Master of Science Thesis

Application of Statistical Life Data Analysis for Cable Joints in MV Distribution Networks

-An Asset Management Approach-

By

Ravish P.Y. Mehairjan BSc

High-Voltage Technology and Asset Management

Delft University of Technology

TU Delft Delft University of Technology

Challenge the future

Application of Statistical Life Data Analysis for Cable Joints in MV Distribution Networks

-An Asset Management Approach-

By

Ravish P.Y. Mehairjan BSc

(Student #: 1530399)

Master of Science Thesis

High-Voltage Technology and Asset Management (HTM)

Faculty of Electrical Engineering, Mathematics and Computer Science (EWI)

Delft University of Technology (TU Delft)

This thesis has been carried out in cooperation with Stedin

Thesis Committee:

Prof. dr. Johan J. Smit

Dr. ir. Dhiradj Djairam

Dr.ir. Marjan Popov

Dr.ir. Arjan van Voorden

Ir. Herman Arts

R.P.Y. Mehairjan:
Application of Statistical Life Data Analysis for Cable Joints in MV Distribution Networks

For further information contact the author: r.mehairjan@gmail.com/rottra@hotmail.com

To my parents
my sister
and my brother

Summary

The greater pressure from both customers and regulators to maintain and enhance service reliability, while at the same time controlling costs, has caused many utility distribution businesses to adopt *Asset Management* (AM) as their framework to balance the financial aspects with the engineering and infrastructure aspects. Therefore, AM is widely being applied in asset intensive industries around the world. Generally, AM consists of data driven decision-making processes with the goal of deriving the most value from utility assets within the available budget. AM provides access to quantitative and qualitative data and allows decision makers to more readily identify and focus on key issues (risks). Asset intensive industries rely on asset data, information and asset knowledge as key enablers in undertaking both strategic AM activities and operational activities. Good asset information (timely, reliable and accurate data) enables better decisions to be made such as determining the optimal asset maintenance or renewal frequency for an asset. Consequently, in the past years utilities have progressively created databases to record asset or business data such as failure, maintenance, operation and cost. However, in many cases, the available data required to track equipment reliability are not sufficiently rich to provide a basis for straightforward decision-making processes. There are a number of reasons why data may not be sufficient. An important reason, among others, is because AM is a fairly new concept and many utilities did not have a reason to collect detailed information to track equipment lifetimes. The determination of equipment reliability required the collection and systematic evaluation of data on equipment failures.

In this context, there is a strong need for people, in practice, to have access to systematic techniques and guidelines on how to deal with information of equipment lifetimes. Recently, CIGRE has also felt the strong need to develop practical solution to deal with asset life data and, in this sense, has approved the creation of a new Working Group *WG D1.39*. The title of this group is *"Methods for Failure Data Collection & Analysis"*.

In this MSc thesis report, a systematic approach for analyzing asset life data (data describing equipment lifetime) in presence of incomplete data by means of statistical analytical methods is introduced. The analysis in this report mainly focus on a statistically based approach which uses data available from the past to predict short term reliability of a specific group of assets.

In chapter 1, a brief introduction of the AM approach is given and the related information aspects which are required to facilitate the asset managers decision-making processes. Subsequently, the importance of failure statistics within the AM framework is explained. Finally, the research description, research objectives & challenges and research approach & scope of this thesis report are explained. In this thesis, the analysis of equipment life data is carried out for three types of 10 kV medium voltage (MV) cable joint populations.

In chapter 2, the fundamental aspects and construction principles of typical Dutch distribution networks are described. It begins by describing the components of which such underground distribution networks are constructed. More specifically, the emphasis in this chapter is on cable systems consisting of three different types of components, which are cable parts, cable termination and cable joints. It continues by describing why defects occur in cable insulation systems with emphasis on MV cable joint failure causes. The chapter concludes by discussing examples of a number of typical cable joint failures and a historical failure data analyses for the 10 kV MV distribution network of the host company, Stedin.

In chapter 3, a theoretical overview of basic modelling concepts such as failure rate functions, probability distribution functions and statistics is discussed. The available Statistical Life Data Analysis (LDA) methods that can be used for analysing component reliability data are explained. Furthermore, this chapter describes available statistical tools, such as the Maximum Likelihood Estimation (MLE) technique, that can be used for analysing incomplete data sets.

In chapter 4, the application of statistical life data analysis for three types of 10 kV cable joint populations are discussed. First, the process of collecting and handling available life data for the cable joint populations is described. Secondly, the collected life data is used for estimating the probability distribution functions for each type of cable joint population. The MLE technique is used for parameter estimation in accordance with Goodness-of-Fit tests and engineering knowledge of the failure mechanisms. Based on these test the best fitted failure probability model is selected. Finally, the selected failure probability models for the three types of cable joints are used to compare the three populations of cable joints with each other. With these fitted probability models the available data from the past can be extrapolated to predict future failure expectations.

In chapter 5, a number of analytical tools which can be used to facilitate the asset manager to make sound decisions with regard to the failure behaviour of the cable joints are described. These analytical tools share a basic framework for decision-making and specify the evolution of the failures of the cable joint population over time.

In chapter 6, several conclusions of this thesis are presented and a number of recommendations are given.

Keywords

- Asset Management
- Underground Distribution Network
- Cable Joint
- Equipment Failure
- Statistics
- Life Data
- Reliability
- Ageing Assets

Samenvatting

Als gevolg van de toenemende druk op de energie- en nutsbedrijven, van zowel afnemers als toezichthouders, om een constante verbetering van de betrouwbaarheid, operationele presentaties en financiële presentaties te demonstreren is er een brede interesse in *Asset Management* (AM) ontstaan. Zodoende kan er een balansoefening tussen de financiële aspecten en de technische aspecten bewerkstelligd worden. Als gevolg hiervan wordt het concept van AM wereldwijd in toenemende mate toegepast in nutsbedrijven. Algemeen gesproken bestaat AM uit besluitvormingsprocessen die gebaseerd zijn op gegevens en feiten met als doel om assets zo effectief en efficiënt mogelijk in te zetten en de kosten binnen de perken te houden. Hierdoor hebben asset managers toegang tot kwalitatieve en kwantitatieve informatie waardoor zij belangrijke kwesties (risico's) makkelijker kunnen identificeren en zich op deze kwesties kunnen richten. Door gebruik te maken van analytische tools kunnen beslissingen worden gebaseerd op gegevens en feiten en niet alleen op subjectiviteit en intuïtie. Een belangrijk aspect voor effectieve AM is het beschikken over goede informatie (tijdige, betrouwbare en nauwkeurige informatie) teneinde deze informatie te voorzien aan de asset managers ter ondersteuning van de besluitvormingsprocessen. Als gevolg hiervan zijn er de afgelopen jaren veel meer databases met asset en bedrijfsspecifieke informatie zoals aantal storingen, onderhoudsbeurten, operationele aspecten en kosten tot stand gekomen. Echter kan het voorkomen dat de beschikbare informatie niet altijd toereikend is om de besluitvormingsprocessen te ondersteunen. Er zijn een aantal redenen waarom de beschikbare informatie niet altijd toereikend genoeg is. Een belangrijke reden is omdat het concept van AM relatief nieuw is voor energie- en nutsbedrijven. De energie- en nutsbedrijven hadden in het verleden meestal geen reden om over te gaan tot het bijhouden van gedetailleerde informatie over de levensduur van componenten.

In de praktijk blijkt er een sterke behoefte te zijn aan mensen die informatie over de levensduur van componenten kunnen beoordelen. In kader hiervan heeft CIGRE onlangs een nieuwe Working Group *WG D1.39* ingesteld. Deze nieuwe Working Group heeft als doel het ontwikkelen van praktische oplossingen voor het beoordelen van informatie over de levensduur van componenten. Deze Working Group is genaamd *"Methods for Failure Data Collection & Analysis"*.

Dit MSc thesis behandelt een systematische methode voor het analyseren van volledige en/of onvolledige beschikbare informatie, die de levensduur van componenten beschrijft, met behulp van statistische methoden.

In hoofdstuk 1 wordt een beknopt overzicht over het AM concept gegeven en laat zien welke informatieaspecten noodzakelijk zijn voor het ondersteunen van de besluitvormingsprocessen van de asset manager. Hierna wordt het belang van storingsstatistieken binnen het kader van AM uitgelegd. Uiteindelijk wordt het onderzoek zelf beschreven. Hierbij wordt ingegaan op de onderzoeksdoelen & uitdagingen en het plan van aanpak & de reikwijdte van het onderzoek. Dit onderzoek buigt zich over de analyse van de levensduur gegevens van drie type 10 kV middenspanning kabelmof populaties.

In hoofdstuk 2 worden de fundamentele aspecten en de constructie principes van typische Nederlandse ondergrondse elektriciteitsdistributie netten beschreven. Allereerst worden de componenten die gebruikelijk zijn in ondergrondse netten besproken. In het bijzonder ligt de nadruk in dit hoofdstuk op kabelsystemen die in drie categorieën kunnen worden onderverdeeld, namelijk kabels, kabel eindsluitingen en kabel moffen. Hierna worden de veelvoorkomende faaloorzaken van isolatiematerialen van kabelsystemen beschreven met de nadruk op middenspanning kabelmof defecttypen. Uiteindelijk sluit dit hoofdstuk af met enkele typische voorbeelden van kabelmof storingen gevolgd door een historische storingsdata analyse voor 10 kV middenspanning distributie netten van Stedin.

In hoofdstuk 3 wordt een theoretisch overzicht van basis concepten voor het modelleren van betrouwbaarheidsparameters zoals faalkansen, kans verdelingsfuncties, en statistieken gegeven. De beschikbare Statistische Levensduur Data Analyse (LDA) methoden die gebruikt kunnen worden voor het analyseren van component gerelateerde betrouwbaarheidsgegevens worden hier besproken. In dit hoofdstuk worden ook de beschikbare statistische technieken die gebruikt kunnen worden voor het analyseren van onvolledige data zoals Maximum Likelihood Estimaltion (MLE) aangehaald.

In hoofdstuk 4 wordt de beschikbare levensduur informatie van drie type 10 kV kabel mof populaties onderworpen aan statistische analyses. Allereerst wordt het proces voor het verzamelen en ordenen van beschikbare informatie van de kabelmoffen toegelicht. Vervolgens wordt de beschikbare data onderworpen aan statistische analyses, waarbij met behulp van statistische kansverdelingen de faalkansen voor elke kabelmof populatie wordt uitgerekend. De MLE methodiek wordt gebruikt voor het schatten van de parameters van de kansverdelingen. Hierbij wordt gebruik gemaakt van Goodness-of-Fit tetst en kennis van faalmechanismen voor het selecteren van een geschikte kansverdeling. De geselecteerde kansverdelingen worden uiteindelijk gebruikt om de drie verschillende kabelmof populaties te vergelijken met elkaar.

In hoofdstuk 5 worden enkele analytische modellen en tools beschreven die gebruikt kunnen worden door de asset manager om weloverwogen besluiten te nemen betreffende het storingsgedrag van de kabelmoffen. Deze analytische tools vormen een basis strategie voor besluitvorming and beschrijven de ontwikkeling van storingen van kabelmof populaties.

In hoofdstuk 6 wordt deze MSc thesis afgesloten door enkele conclusies te trekken en aanbevelingen te doen.

Acknowledgements

I would like to start my thesis report by owing gratitude to those who have contributed to it. Besides my own effort and time management this work would not have been possible without the guidance and assistance of the following people and organisations. Therefore I would like to use this opportunity to express my sincere gratitude to a number of people.

To begin with, I would like to thank Dr. ir. Dhiradj Djairam, who, as my daily supervisor on behalf of TU Delft, guided me in an organized manner in my journey towards this Master of Science degree. Dhiradj guided and supported me from the initial to the final stages of this research and the thesis writing task. I have really enjoyed working with him.

For his enduring support and believe in me, I am thankful to Dr. ir. Arjan van Voorden, my supervisor on behalf of Stedin. It was a pleasure working with Arjan and in particular I am grateful for the opportunity and freedom he gave me to work on this MSc project. I am also very thankful for the regular meetings we had every week and for the valuable inputs I acquired from these meetings.

Two other people of Stedin also had a special role in my supervisory team whom I should thank. Ir. Herman Arts for discussing the practical issues of this work with me and making time free for our weekly meetings, despite his very busy schedule. Hans Elenbaas for sharing his room for 9 months with me and treating me like a son. Hans was always there for a discussion and has always brought me in contact with the right people when I needed more information or assistance.

I deeply express my sincere thanks to, Prof. dr. Johan J. Smit and Dr. ir. Peter F.H. Morshuis, for their contributions and discussions on behalf of TU Delft during this project. Especially, I would like to thank them for giving me the opportunity to participate and present part of this research work during the Cigre Session 2010 in Paris. It was an honour for me to experience the professional atmosphere at Cigre in Paris. My presentation in the new Working Group (WG D1.39) was in some way a test for me and my thinking has certainly improved from this experience.

A special word of thanks goes to Lukasz Chmura MSc, PHD student at TU Delft at the High Voltage Department, for guiding me through the basics of statistics and assisting me with the Reliasoft Weibull ++ Software.

I would like thank Stedin, as a whole, and, particularly, the departments Asset Management and Network Coordination for supporting this research. Especially, the warm, open and friendly atmosphere I experienced during these 9 months. There are a lot of names to mention here, but I would especially like to thank Jan Pellis and Marko Kruithof for the cooperation between the departments Asset Management and Network Coordination.

Performing a statistical study without having the appropriate input data would be meaningless and therefore I want to thank the following people of Stedin: Peter Buys, Ruud Kaufman, Rob Schoovers and Age Westerhoven for providing me with the data that was required for this research. At some moment in time it would seem very hard to acquire the necessary data, however somehow they managed to provide me with essential information.

I would also like to owe my nearest thanks to some industrial specialist, Dr. Richard Brown (Quanta Technology), Dr. ir. Viren Ajodhia (Kema) and Rogier Jongen MSc (Seitz-Instruments). They discussed and answered many questions, which I used to pose via e-mail, from which I have learnt a lot and gained more insight in the topic of failure rate modelling and reliability thinking.

Furthermore, I gratefully acknowledge my friends and families in Suriname and the Netherlands whom have supported and encouraged me during my stay in Delft for last 2 years.

Finally, I wish to express my deepest and nearest gratitude to my father Samuel Mehairjan, my mother Chandra Kalpoe Mehairjan, sister Vandana Mehairjan, brother Ashis Mehairjan and girlfriend Evita Parabirsing for their motivation, support and everlasting love during my entire stay in Delft. Their support provided me with continuous inspiration to finish my studies in Delft and without their support my studies in Delft would never be successful...thank you very much!!

Ravish Preshant Yashraj Mehairjan

Delft, the Netherlands
November 2010

Table of Content

1 Introduction

Nowadays, utilities, especially power distribution utilities, face challenges of satisfying increasingly high standards for reliability and service quality while at the same time reducing costs and improving earnings. To meet the challenges, utilities are adopting *Asset Management* (AM) as their framework to balance the financial aspects with the engineering and infrastructure aspects. Generally, AM consists of data driven decision-making processes with the goal of deriving the most value from utility assets within the available budget. However, in many cases, the data that are available are not sufficiently rich to provide a basis for straightforward decision-making processes. There are a number of reasons why data may not be sufficient. One important reason, among others, is because AM is a fairly new concept, and many utilities did not have a reason to collect detailed information to track equipment lifetimes.

In this thesis, a systematic approach for analyzing asset life data in presence of incomplete or inconsistent data by means of statistical analytical methods is introduced.

This chapter starts, in section 1.1, with a general outline of AM and the importance of failure statistics within this context. Subsequently, in sections 1.2, 1.3 and 1.4, respectively, the research description, research objectives & challenges and research approach & scope of this thesis project are explained. Section 1.5 gives an outline of the chapters for the remainder of the MSc thesis report.

1.1 Background

1.1.1 Asset Management

The term AM is frequently encountered in utilities around the world. AM is a business approach that balances the financial aspects of a utility with the engineering and infrastructure aspects. AM of infrastructures (e.g. power utilities) involves making data-driven infrastructure investment decisions so that the life cycle cost are minimized while satisfying performance, risk tolerance, budget and other operational requirements [1]. Stated simply, AM is a corporate strategy that seeks to balance performance, cost and risk for the infrastructure as a whole. Achieving this balance requires the alignment of corporate goals, management decisions, and engineering decisions. It also requires a corporate culture, business processes, and information system capable of making thorough and consistent spending decisions based on asset-level data. Nowadays, AM heavily relies on the use of information and data to facilitate the decision-making process. Resulting from this the decisions should be based on information from various sources. This involves a large amount of different information aspects such as failure statistics, network operation, condition assessment and reliability & availability. In figure 1, various aspects that typically are concerned with AM are given in more detail [2].

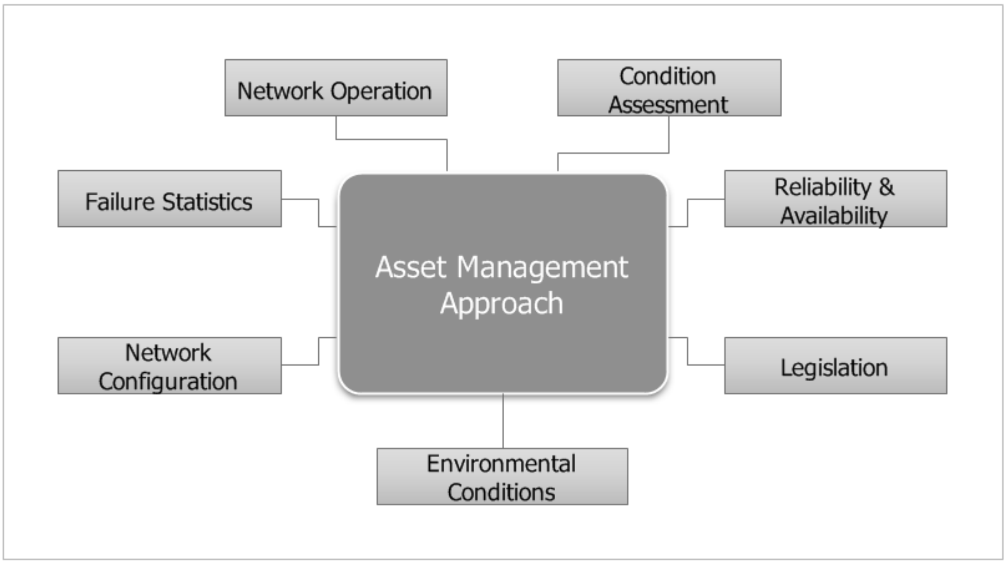


Figure 1.1: Different aspects for life assessment considerations for HV infrastructures **[2]**. *Asset Management is a centralized decision-making process, which uses information from various sources such as failure statistics, network operation, condition assessments etc.*

An important aspect shown in figure 1.1 is failure statistics. In order to incorporate this aspect into the decision-making process of the asset manager, it is required to consider the recorded life data (data describing components life) and perform analysis in order to determine the asset related failure behaviour.

Analysis on which decision can be based requires reasonable and acceptable data. The quality of the data is of fundamental importance for these analyses. However, in practice, it is difficult, expensive and sometimes impossible to collect all required life data. Nevertheless, with the AM heavily relying on asset level data to support sound AM decisions, there is a strong need for methods that are able to analyse life data even in the case of incomplete or inconsistent data. Therefore, developing a systematic approach in which optimal use is made of the available data for supporting the AM decision processes is required. Statistical life data analysis for estimating and predicting the probability of survival, the mean life, failure rate or, more generally, the life distribution of components can be a powerful tool for satisfying such needs. One way to capture the information in the data is to fit the parameters of a hypothesized failure distribution to the data. This process is known as parametric distribution fitting, and will be used in this thesis as part of the statistical analysis.

