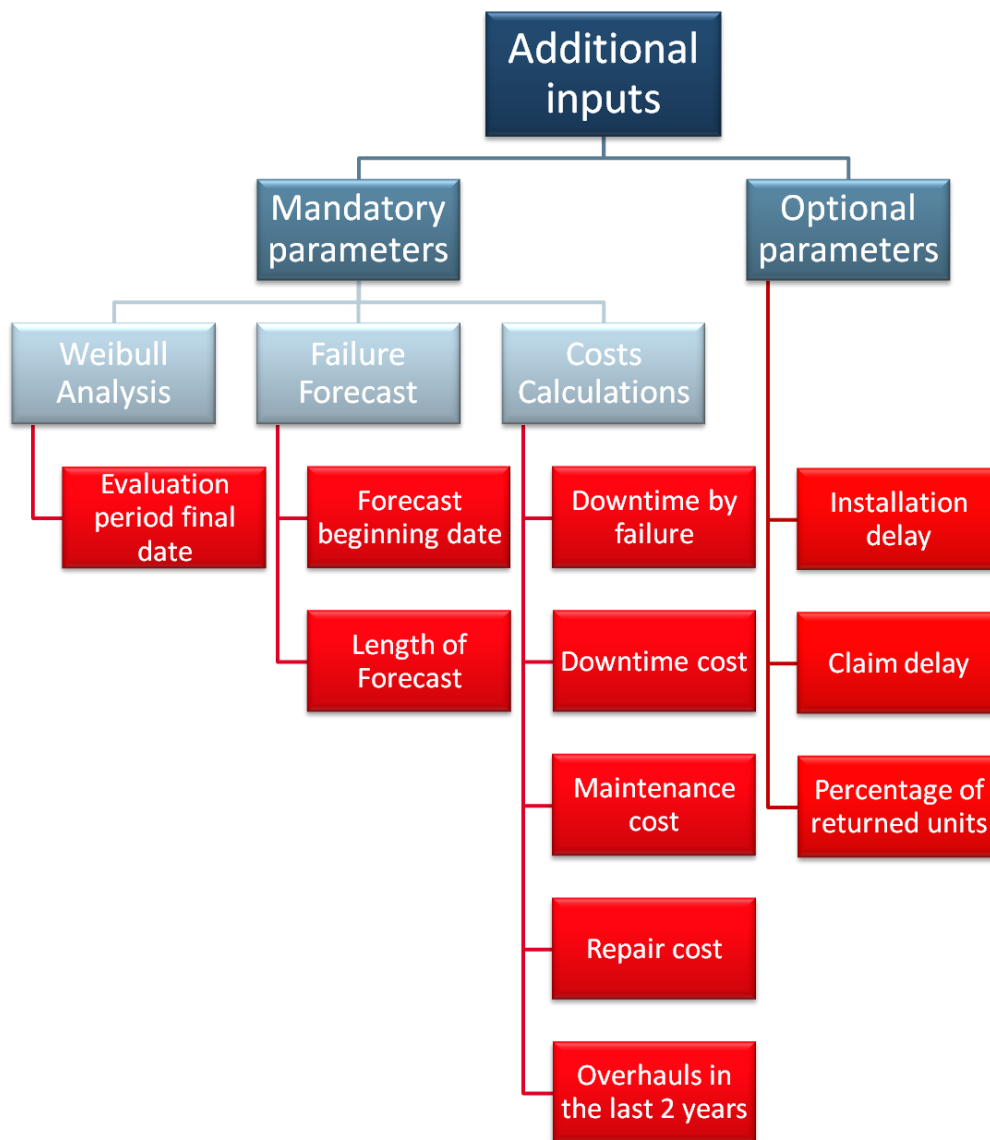


*Figure 10 - Additional data inputs*

For the Weibull analysis, besides the delivery and complaint dates of failures and survivors the evaluation period final date is required. It represents the date where the information was retrieved or updated, meaning that the analyzed information represents the real product data until that day.

Alongside with the scale and shape parameters of the correspondent Weibull distribution and the survivals units two other parameters are required to define and calculate a failure forecast according to Chapter 2:, the forecast beginning date and the length of the forecast. Those values are usually defined according to the customer needs. The forecast beginning date can assume the evaluation period final date as a standard but could also be defined as the potential date to realize a preventive service. The length of the forecast is usually shorter than five years since the accepted amortization time for an investment in Industry is close to this period but it could also be defined concerning the residual lifetime of machine or production plant in which the Rexroth products are installed.

The last mandatory parameters for the evaluations are related directly with the customers analyzed and are used in the calculation of preventive and corrective maintenance costs. These are the amount of downtime caused by a unit failure, how much an hour of downtime costs for the customer, the maintenance cost of a unit, the repair cost of a unit and the number of overhauls made in the last 2 years.

In addition to the mandatory inputs four optional data parameters can be considered in the developed methodology, they are installation delay, claim delay, failure free time and percentage of returned units. If assumed correct these parameters can increase the accuracy level of the evaluation. The installation delay describes how much time was necessary to install the product and start operation and is used to calculate a more precise service time of the products. The claim delay describes how long it took from the failure occurrence to the failure claim and is also used to calculate a more accurate service time of the products.

In some cases these delays can be considerable; the differences in the service time are as seen in Figure 11.
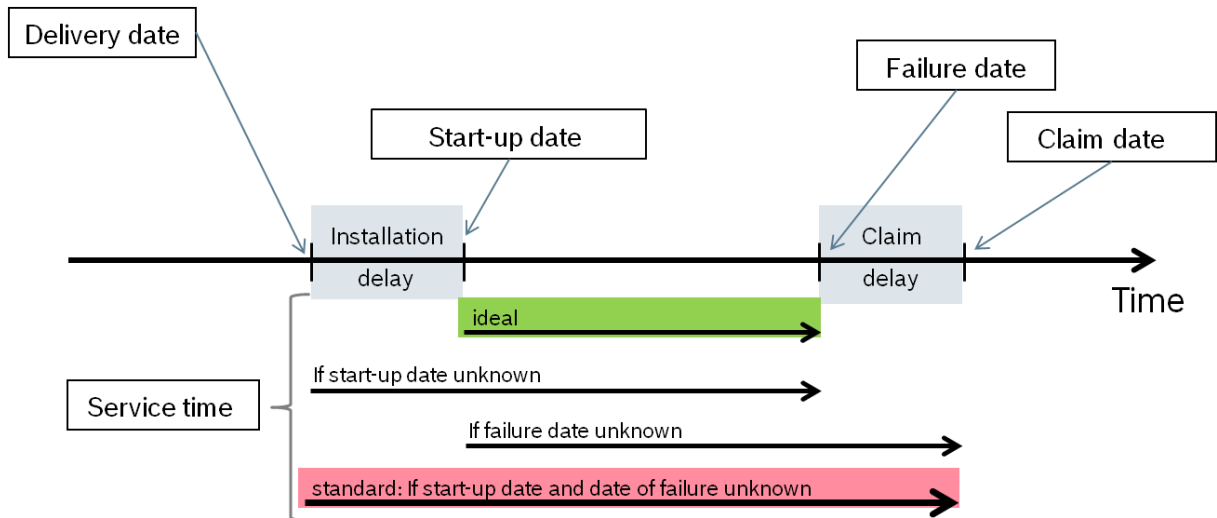


*Figure 11 - Product service time*

Finally if the proportion of complaints among the total occurred failures is known, the so called percentage of returned units, it can be used to calculate a more correct failure forecast. This percentage of returned units reflects the field failures, which are actually claimed by the customer. In cases where the percentage value is smaller than 100%, not considering it will lead to inaccurate results. Possible solutions are the reduction of the delivery quantity or the correction of the failure forecast according to the percentage value.

## 3.3: Lifetime evaluation

After accomplishing the preparation and selection of the delivery and claim data, followed up by the definition of the aforementioned parameters, the lifetime evaluation step of the proposed methodology may begin. The assigned procedure consists of performing a Weibull analysis based on field failure data followed by a failure forecast calculated using as basis the *beta* and *eta* Weibull parameters. The Weibull analysis and failure forecast procedures are respectively explained in the 3.3.1:/3.3.2: and 3.3.3:. Additionally, as it will be explained, a proportional forecast may be needed in cases where data restrictions are found.

### 3.3.1: Weibull analysis

Performing a Weibull analysis (a common reference to a lifetime analysis), consists on the attempt to make predictions about the life of a population by modeling its behavior through a statistical distribution [8]. With the obtained parameterized distribution it is then possible to estimate useful life characteristics of the products such as reliability, MTTF (mean time to failure), probability of failure at a specific time and failure rate.

After the desired data is selected the information contained in the delivery and claim file must be categorized as survivals and failures units as seen in Figure 12. Failed units can be directly obtained from the claim file and are described through its delivery and claim dates. The survivals are the remaining units obtained through the comparison of delivery dates between the delivery and claim file, i.e. delivered units that have not failed.
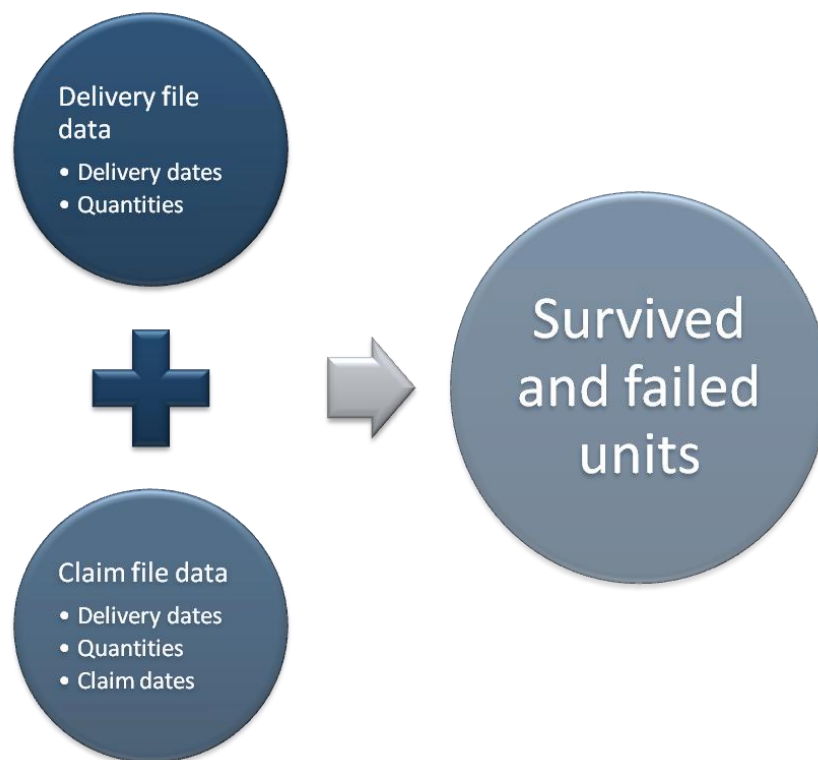


*Figure 12 – Survived and failed units*

In the present methodology the 2-parametric Weibull was the standard lifetime distribution chosen for the life data analysis. The first step in order to fit the distribution to the data set is to estimate the *beta* and *eta* Weibull parameters.