1.1.2 Importance of Failure Statistics

Every installed power system component will eventually fail. Therefore, at a certain point in time, every component will start exhibiting failure behaviour. For each type of component, this failure behaviour can be different. The purpose of applying maintenance is to prevent the occurrence of failures. However, if certain (e.g. critical) components show an increasing number of failures, it could mean that the maintenance or replacement program should be adjusted and be more focussed on these failing components. Therefore, detailed knowledge of failure probability, ageing and failure frequency can contribute to the AM decision process.

The frequency at which components fail, usually denoted as *failure rates*, is an important indicator for selection of an asset for condition assessment or replacement. These failure rates can be analyzed on different levels [3]:

- *Network system level:* Failure patterns on the network system level provide information about failures which are typical for specific areas or groups of components. An example is the comparison of failure rates of HV network level with MV network level. Failure patterns on system level can also give information about the relation between failures and component groups that are responsible for these failures such as transformers, cables or switchgears.
- *Asset system level:* If a specific asset system e.g. a certain cable network shows an increasing number of failures in a specific time span, then it is an indicator to perform assessments to this specific cable system.
- *Component level:* Failure rates on component level are more related to component types such as joints, termination, cable parts etc. Depending on the type of components, manufacturer or year of construction, the deterioration behaviour can be different. If a specific component type shows an increasing number of failures, the condition assessment can be focussed on this group of assets.

In the end, a combination of the above described selection levels will result in the best selection of component groups which are suitable for condition assessment. Therefore, it is important to understand what types of components are likely to fail, and how these failure rates will change over time.

1.2 Research Description

1.2.1 Description of the project

The department Asset Management at Stedin performs trend analysis of the annual outage hours. Based on this analysis and the information from specific failure analysis (size and/or duration), conducted by the department Network Coordination, necessary measures to control the annual outage duration are taken. The total outage duration is the product of 2 factors: average interruption duration and the failure frequency. Therefore, the chosen measures to control the outage duration focus on, both, average interruption duration (by means of organizational measures) as well as on failure frequency (by performing preventive maintenance or adjusting replacement programs). Currently, the average outage duration of the total system is primarily determined by failures in the MV distribution networks and a vast majority of these outages are due to failure in MV cables and especially in their joints.

MV cable and joint failures make up approximately 82% of the total number of component related failures. Also, the MV distribution network accounts for the largest part of assets in the total power network. In the Netherlands, almost 100% of the distribution of power is realized by means of underground medium voltage cables, cable joints and transformers. Most parts of this infrastructure were constructed more than 30 years ago and a preventive strategy, as was often applied in the past, would require replacement. Due to the regulation of the energy market, asset managers are forced to reduce costs, postpone investments, while maintaining or upgrading the reliability of power delivery. This, together with the increasing demand for electricity and customer awareness, leads asset managers to employ various maintenance and replacement strategies by using asset failure forecasting tools. Therefore, it is important to understand which types of components are likely to fail. Furthermore, it is also important to understand how these failure rates will change over time what can be known by determination of the failure probabilities.

Performing failure probability studies to analyse life data of suspect groups of components required the data to be collected throughout the entire life cycle of the components. Information or rather, the lack of information is the main challenge in performing statistical life data analysis. Database systems have undergone many changes and throughout the years many regional Distribution System Operators (DSO) have merged their asset inventory databases causing information discrepancies. Furthermore, distribution infrastructures were constructed many decades ago, and sometimes asset information is not available anymore and, most of the time, the exact age of equipment at the moment of failure is unknown.

This thesis deals with the application of appropriate statistical tools to analyse incomplete data sets for modelling component failure rates. This, in turn, will assist the asset manager in making failure forecasts and acquire more knowledge on the failure behaviour of component groups.

1.2.2 Purpose of the project

The purpose of the project is to develop an extension for this particular purpose of an already used systematic approach, in which all available data of a specific group of assets is used, for forecasting asset failure behaviour. In order to serve this purpose the use of statistical tools to determine the failure behaviour plays a key role. The most dominating component related failures are observed for MV cables and especially in cable joints. This thesis will mainly focus on the development of failure rate models for cable joints and its implication for asset management decision support. This, in turn, will facilitate the asset manager in the decision support regarding asset failure behaviour and possible maintenance or replacement policies for cable joints. Consequently, by creating more insight into the failure behaviour of these assets, more knowledge of the system reliability can be extracted.

1.3 Research Objectives & Challenges

1.3.1 Research Objectives

This thesis will focus on the statistical failure rate modelling of distribution systems on component level. The objective of the present study is to obtain more knowledge about the feasibility of statistical methodologies by using available asset life data to develop failure rate models for 10 kV cable joints. In summary the objectives are:

- Acquiring more knowledge with regard to the failure behaviour of 10 kV cable joint populations
- The development of failure rate models for 10 kV cable joint populations by using available cable joint life data in combination with parametric distribution fitting tools
- The prediction of future failures of 10 kV cable joint populations by using the developed failure rate models together with the service history
- The development of analytical tools, based on the statistical failure models, which can be used to acquire more knowledge with regard to the failure behaviour of the 10 kV cable joint populations

1.3.2 Scientific Challenge

AM for distribution networks are usually considered a different case when compared to AM for transmission networks. First of all, the number of installed equipment is much larger for distribution networks than for transmission networks. Consequently, monitoring and processing data of individual equipments and its sub-components would likely be very costly and time consuming. Due to the numerous types of different components installed in the network it would be impossible to implement condition assessment for the whole population at once. Structural analyses are required to select those components which could benefit from e.g. condition assessment. Since it is also expensive to replace a large number of the asset population, it is still important to understand the impact that age has on the failure rates of the asset groups. By using these statistical methods more knowledge about the failure behaviour, especially age related failure behaviour, can be extracted. Developing failure rate models (in presence of limited information) for specific groups of assets in distribution networks is necessary, as these form important indicators for selecting a network, component or sub-component for condition assessment.

1.4 Research Approach & Scope

1.4.1 Research Methodology

The thesis can be divided into two main parts namely a theoretical part and an analysis part. The theoretical part corresponds to a literature overview while the analysis part corresponds to a statistical parametric distribution fitting analysis section.

In the theoretical part, the underlying concepts for underground distribution networks in the Netherlands, impact of ageing, and statistical life data analysis are explored and discussed. In this part, literature review plays an important role. This review includes academic literature and consists mainly of publications, books and magazines on the engineering issues of component failure probabilities. In figure 1.2 a schematic overview of the theoretical part of the thesis is shown.

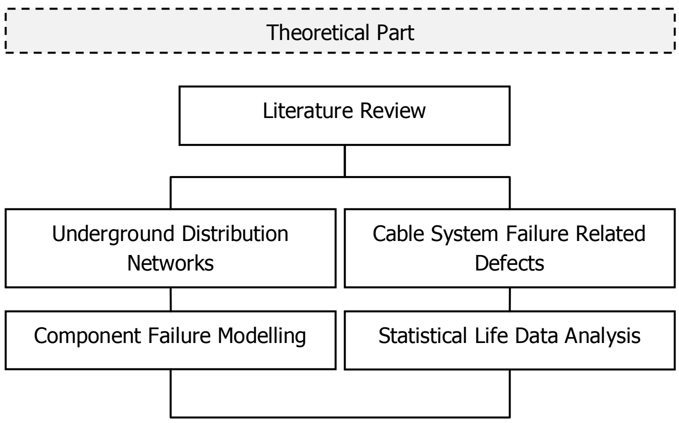


Figure 1.2: Schematic representation of the theoretical part of the thesis

The analysis part provides an application of statistical life data analysis for 10 kV cable joints. The first step in the application of statistical life data analysis is the collection of available cable joint data. The second step is to perform a parametric distribution fitting method for the 10 kV cable joint data. The statistical calculations and evaluations are performed in *Reliasoft's Weibull ++ 7* tool for life data analysis. The software is capable of performing life data analysis utilizing multiple lifetime distributions. Subsequently, after the parametric distribution fitting, an appropriate failure distribution is selected for further analysis. With the selected failure distribution and their related statistical parameters the failure probability characteristics of the cable joint populations can be determined. Finally, the developed failure models are used to facilitate the asset manager with tools in order to make sound decisions. In figure 1.3 a schematic overview of the analysis part of the thesis is shown.

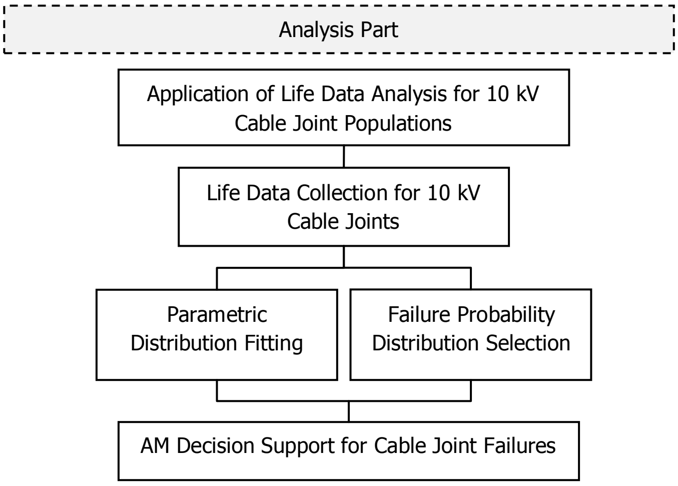


Figure 1.3: Schematic representation of the analysis part of the thesis

1.4.2 Scope of the Project

Before proceeding further, it is helpful to first define the scope of this research. The available data that will be used during this study comes from the recorded databases of Stedin. The failure rate models that are developed are based on component level failure statistics and deals with 10 kV distribution cable joints of three different type of insulation material namely:

- Oil insulated joints
- Mass insulated joints
- Synthetic insulated joints

The part of the system that is investigated is a 10 kV distribution network, a common applied distribution voltage level in the Netherlands, and this area is known as "Region X" in this report.

Figure 1.4: Geographical overview of the service area for electricity of Stedin

A geographical representation of the service area for electricity of Stedin is shown in figure 1.4. In combination with statistical analysis, the available data is used for failure rate modelling and predicting future failure rates. The available data, as provided by Stedin, will be used and, where necessary, data enhancement techniques and estimation are made based on expert knowledge.

1.5 Outline of the Thesis

The report is described in several chapters and is in line with the research approach highlighted in figure 1.2 and 1.3.

Chapter 2 performs a literature review of the construction of general Dutch distribution networks and its characteristics. The emphasis in this chapter is especially on underground distribution networks. Furthermore, this chapter highlights the typical defects that can occur in cable insulation and cable accessories. In this chapter special notice is given to cable joint failures.

Chapter 3 deals with the general background of component failure models and how they can be represented. In this chapter, the theoretical review of statistical failure distributions is described. Subsequently, the procedure for statistical LDA is described and, here, two methods are introduced which can be used for analysing continuous life data. These methods are the parametric distribution method and the non-parametric distribution method.

Chapter 4 considers the application of statistical LDA for three different types of 10 kV cable joint populations. In this chapter, the process of data collection and representation is described. The collected data is then prepared for the LDA analysis, which follows the parametric distribution fitting method. The MLE is used for parameter estimation of the fitted failure distribution.

Chapter 5 develops various tools to facilitate the asset manager in the decision-making process regarding the failure behaviour of the 10 kV cable joint populations. B(x)-lives, forecasting future failures and the failure count diagram are described in this chapter.

Finally, chapter 6 presents the conclusions of the thesis and makes recommendations.

2 Distribution Networks

In this chapter, an overview of MV distribution networks in the Netherlands is provided, to help understand the construction of such networks. This chapter starts, in section 2.1, with an explanation of the construction of typical Dutch distribution networks. In the Netherlands, MV distribution networks are almost 100 % underground constructed, and therefore, a large part of the investment cost of the distribution network is taken by the MV cable systems.

In section 2.2, factors which can result in a defect of the different insulation systems of the cables and cable accessories are described. Three categories of important stresses that can cause defects in insulation systems are explained in this section. This section proceeds in sub-section 2.2.3, with special emphasis on MV power cable joint failures and highlights the most significant reasons why these cable accessories are subjected to more failures. Furthermore, a brief overview of typical cable joints used in MV distribution networks is provided. In order to illustrate all of the above mentioned aspects, this section, presents a number of practical examples of real cable joint failures.

In section 2.3, we will discuss the recorded historical failure statistics for 10 kV MV distribution networks of Stedin. From the investigated failure patterns, it becomes clear that cable systems are associated to the major part of MV distribution network failures.

At the end of this chapter, in section 2.4, an overview of the main conclusions is provided. The knowledge gained in this chapter will then help to understand the basic principles of failure modelling, which is explained in chapter 3.

2.1 Background

2.1.1 Distribution Network in the Netherlands

In the early days of electrification, cable technology was just in development and overhead lines were used mainly. Today, power cables are available for a variety of voltage levels and power levels. From an investment viewpoint, overhead lines are usually still preferred. However, due to public awareness for aesthetics, environmental reasons and safety issues, but also technical matters, power cables can be a viable alternative. If operation and maintenance (O&M) costs are taken into account, power cables can also be competitive, especially when applied to MV distribution networks.

Nowadays, distribution networks are constructed more frequently using underground power cables, especially in areas with a dense infrastructure. In the Netherlands, MV and LV networks are constructed for almost 100% of underground power cables. In the rural areas, where soil conditions are too wet, overhead lines are used. The commonly used nominal voltage level for the distribution network in the Netherlands is 10 kV. However, other voltage levels also exist such as 25 kV, 20 kV, 12.5 kV, 6 kV and 3 kV. The total MV cables length in the Netherlands alone is approximately 100.000 km and a large part of the investment cost of the distribution network is taken by the MV cable network. Because the MV distribution network is almost completely constructed underground, the reliability and availability in the Netherlands are among the highest in Europe [4]. When the total power grid is considered, the distribution power grid is responsible for the major part of power-delivery outages. A vast majority of the distribution grid outage times is due to failures in MV cables and more specifically their joints.

Typically, the high failure rate is related to external causes of non-electrical nature, e.g. excavator digging. Besides this, insulation deterioration in cable systems, due to a wide variety of different defects [5], is causing the major part of the functional losses in the distribution network. Typical ageing stresses affecting cable insulation systems are thermal influences, mechanical influences, environmental influences and electrical influences. Because the power cable system forms a substantial part of the total MV distribution network system, and due to the high probability of functional losses, strategies are implemented to reduce the failure rate in the distribution cable network. Consequently, the continuity of the energy supply can be improved. To prevent external causes, such as dig-ins, from occurring, utilities should encourage the public agencies to have cable routes indentified before giving permits for initiating site excavation. Also, to prevent dominant failures due to internal causes from occurring, utilities should perform inspections of the insulation quality. However, as mentioned in chapter 1, due to the numerous types of different components installed in the network it would be impossible to implement condition assessment for the whole population simultaneously.

2.1.2 Underground Distribution Networks

In order to connect a cable to a protection fuse, a transformer or a circuit breaker, terminations are required. As result of limited maximum lengths of MV cables, or partial replacement, cable joints are used to interconnect these cable sections. As a consequence, distribution power cable networks are a building block of multiple links of several cable systems connected to the feeding substation. In general, each of these cable systems has an insulation system, consisting of three different types of components, which are cable parts, terminations and cable joints. In principle, a cable system is constructed with two terminations, N cable parts and $N-1$ cable joints. This is illustrated in figure 2.1.

The reliability of the cable system depends on the reliability of all individual components. When a failure occurs in a cable system the failed part can be repaired by replacing several meters of cable and two new joints (see figure 2.1.b). Due to these changes in the topology, a cable system generally consists of various types of joints, terminations and cable parts. Resulting from this, the age of the different components will vary and different kinds of defect will occur. This means that throughout the service years different deterioration mechanisms may be active for power cable systems and especially for the cable components [5], [2].

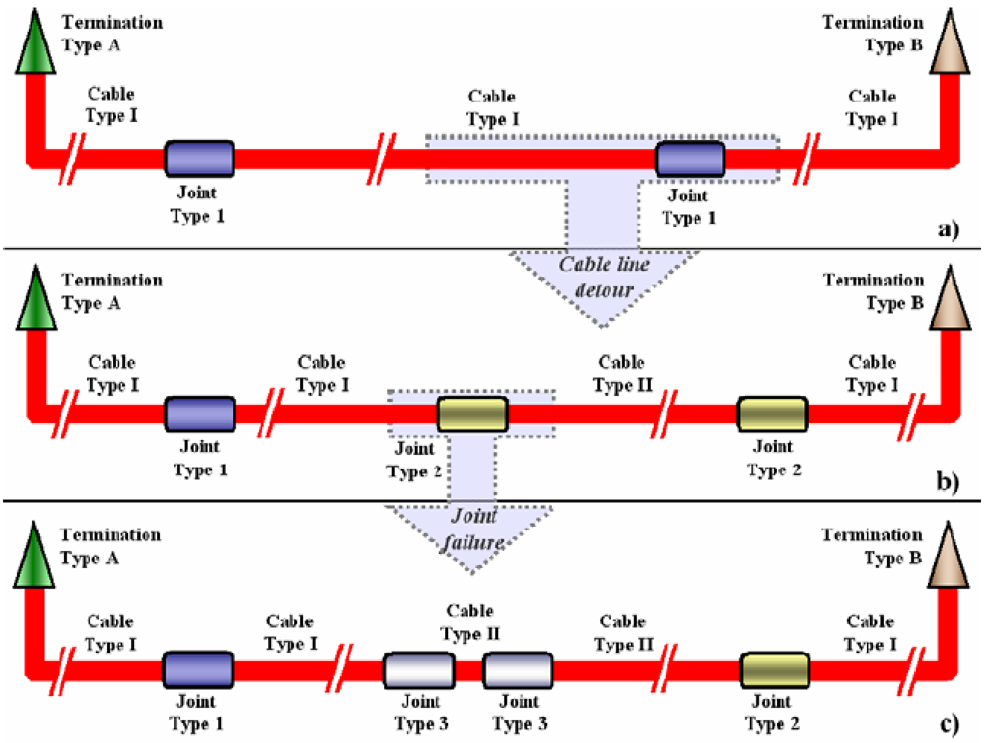


Figure 2.1: Construction of cable-section from different types of components, the age of components as well as the maintenance history may vary **[2]**.

The cable systems quality depends on the quality of its individual component(s). In the average service aged cable system, a mixture of old and new defect types occurs with different deterioration behaviour. This implies that throughout the service years of cable system the failure rates may vary according to many aspects.

2.2 Failure Related Defects

2.2.1 Factors Resulting in Defect

Besides external damages, many references show that more than 60 – 70% of the breakdowns in the MV cable network are caused by internal defects in the insulation system of the cable network [5], [6], [7]. There are many reasons why defects occur in cable insulation systems. In general, there are three categories of stresses that can cause internal defects. These categories are:

- Operational stresses
- Environmental stresses
- Human handling

In order to understand above mentioned stresses, a brief explanation of each category is given.

Operational Stresses – Load changes in cable systems cause temperature fluctuations and result in movement (transversal strengths) on the cable system. As a result of axial forces on the cable connectors, the connectors inside the cable accessories can move and cause increased electrical stresses inside. This may happen for instance in a cable joint. Also, load changes and short circuit currents in the network can cause mechanical stress on the cable connectors in joints. Due to these mechanical stresses, the connectors may lose their connecting strength, which will cause locally increased heat production into the insulation [5]. Moreover, daily load cycles can make the power cable expand and shrink thermo-mechanically. For example in paper oil insulation a fluid pressure is build up during high loads, which results in an expansion in the radial direction. When the load decreases, the cable shrinks starting at the conductor and an under-pressure occurs. The insulation fluids present in cable components will fill up the voids in the paper insulation and the oil level in the cable and cable accessories will decrease resulting in impaired insulation strength in these components. In worst cases, this may lead to implosions of cable joints or terminations [5].

Table 2.1: Important operational stresses, which induce defects in power cables and cable accessories **[5]**.

Stresses	**Induced factors**	**Induces defect**
Daily load cycles →	Thermal expansion →	• Increase of migration of materials
High temperatures →	Chemical reactions →	• Drying out of paper insulation • Depolymerisation of paper • Voids creation • Embrittlement of material • Gas formation

In most cases, the above mentioned operation effects cannot be prevented and the development of the network loads are not always known at the start of a cable system life-time. Appropriate diagnosis methods can be applied in order to gain more knowledge about these operational stresses. More information on the topic of appropriate diagnosis methods for MV cables can be found in [5].

Environmental Stresses – The condition or environment where cable systems are located has a large influence on the initiation of defects in the insulation. Such environmental stresses are, amongst other, moisture penetration (ground humidity), corrosion of cable covers (ground pollution and acidity content of the ground) and tensions on connections in accessories (mechanical stress). Usually, power cable systems are positioned under groundwater level and water could enter into the insulation (e.g. if the water blocking of the cable system is faulty). In wet grounds, which is common in the Netherlands, cable movement may impose mechanical stresses on a cable system, for instance on the cable joints. This, in turn, can again lead to breaking of the water blockings and impaired insulation breakdown strengths.

Table 2.2: Important environmental stresses which introduce defects in power cables and cable accessories **[5]**.

Stresses	Induced factors	Induces defect
Water/humidity →	Moisture penetration →	• Decrease of insulating qualities
Ground pollution →	Corrosion →	• Lead sheath penetration, water penetration
Mechanical stress →	Dielectric losses →	• Local overheating • Forces on cable accessories
	Tensions →	• Thermal ageing • Loss of adhesion • Cracking of materials

Human Handling – When replacements or new installations are made on-site, defect can be introduced by means of human induced errors. This is especially the case for cable accessories, which are most of the time assembled on-site. Improper assembly can cause defects. The installation of a cable part in an accurate manner depends on several external factors of which some are weather, time pressure, dedications, experience, tools and proper training. Even if the work is performed correctly, small defects can lead to breakdowns on the mid-long term [5]. Such human induced influences can be prevented by means of improving installation and repair instructions.

Table 2.3: Important human induces stresses which introduce defect in power cables and cable accessories **[5]**.

Stresses	Induced factors		Induces defect
Manuals/Instruction →	Critical construction	→	• Cavities
Mounting error →	Mixtures of insulation materials	→	• Bad hardened resin
Design →	Imperfect water blocking of accessories	→	• Moisture penetration (decrease of insulation properties)
	Introduction of PD related defects	→	• Erosion of insulation
	Damaging of sheaths during laying	→	• Lead sheath perforation, leakage of oil, water penetration
	Bad connection between conductors	→	• Local overheating

2.2.2 Overview of Insulation Defects

In general, visual inspections and forensic investigations of failed components provides insight into the different types of fault causes that result in breakdown. The most common sources for the initiation of breakdown in cables and their accessories are the presence of voids, cavities, sharp edges, metallic impurities and other defect in insulation. From many years of defect analysis, which include visual analysis and forensic studies, a number of repetitive fault causes can be listed for cable systems [5], [8], [9], [10]. Typical insulation defects for different type of power cables, joints and terminations are listed in the table below.

Table 2.4: Overview of typical insulation defects for different types of power cable components **[5]**, **[8]**, **[9]**.

Cable Type	Joint & Termination	Cable Insulation
Paper Insulated Lead Covered (PILC)	Lowered oil level	Corrosions of sheath
	Sharp edges on connectors	Perforations of sheath
	Moisture penetration	Cavities
	Air/Gas bubbles	Internal damages (bending)
	Bad hardened resin	Drying out
	Cavities/ Voids	
	Impurities	
Crosslinked polyethylene (XLPE)	Bad field grading (interface)	Moisture penetration
	Sharp edges on conductors	Cavities
	Moisture penetration	Damaged outer sheaths
	Remaining semi-conductor layer	Water treeing
	Air gaps	Electrochemical degradation (electrical treeing)

The typical insulation defects mentioned in table 2.4, are in general induced by either operational stresses, environmental stresses, human induces stresses or a combination of these three stresses.

2.2.3 MV Power Cable Joint Failures

This thesis work mainly focuses on the failure statistics of cable joints and therefore more attention will be paid to cable joint failures in this section. Although the cables are far more expensive than the related cable accessories, it is usually the accessories that affect the reliability of the cable system more often. The most significant reasons why cable accessories are subjected to more failures are because:

- they are subjected to higher electrical, mechanical and thermal stresses
- they are mounted in the field and most of the time under non-ideal circumstances, particularly during outage situations
- they are not subjected to expensive reliability testing procedures like the cable itself
- the quality of mounting the accessories is quite sensitive to workmanship, experience and care of the involved employee.

It can be concluded that especially cable joints form the weakest link in the total cable system. The complicated insulation interfaces in joints form a possible weak point in its construction. This is due to the insulation interfaces, which are at an angle to the electric field present in the cable system. As a consequence, when designing cable joints special precaution has to be taken in order to keep the electric field strength (tangential component) along the interface constant and sufficiently low [11]. Therefore, considering the above mentioned susceptibilities, the likelihood for a cable joint failure is high.

Type of MV cable joints

Various types of MV cable joints exist and are used by network operators. Amongst other, some typical types of cable joints are:

- straight joints (i.e. joints connecting the same type of cable)
- transition joints (i.e. joints connecting different type of cable)
- cross bonding joints (used in cable systems to minimize screen losses and to limit voltage rise).

Straight joints are mostly used when two cables with the same insulation type need to be jointed together at a certain point in the cable system. There are two main types of straight joints, namely filled joints and massive joints. Within the category of filled joints a further distinction can be made according to the type of filling used. An overview of straight cable joint types and the related joint filling material is given in figure 2.2:

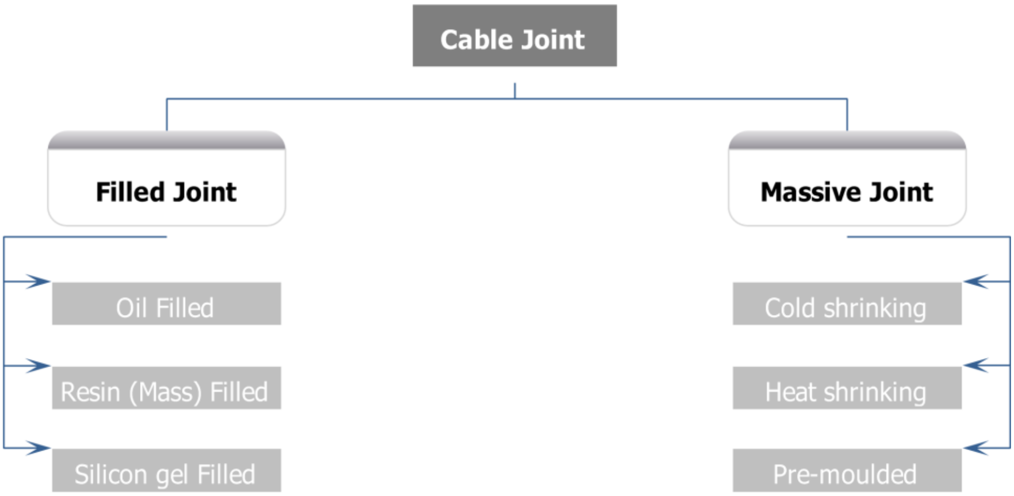


Figure 2.2: Schematic overview of straight MV cable joint types

In filled joints, some type of insulating materials, such as oil, silicon gel or resin/bitumen, is used to insulate the casing from the cable conductors. In massive joints, the insulation material being is used is solid from the beginning. Processing of the massive joint can be done through heat or cold shrinking or pre-moulded joints can be used. In the figures 2.3 and 2.4, two examples of MV cable joints are shown.

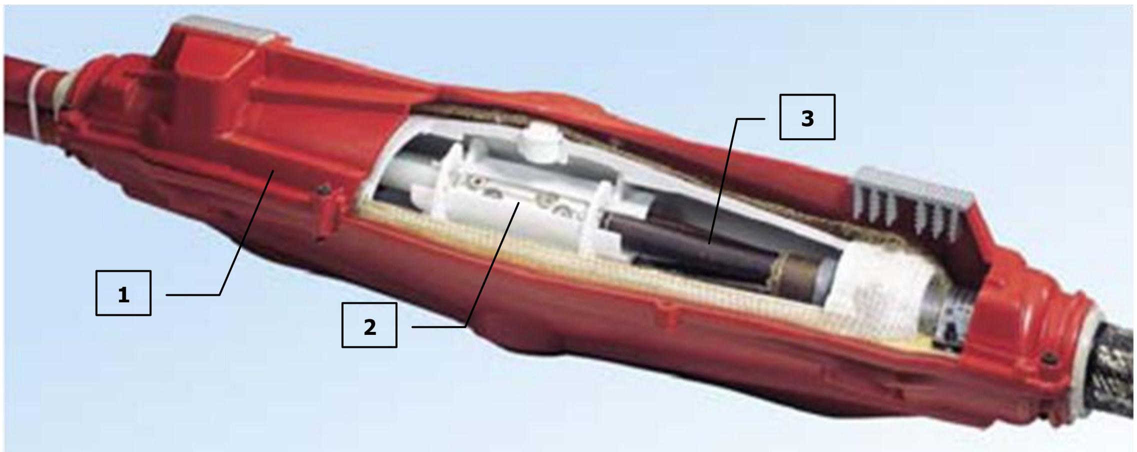


Figure 2.3: Example of a straight filled MV cable joint (source: Lovink Enertech).
1: Joint body shield
2: Connector
3: Cable insulation

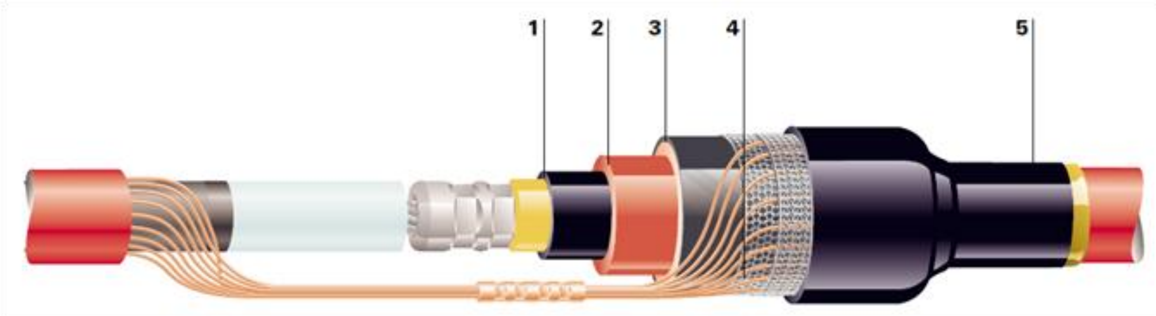


Figure 2.4: Example of a synthetic heat shrink cable joint (source: Tyco Electronics).
1: Electric field control
2: Insulation
3: Insulation Screen
4: Metallic shielding
5: Outer sealing

2.2.3.1 Examples of Cable Joint Failures

Different ageing mechanisms exist for both filled and massive joints. The most important ageing mechanisms for each of these joint types are discussed briefly here [12].

Filled Joints – As mentioned in the previous section, filled joints can be filled with a material that stays viscous (i.e. oil or silicon gel) or with a material that hardens (i.e. resin of bitumen). A dominant ageing mechanism which exists for joints filled with viscous material is a lowered liquid level. Most of the time this is caused by the connected MV cable which soaks up the liquid from the cable joint due to thermal heat cycles as a result of daily load cycles as mentioned earlier in section 2.1.1. The liquid also gets contaminated and both of these causes result in impaired breakdown strength. Both, lowered liquid level and contaminated liquid can lead to partial discharges that may ultimately lead to a breakdown. Asymmetric conductor positioning inside a joint will create areas with higher field strength and this can cause heating of the conductor. Inclusions of air present in joints filled with materials that harden form the most important ageing mechanism. Consequently, partial discharges can occur, which in turn can lead to a breakdown.

Massive Joints – Massive joints, either heat shrink, cold shrink or pre-moulded, are subjected to ageing mechanisms which are caused either by contamination (conductive parts) of the joint during the assembly in the factory or by voids between the conductors and the joint during the installation in the field. On-site installation of cold and hot shrink joints usually results in long flat cavities on the boundary between the joint and the cable insulation. These defects are caused by improper installation and inaccuracies, such as insufficient shrinking. In these defects (cavities) partial discharges can occur, causing degradation of the insulation material that ultimately can cause breakdown.

Throughout the literature, examples of typical MV cable joint failures can be found and will be used here in order to illustrate the above mentioned failure causes for these cable joints. These examples are according to [5], [8], [10] and more examples for cable terminations and cable failures can be found in these references.

Decreased Insulation Level

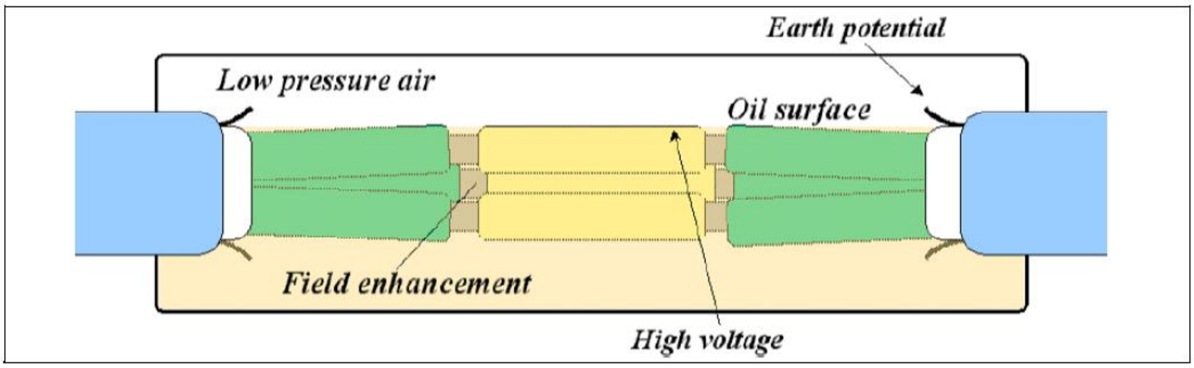


Figure 2.5: Liquid level decrease in a fluid filled cable joint. This occurs when the insulation fluid in fluid filled cable joints migrate to other places in the cable system **[5]**.

This example according to reference [5] illustrates the problem in a filled cable joint when the insulation fluid (permanently viscous material) from the joints migrates into the cable insulation due to load cycles. As a result, the liquid level in the joint is decreased and a gas under pressure situation in the upper part of the joint (see figure 2.5) is created. Consequently, the upper high voltage connector will be located at the edge between the low pressure gas and insulation liquid. A situation occurs where a low pressure gas is present between the connector and the wall of the joint. This may result in the following undesirable situations:

- The partial discharge inception voltage of the fluid surface decreases with the decreased gas pressure [11]
- At the edge of the connector a field concentration is present and discharges will take place over the fluid surface and result in floating carbonised oil parts
- A carbonised particle path may be formed between the connector and the earthed lead screen until the oil breakdown value is reached.

Displacement of conductors

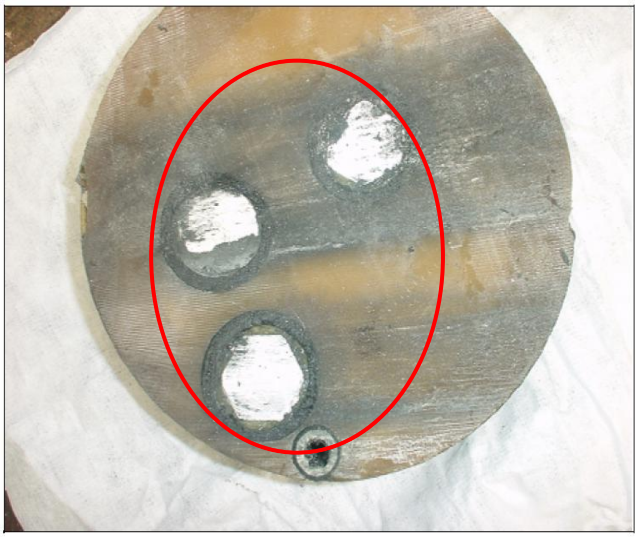


Figure 2.6: Conductor displacement inside a resin insulated joint **[5]**.

Another example according to reference [5] is the case when conductors are displaced during assembly of a joint. Because the conductors are asymmetrically placed, regions inside the joint are present where the electric field strength is increased. Due to this, local heating of the conductor may occur. Usually, it is common in distribution networks to place over dimensioned accessories, so that the effect of increased field is minimal. Still there is a risk present for breakdown when multiple defects would occur, for example, asymmetric conductor poisoning in combination with cavities or moisture penetration.

XLPE Joint Failure

Figure 2.7: A tree like erosion on the surface of polymeric insulation material in a cable joint **[5]**, **[13]**.

This example according to [5] and [13] gives an indication of the effect on breakdown by chemical changes. Chemical changes may occur in polymer/air interfaces due to moisture ingress and is accompanied by partial discharges which degrade the interface. Chemical changes can result in the formation of crystals on the surface and at these points field intensification will occur, increasing the discharge activities. Due to this increased breakdown strength a tree like erosion (see figure) will be noticeable on the insulation material surface. At the tree tip, the field strength is increased and new breakdowns over short distances of the insulation material occur. The tree will grow over the surface of the polymeric material, bridging the electrode. The formation of a tree may take months or years, however, once a tree is initiated breakdown may follow within hours or minutes.

Probable Moisture Ingress

Figure 2.8: Fluid filled joint from a certain circuit containing electrical tracking marks (left) and probable moisture ingress **[14]**.

This example according to reference [14] is a filled cable joint which revealed traces of electrical tracking as result of probable moisture ingress. The joint was replaced and examined after excessive partial discharge activities were spotted. The probability of failure by moisture penetration in the cable insulation depends on the regional circumstances. In the Netherlands the ground water level can be relative high. Defective water barriers in combination with lowered insulation fluid levels (the case in the first example) will result in easy moisture penetration in cable joints.

2.3 Historical Failure Statistics

2.3.1 Failure Statistics of 10 kV Distribution Networks

The service area for electricity of Stedin can be divided in five areas. The geographical representation of these areas is given in figure 2.9.

Figure 2.9: Geographical representation of the service area of electricity of Stedin

In this section, the historical failure statistics in the three largest areas of the 10 kV distribution networks of Stedin for the period 2004 until 2009 are described.

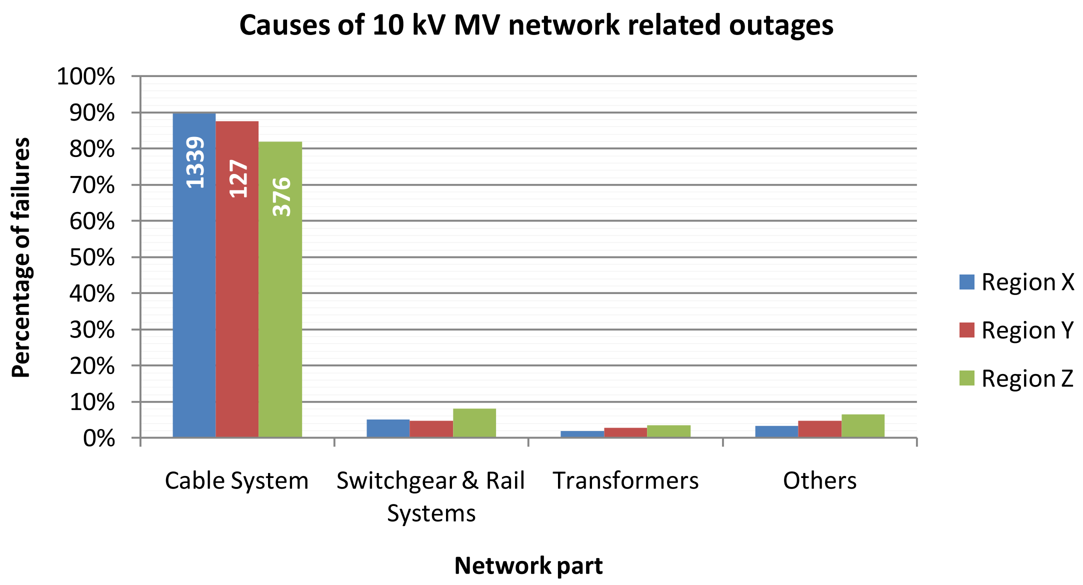


Figure 2.10: Failure analysis of three 10 kV network related outages

Outages find their origin in a number of components such as: cable systems, switchgears, transformers, rail systems and others. The relative number of outage causes related to these categories is shown in figure 2.10. When these failure patterns of the three largest service areas of Stedin are investigated and compared with each other, it becomes clear that cable systems are associated to 88% of the outages in the MV networks. Under the label "others" outage related to switching actions, protection malfunction, secondary installations and unknown reasons are reflected.