Therefore it is necessary to calculate the ages of all analyzed products, for both the failures and survivors described by the selected date, from the delivery and claim files. Products ages are calculated in months between the start-up and the failure date for failed units and in months between the start-up date and the final date of the evaluation period for survivors units.

Following the steps of the RRY and MLE parameter estimation methods previously described on Chapter 2: a Weibull distribution is fitted to the data, *beta*, *eta* and correlation coefficient are calculated. A Weibull plot showing the graphic visualization of the data and the Weibull curve are also obtained. Also, based on the estimated parameters it is now possible to calculate the MTTF, B10 and failure rate of the analyzed product.
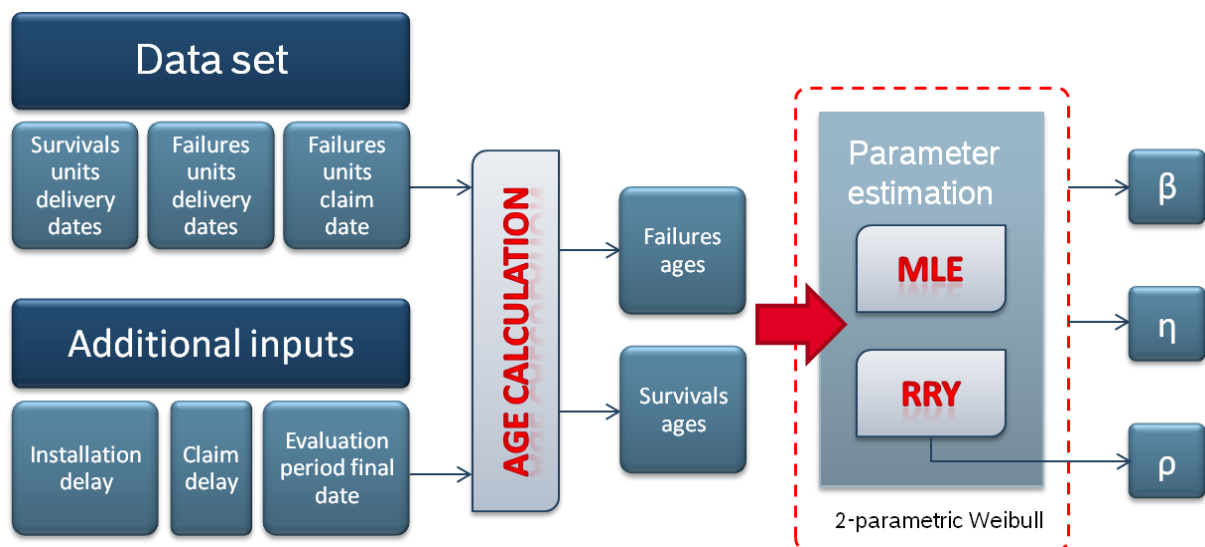


*Figure 13 - Weibull analysis*

## 3.3.2: Quality Indicators

Before proceeding with the failure forecast and determining the maintenance advice it is necessary to verify the quality of the Weibull distribution. Without a reliable analysis the results obtained cannot be trusted and the maintenance advice can be misdiagnosed. In order to identify sources of data problems in the analyzed data five quality indicators were defined and can be seen in Figure 14.

Each indicator is similar to a traffic light, when green it is possible to proceed, if yellow the results should be carefully employed knowing that the level of accuracy is not the ideal and finally a red indicator implies that there are evidences of problems in the data set.
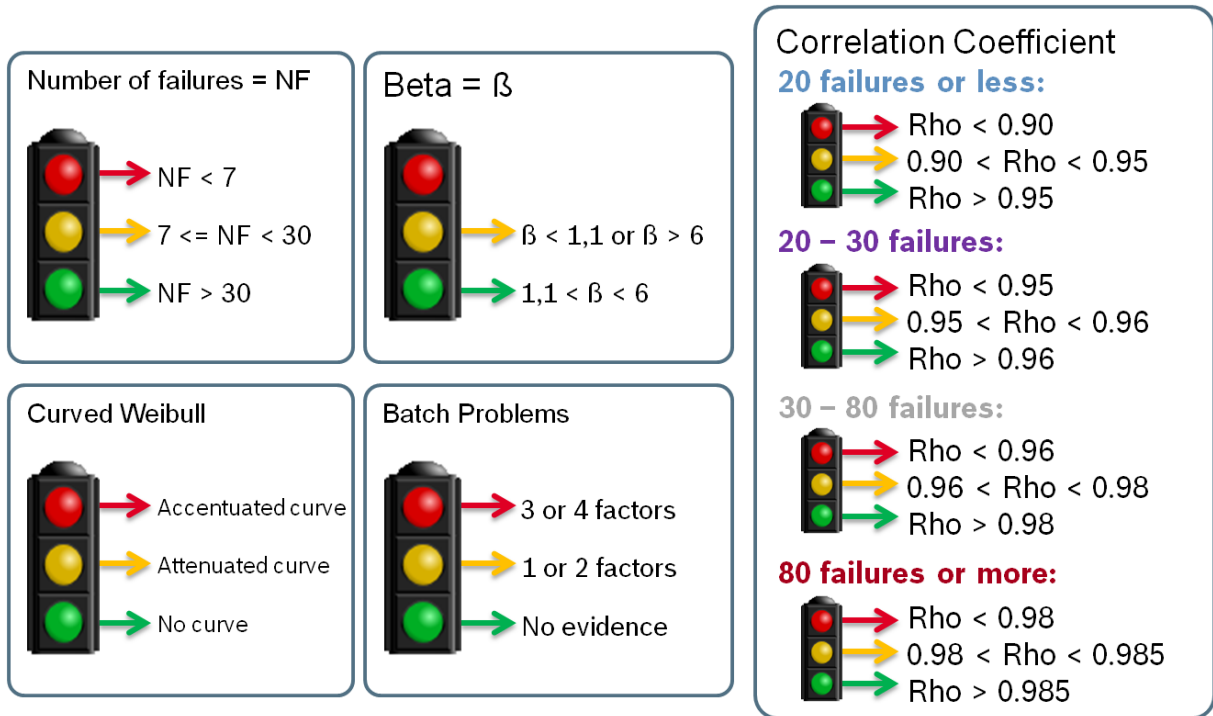


*Figure 14 - Quality indicators*

The first evaluated indicator is the number of failed units. For statistical purposes a minimum of 30 failures is recommended although 20 failures would be acceptable [7]. Based on literature and previews experience with Weibull analyses it was defined that if less than 7 failures are available, the indicator is red. From 7 to 29 failures the indicator is yellow and if at least 30 failures are available for analysis the indicator is green.

The next indicator is related to the range of the Weibull shape parameter (beta). For wear out failures there is a drastically enhance in failure rates as the product lifetime increases, this behavior can be described with betas values greater than 1 [9]. Therefore the minimum value for beta was defined as 1.1 and the maximum as 6 since steep slopes in the Weibull plot can often hide curves, outliers and other Weibull data problems [7]. If beta is not in the defined range the indicator is

considered yellow. There are no red values for this indicator since values smaller than 1.1 or higher than 6 do not necessarily indicates problems but early or random failures in the first case and can simply hide other problems in the second case.

The third quality indicator is related to the goodness of fit of the Weibull distribution and is defined based on the correlation coefficient. The correlation coefficient measures how strong the linear relation between the median ranks (calculated in the RRY method) and the age of the units is [8]. The closer the correlation coefficient is to 1 the higher is the probability that the analyzed data can be represented by the Weibull distribution. The values for a good fit are defined based on the 90% critical correlation coefficient (CCC) found by ranking the correlation coefficient of 1000 simulated Weibull distributions trials and choosing the highest value of the lowest 100 values [7]. The results obtained are shown in Figure 15 and were used as basis for the correlation coefficient indicator's limit values. As the limit values varies with the number of evaluated failures they can be easily visualized in Figure 14.



*Figure 15 - CCC values based on [7]*

Curved Weibulls can express a wide range of problems in the analyzed data such as mixture of failure modes and mismatch of the Weibull distribution (another distribution can be more appropriate for the data set) [7]. Hence the curved Weibull quality indicator was created to evaluate the presence of curves in the plot. The
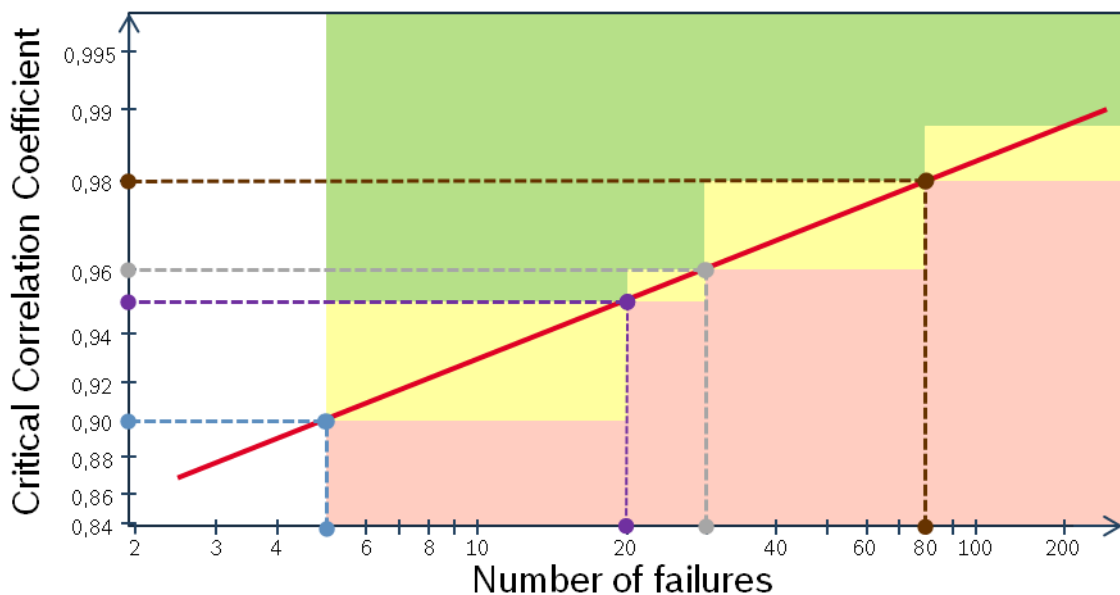
green light is assigned when no curves are noticed, few and attenuated curves will make the indicator yellow while accentuated curves will make the indicator red.