The number of cable system failures for the area "Region X" (1339 cable system related failures) is significantly higher than the other areas. It is interesting to investigate the historical failure statistics associated to cable system failures for the area "Region X", especially if the focus is on components level failure patterns. By pursuing a component level analysis of the historical failure statistics, different failure behaviour of particular component groups can be found.

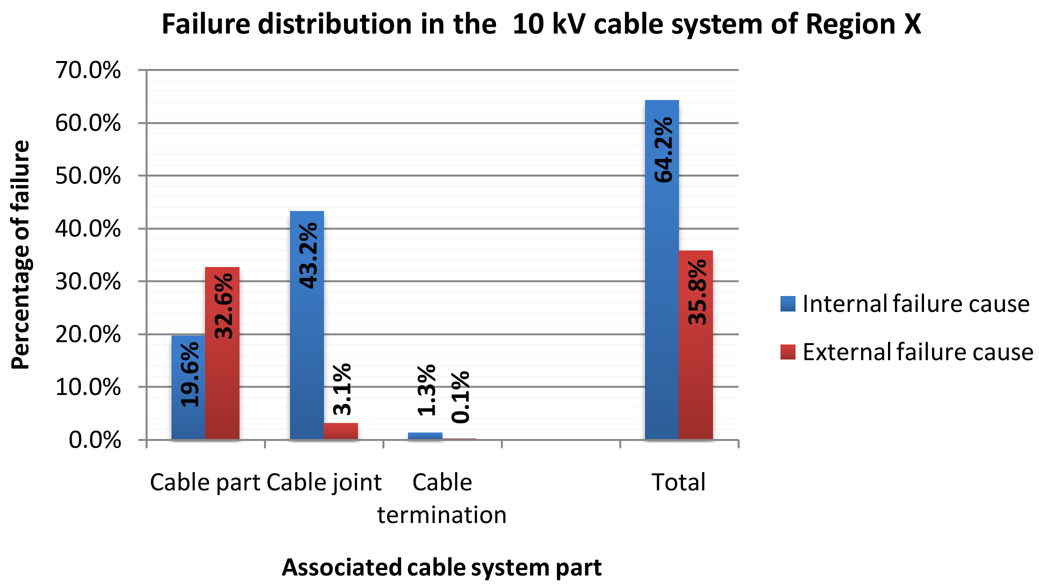


Figure 2.11: Failure distribution for particular MV cable system parts. The failures are divided according to: internal and external failure causes and the associated cable system parts are: cable part, cable joint and cable termination

From the historical failure statistics shown in figure 2.11, for the area "Region X", it is indicated that external, third party, damages are causing 35.8% of the failures in the 10 kV cable network of Stedin. Furthermore, we found that cable parts are related to the most external causes, namely 32.6%. This is due to the extensive lengths of the cable infrastructure. Contrary to cables, cable joints are discrete confined components, and the probability of striking a joint during digging activities is by far smaller.

Besides the external damages, the majority of breakdowns, which is 64.2%, are still caused by internal component related defects of the cable network. Noticeably, the majority of failures caused by internal defects occur in the cable joints (43.2%). The historical failure statistics for the 10 kV cable joints of the particular network "Region X" is further analyzed in the next section.

2.3.2 Cable Joint Failures

In the records of the failure database the cable joint types can be divided into three categories for the 10 kV distribution network. The three categories are based on the principle of joint insulation used. These three categories are:

- Mass insulated joints (liquid mixture of oil and resin)
- Oil insulated joints
- Synthetic insulated joints (shrink or resin)

For the area "Region X", the failure distribution for the three cable joint categories is depicted in figure 2.12.

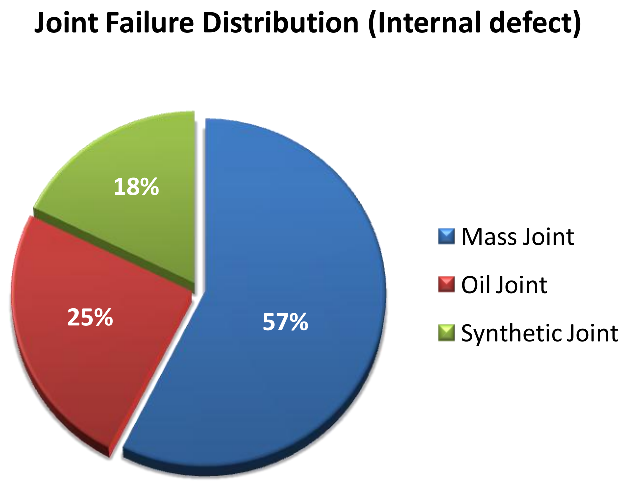


Figure 2.12: Recorded failure distribution overview related to three categories of cable joints. The three categories are: mass filled, oil filled and synthetic insulated joints

The mass insulated joints have the highest amount of failures (57%), followed by the oil insulated joints (25%) and the synthetic insulated joints (18%) respectively. For the mass and oil insulated joints the typical cause of failure is due to ageing and ground conditions (environmental stresses). The failure data records also indicate that moisture ingress plays a key role in the failures of oil insulated joints. As result of water penetration, the insulation quality of the material decreases tremendously [15], as was already highlighted in section 2.2 on cable joint failure examples.

It is also important to touch on the fact that the joint population for mass insulated joints is larger and older than the oil and synthetic insulated joints. This may explain the higher share of joint failures for mass filled joints. Likewise, it is worth mentioning that the synthetic insulated joints have the youngest population, but still turn out to have a considerable share in the joint failure statistics. This could be due to early failure caused by poor workmanship or manufacture errors (human induced stresses).

2.4 Conclusions

We can conclude that from a general point of view, the MV distribution networks in the Netherlands are almost 100% constructed of underground power cables. It is common perception that underground systems are more reliable than overhead systems. Therefore, the reliability in the Netherlands is among the highest in Europe. However, we have found that, when the total (HV, MV and LV) power grid is considered, the MV distribution network is responsible for the major part of power-delivery outages. From many consulted literature we have found that the vast majority of distribution grid outages is due to failures in MV cables and more specifically their joints.

We can conclude that, besides, external defect of non-electrical nature, more than 60-70% of the breakdowns in the MV cable systems are caused by internal defects in the insulation systems. Furthermore, we can conclude that cable joints have higher likelihood of failure in a cable system due to various reasons. The most significant reasons why cable joints are subjected to higher likelihood of failure are because:

- they are subjected to higher electrical, mechanical and thermal stresses
- they are mounted in the field and most of the time under non ideal circumstances, particularly during outage situations
- they are not subjected to expensive reliability testing procedures than the cable itself
- the quality of mounting the accessories is quite sensitive to workmanship, experience and care of the involved employee.

These conclusions, which are mostly based on literature review, are confirmed by findings from a historical failure database analysis for 10 kV MV distribution networks of Stedin. From this analysis we have found that cable systems are associated to 88% of the outages in these MV networks. A further analysis of cable system failures for an area named "Region X" indicates that 35.8% of failure causes are due to external (third party) influences, while 64.2% are due to internal (component) related causes. From this, we can conclude that a large portion of MV cable systems failures are due to "real" component related issues.

Furthermore, from an in-depth analysis on the historic failure statistics for cable joints we found that mass insulated joints have the highest amount of failures (57%), followed by the oil insulated joints (25%) and the synthetic insulated joints (18%).

3
Statistical Failure of Components

In chapter 2, an overview of the construction of general Dutch distribution networks and its characteristics was provided, with special emphasis on underground distribution network. Added to this, chapter 2 highlighted the most common defects that can occur in cable insulations and cable accessories, as it is these types of components that form the backbone of an underground distribution network. From the investigated failure data, in the last section of chapter 2, it became clear that cable systems contribute to the majority of outages in distribution networks. More specifically, cable joints are found to have the highest contribution to outages in the distribution network. In this chapter, the attention is shifted to the statistical modelling of component failure data. Statistical modelling can be used for developing failure rate models which can give insight into the failure behaviour of a certain population of components.

Chapter 3 starts in section 3.1 with a general background of component failure rates and how they are related to the well-known bathtub curve. A distinction is made between the standard bathtub curve and the saw-tooth bathtub curve, and the latter describes the failure behaviour of a population of components when subjected to ageing and maintenance activities. This section ends with an example of deteriorating failure rates for different MV cable joints.

In section 3.2, the theoretical background of statistical failure distributions is described. These statistical failure distributions are used for modelling component life data. Furthermore, statistical functions used for describing the failure distributions are provided, together with the most commonly used failure distributions.

Statistical Life Data Analysis (LDA) is described in section 3.3 and, here, two methods are introduced which can be used for analysing continuous life data, namely the *parametric method* and the *non-parametric method*. The difference between these two methods lies in the assumption whether or not to use a predetermined failure distribution for describing the underlying data. Parametric methods are used when the purpose of life data analysis is to describe large amounts of information, characteristics and behaviours by a small number of parameters. Non-parametric methods are most appropriate when the data sets are small. In this section a straightforward procedure for applying parametric life data analysis is provided.

Finally, section 3.4 gives an overview of the main conclusions of this chapter and also indicates how the knowledge gained in this chapter will contribute in chapter 4.

3.1 Background

3.1.1 Component Failure Rate and Bathtub Curve

Generally speaking, a distribution system is a building block of many components such as transformers, underground cables, cable accessories, circuit breakers and fuse cut-outs. These components can be installed and interconnected together in different ways to form a working power delivery system. Usually, the occurrence of a power outage means that the system has failed in doing its job, which is delivering continuous and reliable power to the customer. Basically, outages occur because one or more of these building blocks have failed to fulfil its role in the system. In section 2.3, we found that many distribution system outages are due to component failures.

Every distribution system component can be described by a set of reliability parameters [16]. Component failure rates and component repair times are examples of simple reliability models. Other sophisticated models are, amongst others, Probability of Operational Failures (POF), Permanent Short Circuit Failure Rate (PSCFR), Mean Time To Switch (MTTS) etc. These mentioned reliability parameters are important, but component failure rates have historically received the most attention [16]. The reason for this is that failure rates have unique characteristics and are essential for all types of reliability analysis. Before proceeding with the aspects behind failures and failure rates, it is necessary to first give a definition for these terms [16], [17].

Failure – The definition of failure as used in the context of this thesis will be an event that ends the life of a device.

Failure rate – Failure rate is defined as the annual rate of failure. This can be either, the likelihood of failures (when predicting the future) as a portion of a population, or, the actual rate of failures as a portion of a population (if analyzing the past).

Using scalar values, such as 0.025 failures per year for a transformer, might indicate that the failure rate is constant. But, it is commonly observed that the failure rates of certain components tend to vary with time (thus with age). In many reliability studies, the failure rate is assumed to be constant during the useful life thereafter it increases with component age. A graph that is commonly used to represent how components failure rate changes with time is the well know bathtub curve, which is shown in figure 3.1 [18], [19], [20].

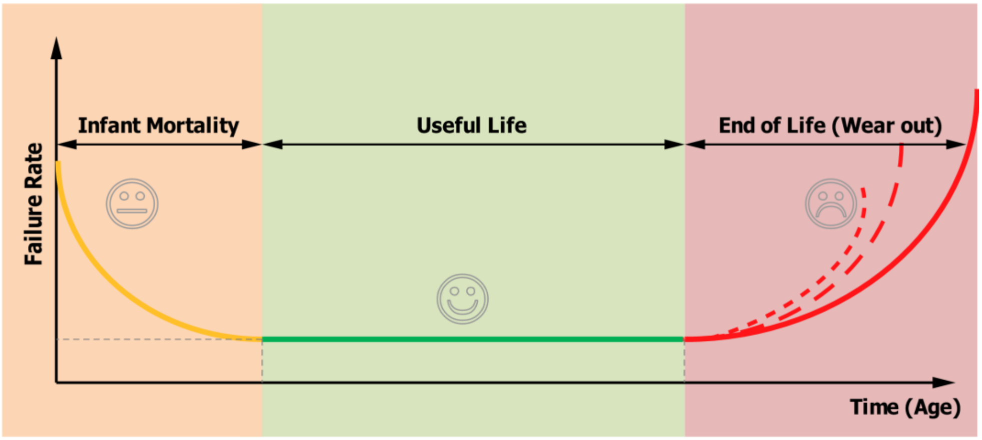


Figure 3.1: A standard bathtub curve that is characteristic for failure rates of many electrical components that are prone to manufacturing errors, installation flaws and wear out

It is important to note that the bathtub curve does not depict the failure rate of a single item, but the relative failure rate of an entire population of components over time. Figure 3.1 shows that in the early life time stages, newly installed electrical equipment has a relative high failure rate. This could be attributed to the possibility that the equipment may have manufacturing errors, was maybe damaged during transportation, or damaged during installation or installed poorly. This period, with monotonically decreasing failure rate, is also called the *infant mortality period*, the *break-in period* or the *early failure period*.

The intermediate period, after the infant mortality period, is characterized by a nearly constant failure rate for a relatively long period. This period is also called *useful life* or *normal life period*. In this period failure rates can be modelled by single scalar numbers. All equipment age during their lifetime, and will eventually reach the end of useful life. This period is accompanied by an exponential increase in the failure rate due to ageing and wear-out. This is why this period is referred to as the *wear out period* or *end-of-life period*.

Different causes of failure could lead to several levels of rate of rise of the failure rate. Therefore, there is not a general consensus on the best way for modelling the wear-out period. However, it is clear that depending on the degree of ageing the slope of the failure rate curve can be large or small. This implies that failure rates tend to vary with time [16].

3.1.2 Impact of Ageing on Component Failure Rate

The degree of ageing will eventually result in an increase of the failure rate of components. Much of the available data [17] indicate that, inevitably, failure rates will increase when equipment are in the final stages of ageing. This increase is attributed to wear and ageing, and can be mitigated by maintenance. In practice, maintenance cannot be performed perfectly and often create their own temporary "infant mortality" increase in the failure rate. This can be represented by using a saw-tooth bathtub curve [16], [17], [19]. The reason for the temporary increases in the failure rate is due to the possibility of maintenance crews causing damage or making errors during assembly. Equipment surviving this short period of time is actually maintained properly and, therefore, the failure rates decrease accordingly. This situation is depicted in figure 3.2.

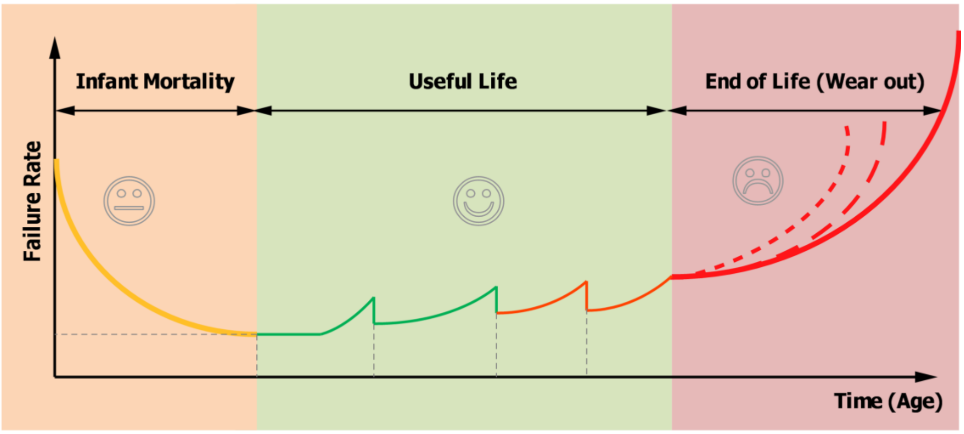


Figure 3.2: The saw-tooth bathtub curve, which models the increasing failure rate of a component between maintenance services and indicates the reliability improvement after each maintenance service. Note however, that the failure rate still increases over time and that maintenance services create their own infant mortality

In figure 3.3 an example, according to [16], for cable joint failure rates of different types is shown. In typical cases the failure rate increases slowly over time, while in other cases the failure rate can increase exponentially. It can also occur that in some cases the failure rate does not increase over time; it might even decrease over time. Unfortunately, these specific situations are rare and are not seen for power system equipment. In general, electrical equipment experience sufficient deterioration with time and therefore, failure rates does increase over time.

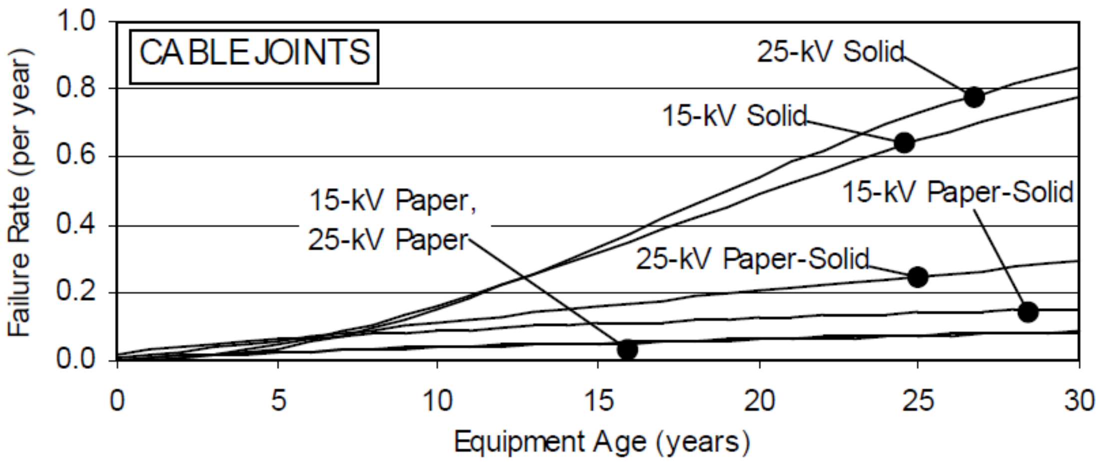


Figure 3.3: This graph show an example of deteriorating failure rates for different MV cable joints. It also shows that joints of different voltage classes and with different features can have completely different failure characteristics.

3.2 Component Modelling

3.2.1 Statistical Failure Distribution

In reliability engineering and life data analysis, many component reliability parameters vary from component to component or from situation to situation. These parameters depend on many physical causes that individually or collectively might be responsible for the failure of a component at any particular instant [19]. For example, such a parameter is the time-to-failure of a component. Since the exact time-to-failure varies and cannot be known beforehand, it is considered a random variable. Random variables can be either discrete or continuous. Life time data analysis will deal almost exclusively with continuous random variables [21]. Random variables are represented by *probability distribution functions*, also known as *failure distributions*. Failure distribution functions are mathematical equations allowing a large amount of information, characteristics and behaviours to be described by a small number of parameters.

In general, a certain failure distribution for a component is chosen based on one of the following considerations [19]:

- The dominant failure mechanism satisfies most or all assumptions which underlie a certain statistical distribution
- The choice is limited to the failure distribution that best fits the life time data.
- A simple distribution, which is well suitable for analytical computations.

A rough estimate of the reliability can be achieved in this way. The estimate is more accurate when the first two above mentioned considerations for choosing a distribution are better fulfilled. However, it is not possible to isolate all physical causes that might be responsible for a component failure and mathematically account for all of them, and therefore, the choice of a failure distribution is not simple and requires expertise and experience.

Often used statistical functions, which describe the failure distributions are the *probability density function (pdf),* the *cumulative distribution function (cdf)*, the *reliability function (R)* and the *failure rate function (λ)*. These functions are described briefly here.

From statistical reliability engineering, given a continuous random variable *X*, the following statistical failure functions can be denoted [22]:

Probability Density Function – The probability density function *(pdf)* of a continuous random variable, *X*, is a function that describes the probability that *X* assumes a value in the interval *[a, b]*. If *X* is a continuous random variable, then the *pdf*, of *X*, is a function *f(x)* such that for two numbers, *a* and *b* with *a<b*:

$$P(a \leq X \leq b) = \int_a^b f(x)dx \text{ and } f(x) \geq 0 \; for \; all \; x \tag{3.1}$$

That is, the probability that *X* assumes a value in the interval *[a, b]* is the area under the density function.

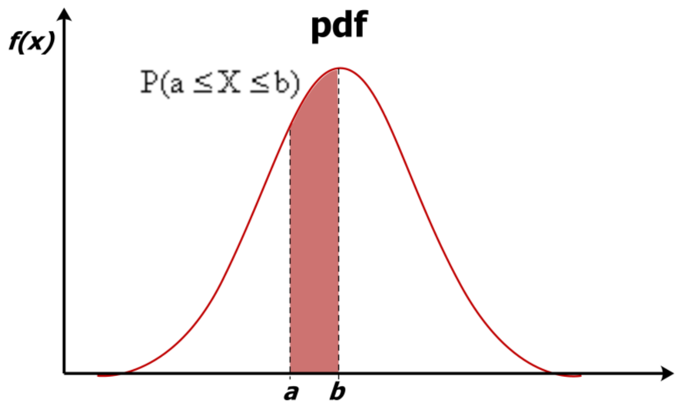


Figure 3.4: Probability density function (pdf), which indicates the probability that X takes on a value in the interval [a, b] is the area under the density curve from a to b

Cumulative Density Function – A cumulative distribution function*, cdf*, refers to the probability that the value of a random variable falls within a specific range. The *cdf* is a function *F(x)* of a random variable, *X*, and is defined for a number *x* by:

$$F(x) = P(X \leq x) = \int_{0,-\infty}^{x} f(s)ds \qquad (3.2)$$

The *cdf* is the integral of the *pdf*, and reflects the probability that *f(x)* will be equal to or less than *x*.

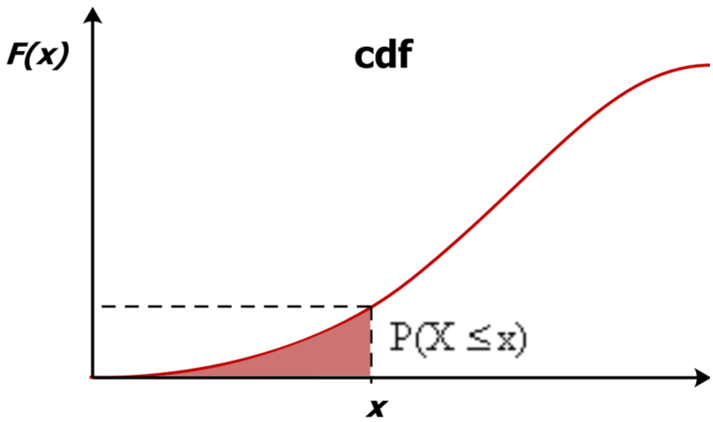


Figure 3.5: A graphical representation of the cumulative distribution function (cdf), which describes the probability that a continuous random variable will be equal to or less than x

Reliability Function – The reliability function, *R (t)*, can be derived using the previous definition of the *cdf* function. This *cdf* (3.2) is also called the unreliability function, *Q (t),* which is the probability of failure in the region of 0 and *t*. From equation (3.2) the association between *F(t)* and *Q(t)* becomes:

$$F(t) = Q(t) = \int_{0}^{t} f(s)ds \qquad (3.3)$$

In general, there are only two states that can occur: success or failure. These two states are also mutually exclusive (they cannot occur at the same time). Since reliability and unreliability are the probabilities of two mutually exclusive states, the sum of both is always equal to unity. Therefore:

$$Q(t) + R(t) = 1$$
$$R(t) = 1 - Q(t)$$
$$R(t) = 1 - \int_0^t f(s)ds$$
$$R(t) = \int_t^\infty f(s)ds \tag{3.4}$$

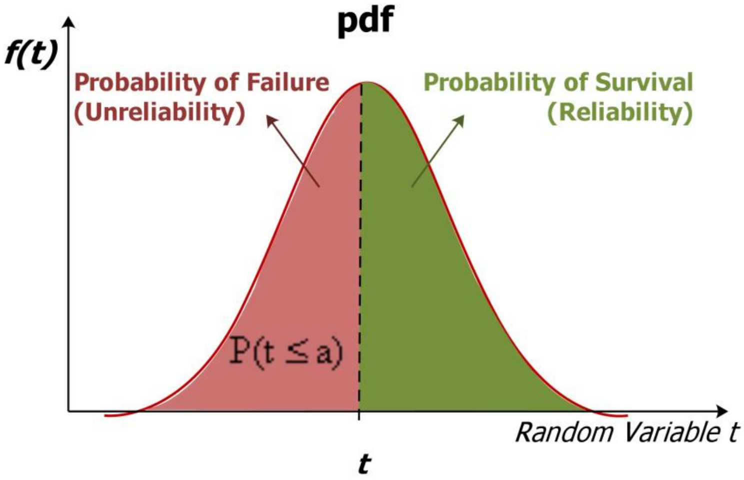


Figure 3.6: A graphical illustration of the reliability, which may be expressed through the pdf as the probability of survival and the probability of failure

The Failure Rate Function – The failure rate function, *λ(t)*, is a function of time and has a probabilistic interpretation, namely *λ(t).dt* represents the probability that a device of age t will fail in the interval *(t, t+dt).* The failure rate function is equal to the probability of a component failing if it has not yet failed. Since the *pdf* is the probability of a component failing and the *cdf* is the probability that it has already failed, the failure rate can be mathematically characterized as follows;

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{1 - \int_0^t f(s)ds} = \frac{f(t)}{R(t)} \tag{3.5}$$

The failure rate function can be expressed as failures per unit time, e.g. 2 failures per year.

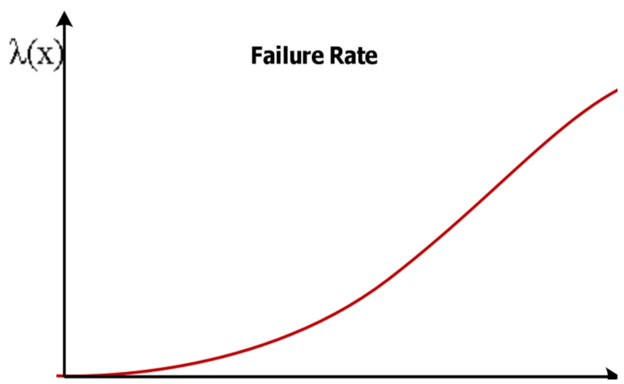


Figure 3.7: This graph show an increasing failure rate function

The above mentioned quantities *f(t)*, *F(t)*, *R(t)* and *λ(t)* can be converted into each other. Therefore, they contain all information about the failure process of the system under consideration.

The Mean Life Function – Another often used statistical measure is the mean life function, which provides a measure of the average life to failure of a component.

$$\mu = \theta = \int_0^\infty t \cdot f(t) dt \tag{3.6}$$

The mean life is a global quantity that does not anymore contain time-dependent information previously stored in *f(t)*, *F(t)*, *R(t)* and *λ(t)* contain.

Now that the basic theory and vocabulary for failure distribution functions have been discussed, the next section will present some common and useful failure distribution functions, which are applicable to life data analysis.

3.2.2 Most Commonly Used Failure Distributions

In general, a statistical distribution is fully described by the *pdf*. Throughout the literature many statistical probability distribution can be found and entire texts are dedicated to defining probability distributions. Some of these distributions seem to better represent life data and each one of them has a predefined form of *pdf*. The following failure distribution will be described briefly in this section and more detail can be found in [16], [19], [21], [22], [23]:

- Normal distribution
- Lognormal distribution
- Exponential distribution
- Weibull distribution

3.2.2.1 Normal Distribution

The normal distribution is one of the best known distributions and is the most widely used general purpose distribution. This distribution is used for several problems in HV technology especially regarding air insulated systems. The normal distribution is commonly referred to as a "bell curve" because its *pdf* resembles the cross section of a bell. A normal distribution describes components which only fail as a result of wear process and, therefore the failure rate increases monotonically. The normal distribution is mathematically characterized by two parameters, which are the mean life, μ, and its standard deviation, σ. The standard deviation gives the extent of a certain spread. Formulae corresponding to the normal distribution are:

$$cdf = F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \tag{3.7}$$

so:

$$pdf = f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \tag{3.8}$$

and:

$$failure\ rate = \lambda(t) = \frac{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}{\int_t^{\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt} \tag{3.9}$$

One disadvantage of the normal distribution for modelling lifetime data is the possibility to allow negative time at the left hand limit of the distribution. Especially in the area of electric distribution system component modelling, random variables characterizing component reliability parameters are typically constrained to positive time to failures. However, provided that the normal distribution has a relatively high mean and relatively small standard deviation, the issue of negative failure times should not present itself as a problem.

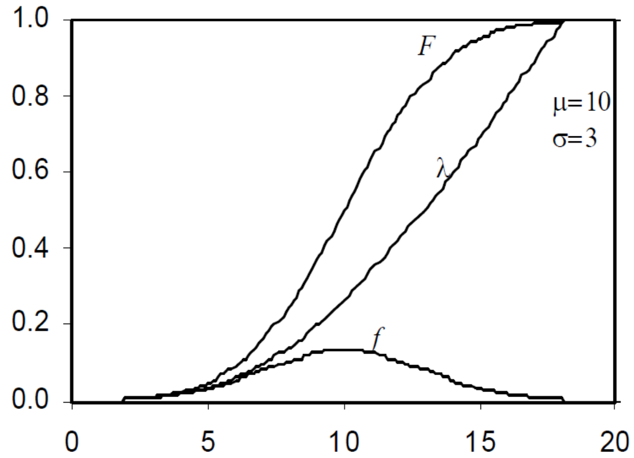


Figure 3.8: This figure shows the pdf, cdf, and failure rate function for the normal distribution for specific distribution parameters. In this example the mean value is 10 and the standard deviation is 3

3.2.2.2 Exponential Distribution

The exponential distribution is the most common failure distribution function used in the field of reliability analysis. This is due to the characteristic of a constant failure rate, which is representative of electrical components during their useful life. Mathematically, it is a simple distribution that is fully characterized by a single parameter, λ. The exponential distribution is used to model the behaviour of components with a constant failure rate or units that do not degrade with time or wear out. Formulae related to this distribution are:

$$cdf = F(t) = 1 - e^{-\lambda t} \tag{3.10}$$

so:

$$pdf = f(t) = \lambda e^{-\lambda t} \tag{3.11}$$

and:

$$failure\ rate = \lambda(t) = \lambda \tag{3.12}$$

The exponential distribution is not symmetric, thus it has no symmetric probability density function.

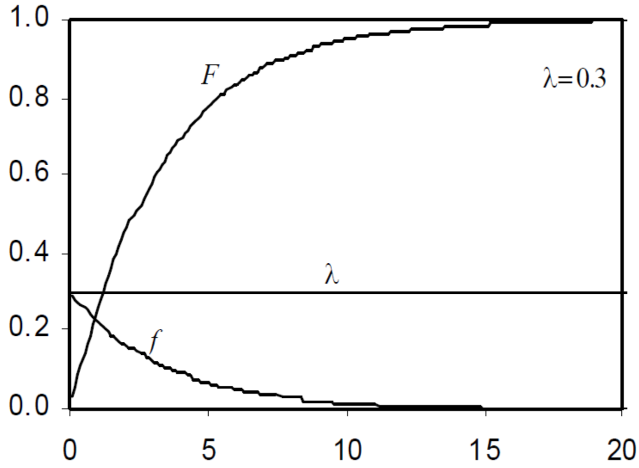


Figure 3.9: This figure shows the pdf, cdf and failure rate function for the exponential distribution for specific (λ=0.3) distribution parameters

3.2.2.3 Log-normal Distribution

The log-normal distribution is closely related to the normal distribution and is a transformed distribution of the normal distribution. The log-normal distribution differs from the normal distribution because it uses the natural logarithm of the random variable, and the random variable is constrained to be nonnegative (it only exists for $t>0$). In this sense, it is assumed that, instead of the random variable t, the random variable $g(t)$ is distributed normally. The lognormal distribution is commonly used to model lives of units whose failure modes are of fatigue-stress nature. Formulae related to the log-normal distribution are:

$$cdf = F(t) = \frac{1}{\sigma'\sqrt{2\pi}} \int_0^t \frac{1}{t} e^{-\frac{(\ln kt-\mu')^2}{2(\sigma')^2}} dt \tag{3.13}$$

Here μ' is the mean of the natural logarithm of the time to failure and σ' is the standard deviation of the natural logarithm of the time to failure. The constant k is related to the transformation $g(t)=\ln kt$. The *pdf* is:

$$pdf = f(t) = \frac{1}{\sigma'\sqrt{2\pi}} e^{-\frac{(\ln kt-\mu')^2}{2(\sigma')^2}} \tag{3.14}$$

The failure rate can be calculated with the expression of equation (3.5):

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{1 - \int_0^t f(s)ds} = \frac{f(t)}{R(t)}$$

The log-normal distribution is a distribution skewed to the right. Density functions of this shape are useful for characterizing reliability parameters such as repair times and switching times.

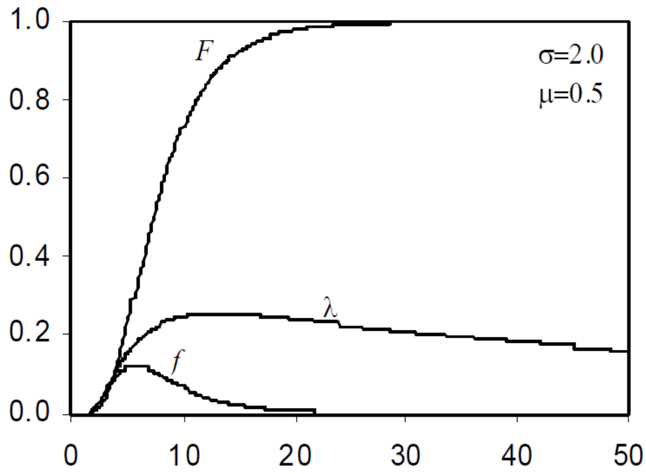


Figure 3.10: This figure shows the pdf, cdf and failure rate function for the log-normal distribution for specific (σ=2 and μ=0.5) distribution parameters

3.2.2.4 Weibull Distribution

The Weibull distribution, named after Waloddi Weibull[1], is the most generally used distribution for ageing failures and often used for power system components [23]. It is a flexible distribution that assumes various shapes to fit varying data sets, based on the value of the shape parameter, β, and the scale parameter, η. The Weibull distribution is related to the exponential distribution and is a transformed version of it. In this sense, it is assumed that, instead of the random variable t, the random variable $g(t)$ is distributed exponentially, where $g(t) = t^{\beta}$ and β changes the shape of the distribution. The formulae related to the Weibull distribution are:

$$cdf = F(T) = 1 - e^{-(\frac{T-\gamma}{\eta})^{\beta}} \tag{3.15}$$

so:

$$pdf = f(T) = \frac{\beta}{\eta}(\frac{T-\gamma}{\eta})^{\beta-1}\, e^{-(\frac{T-\gamma}{\eta})^{\beta}} \tag{3.16}$$

where, $f(T) \geq 0, T \geq 0, \beta > 0, \eta > 0, -\infty \leq \gamma \leq \infty$

[1] Ernst Hjalmar Waloddi Weibull (18 june 1887-12 oct 1979) was a Swedish engineer, scientist and mathematician

and:

$$failure\ rate = \lambda(T) = \frac{\beta}{\eta}\left(\frac{T-\gamma}{\eta}\right)^{\beta-1} \tag{3.17}$$

Equations (3.15) through (3.17) are the probability equations for the three-parameter Weibull distribution. In these equations the parameter *β*, *η* and *γ* are respectively the shape parameter, scale parameter and the location parameter.

If the location parameter, *γ*, is assumed to be zero, the distribution then becomes the two-parameter Weibull and the *pdf* equations then is:

$$pdf = f(T) = \frac{\beta}{\eta}\left(\frac{T}{\eta}\right)^{\beta-1} e^{-(\frac{T}{\eta})^{\beta}} \tag{3.18}$$

The power of the Weibull distribution lies not so much in a certain theoretical failure model, but in its flexibility to describe the different life stages of a population of components. The Weibull distribution is often used to model the bathtub curve. The shape parameter is able to model the different service lifetimes modes of the bathtub curve as follows:

- $0<\beta<1$ represents infant mortality
- $\beta=1$ represent random failures (similar to the exponential distribution)
- $\beta>1$ represents ageing failure modes

Three examples of Weibull distributions are shown in figure 3.11, in which the effect of varying shape and scale parameters on the distribution shape is shown.

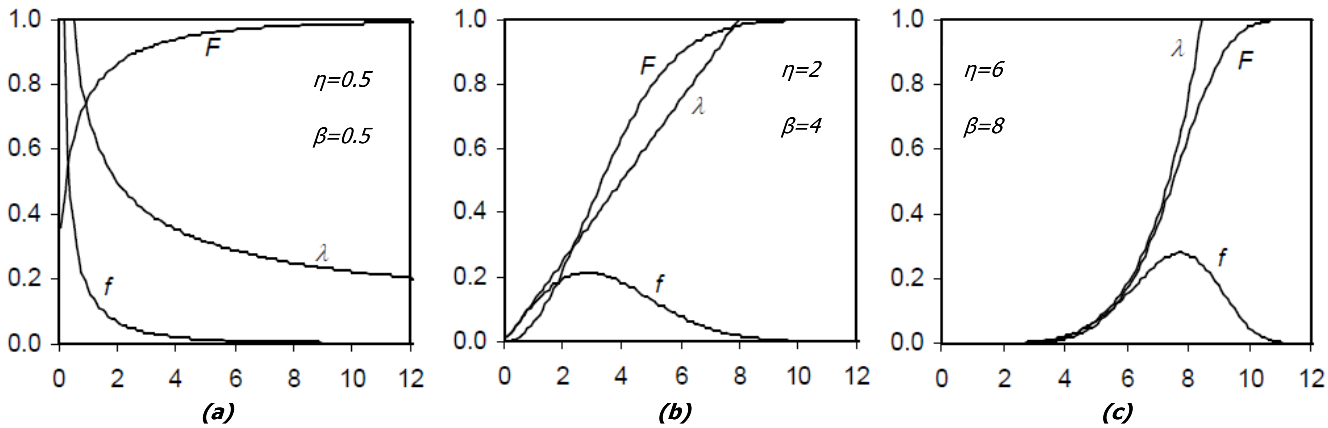


Figure 3.11: An illustration of three different 2-parameter Weibull distributions. By varying the scale parameter, η, and the shape parameter, β, a wide variety of distribution shapes can be modelled. Figure (a) represents an infant mortality period with decreasing failure rate. Figure (b) represents an ageing period with β>1. Figure (c) also represents an ageing period, however with a higher rate of rise of the failure rate. When figures (b) and (c), both describing ageing periods, are compared witheach other, we find that the peaks of the pdf function are situated around the scale parameter. Additionally, in figure (c) the failure probability is less spread (low variability) than in figure (b).

3.3 Statistical Life Data Analysis (LDA)

3.3.1 Parametric Distribution Fitting Method

Methods for analysing continuous life data fall into two classes, distinguished by whether or not they make assumptions about the distribution of the data [24]. The failure distributions mentioned in the previous section are theoretical distributions that are described by their related parameters (shape parameter, scale parameter, mean, standard deviations etc). Methods that use distributional assumptions are called *parametric methods*, because parameters of the distribution are estimated based on the life data. All of the commonly used parametric methods assume that in some way the data follows the chosen failure distribution. Parametric methods are most appropriate for large data samples.