The last devised indicator is meant to identify signs of batch problems in the analyzed data. There are well known methods to recognize batch failures and for this indicator four factors based on them were used. The first is the number of late suspensions (units that survive much longer than the last failure). Many late suspensions, especially in combination with early failure of new units may suggest a batch problem. The comparison of the beta value for MLE and the beta value for RRY is another used method. If beta derived from MLE is much less than the beta derived from RRY, this may indicate a batch problem. Also if the expected number of failures (calculated by the "Now Risk") is much larger than the observed number of failures there is evidence of a batch problem [7]. Finally the previous curved Weibull quality indicator is also a sign of batch failures.

When together analyzed the quality indicators are a measure of the Weibull analysis reliability and indicate if the reality will be reflected in the final maintenance advice. If the quality indicators suggest that the quality of the analysis is not satisfying the data must be manually reevaluated to find the problem.

### 3.3.3: Failure forecast

Once an adequate Weibull distribution is obtained to represent the product data it is possible to perform a forecast of the expected future failures. Using the preset parameters forecast beginning date and length, the forecast period is defined and the forecast is calculated based on the procedure described in Chapter 2. An overview of the process with its inputs and outputs is shown in Figure 16

If the forecast beginning date is not the same as the evaluation period final date the forecasted failures for the period in between these dates must be discarded. This can occur in cases where the customer cannot close a maintenance contract immediately but only in a few months or years. The failures occurred until then are irrelevant for obtaining the maintenance advice.

There are cases in which the customer has not sent all failed units to repair. If the percentage of returned units is known in relation to the total number of failures

the failure forecast can be corrected to better represent the future failures of the customer.

Another possible scenario is that the customer did not send any product for repair or they were sent through the OEM. Small companies usually do not build their own machines; therefore they buy Rexroth products installed in a machine built up by an OEM. Likewise instead of sending their units to repair straight to Rexroth they send it to the OEM. In this cases no record of the customer's failed units will be found in the claim files, they will all be described as OEM units. The procedure in this case is to make a proportional forecast. A Weibull analysis and a failure prediction are done with the OEM data and based on the percentage of deliveries between the customer and the OEM a proportional correction is made in the failure forecast. This is only possible because both the OEM and the final customer can be retrieved from the delivery data files. Unfortunately this is not yet available in the claim data files.
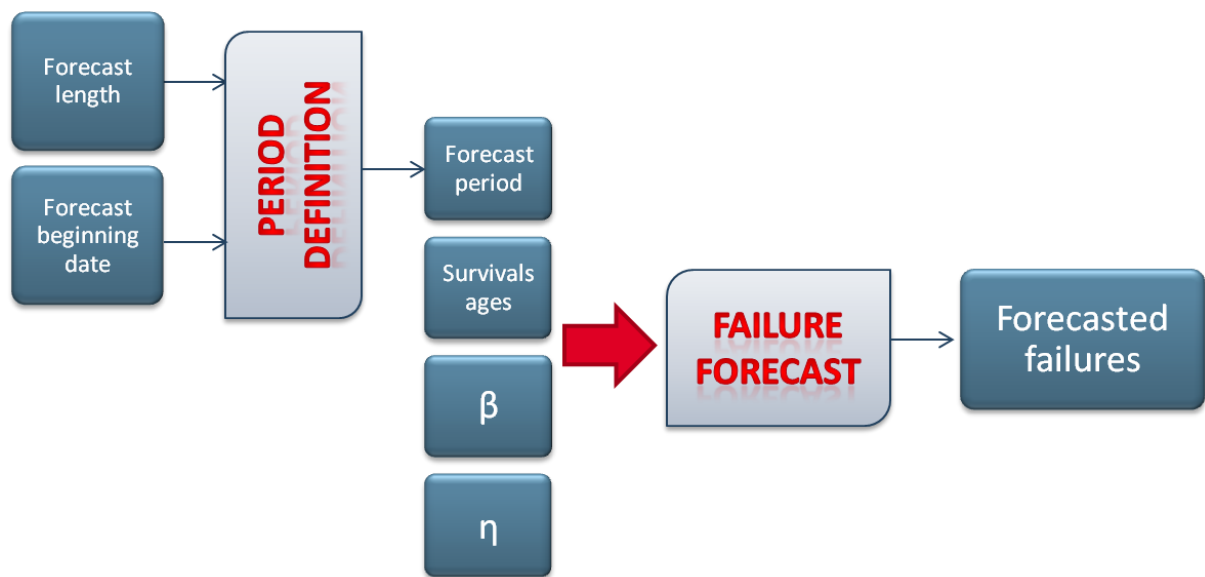


*Figure 16 - Failure forecast*

## 3.4: Costs calculation

The final step to obtain the maintenance advice is to estimate the costs of corrective and preventive maintenance based on the failure forecast. Different factors have influence in the calculations and are represented by the mentioned parameters: amount of downtime caused by a unit failure, cost of an hour of downtime, overhaul cost of a unit repair cost of a unit and number of overhauls made in the last 2 years.

The methods to compute the costs of corrective and preventive maintenance are similar as seen in Figure 17; its differences are explained in the following subsections.
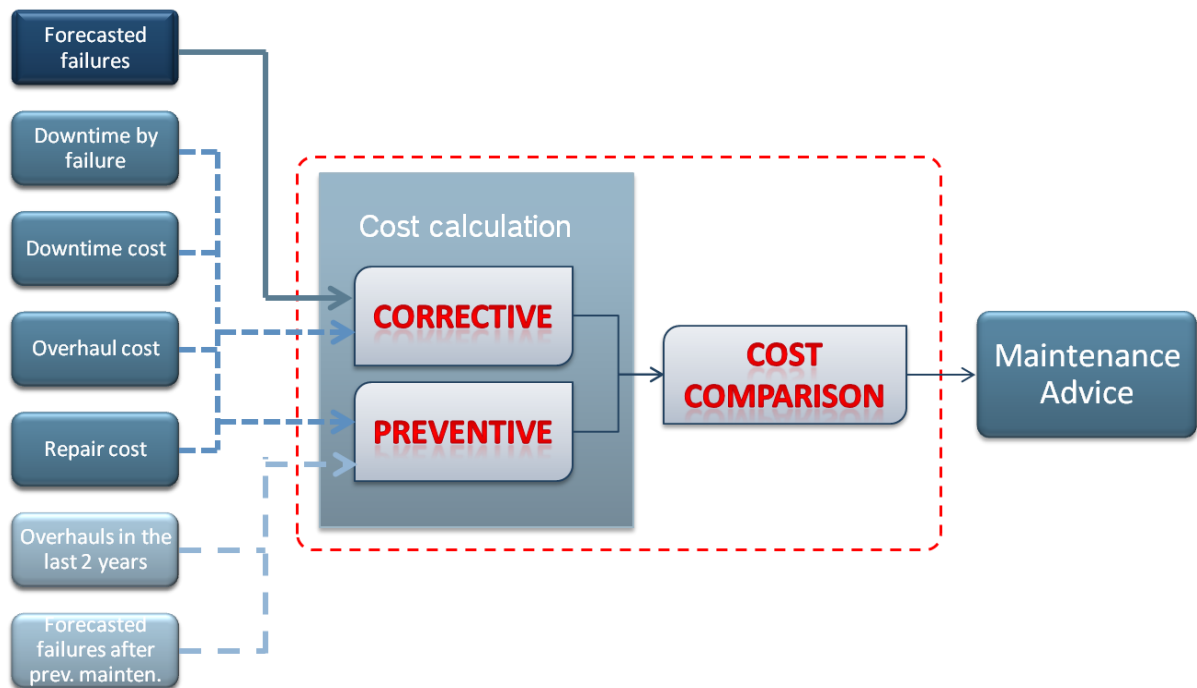


*Figure 17 - Corrective and Preventive maintenance*

### 3.4.1: Corrective Maintenance

To calculate the costs of corrective maintenance during a certain period of time not only the repair and overhaul costs of forecasted failures are considered. When a unit fails, the production depending on it stops resulting in losses. These losses must be taken in account in the calculation as well. Therefore the total cost of corrective maintenance is expressed by the sum of the costs to repair and overhaul all failed units plus the costs of downtime generated by the same failed units. A simple formula can be used to calculate and understand how the corrective maintenance costs are estimated:

$$Total\ cost\ of\ corrective\ maintenance = C_{cm} = \sum_{i=1}^{y}(f_f[i]*(C_o + C_r) + f[i]*t_d*C_d$$

$y = lenght\ of\ forecast\ in\ years$

$f_f = array\ containing\ the\ failure\ forecast\ for\ each\ of\ the\ y\ years$

$C_o = overhaul\ cost\ of\ a\ unit$

$C_r = repair\ cost\ of\ a\ unit$

$t_d = downtime\ by\ failure\ in\ hours$

$C_d = downtime\ cost\ for\ a\ unit\ in\ one\ hour$

To clarify the corrective maintenance costs are calculated with the data of Table 3:

| Param. | Value |
|:---:|:---|
| y | 7 years |
| $f_f$ | [5 7 10 14 19 25 32] |
| $C_o$ | €700,00 |
| $C_r$ | €500,00 |
| $t_d$ | 2 hours |
| $C_d$ | €5.000,00/hour |

*Table 3 – Example of parameters for the calculation of the corrective maintenance costs*

The results obtained can be seen in Table 4.

| Year | Failures | Overhaul + repair costs | Downtime costs | Total costs |
|:---:|:---:|:---:|:---:|:---:|
| 1 | 5 | € 6.000,00 | € 50.000,00 | € 56.000,00 |
| 2 | 7 | € 8.400,00 | € 70.000,00 | € 78.400,00 |
| 3 | 10 | € 12.000,00 | € 100.000,00 | € 112.000,00 |
| 4 | 14 | € 16.800,00 | € 140.000,00 | € 156.800,00 |
| 5 | 19 | € 22.800,00 | € 190.000,00 | € 212.800,00 |
| 6 | 25 | € 30.000,00 | € 250.000,00 | € 280.000,00 |
| 7 | 32 | € 38.400,00 | € 320.000,00 | € 358.400,00 |
| Total | 112 | € 134.400,00 | € 1.120.000,00 | € 1.254.400,00 |

*Table 4 – Example of the resulting calculation of the corrective maintenance costs*

### 3.4.2: Preventive Maintenance

As previously explained, the goal of a preventive maintenance contract is to overhaul all the customer's operating units so they become as good as new ones. With that in mind, for the preventive maintenance costs calculation it is defined that

all units that were not overhauled in the last two years will be overhauled in the first year of the contract. The costs for that year can be simply defined as:

$$Preventive\ cost\ of\ the\ first\ year\ of\ contract = C_{pm1} = (N - N_o) * C_o$$

$N = total\ number\ of\ units$

$N_o = units\ overhauled\ in\ the\ last\ 2\ years$

$C_o = overhaul\ cost\ of\ a\ unit$

Units overhauled in the last two years will not be overhauled because they are considered too young to undergo preventive maintenance again.