An alternative method is the *non-parametric method*, which allows the user to analyze data without assuming an underlying failure distribution, and therefore, it is also known as the *distribution free method*. With this method, the data can be analysed based on histograms of the life data, and the cumulative distribution function *(cdf)* is based on the histogram and has no underlying mathematical distribution [25]. This method can have certain advantages as well as disadvantages. Non-parametric methods avoid potentially large errors caused by making incorrect assumptions about the distribution. However, the confidence intervals associated with non-parametric analysis are usually much wider than those calculated by parametric methods [21]. Another disadvantage of the non-parametric method is that making predictions outside the range of the observation is not possible. Non-parametric methods are most appropriate when the data sizes are small.

In general, non-parametric methods are less statistically powerful than their parametric counterparts, especially when the life data analysis is intended for creating more knowledge with regard to the failure behaviour of a population. Typically, by using the parametric method, the information associated with many data points can be reasonably modelled with one or two parameters. The parametric distribution fitting method will be used in this study for the life data analysis. Parametric distribution fitting comprises a number of steps and a straightforward procedure is depicted in figure 3.12 [18], [26]. In the following paragraphs, the process of parametric distribution fitting is discussed.

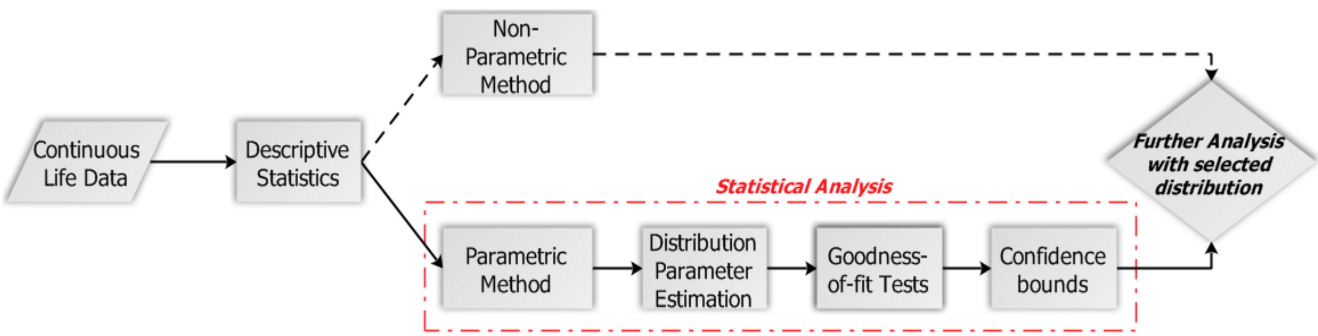


Figure 3.12: Evaluation flowchart for life data analysis

3.3.2 Data Types

The first step in the process of parametric distribution fitting is the collection of life data. Statistical failure distribution models rely extensively on the data, *life data* or *time-to-failure* of a component, to make predictions. The accuracy of any prediction is directly proportional to the quality and completeness of the supplied data. The combination of good data and appropriate model choice, will usually results in acceptable predictions [21]. Typical for life data is that failure data is considered as a failure while the un-failed components (in-service data) are used as suspended data. Suspended data means that the units are still operating at the time the reliability of these units is to be determined. The life data is gathered during the whole life of a technical component, starting with the installation and ending with its disposal. Furthermore, the collected life data for statistical analysis should have the following properties [18]:

- Randomness
- Independency
- Homogeneity
- Sufficient amount of data

In the analysis of life data it is deemed to be advisable to use all available data. In practise, however, it is hard, expensive and sometimes impossible to collect all required life data. Therefore most of the time, the available data is incomplete or includes uncertainties as to when a component failed exactly. To interpret this, life data can be separated into two categories [21], [22]:

- Complete Data (all units have failed)
- Censored Data (not all units have failed)

Complete Data – Complete data is used when the value of an observation is known completely. For example, if the time-to-failure for a cable joint population with 200 units is observed and all 200 units have failed (and the time-to-failures has been recorded), then the complete information as to the time of each failure is known. It goes without saying that processing complete data is much more efficient and easier than censored data.

Censored Data – Censoring occurs when the value of an observation is only known to some extent. Censored data is often encountered when analysing practical life data, especially in case of electrical power systems where the majority of installed equipment is still in-service, and most of the time the exact age of equipment at the moment of failure is unknown. Three censoring schemes are possible, which are:

1. Right-censored data (suspended data): When a data set is composed of components that did not fail, it can be referred to as right-censored data or suspended data. The term "right-censored" indicates that the event is located to the right of the data set, which implies that certain components are still operating.

2. Interval-censored data: This reflects uncertainty as to the exact times the equipment failed or exact age of an equipment upon failure. Interval data is often encountered

in asset related databases when components are not constantly monitored.

3. Left-censored data: This censoring scheme is a special case of interval-censored data. With left-censored data the time-to-failure for a particular component is known to occur between time zero and some inspection time.

In the present study, life data analysis is performed for 10 kV cable joints, and the consulted databases contain right-censored as well as interval-censored data.

3.3.3 Failure Distribution Fitting and Parameter Estimation

After life data is collected and prepared (categorized), the statistical analysis can be performed, which addresses the question of determining which failure distribution function to use and what the estimated parameter value is. Distribution fitting can be seen as a process that fits the data points from the life data with an appropriate distribution. After a certain distribution is selected to fit the data, the next step is to estimate the parameters of this distribution. Subsequently, these estimates can be used to construct reliability functions and plots. Estimated distribution parameters can be calculated on several methods. Three methods that are applicable to life data analysis are presented [21]. These are:

- Probability Plotting
- Rank Regression Analysis (Least Squares Estimation, LSE)
- Maximum Likelihood Estimation (MLE)

Probability Plotting – This is the simplest and least mathematically intensive method for parameter estimation and uses a specially constructed plotting paper. Based on the scales of the plot, the parameters can be estimated. This is illustrated by using an example according to reference [21] for the well-known Weibull distribution. In order to construct a probability plotting paper, the cumulative distribution function *(cdf)* of the Weibull distribution needs to be linearized to the form $y = ax + b$.

The *cdf* or unreliability function of the 2-parameter Weibull distribution is given by:

$$cdf = F(T) = 1 - e^{-(\frac{T}{\eta})^{\beta}} \tag{3.19}$$

The function can be linearized and becomes:

$$ln\left(ln\left(\frac{1}{1 - F(t)}\right)\right) = \beta \ln(T) - \beta \ln(\eta) \tag{3.20}$$

Setting the term in front of the equal sign to:

$$y = ln\left(ln\left(\frac{1}{1 - F(t)}\right)\right) \tag{3.21}$$

and:

$$x = \ln(T) \tag{3.22}$$

Then equation (3.20) can be rewritten as:

$$y = \beta \mathrm{x} - \beta \ln(\eta) \tag{3.23}$$

Equation (3.23) is now a linear equation with slope *β* and an intercept with the y-axis of *βln(η).* The next step is to construct the x and y-axis of the Weibull probability plotting paper. Such probability papers are available in standard formats for different distributions. The a-axis value will correspond to the failure time, since *x=ln(T)*. Each x-value is simply the natural logarithm of each time-to-failure. For the y-axis value an additional step is required because we see from equation 3.21 that the y-coordinate is based on the unreliability, *F(t)*. Therefore, we need to come up with an estimate for the unreliability for each time to failure in order to plot the data points. Usually, these unreliability estimates are accomplished with *Median Ranks*. Detailed explanation of the median ranks can be found in [21]. An approximation that can be used to estimate median ranks is called the *Benard's* approximation [21] and has the form:

$$MR = \frac{j - 0.3}{N + 0.4} \tag{3.24}$$

where *N* is the total number of failures and *j* is the failure order number. Based on the Benard's approximation, we can calculate unreliability estimates for each time-to-failure as shown here:

Table 3.1: This table shows the calculated unreliability estimates with the Benard's approximation

Units	**Time-to-Failure (hours)**	**Median Ranks (%)**
1	16	10.91
2	34	26.44
3	53	42.14
4	75	57.86
5	93	73.56
6	120	89.10

Then, given the *y* and *x* value for each point, the points can easily be put on the plot as shown in figure 3.13. Once the points have been placed on the plot, a straight line is drawn through these points. Afterwards, the slope of the line can be obtained and this is the parameter *β*. To determine the scale parameter, *η*, a mathematical manipulation of the *cdf* is required by setting *T=η*.

Substituting *T=η* in equation (3.19) results in:

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

$$cdf = F(T) = 1 - e^{-1}$$

$$cdf = F(T) = 0.632 = 63.2\%$$

Thus, if *F(t)=63.2%* is entered for the y-axis, the corresponding value of *T* will be equal to *η*.

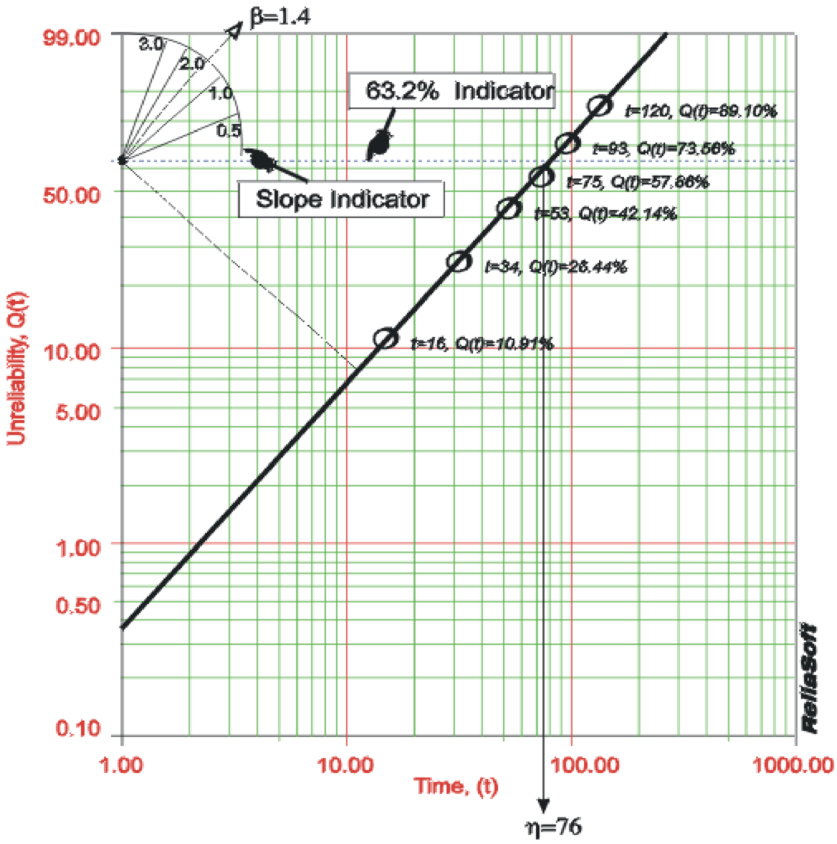


Figure 3.13: Probability plot for estimating Weibull distribution parameters by means of linearizing the cdf function. The slope of the probability plot is the shape parameter, β, and the x-axis intercept for F(t)=63.2% is the scale parameter, η. In this example β=1.4 and η=76 **[21]**

This example has illustrated a simple but rather time consuming probability method for the 2-parameter Weibull distribution for a complete data set. The methodology can be more difficult, for example, if the data set contains suspensions. This method of parameter estimation requires much effort and is not consistent in the results. Furthermore, this method was used primarily before the widespread use of computers that are well capable of performing the calculations for more complicated parameter estimation methods, such as least squares and maximum likelihood methods [21].

Rank Regression Estimation – Rank regression parameter estimation or least squares method is, in essence, a more formalized method of probability plotting, in that it is a mathematically based version of probability plotting. The process of using rank regression to analyse life data requires a straight line to be fitted to a set of data points, in such a way that the sum of the squares of the distance of the points to the fitted line is minimized. This minimization can be performed in either the vertical or horizontal direction. If the regression is on the x-axis, then the line is fitted so that the horizontal deviations from the points to the line are minimized. If the regression is on the y-axis, then this means that the distance of the vertical deviations from the points to the line is minimized. This is illustrated in figure 3.14.

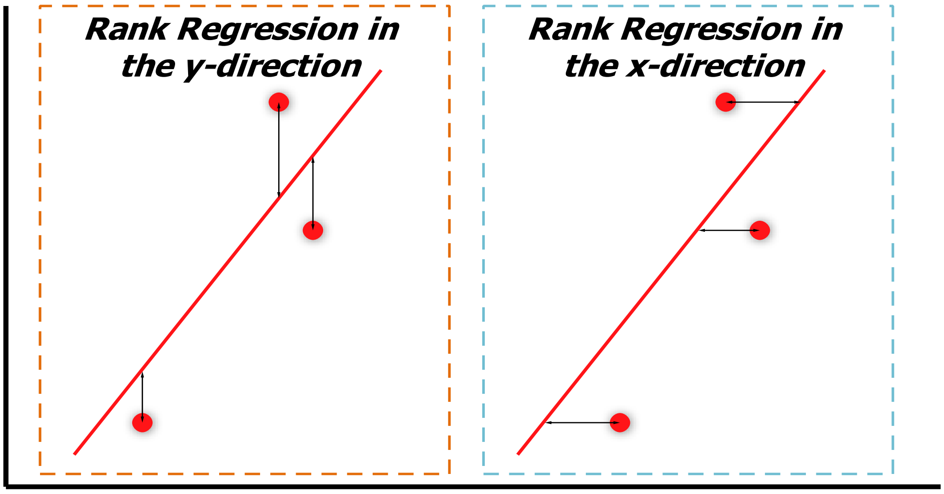


Figure 3.14: This figure illustrates the rank regression method for parameter estimation. Rank regression in the y-direction minimizes the distance in the y-direction, while rank regression in the x-direction minimizes the distance in the x-direction

Detailed information on the equations for rank regression can be found in reference [21]. Essentially, the rank regression estimation method is quite good for functions that can be linearized and is generally best used with data sets containing complete data. For data sets containing large quantities of suspended data points, the next method called Maximum Likelihood Estimation can be a preferable estimation technique.

Maximum Likelihood Estimation (MLE) – The MLE process is a method for estimating the most likely parameters, for a given distribution, that will best describe the data. This is achieved for a set of data by maximizing the value of the so called "likelihood function" [21], [22]. The likelihood function is based on the probability density function *(pdf)* for a given distribution. For instance, consider a generic *pdf:*

$$f(x;\ \theta_1\theta_2 \ldots, \theta_k) \tag{3.25}$$

where, x represents the time-to-failures, and $\theta_1, \theta_2,\ldots,\theta_k$ are the parameters to be estimated. In case of 2-parameter Weibull distributions, the parameters would be β and η. The likelihood function for complete data is the product of the *pdf* functions, with one element for each data point in the data set. The likelihood function is:

$$L = \prod_{i=1}^{R} f(x;\ \theta_1\theta_2 \ldots, \theta_k) \tag{3.26}$$

where R is the number of failure data points in the complete data set. According to [21] it is often mathematically easier to manipulate the likelihood function by taking the logarithm of the function. Using the log-likelihood function does not affect the validity of the results. The log-likelihood function is:

$$\Lambda = lnL = \sum_{i=1}^{R} f(x;\ \theta_1\theta_2 \dots, \theta_k) \tag{3.27}$$

It is required to find the values for the parameters that result in the highest value (maximum) for the log-likelihood function. This is commonly done by setting the partial derivative of each of the log-likelihood function of each parameter to zero:

$$\frac{\partial \Lambda}{\partial \theta_j} = 0, \qquad j = 1,2, \dots, k \tag{3.28}$$

Equation (3.28) results in a number of equations with an equal number of unknowns, which can be solved simultaneously.

When dealing with censored data, be it right-censored or interval-censored data, another term should be included in the likelihood function. The term that is included for censored data includes the cumulative density function *(cdf)*. The extended likelihood function has the form:

$$L = \prod_{i=1}^{R} f(x;\ \theta_1\theta_2 \dots, \theta_k) \cdot \prod_{j=1}^{M} \left[1 - F\left(S_j;\ \theta_1\theta_2 \dots, \theta_k\right)\right]$$
$$\cdot \prod_{l=1}^{P} \left[F(I_{l2};\ \theta_1\theta_2 \dots, \theta_k) - F(I_{l1};\ \theta_1\theta_2 \dots, \theta_k)\right] \tag{3.29}$$

with, M is the number of suspended units, P is the number of units with interval time-to-failures. And, S_j, is the j^{th} time of suspension, while I_{l1} and I_{l2} are, respectively, the beginning and ending of the time interval.

With this function (3.29), the analysis procedure proceeds as described in equations (3.27) and (3.28). The ability to take into account large number of suspensions is the foremost advantage that MLE analysis has over other parameter estimation techniques. MLE analysis is preferred over the other parameter estimation techniques when large number of suspensions are present and also when the data set gets larger. MLE is asymptotically consistent, which means that as the data set gets larger, the estimates converge to the true value.

3.3.4 Goodness-of-fit Test

When modelling life data, it is often desirable to diagnose the fitted model in order to assess whether the assumed model matches the data that it is supposed to represent. While engineering knowledge should always govern the choice of a failure distribution, nevertheless, there are many statistical tools that can help in deciding whether or not a distribution model is a good choice from a statistical point of view. The method for calculating the best fitting distribution is called the Goodness-of-fit Tests [21], [25], [27]. These methods can also be used to compare different distribution with each other.

Graphical methods, e.g. probability plots, can give a visual assessment of the models fit. This method can be used with rank regression estimation, but should not be used for MLE. Another widely used goodness-of-fit test, when the rank regression estimation is applied, is the use of the correlation coefficient. This describes the distance between the data points and the fitted distribution and is usually denoted by ρ (rho). The closer the value of ρ is to 1 or -1, the better the linear fit is assumed to be. If the value of ρ is zero, it means that the data are randomly scattered and have no pattern or correlation in relation to the regression line model. When using the MLE method to estimate parameters of the distribution model, the likelihood value L can be used to assess the fit to the distribution data set. Contrary to the correlation coefficient, the likelihood value is not constrained by a certain range of possible values. The likelihood value L can be used to compare the fit of multiple distributions and the distribution with the largest L value is the best fit statistically. Other methods which can be used for rank regression and MLE are e.g. the Kolmogorov-Smirnov (K-S) test, chi-squared test, Anderson-Darling test [21] [27].

Finally, the failure model that, according to engineering knowledge and statistical tests, is the best fit should be selected for further analysis.

3.3.5 Confidence Bounds

Estimating the precision of an estimate can be confusing, however, this is an important concept in the field of reliability engineering, leading to the use of confidence bound (or intervals). A confidence bound can be seen as an estimated range of values which are likely to include an unknown population parameter. To illustrate this concept, consider independent samples that are taken repeatedly from the same population. When the confidence interval is calculated for each repeated sample, then a certain percentage (confidence level) of the intervals will include the unknown population parameters. For example, when taking independent samples from a population, it may be noticed that 90% of the time the estimate is between X1% and X2%. The width of the confidence interval gives an indication of how uncertain the estimate of an unknown parameter is, and what the range of the plausible value could be.

Confidence limits are the lower and upper boundaries of a confidence bound, which is the value that defines the range of a confidence bound. The confidence limits are generally described as being one-sided or two-sided.

There are several ways to calculate reliability confidence bounds for different distribution [21]. The Fisher matrix (FM) described in [28] and the likelihood ration bound method (LRB) method are both used very often. Detailed description and mathematical derivations of the different goodness-of-fit test can be found in reference [21].

3.4 Conclusions

Failure rates and statistical modelling of life data has been the central theme of this chapter. Overall, it can be concluded that a failure is an event that ends the life of a device. In general, failure rates are defined as the annual rate of failure. Moreover, it can be concluded that failure rates tend to vary with time. The levels with which time varying failure rates changes depend on the degree of ageing, maintenance actions and the actual age of components. In general, it is found that electrical equipment experience deterioration with time, and as a consequent, failure rates will increase over time.