For the subsequent years of the intended lifetime the costs are calculated as the corrective maintenance costs with two main differences. The first is the failure forecast used, as all units undergo an overhaul the new failure forecast must be recalculated based on the same beta, eta and all units with age zero. The second relates to the maintenance cost of a unit. Until five years after overhaul the units that fail do not need to be overhauled, instead they are simply repaired and repair costs are lower than overhaul costs. The costs for the remaining intended lifetime are divided then in two:

$$Preventive\ cost\ of\ the\ first\ five\ years\ after\ overhauls = C_{pm5}$$

$$= \sum_{i=1}^{5} (f_{fn}[i] * C_r + f_n[i] * t_d * C_d$$

$$Preventive\ cost\ for\ the\ remaining\ intended\ lifetime\ = C_{pmr}$$

$$= \sum_{i=6}^{y} (f_{fn}[i] * (C_o + C_r) + f_n[i] * t_d * C_d$$

$y = lenght\ of\ forecast\ in\ years$

$f_{fn} = forecasted\ failures\ after\ preventive\ maintenance$

$C_o = overhaul\ cost\ of\ a\ unit$

$C_r = repair\ cost\ of\ a\ unit$

$t_d = downtime\ by\ failure\ in\ hours$

$C_d = downtime\ cost\ for\ a\ unit\ in\ one\ hour$

As an example, the preventive maintenance costs are calculated based on the same data used for the corrective maintenance costs, the data is shown in Table 5:

| Param. | Value |
|--------|-------|
| y | 7 years |
| N | 250 units |
| $N_o$ | 0 units |
| $f_f$ | [5 7 10 14 19 25 32] |
| $f_{fn}$ | [0 0 1 3 6 10] |
| $C_o$ | €700,00 |
| $C_r$ | €500,00 |
| $t_d$ | 2 hours |
| $C_d$ | €5.000,00/hour |

*Table 5 – Example of parameters for the calculation of the preventive maintenance costs*

The results obtained can be seen in Table 6:

| Year | Overhauls /Failures | Overhaul + repair costs | Downtime costs | Total costs | Maintenance description |
|------|---------------------|-------------------------|----------------|-------------|-------------------------|
| 1 | 250 | € 300.000,00 | € 0,00 | € 300.000,00 | Overhauled units |
| 2 | 0 | € 0,00 | € 0,00 | € 0,00 | Repaired units |
| 3 | 0 | € 0,00 | € 0,00 | € 0,00 | Repaired units |
| 4 | 1 | € 500,00 | € 10.000,00 | € 10.500,00 | Repaired units |
| 5 | 3 | € 1.500,00 | € 30.000,00 | € 31.500,00 | Repaired units |
| 6 | 6 | € 3.000,00 | € 60.000,00 | € 63.000,00 | Repaired units |
| 7 | 10 | € 12.000,00 | € 100.000,00 | € 112.000,00 | Repaired and overhauled units |
| Total | 270 | € 317.000,00 | € 200.000,00 | € 517.000,00 | |

*Table 6  - Example of the resulting calculations of the preventive maintenance costs*

### 3.4.3: Maintenance Advice

After the calculation of preventive and corrective maintenance the only remaining step to determine the most adequate maintenance advice to a customer is to compare the abovementioned costs. The maintenance with lower calculated costs will be the one advised since the advice is based on monetary saving potential.

In the used example the advice would be preventive maintenance since it has a cost of €517.000,00 for the next seven years while corrective maintenance has a cost of €1.254.400,00, a substantial difference of €737.400,00.

## 3.5: Summary

The methodology explained in this chapter makes it possible to go from raw SAP data to a specific customer maintenance advice based on a Weibull analysis, failure forecast and maintenance costs comparison. An overview of the method is exposed in the following Figure 18, showing its methods, inputs and outputs.
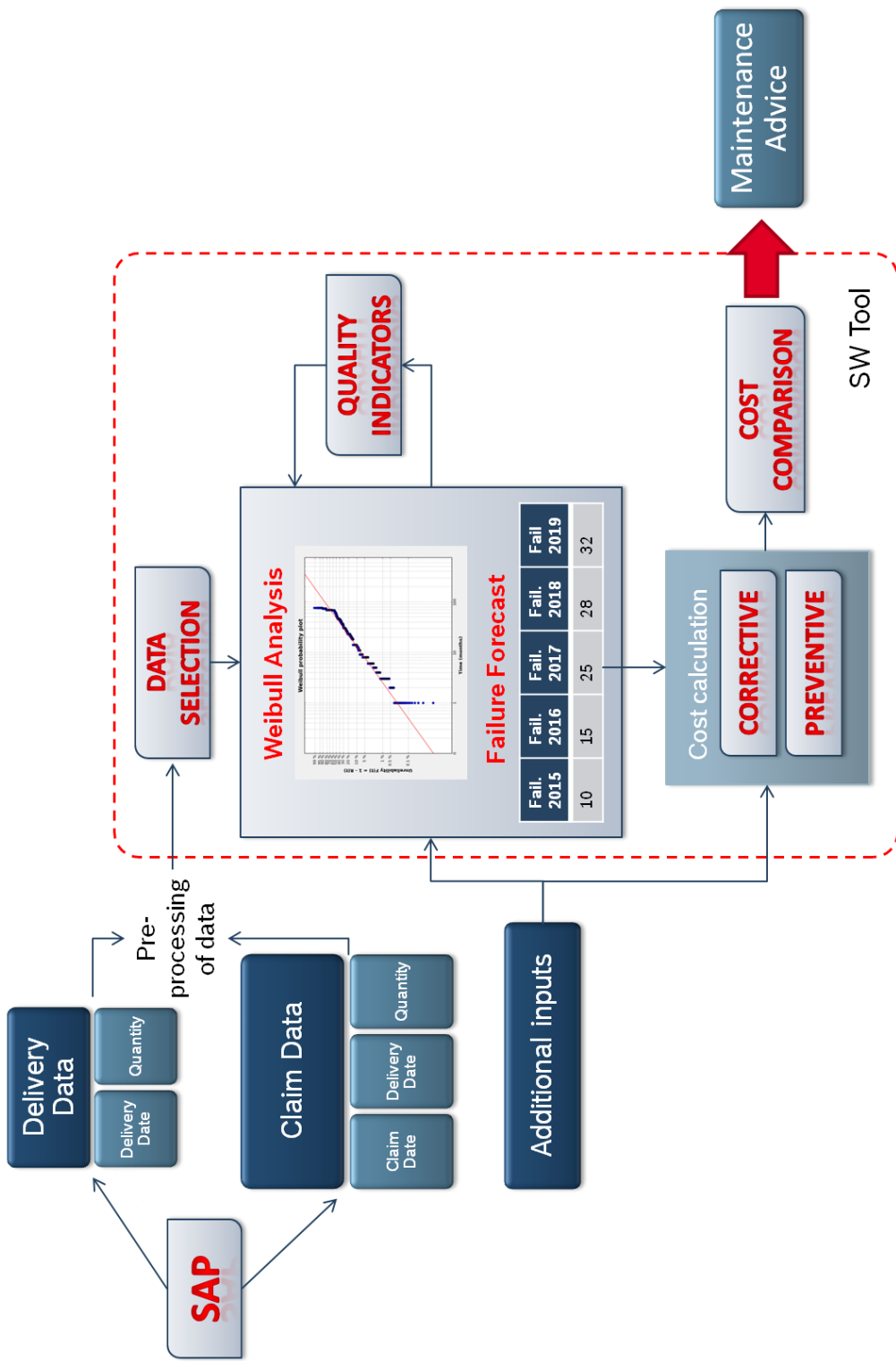
Figure 18 - Method overview

# Chapter 4: Implementation

The final assignment of the project was to implement the developed methodology of Chapter 3 in a software tool to offer more flexibility, agility and require less expert knowledge in the lifetime data evaluation. A software tool can also provide a better visualization of the advised maintenance strategy advantages and facilitate the dialog with the customer.

In this chapter an overview of the software solution implementation will be given beginning with the reasons for the use of Python as programming language. Then, in the subsections that follow, the tool is described based on its functional modules, also the graphic user interface for each one of the functionalities is presented.

## 4.1: Selection of programming language for tool implementation

The development of tools such as the one issued in this project is usually done using Excel Macros and Visual Basic (VBA) functions for two main reasons. The first one is that Excel is available to all employees of Bosch, making it easier to share the tools with other departments of the company. Also Excel Macros and the VBA language are reasonably easy to understand and implement even for persons with barely any programming knowledge.

Nevertheless, recently Bosch Rexroth faced complications with some of their Excel-based tools. Each time a new version of Office is released, considerably effort is necessary to tailor the functionalities of the tools to the changes implemented.

Thereby, Python was chosen for being an open source language free of charge and independent of new releases. Despite being an open source language, many of the packages and libraries developed for Python are under specific licenses (e.g. GLP licenses) which require that whoever uses them must open the source code of the final program. It is not of Rexroth interest to open the source code of the final program therefore, it was necessary to ensure that all packages and libraries used were under adequate licenses. In order to do so a list of useful packages and

libraries was created and its licenses where verified. More than one option of packages and libraries where searched for each desired functionality to guarantee that at least one option is available under adequate licenses. This was done with support of the Software Licenses department of Bosch Rexroth that explained and made clear what types of licenses could be used in the project. No critical licenses were found among the packages and libraries required for the project as long as the final tool is not sold. As the intention of DC is to use the software only internally and not sell it, there are no license problems. The final list can be found in Appendix A: List of packages and libraries licenses and is an important resource for further development of the tool.

## 4.2: Implementation of functionalities

To simplify the implementation, the developed methodology was structured in five different modules, each one comprising a major functionality: load and selection of data, Weibull analysis, quality indicators, failure forecast and costs calculation. The tool implementation was divided in four phases. First, the functionality of each module was individually implemented and tested. Next the modules were integrated with each other, in this phase the complete functionality of the tool was tested and verified. Then, the user interface was developed and test functionalities were attributed to the widgets and buttons to see if the connections were working properly. Finally the user interface was integrated to the functionalities closing the development cycle.