It is found that life data describes the entire life of a given component and qualitatively can be explained best by referring to the bathtub curve. It can be concluded that it is possible to model life data of components by applying parametric or non-parametric distribution fitting methods. The choice of an appropriate method is based on the goal of the application. When the purpose of the statistical life data analysis is to create knowledge with regard to the failure behaviour of a population, the use of parametric distribution fitting methods is to be preferred. In that way, large amounts of information, characteristics and behaviours can be described by a small number of parameters.

In general, a statistical distribution is fully described by the *pdf*. From the *pdf* other functions such as cumulative distribution function $F(t)$, reliability function $R(t)$ and failure rate function $\lambda(t)$ can be derived. These functions play an important role in reliability and life data analysis. Many failure distributions are possible and described thoroughly in literature. Some of these distributions tend to better represent life data and each one of them has a predefined form of $f(t)$.

3.4.1 Parametric Distribution Fitting Method Procedure

The parametric distribution fitting method will be used for this study. Parametric distribution fitting comprises a number of steps and a straightforward procedure is illustrated in this chapter. In summary, the procedure steps for parametric distribution fitting are:

- Data collection and preparation
    - Complete data
    - Censored data
- Distribution parameter estimation
    - Probability plotting
    - Rank Regression
    - Maximum Likelihood Estimation (MLE)
- Goodness-of-fit Test
    - Visual inspection
    - Correlation coefficient
    - Likelihood Value
    - Other methods, Kolmogorov-Smirnov test, Anderson-Darling test etc
- Confidence bound

Finally, a statistical failure distribution is selected for the associated life data based on engineering knowledge and additional statistical tests. In the way forward, the results of the statistical analysis can be adopted for creating different information about the failure behaviour of components in the near future. With the failure rate function of the selected distribution and the population of components still in service a failure prediction for the coming years can be estimated. This will be the central theme in the following chapters.

4 Application of Statistical Life Data Analysis for 10 kV Cable Joints

In chapter 3, the background of life data analysis and the corresponding statistical theory was described. In this chapter the application of statistical life data analysis for 10 kV cable joints is described.

Section 4.1 starts with the background of the case study. The analysis is carried out for a particular area within the total service area for electricity of Stedin. The reasons for selecting the 10 kV cable joints from this particular area will be listed in this section.

Subsequently, in section 4.2, the process of collecting life data for the cable joints, which is a combination of failure data and information of components that are still in service, is described. The available data from the dedicated databases are shown and, when required, additional assumptions for improving missing data are described in this section. It should be noted that the process of collecting and sorting data for the life data analysis has been the most time consuming process throughout this study.

The core of this chapter is discussed in section 4.3. In this section, the application of parametric distribution fitting is performed, followed by the selection of appropriate failure distribution for each type of cable joint population. The reliability software tool, Weibull++ 7, is used for performing the parametric distribution fitting analysis. Maximum Likelihood Estimation (MLE) is used for parameter estimation in accordance with Goodness-of-fit Tests and engineering knowledge of the failure mechanisms. The Goodness-of-fit Tests includes the Kolomogorov-Smirnov (K-S) test, a correlation coefficient and likelihood value test.

Several conclusions of this chapter are drawn in section 4.4. These conclusions together with the results from this chapter will be used in chapter 5 to provide a number of tools required to support typical AM decision making processes.

4.1 Background

4.1.1 Case Study: Area "Region X"

A case study for the application of statistical life data analysis is carried out for a particular 10 kV distribution network of Stedin. This area is called "Region X". The reason for choosing this area as case study is twofold.

First of all, the frequency of 10 kV cable joint failures resulting in power delivery outages in this region is higher when compared to the other areas. In terms of statistics, this ensures that the statistical estimations are based on sufficient number of failure events (data points). Secondly, the available in-service population data for this particular analysis has been recorded with more accuracy and consistency for this area. These installation records constitute the historical population data, from which the age distribution for the overall cable system can be derived.

Furthermore, in chapter 2 the historic failure statistics for 10 kV distribution network were discussed. From the component level failure analysis it was found that failures caused by internal defects in cable joints make up the majority of component related failures. Therefore, cable joint failure data together with their population data will be used for the application of statistical life data analysis.

4.1.2 Analysis Steps

In chapter 3.3, methods for analysing continuous life data were discussed. Out of these methods, the parametric distribution fitting method will be used and a straightforward procedure for applying this method is depicted in figure 3.12. A more detailed, case specific, procedure of the analysis is illustrated in figure 4.1 and 4.2. The different steps of the procedure will be discussed in this chapter, starting with the collection of available life data followed by the statistical analysis itself.

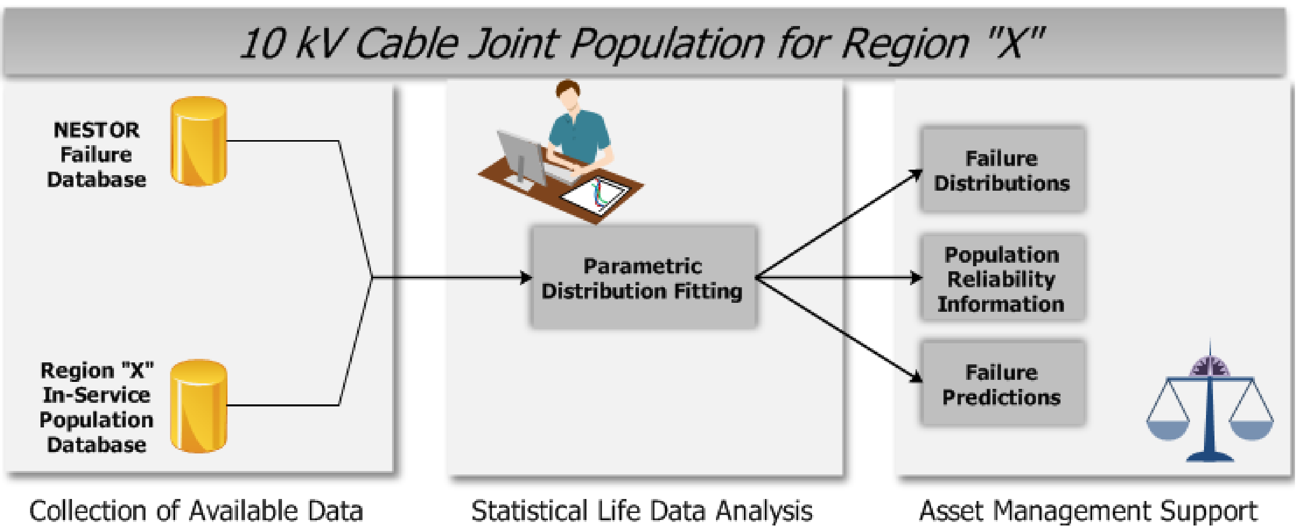


Figure 4.1: Application of statistical life data analysis for 10 kV cable joints. This figure illustrates the procedure for performing the life data analysis. The relevant databases used for extracting failure and in-service data are shown

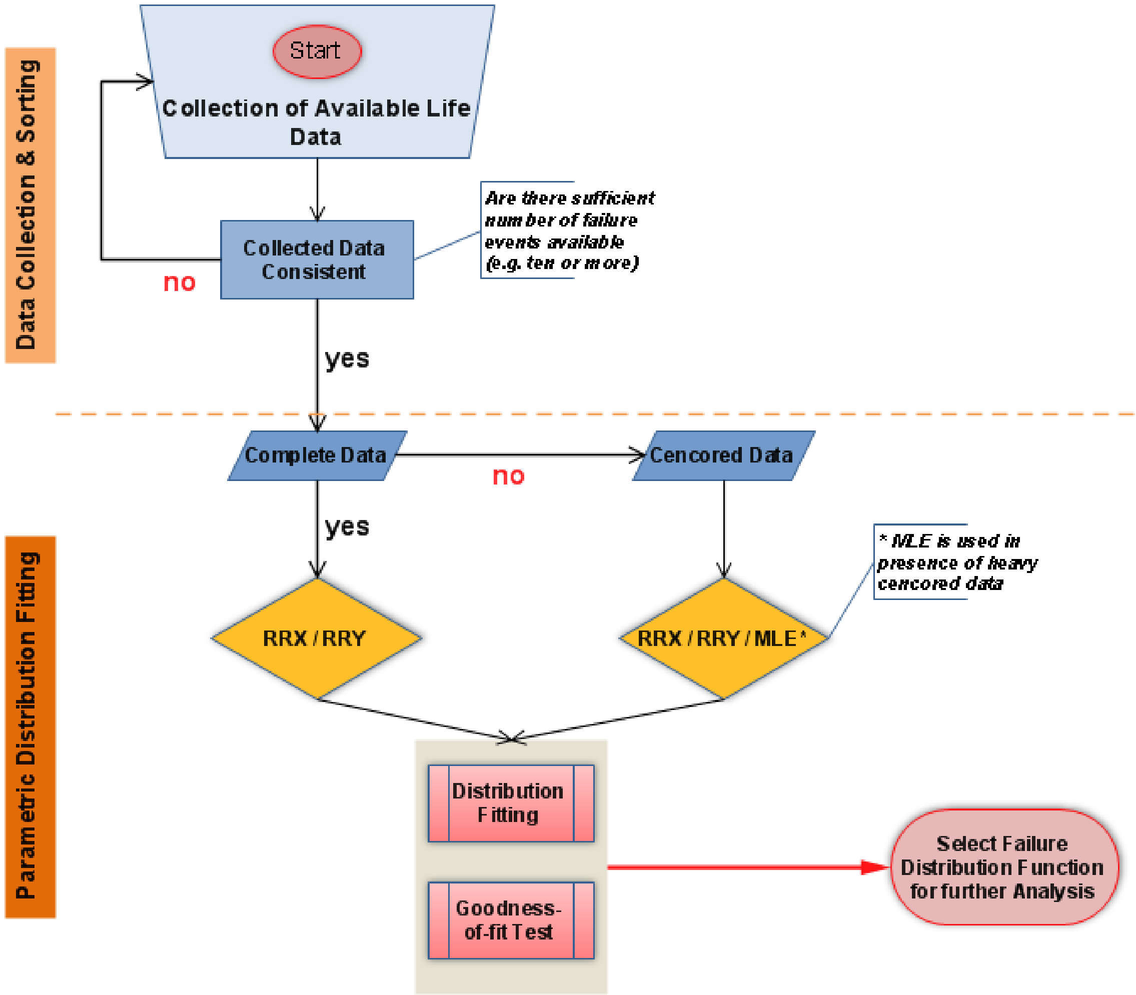


Figure 4.2: Detailed procedure flowchart for the life data analysis. In general the process follows two steps, which are: 1. Data Collection & Sorting and 2. The Parametric Distribution Fitting. When censored data is present, MLE parameter estimation may be used. The choice for using MLE depends on where the suspensions are located in the population and how many of them are available. MLE must be used when the suspensions are at the end of the population. In the study case for the 10 kV cable joints the suspensions are almost always located at the end of the population

Note:

In the past, the Dutch power distribution networks were owned and operated by many small regional utilities. These utilities managed and operated the network in different ways, and sometimes used different voltage levels. Later, during the regulation of the power sector, many smaller regional network operators merged into larger consolidated Distribution System Operators (DSO). For similar reasons the area "Region X" can be further divided into two smaller areas. These areas are known as "Region X1" and "Region X2" in this report. In the remainder of this report only 10 kV cable joint data of "Region X1" is considered. This choice is made because the in-service population data for the "Region X2" was not completely available at the time of this study.

4.2 Collection of Available 10 kV Cable Joint Life Data

4.2.1 Available Failure Data

The Dutch utilities have been collecting outage data since 1976. This was the case for HV, MV and LV systems. In these early days, the failure data collection was primarily a paper matter and not all utilities participated in the reporting. Since 1991, a specific data collection tool is in use, developed by KEMA, and named "KEMA Nestor". This failure reporting database has developed throughout the years and has improved.

During the past years, the type of network components have changed, new voltage levels have been introduced, the utilities involved have changes and merged, the way of data collection has changed and the data definition has been adapted etc. More important are the changes imposed by the liberalization of the power sector, causing the importance of such failure databases to grow, especially as benchmarking tools. In Stedin, much effort is put into recording accurate failure information into the Nestor database. However, as result of the above mentioned changes, useful failure data has been lost or became incomplete. At the same time, discrepancies in databases are caused by changes in the way data is recorded. As a result, the available MV failure data for the period 2004 until 2009 could be used in a viable way for the present study.

This implies that incomplete failure data is available, and, it should be addressed if and how this can be taken into account when performing distribution parameter estimation. The situation of incomplete failure data is schematically illustrated in figure 4.3.

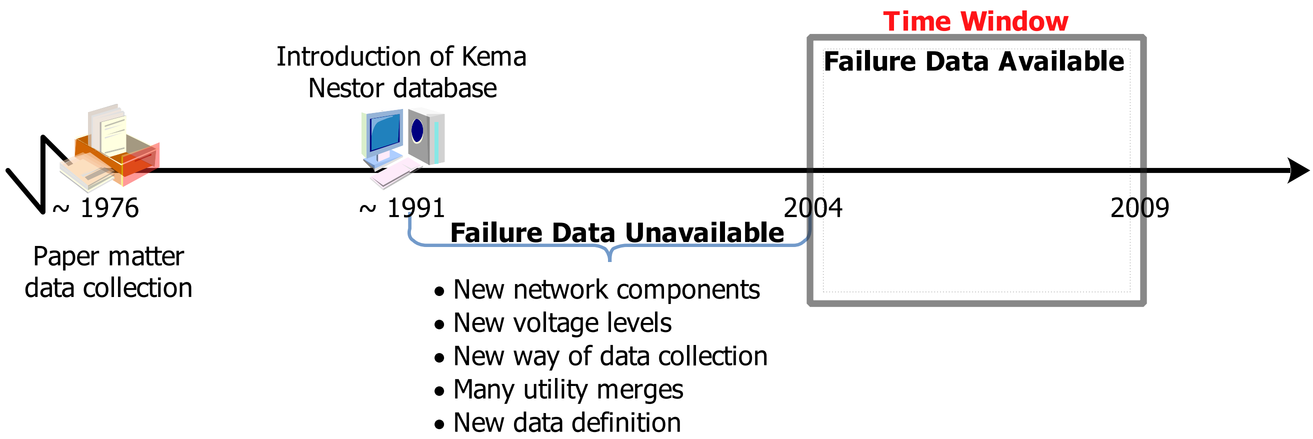


Figure 4.3: Typical availability of failure data in distribution network for the present study. The time window reflects the period where failure data is available. In the intermediate period, starting in 1991, failure data is often unavailable or incomplete

In section 2.3.2 the failure statistics for cable joints have been discussed. It was mentioned that three categories of 10 kV cable joints can be distinguished. The differences between these three categories are in the principle of joint insulation used. The three categories of cable joints and the related share of failures, as described in section 2.3.2, are:

- Mass insulated joints (57% of internal defect related failure share)
- Oil insulated joints (25% of internal defect related failure share)
- Synthetic insulated joints (18% of internal defect related failure share)

Besides for information regarding the cause of failure and the number of joints failed, additional information about the age of the joints at the moment of failure is available. Most of the time, the exact age of the joints at the moment of failure is not known to the utility. To circumvent this problem, rough estimations of the age are reported by using age intervals (age bins). The total number of reported failures for the period 2004 until 2009 for each category of 10 kV cable joint together with the reported age intervals are shown in figure 4.4.

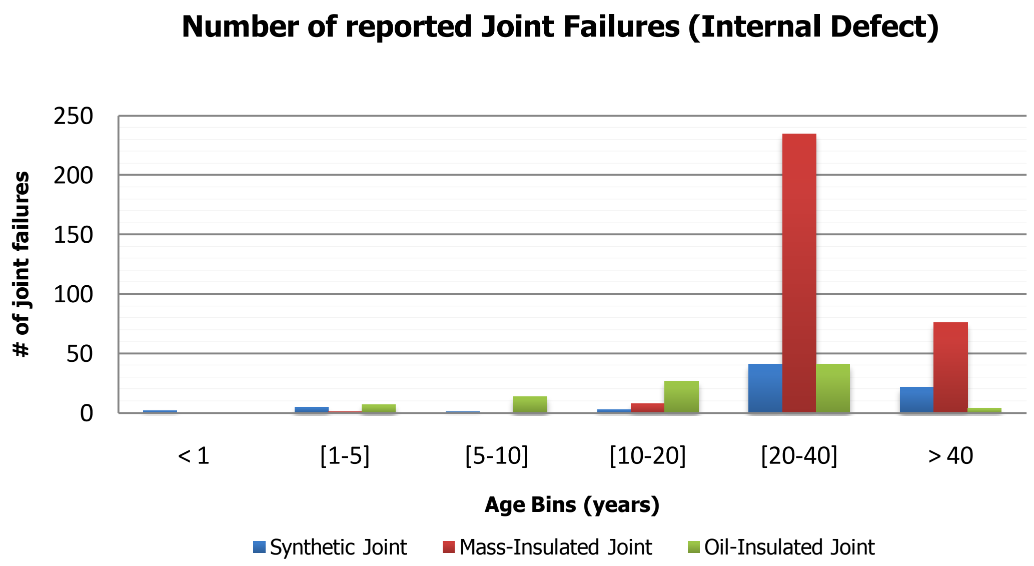


Figure 4.4: 10 kV joint failure records for the period 2004-2009 for three categories of cable joints. As result of unknown exact age at the moment of failure of a component, age intervals are used to estimate the age of the failed components

It becomes evident from figure 4.4, that interval-censored data is available and should be taken into account when performing the failure distribution parameter estimation. This is due to the fact of the age bins, which are recorded to estimate the age of the joint at the moment of failure.

In order to perform statistical analysis, the data should have properties such as, independency and homogeneity. However, due to the lack of detailed recorded information, it is necessary to make certain assumptions. From a statistical point of view, it is necessary to assume that the cable joint constructions are comparable with each other and operated under equivalent conditions (i.e. the same load and ambient temperature).

4.2.2 Available In-Service Data

Prior to the mergers of small network companies, different database systems were used for collecting and keeping asset information. Nowadays, the intention is to move the different database systems to more centralized asset information inventories. The information required for this study, particularly the in-service data for the area "Region X", is available in a data inventory system named TKV[2]. In this system, asset specific information can be found. The features recorded for the joint data in the TKV asset inventory database are used for analyzing the age distribution of the joints populations.

Although Stedin is presently keeping complete and accurate asset installation data in dedicated databases, such records are often missing for assets that were installed more than 20 to 30 years ago. Added to this, as mentioned earlier, the various mergers of smaller regional network operators to single distribution operators has caused losses of data records. While tracing back the lost data is a challenging task, the regulators however are forcing the utilities to manage and improve the asset data inventories more accurately. In this context, Stedin is applying data enhancement methods to improve asset inventory. Unfortunately, these data quality enhancement projects are still in progress and the results were not yet available for the present study. To circumvent this problem, assumptions were often made to compensate for the missing data. These assumptions were based on expert knowledge from the utility itself.

The features captured for the joint data in the TKV database are used for analysing the age distribution of the joint population. However, for a large portion of the joint population the exact age (year of installation) is not specified, as result of missing data records. Likewise, for some part of the joint population the corresponding joint type is unknown.

The first shortcoming is dealt with by dividing in terms of percentages, the number of joints without age proportionally and adding these joint to the joints installed in particular years. This way a new estimation of the number of joints with a particular age is achieved. A formula related to this procedure is:

$$New\ \#\ of\ joints_{age,i} =$$

$$\left(\frac{recorded\ \#\ of\ joints_{age,i}}{total\ \#\ of\ joints\ with\ age} \times Total\ \#\ of\ joints\ without\ age\right) + \ recorded\ \#\ of\ joints_{age,i} \qquad (4.1)$$

With *age,i* being the age in years.