In the following subsections an overview of each functionality module and its implementation is given such as the use of external libraries and the graphical user interface developed.

### 4.2.1: Basic graphical user interface

The implementation of the graphic user interface in the tool was made using the built-in GUI interface for Python, referred as the Tkinter module. Compared to other Python modules for GUI programming Tkinter is simpler to learn, effective, meaning that short blocks of code can produce powerful GUI applications, and easily accessible as it comes pre-installed with standard Python distributions [15].

When the program is started the GUI displayed is the one of Figure 19. The three main groups of widgets can be seen in this interface.



*Figure 19 – Initial GUI*

The first is the menu bar, which contains the tool options and functions. The implemented functions are seen enabled in Figure 20.
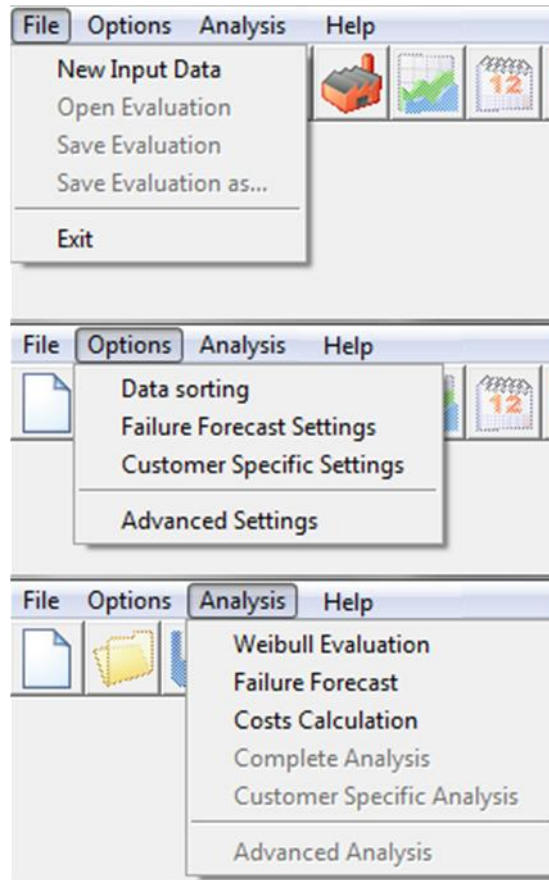


*Figure 20 – Menu bar*

Another group of widgets is the shortcut toolbar, composed by buttons that provide fast access to the tools main functionalities as showed in Figure 21:
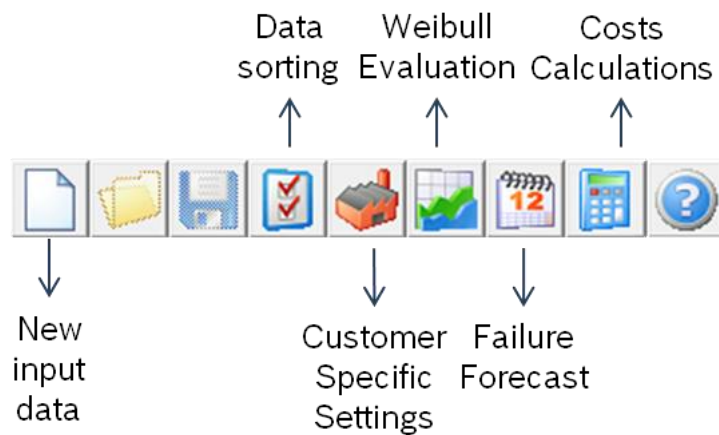


*Figure 21 – Shortcut toolbar*

The last widget group corresponds initially to a set of blank frames, signalized on Figure 22. Each one of them will display different information about the evaluation, its contents are described in the next subsections.



*Figure 22 – Information frames*

In the tool initialization not all menu options and shortcut buttons are enabled. Some functions can only be raised after the completion of previous steps. The first one is the load and selection of data explained in the next subsection.

### 4.2.2: Load and selection of data module

The main goal of this functionality is to read the product data of the delivery and claim Excel files and give the user the possibility to choose the required data to analyze. A secondary objective is to verify if the chosen data is in the right format and adequate to perform the next step, a Weibull analysis. If not, the user is warned with an error message. By default all errors are written in log files so the user can identify the problems in the data.

From the main screen of the tool, to select the data files the user can choose the option "New input data" in the file menu or press the first button of the shortcut toolbar. Then, the file selection widget seen in Figure 23 is displayed and the user can intuitively choose the desired files.
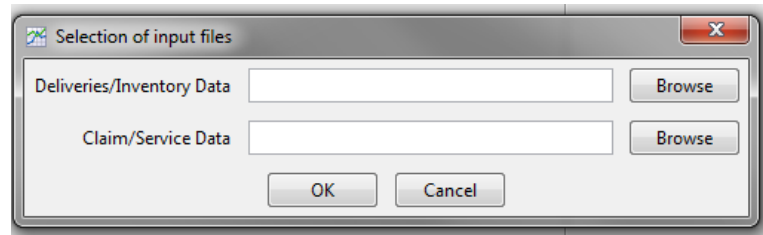
*Figure 23 – File selection widget*

After the completion of this step the information of the files are loaded with help of the external library openpyxl. This library provides functions to read and write Excel files in the .xlsx format. If no problems occur during the reading of the files a second widget is presented to the user, showed in Figure 24. Within this widget the user must identify the useful headers of the data so that the tool can recognize where the same information is located in both files. The first two headers of the delivery file and the first three headers of the claim file are the ones containing information about the data used in the Weibull analysis and must, therefore, be fulfilled. The other headers are optional and are used for selecting the data (e.g. the header Final Customer can be specified if the analysis desired is restricted to one customer).
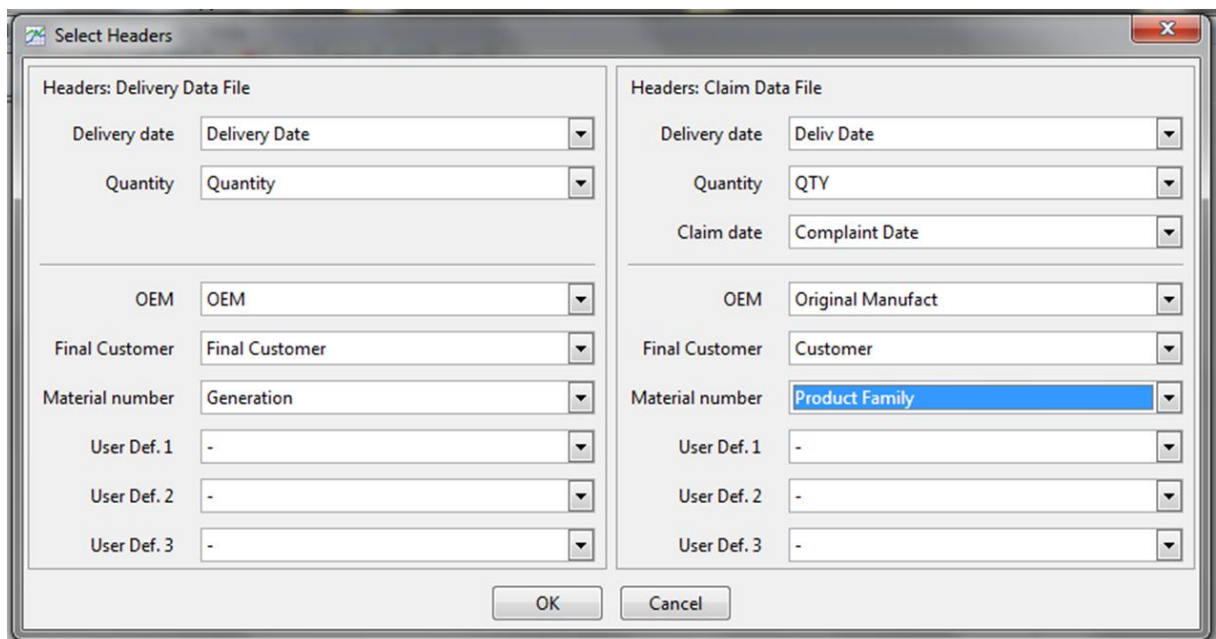


*Figure 24 – Headers selection widget*

After the specification of mandatory and desired optional headers the tool verifies if the assigned delivery dates, quantities and claim date are adequate for being analyzed. That means all dates must be in a valid date format, the quantities

60

must be valid integers greater than zero and all complaint dates in the claim files must have occurred after the correspondent delivery date of the claim file. If any of these requirements are not fulfilled, an error widget is presented to the user and all errors are saved in external files. If desired the error log containing all errors messages can also be displayed to the user.

If no problems in the data were found a third widget is displayed so that the user can select the data to be analyzed according to the headers that were selected, as seen in Figure 25. Restricting the data according to different categories (e.g. one customer or one product generation) can result in a more accurate result since the population analyzed tends to be more homogeneous.
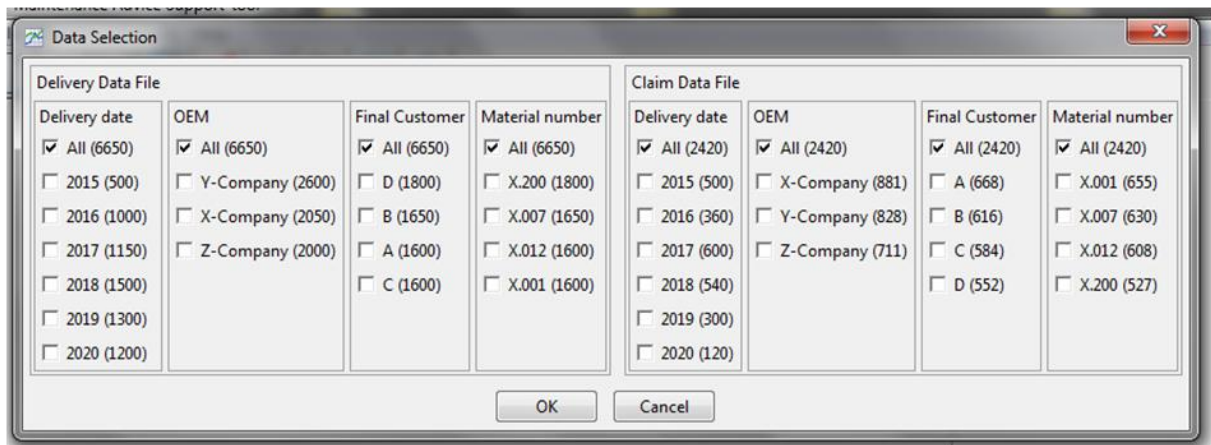


*Figure 25 – Data selection widget*

Once the user chooses the values to be analyzed the final data set is created, the load and selection functionality is completed and the Weibull analysis functionality is enabled.

### 4.2.3: Weibull analysis module

Basis for all the subsequent functionalities, the main goal of the Weibull analysis is to determine the parameters of the distribution *(beta, eta and gamma)* that describes the data set. Secondary goals are the calculation of the correlation coefficient, MTTF and B10 of the analyzed products and also the graphic display of the Weibull plot, parameters and other related information. To simplify the calculations two external packages were used, Numpy and the Scipy. To display the Weibull plot the external graphic library Matplotlib was used.

To perform the Weibull analysis the user must, after loading and selecting the desired data, choose the analysis menu option Weibull Evaluation or click the Weibull shortcut button.

Before deciding how this module would be implemented several Weibull analysis where manually performed. The goal was to find a systematic way of performing Weibull analysis. Parameters such as evaluation level (which analysis are more accurate, the final customers analysis, the OEM analysis or a industry sector analysis) and parameter estimation method (RRY, RRX or MLE) where tested.

Due to its versatility the implemented distribution was the 2-parametric Weibull and the parameter estimation method used was the rank regression from Y to X already explained in Chapter 2. Before fitting the data to a distribution all delivery dates in the claim data file must be verified to confirm if they all have a match in the delivery data file. Claims with no match are not used in the analysis, the user is warned and again an error log file is created.

Next, the evaluation period final date and the units' ages are calculated considering parameters and the delays described in section 3.2.6:. Then the units' ages are ranked according to the median rank method and the unreliability of each survival and failure unit is calculated. Here the Scipy library is used to perform a linear regression and the Weibull parameters beta and eta are calculated. Scipy also generates the curve correlation coefficient to the data points. The MTTF and B10 values are calculated in months based on the formulas of Chapter 2.

After recording *beta, eta* and the data points it is now possible to generate a visual display of the Weibull distribution using the Matplotlib functions. The generated

Weibull plot is shown in frame A, an example of the plot can be seen in Figure 26. Then all the information regarding the analysis is shown in frames B and C of the same figure and the functionality tasks are completed.
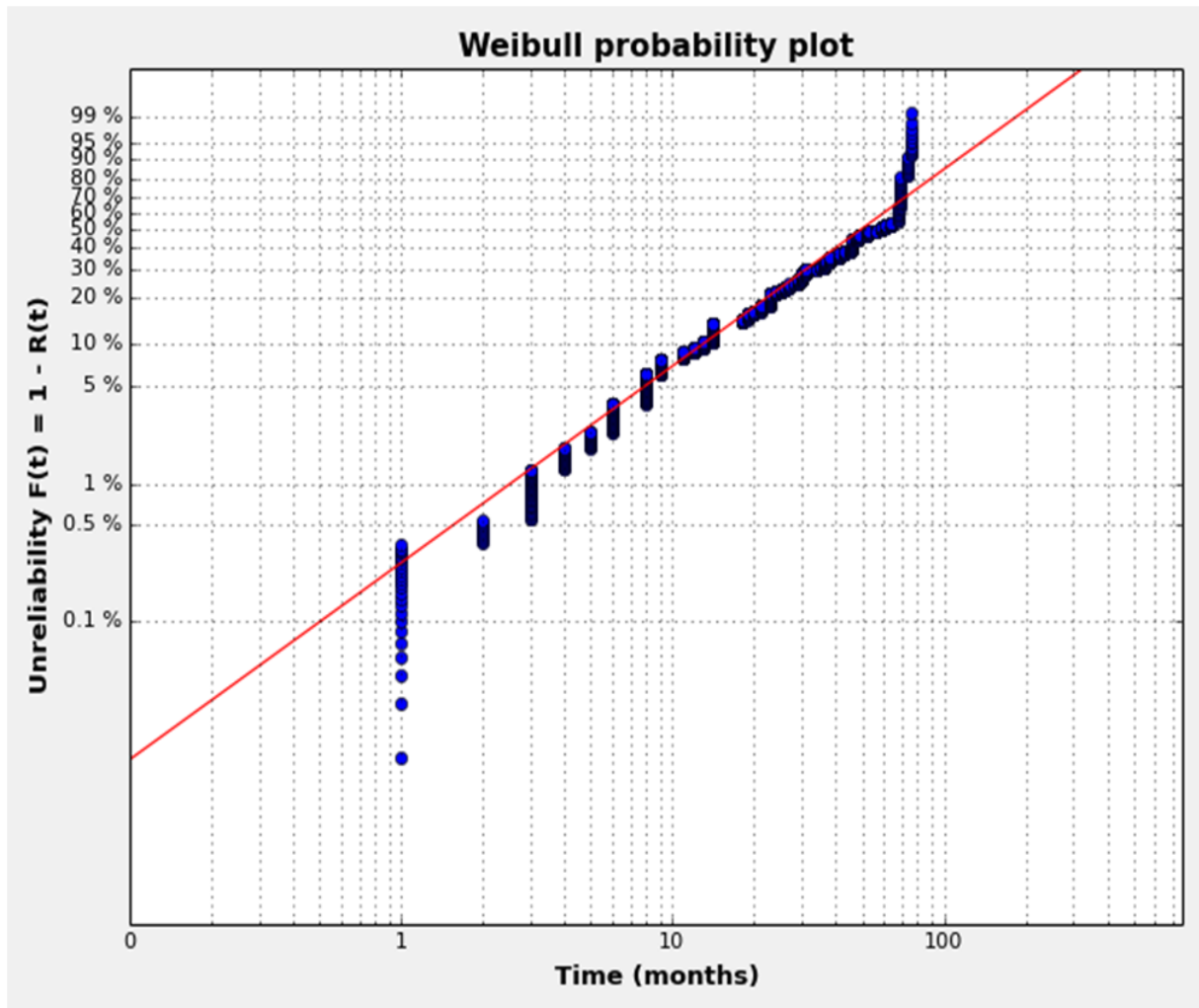


*Figure 26 – Graphical visualization of the Weibull analysis*

### 4.2.4: Quality indicators module

In the scope of this project four quality indicators (QI) were implemented, the number of failed units, goodness of fit, beta value and curved Weibull. It is planned to implement the batch problems indicator in the future. As this QI requires the comparison of beta values obtained with the MLE and RRY parameter estimation methods, the MLE must be also implemented in the tool.

The evaluation of the quality indicators is done automatically after the Weibull analysis so that the user has immediately a feedback concerning its quality. The indicators values are displayed in Frame D in the format seen in Figure 27.



*Figure 27 – Graphical visualization of the Quality Indicators*

The qualities indicators Number of Failures, Correlation Coefficient and Beta value are implemented through the comparison of the corresponding values to the minimum and maximum values defined at Chapter 3.3.1:. For the evaluation of the Curved Weibull the curve obtained is divided in two. The slopes of each half are compared and according to the difference between them the value of the quality indicator is defined.

Each quality indicator receives a weight according to its relevance. The combined evaluation of the weighted individual indicators values (green, yellow or red) generate the global quality indicator, represented by the traffic light in Figure 27.

### 4.2.5: Failure forecast module

In this module the failure forecast based on the previous calculated *beta* and *eta* parameters is performed according to the defined forecast beginning date and length forecast. The results are displayed both as a graphic and a table in Frames F and G respectively. The graphics gives the user a visual idea of the failures behavior in the next years, while the table provides more precisely the calculated number of failures as seen in Figure 28. Here again, to display the forecast plot the external Matplotlb graphic library was used.
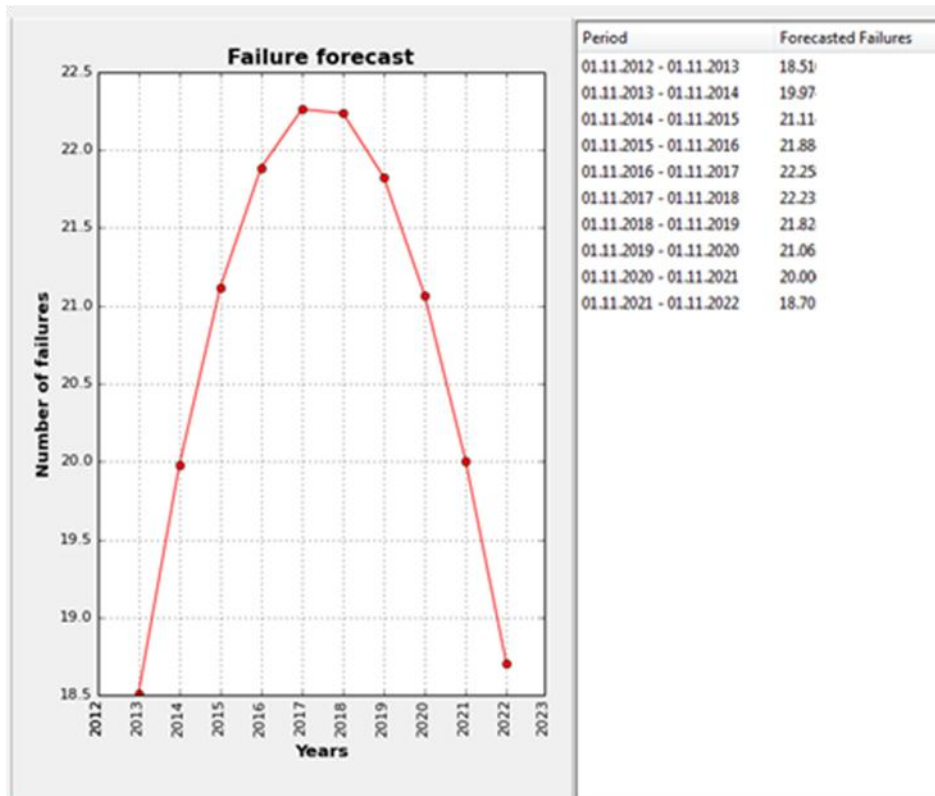
*Figure 28 – Failure forecast graphic and table*

To perform the failure forecast the user must, after the Weibull analysis, choose the analysis menu option Failure Forecast or click the Failure Forecast shortcut button. Once this is done the forecast settings widget shown in Figure 29 appears so that the user can define the forecast beginning date and the intended residual lifetime of units.
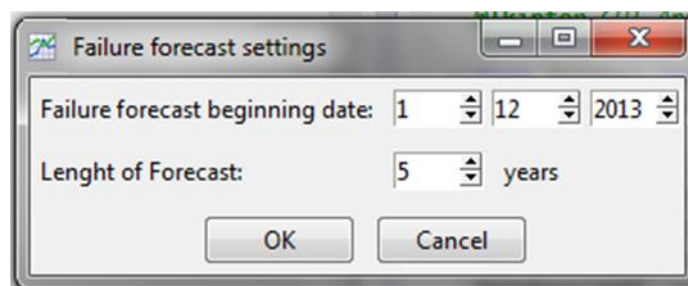


*Figure 29 – Failure forecast settings widget*

The proportional forecast mentioned in the methodology description is one of the future planned implementations.

## 4.2.6: Costs calculations module

The last module described here has as main goals the calculation of the corrective and preventive costs to evaluate which is the most adequate maintenance strategy for the analyzed customer. The resulting maintenance advice is displayed in Frame E in the format of a simple label. The costs and the comparison between them are shown as tables in Frame H as seen in Figure 30.

| Corrective Cc | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Failures | 1254.0 | 1137.0 | 975.0 | 803.0 | 641.0 | 498.0 | 378.0 | 281.0 | 205.0 | 147.0 | 6318.0 |
| Costs (Euros) | 5204100.0 | 4718550.0 | 4046250.0 | 3332450.0 | 2660150.0 | 2066700.0 | 1568700.0 | 1166150.0 | 850750.0 | 610050.0 | 26223850.0 |
| Preventive Cc | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
| Overhauls | 6638 | 587.17884760 | 373.34070498 | 215.05986718 | 162.75532756 | 133.68250809 | 114.63089164 | 100.97467764 | 90.607976184 | 82.416545322 | 8498.6473462 |
| Costs (Euros) | 10952700.0 | 2436050.0 | 1547950.0 | 892250.0 | 676450.0 | 556100.0 | 477250.0 | 419150.0 | 377650.0 | 340300.0 | 18675850.0 |
| Comparative | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
| Saving Potential | -5748600.0 | 2282500.0 | 2498300.0 | 2440200.0 | 1983700.0 | 1510600.0 | 1091450.0 | 747000.0 | 473100.0 | 269750.0 | 7548000.0 |

*Figure 30 – Costs tables*

To accomplish the costs calculation functionality, the user must, choose the analysis menu option Costs Calculation or click the Costs Calculation shortcut button. Once this is done the customer specific settings widget shown in Figure 31 appears. There the user can define the amount of downtime caused by a fault, the cost of an hour of downtime, the amount of overhauls performed in the last two years, the cost of a product repair and the cost of a product overhaul.
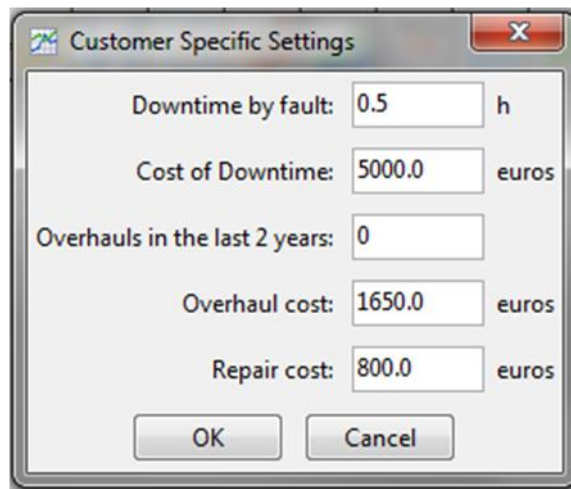


*Figure 31 – Customer specific settings widget*

# Chapter 5: Validation and Verification

According to the IEEE Standard Glossary of Software Engineering Terminology [16], verification is defined as:

"The process of evaluating a system or component to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase."

Validation, on the other hand, is defined as:

"The process of evaluating a system or component during or at the end of the development process to determine whether it satisfies specified requirements."

In this context verifications are done to demonstrate that the output of a phase is correct according to the phase's input, i.e. calculations are done properly. Validation is necessary to demonstrate that the system is operational, i.e. the right types of calculations are being used and the expected results are achieved.

This chapter describes the validation methods used to make sure that the processes were implemented correctly. Also the methodology and the tool requirements described in Chapter 1: are verified to see if the project has achieved its goals.

## 5.1: Methodology and tool verification

After its implementation the functionality modules of the methodology were tested to see if the expected results were obtained. Three functionalities were verified by comparing the tool results with the results obtained with other commercial software solutions, the selection functionality, the Weibull analysis and the Failure forecast. Some inputs were used to generate either corrective or preventive maintenance so that the costs calculation functionality could be verified through the maintenance advice generated.

The selection functionality was the first to be verified. The data selected in the tool was compared to the data resultant from Excel table filtering. The same filtering parameters were used and the same values were selected in both cases. As an

example customer A, OEM B-Company and product generation 2 were selected both in the tool and in Excel. The results obtained in the tool were the same of the ones obtained with Excel.

To verify the Weibull analysis and the Failure forecast the tool results were compared to the results of the commercial software Weibull++ from ReliaSoft. Weibull++ is a wide-ranging lifetime analysis software that can be used to estimate distributions parameters based on life data through a variety of methods including RRY. The commercial software is also able to calculate a failure forecast based on distributions parameters.

The comparisons of the 2-parametric Weibull analysis through the RRY method between Weibull++ and the developed tool have verified the tool implementation since the differences in the obtained results were negligible. The result of a validation tests for the Weibull analysis can be seen in Figure 32:



*Figure 32 - Weibull analysis validation test*
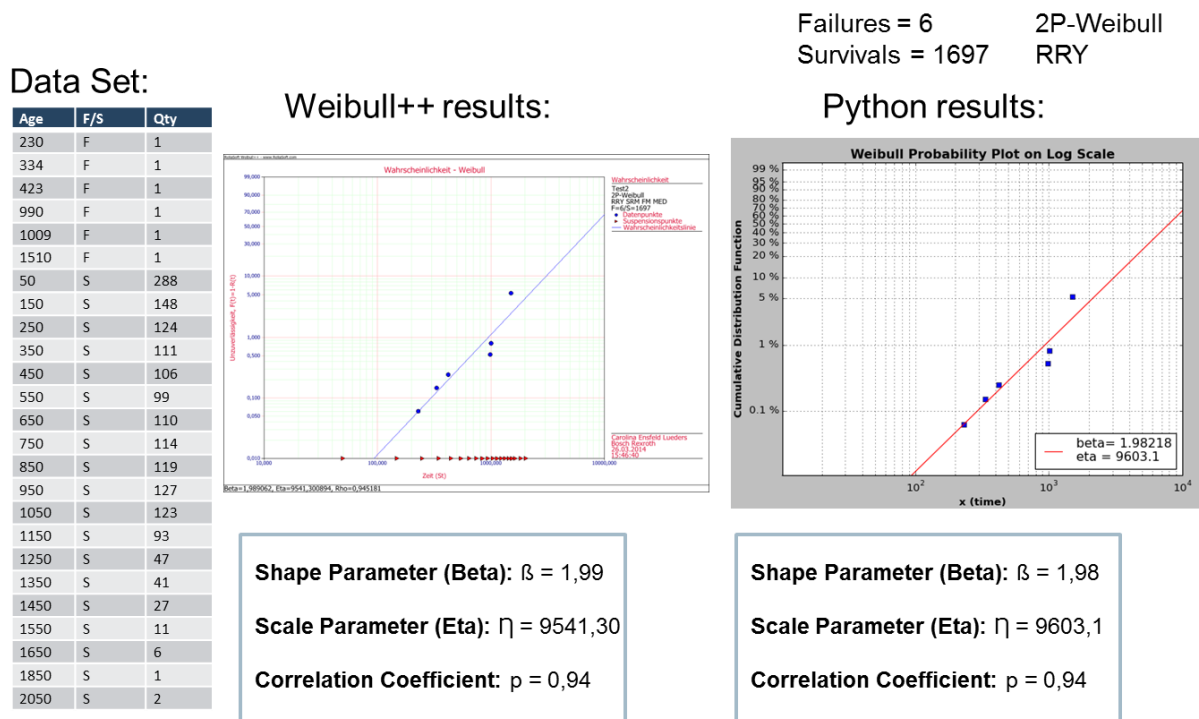
The same data was used to verify the Failure forecast module. A forecast for one year was done and once more the comparisons of the results from Weibull++ and the results from the tool demonstrate that the functionality was implemented correctly as seen in Figure 33:

Data Set:

| Age | F/S | Qty |
|---|---|---|
| 230 | F | 1 |
| 334 | F | 1 |
| 423 | F | 1 |
| 990 | F | 1 |
| 1009 | F | 1 |
| 1510 | F | 1 |
| 50 | S | 288 |
| 150 | S | 148 |
| 250 | S | 124 |
| 350 | S | 111 |
| 450 | S | 106 |
| 550 | S | 99 |
| 650 | S | 110 |
| 750 | S | 114 |
| 850 | S | 119 |
| 950 | S | 127 |
| 1050 | S | 123 |
| 1150 | S | 93 |
| 1250 | S | 47 |
| 1350 | S | 41 |
| 1450 | S | 27 |
| 1550 | S | 11 |
| 1650 | S | 6 |
| 1850 | S | 1 |
| 2050 | S | 2 |

Forecast for 1 year (300 hours):

Weibull++ results:

| Amount of survivors | Age | Forecasted number of failures for the next 300 hours: |
|---|---|---|
| 288 | 50 | 0,393160207 |
| 148 | 150 | 0,301805747 |
| 124 | 250 | 0,336049345 |
| 111 | 350 | 0,375015695 |
| 106 | 450 | 0,428769794 |
| 99 | 550 | 0,466268705 |
| 110 | 650 | 0,591037769 |
| 114 | 750 | 0,68798948 |
| 119 | 850 | 0,79678439 |
| 127 | 950 | 0,934106761 |
| 123 | 1050 | 0,985670064 |
| 93 | 1150 | 0,806397817 |
| 47 | 1250 | 0,438384029 |
| 41 | 1350 | 0,409292591 |
| 27 | 1450 | 0,287206008 |
| 11 | 1550 | 0,12419981 |
| 6 | 1650 | 0,071662025 |
| 1 | 1850 | 0,013246766 |
| 2 | 2050 | 0,029093433 |
| Total | | **8,476140435** |

Tool results:

| Amount of survivors | Age | Forecasted number of failures for the next 300 hours: |
|---|---|---|
| 288 | 50 | 0.39697954771 |
| 148 | 150 | 0.304005325546 |
| 124 | 250 | 0.33787189882 |
| 111 | 350 | 0.376492898889 |
| 106 | 450 | 0.429932719609 |
| 99 | 550 | 0.467046938752 |
| 110 | 650 | 0.591488581179 |
| 114 | 750 | 0.687963412766 |
| 119 | 850 | 0.796183176885 |
| 127 | 950 | 0.93279627924 |
| 123 | 1050 | 0.98370393667 |
| 93 | 1150 | 0.804350494132 |
| 47 | 1250 | 0.437050322471 |
| 41 | 1350 | 0.407855688512 |
| 27 | 1450 | 0.286071986542 |
| 11 | 1550 | 0.123658385637 |
| 6 | 1650 | 0.0713218974612 |
| 1 | 1850 | 0.0131744588391 |
| 2 | 2050 | 0.0289159218446 |
| Total | | **8.4768638715** |

Failures = 6    2P-Weibull
Survivals = 1697    RRY

**99,9%**

*Figure 33 - Failure forecast validation test*

Finally to verify the maintenance costs and advice obtained with the tool a specific customer data set and parameters were defined so that a preventive maintenance advice would be obtained with a high downtime costs. Using the same data and parameters a low downtime costs would imply in a corrective maintenance advice. In fact, the different advices were obtained only by changing the downtime costs.

## 5.2: Methodology and tool validation

To validate the developed methodology and its implementation in a software tool the accomplishment of the main proposed requirements of the project were divided in three stages analyzed as following:

First stage: A standard methodology to determinate the most adequate maintenance strategy for a customer was developed using a valid statistical method for the evaluation of field data (the Weibull analysis). The methodology was used by the Service department to analyze and advice five customers in the automotive and print industry. The feedback obtained was positive and the customers appreciated the obtained cost transparency for the remaining production time of their machines.

The use of reliability techniques to determine maintenance advices for specific applications has strengthened the customer trust.

Second stage: The implementation of the conceived methodology as a new software solution resulted in a fully functional tool in which the fundamental product analysis regarding reliability can be performed. Furthermore the tool has potential to eliminate the necessity of expert knowledge on the Weibull theory through internal quality insurance (the quality indicators). To fulfill this requirement the quality indicators should be further improved and verified. Summarizing, the flexibility supplied by the tool for the data selection and the definition of standard values for the evaluation parameters simplify the whole process of field failure analysis. The evaluation is done in fewer steps compared to the manual evaluation as well as less error-prone and the user only sees the relevant information.

Third stage: The methodology and the tool are a solid basis structure for further developments. To extend the tool usage to a wider range of customers, some additional procedures need to be implemented, such as the proportional forecast for small customers. Several other new functionalities have been envisioned both to enhance the software flexibility and to realize additional reliability analysis. An overview of these functionalities is done in the conclusion of this report.

# Chapter 6: Conclusion

During the accomplishment of this final course project the development of a methodology to evaluate customers' optimal maintenance strategy according to earnings-save potential was realized based on products lifetime analysis. The final result was the implementation of this methodology in a software tool developed with the open source programming language Python. The implemented tool eliminates the need of expert knowledge about Weibull theory, the base of the lifetime analysis. This is achieved by the systematic method used based on field failure data and statistic methods to obtain reliable results.

The methodology and the tool are part of an innovative business model recently implemented to intensify the DC relationship with customers and enhance their trust in the company. As the project stood out for its flexibility and pioneering use of field failure data it is running for the 2014 Bosch Innovation Award.

One of the major difficulties faced during the project was the deficient quality of the data available for the evaluations. Since the data is pre-processed before being used in the tool this problem does not affect the final users. This is done every half year to add new data registers and keep the tool updated. Improvements in the quality of field data are expected once this data is now used with extern purposes.

The presentation of the software tool took place on 10$^{th}$ July 2014 in the presence of the head of the Service department in Lohr. As a result an online trial version of the tool is to be developed and made available in the online portal of the company so that the customers can envision its functionalities an innovation potential. Another aim is to distribute the tool among the Service collaborators and train them to make use of the tool. In the long-term the goal is to distribute the tool among other departments of DC worldwide as part of a standard approach of reliability analysis.

Several additional reliability analyses have been envisioned as further modules of the tool. Techniques to analyze plants availability, plan stocking quantities, verify the effectiveness of product overhauls, determinate the optimal

maintenance interval and products total cost of ownership (TCO) are among the functionalities that can be aggregated as modules to the basis tool.

# Bibliography

S. Schlott, „Industrie 4.0 — Abschied von der Automatisierungspyramide,"
1] *Automobiltechnische Zeitschrift,* pp. 306-307, April 2013.

A. Jung, "Wettlauf der Systeme," *Der Spiegel,* pp. 52-59, 2014.
2]

Bosch Rexroth AG, „About Bosch Rexroth - Bosch Rexroth AG," [Online].
3] Available: http://www.boschrexroth.com/en/xc/company/about-bosch-
rexroth/about-bosch-rexroth. [Zugriff am 04 July 2014].

Bosch Rexroth AG, „Market Segments - Bosch Rexroth AG," [Online]. Available:
4] http://www.boschrexroth.com/en/xc/company/about-bosch-rexroth/market-
segments/market-segments. [Zugriff am 04 July 2014].

Bosch Rexroth AG, „Facts and Figures - Bosch Rexroth AG," [Online]. Available:
5] http://www.boschrexroth.com/en/xc/company/about-bosch-rexroth/facts-and-
figures/facts-and-figures. [Zugriff am 04 July 2014].

R. B. Abernethy, The New Weibull Handbook, Florida: Robert B. Abernethy,
6] 2000.

ReliaSoft Corporation, *Life Data Analysis Reference,* Tucson, 2014.
7]

B. Bertsche, Reliability in Automotive and Mechanical Engineering, Berlin:
8] Springer, 2008.

D. W. Boehner, *Bosch Product Engineering System Glossary,* 2010.
9]

Robert Bosch AG, Series Quality Management in the Bosch Group, No. 13
10] Methods od design for reliability, verification and validation, 2011.

Robert Bosch AG, Series Quality management in the Bosch Group, No. 6
11] Evaluation of Field Data, 2001.

R. Isermann, Fault-Diagnosis Systems: An introduction from Fault Detection to

12] Fault Tolerance, Berlin: Springer, 2006.

*European Standard DIN 13306: Maintenance terminology,* 2010.

13]

B. Chaudhary, Tkinter GUI Application Development, Birmingham: Pack

14] Publishing, 2013.

The Institute of Electrical and Electronics Engineers, Standard Glossary of

15] Software Engineering Terminology, New York, 2002.

# Appendix A: List of packages and libraries licenses

| Python Package/Library/others | relevance (1 high .. 3 low) | critical for use in Bosch | critical for external use | License |
|---|---|---|---|---|
| cx_Freeze | 1 | not critical | not critical | PSF |
| PyInstaller | 2 | not critical with no modification at PyInstaller | not critical with no modification at PyInstaller | GPL |
| py2exe | 1 | not critical | not critical | MIT |
| SciPy | 1 | not critical | not critical | BSD |
| ScientificPython | 2 | not critical | not critical | CeCILL |
| plotly | 2 | Critical, rights will be transferred | Critical, rights will be transferred | own |
| matplotlib | 1 | not critical | not critical | BSD/PSF |
| PyQtGraph | 2 | not critical | not critical | MIT |
| veusz | 2 | Critical (disclosure requirement) | Critical (disclosure requirement) | GPL |
| numpy | 1 | not critical | not critical | BSD |
| SymPy | 2 | not critical | not critical | new BSD |
| pythonxy | 2 | Critical, rights will be transferred | Critical, rights will be transferred | GPL |
| scikit-learn | 3 | not critical | not critical | BSD |
| pattern | 3 | not critical | not critical | BSD |
| SQLAlchemy | 3 | not critical | not critical | MIT |
| SQLObject | 3 | Critical (disclosure requirement for OpenSource SW) | Critical (disclosure requirement for OpenSource SW) | LGPL |
| PyGtk | 3 | Critical (disclosure requirement for OpenSource SW) | Critical (disclosure requirement for OpenSource SW) | LGPL |
| PyQt | 3 | Commercial license must be purchased | Commercial license must be purchased | different available |
| wxPython | 3 | Critical (disclosure requirement für Open Source SW) | Critical (disclosure requirement für Open Source SW) | own/LGPL |

| | | | | |
|---|---|---|---|---|
| openpyxl | 1 | not critical | not critical | MIT |
| xlrd | 3 | not critical | not critical | BSD |
| xlwr | 3 | not critical | not critical | BSD |
| pdfrw | 1 | not critical | not critical | MIT |
| PyPDFlite | 1 | not critical | not critical | MIT |
| Reportlab | 1 | not critical | not critical | BSD |
| PDFMiner | 1 | not critical | not critical | MIT |
| Python Imaging Library | 2 | not critical | not critical | own |

# Appendix B: Tool GUI

**Failure forecast**

| Period | Forecasted Failures |
|---|---|
| 01.11.2021 - 01.11.2022 | 1253.81056549 |
| 01.11.2022 - 01.11.2023 | 1136.8218331 |
| 01.11.2023 - 01.11.2024 | 974.807535448 |
| 01.11.2024 - 01.11.2025 | 803.165101195 |
| 01.11.2025 - 01.11.2026 | 641.005080863 |
| 01.11.2026 - 01.11.2027 | 498.030052409 |
| 01.11.2027 - 01.11.2028 | 377.978662071 |
| 01.11.2028 - 01.11.2029 | 280.916259217 |
| 01.11.2029 - 01.11.2030 | 204.837957472 |
| 01.11.2030 - 01.11.2031 | 146.764945661 |

| Corrective Co | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Failures | 1254.0 | 1137.0 | 975.0 | 803.0 | 641.0 | 498.0 | 378.0 | 281.0 | 205.0 | 147.0 | 6318.0 |
| Costs (Euros) | 5204100.0 | 4718550.0 | 4046250.0 | 3332450.0 | 2660150.0 | 2066700.0 | 1568700.0 | 1166150.0 | 850750.0 | 610050.0 | 26223850.0 |

| Preventive Co | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Overhauls | 6638 | 587.17884760 | 373.34070498 | 215.05986718 | 162.75532756 | 133.68250809 | 114.63089164 | 100.97467764 | 90.607976184 | 82.416545322 | 8498.6473462 |
| Costs (Euros) | 10952700.0 | 2436050.0 | 1547950.0 | 892250.0 | 676450.0 | 556100.0 | 477250.0 | 419150.0 | 377650.0 | 340300.0 | 18675850.0 |

| Comparative | 01.11.2022 | 01.11.2023 | 01.11.2024 | 01.11.2025 | 01.11.2026 | 01.11.2027 | 01.11.2028 | 01.11.2029 | 01.11.2030 | 01.11.2031 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Saving Potential | -5748600.0 | 2282500.0 | 2498300.0 | 2440200.0 | 1983700.0 | 1510600.0 | 1091450.0 | 747000.0 | 473100.0 | 269750.0 | 7548000.0 |