The second shortcoming is dealt with by using information, based on expert knowledge, regarding the historic application of certain joint types. These experts still have knowledge regarding the history of when a certain type of joint was taken into operation. Using the expert knowledge, a rough estimation of the type of joints, for the unknown group, can be made.

[2] TKV: Ten Kilo Volt system (in Dutch: 10 kV systeem)

Table 4.1 illustrates the amount and quality of the available data for each category of 10 kV cable joints population.

Table 4.1: Installed number of joints for 3 categories of joint types. Also shown are the number of joints of which the age is not specified and the type of insulation material is unknown

	Joint Data from TKV database			
Joint type	*# of joints without specified year of installation*	*# of joint with specified year of installation*	*Total number of recorded joints*	*Percentage of "bad" data*
Synthetic insulated	10	754	764	***1.3 %***
Mass insulated	8295	5351	13.646	***60.8 %***
Oil insulated	785	14.525	15.311	***5.1 %***
Unknown	994	803	1797	***55.3 %***

Using the installation records which constitute the historical population data together with the appropriate estimated values for missing data, the age distribution of the overall cable joint populations can be depicted. This is illustrated in figure 4.5 for mass and oil insulated joints. In figure 4.6, the age distribution for the synthetic insulated joint population is shown. Both figures correspond to the area "Region X".

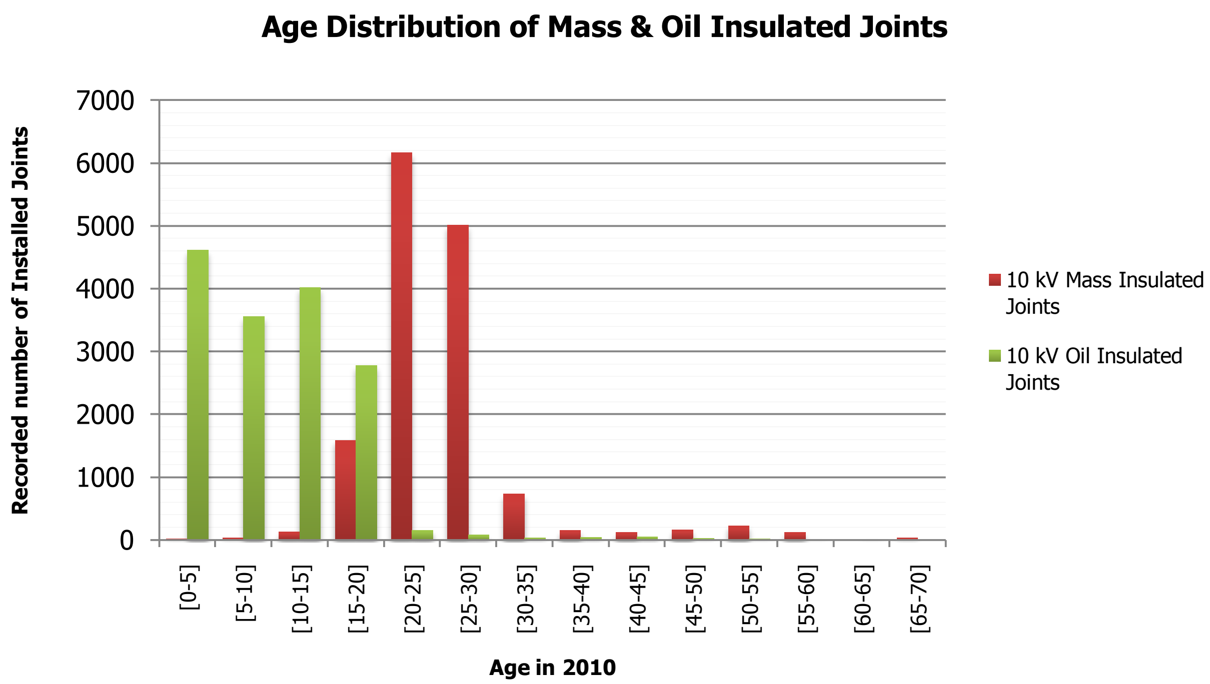


Figure 4.5: Age distribution in years of 10 kV mass & oil insulated cable joints population

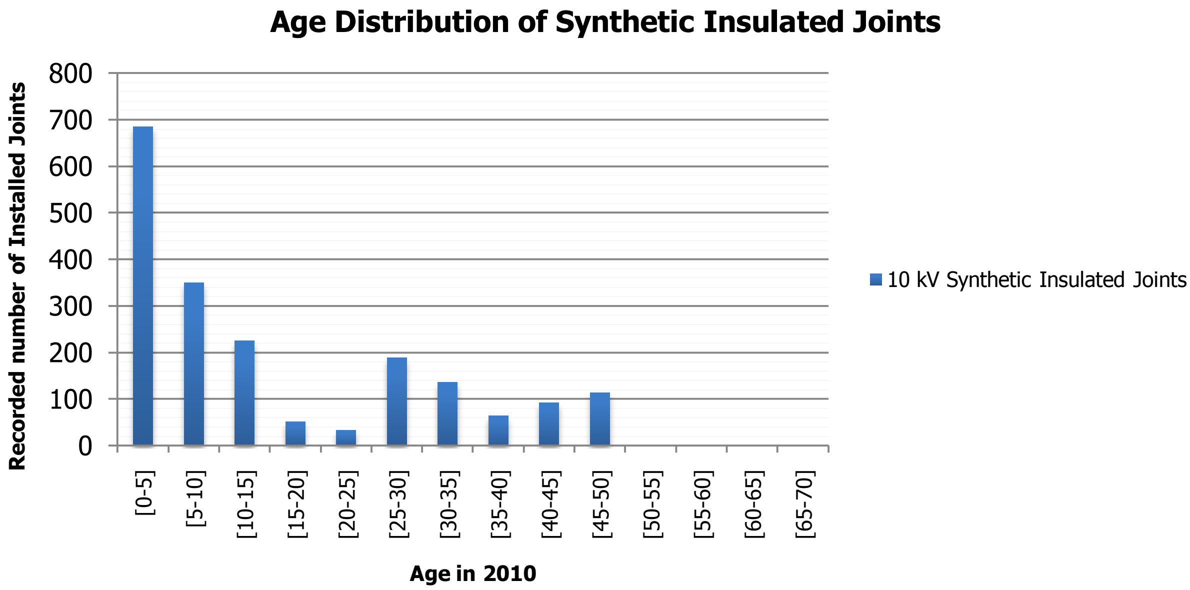


Figure 4.6: Age distribution in years of 10 kV synthetic insulated cable joints population

Figure 4.5 shows that a large portion of the mass insulated 10 kV joint population is older than 20 years. The oil insulated joints are still used frequently and the population is concentrated between 5 and 20 years. Figure 4.5 also reflects that mass insulated joint are not used anymore in "Region X". The policy of Stedin in this area is to replace failed mass insulated joint with oil insulated joints. The numbers of installed mass and oil joints are larger than synthetic joints. The reason for this is that paper-insulated lead-covered (PILC) cables were, historically, applied for the majority of MV cable networks. Mass and oil insulated joints are usually used with PILC cables. Although polymeric cables have already been available since the early 1950s, it was only until several tens of years later that they could commercially compete with PILC cables [15]. Polymeric cables (usually XLPE cables) are often jointed by using synthetic insulated joints. From figure 4.6 it can be seen that the majority of synthetic joints are between 5 and 15 years old.

The failure data, as well as the population still in service, are used as input for the statistical analysis. This analysis is carried out in the following section.

4.3 Statistical Life Data Analysis

4.3.1 Case 1: 10 kV Synthetic Insulated Cable Joint

The parametric distribution fitting method, which is described in chapter 3, is used for modelling the life data of the 10 kV synthetic insulated joints. The proposed procedure which is described in section 3.3.1 is followed. In this case study, the failure data and in-service data for synthetic insulated joints is used.

4.3.1.1 Distribution Parameter Estimation & Goodness-of-Fit Test

The reliability software tool, Weibull++ 7, is used for performing the parametric distribution fitting analysis. The Maximum Likelihood Estimation (MLE) is used for estimating the parameters of the failure distributions. During the life data analysis, Weibull++ 7 provides guidance in selecting a distribution based on statistical tests. The program uses three factors in order to rank distributions. These factors include

1. the Kolmogorov-Smirnov (K-S) test,
2. a correlation coefficient (ρ)
3. the likelihood value (LKV)

In order to determine the ranking, these three tests are used in conjunction with weights assigned to each test. Each one of the above tests are weighted and combined into an overall value (OV), as shown in equation 4.2:

$$OV = (KS\ Rank \times KS\ Weight) + (\rho\ Rank \times \rho\ Weight) + (LKV\ Rank \times LKV\ Weight) \tag{4.2}$$

The weight (or importance) assigned to each test can be defined by the user and depends on whether the parameter estimation method is rank regression or MLE. Once all test results have been calculated for each distribution, distributions are ranked for each test. From a statistical point of view, the distribution that is ranked first is the best fit for the data.

This method as described above will be used for analysing the life data of the 10 kV synthetic insulated cable joints. From the analysis, it was found that more than one distribution seem to represent the data. In such situations, engineering knowledge is used together with the statistical test results to select an appropriate distribution.

Table 4.2: Analysis details (Goodness-of-Fit) results for two competing distribution. With the K-S Test and the Normalized Correlation Coefficient, the higher the number, the worse is the fit. With the likelihood value, the closer the value to zero, the better the fit

	K-S Test	***Normalized Correlation Coefficient***	***Likelihood Value***
Normal Distribution	99.994	7.326	-232.638
2-Parameter Weibull Distribution	97.844	11.214	-214.004

Generally, some rule of thumbs can be applied to judge the values of table 4.2. According to [21], the values in table 4.2 in themselves for each test do not mean much. These values are only useful when they are used for comparing two models of the same population. When the values for the K-S test and the Normalized Correlation Coefficient are high numbers, it is an indication that the distribution is not the best fit, statistically. In case of the Likelihood Value, the closer the value is to zero, indicates that the fit to the data set is better. From table 4.2 it can be seen that the 2-parameter Weibull distribution forms a better fit to the data set in compliance with the K-S test and the Likelihood Value. A visual assessment, shown in figure 4.7, of the failure rate plots reveals that the two models, in fact, have no significant difference for the time interval of interest (0 and 50 years).

Figure 4.7 shows a visual comparison of the normal distribution and the 2-parameter Weibull distribution. It is noticeable that the two models differ after roughly 55 years, after which the failure rates for the Weibull distribution are higher. However, from figure 4.6 it becomes clear that the oldest population of in-service synthetic joints are 50 years. Based on this, the choice for the Weibull distribution would not have significant impact on further analysis of the joint population. Furthermore, according to reference [29], Weibull distributions are found to represent breakdowns in polyethylene power cables insulation & accessories. Resulting from these two arguments, it is chosen to select the 2-parameter Weibull distribution for this case.

It is important to note, as seen in this case, that visual verification is very useful in choosing the best fitting distribution. More specifically, if the there is little difference between the distributions, then visual inspections are necessary to make the final decision.

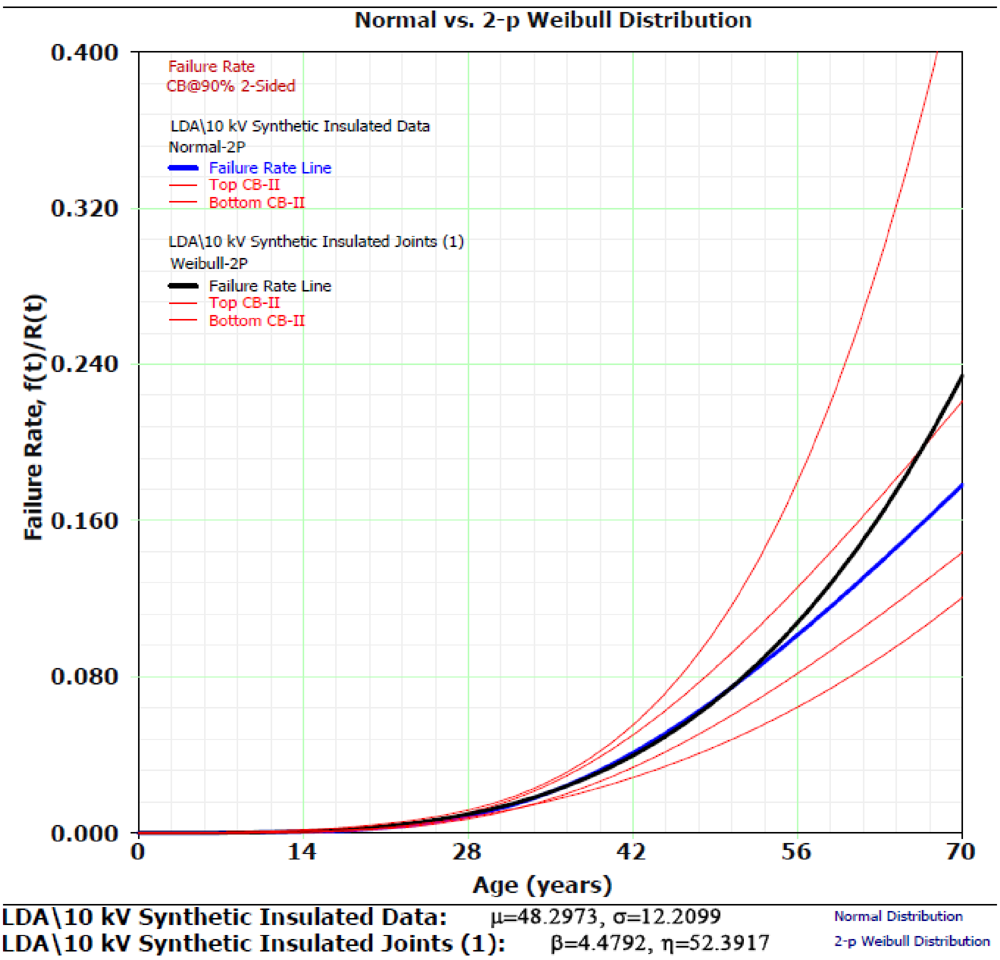


Figure 4.7: Visual comparison of the failure rate curves for the same data set. The normal distribution is compared with the 2-parameter Weibull distribution. After 56 years the failure rates start to differ from each other

4.3.1.2 Statistical Analysis

Based on the selected failure model, it is possible to describe the behaviour and characteristics of the 10 kV synthetic insulated cable joint population in terms of failure probability, failure rates and other related variables. The distribution functions for the 2-parameter Weibull distribution are used for calculating the related failure probabilities.

The fitted Weibull distribution results in a value for the shape parameter of *β=4.48* and for the scale parameter (characteristic life) of *η=52.39*. The high value of the shape parameter, β>1, indicates that ageing is a predominating cause of failure. Additionally, a high shape parameter also indicates that failures are occurring close together, what simply implies a low spread (low variability) in the failures. This should not be a problem as long as the onset of the failures is relatively far on the time scale (scale parameter). This onset of the failures is related with the value of the scale parameter.

The corresponding probability plots for the 2-parameter Weibull distribution of the synthetic joints are shown in figures 4.8, 4.9 and 4.10 for the cumulative distribution function *(cdf),* probability density function *(pdf)* and failure rate function, respectively.

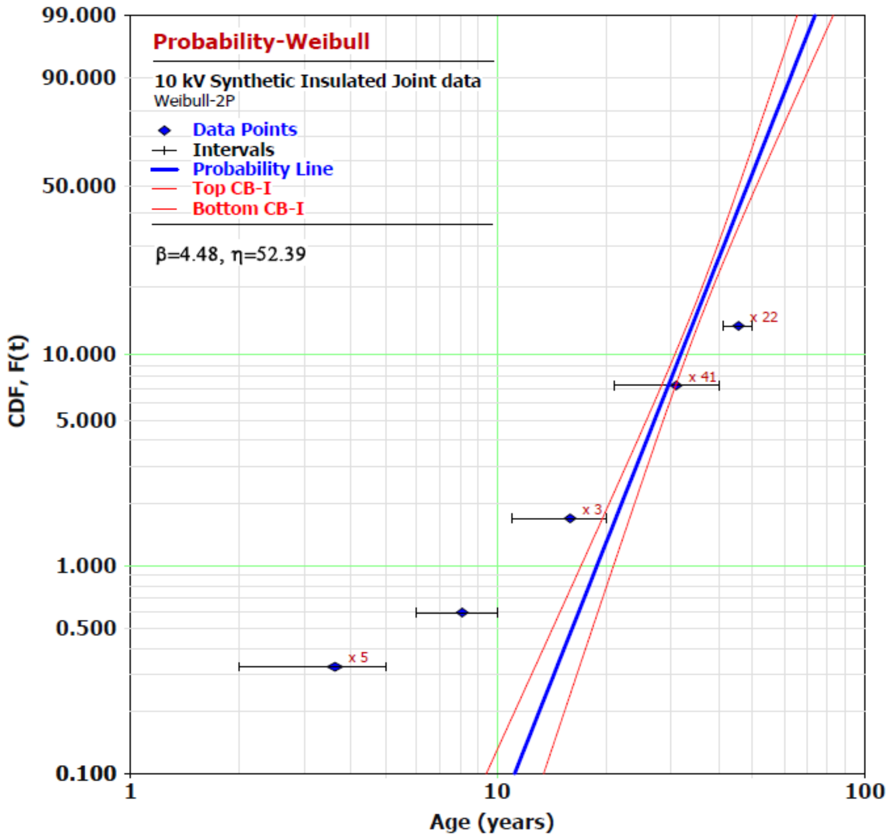


Figure 4.8: Probability plot of the 2-parameter Weibull Cumulative Distribution Function (cdf) for synthetic insulation joints with its corresponding 90 % confidence bounds (red lines).
blue line: probability line
red lines: upper and lower 90% confidence bound
blue points: ranked failure data points

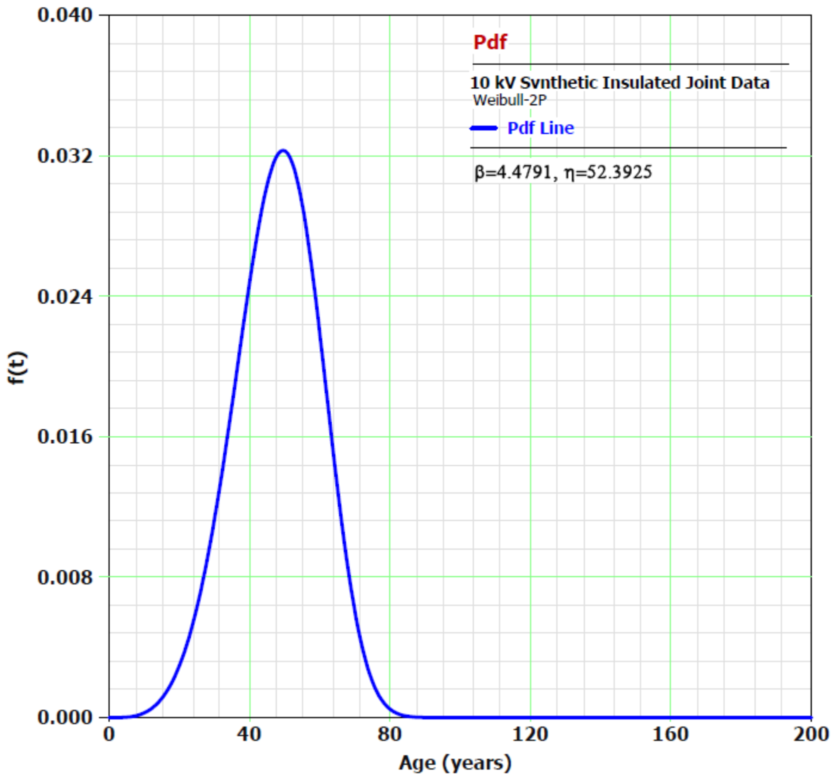


Figure 4.9: Probability density function (pdf) for the 2-parameter Weibull distribution. The probability that a synthetic insulated cable joint will fail within a certain range of age is indicated by the pdf plot

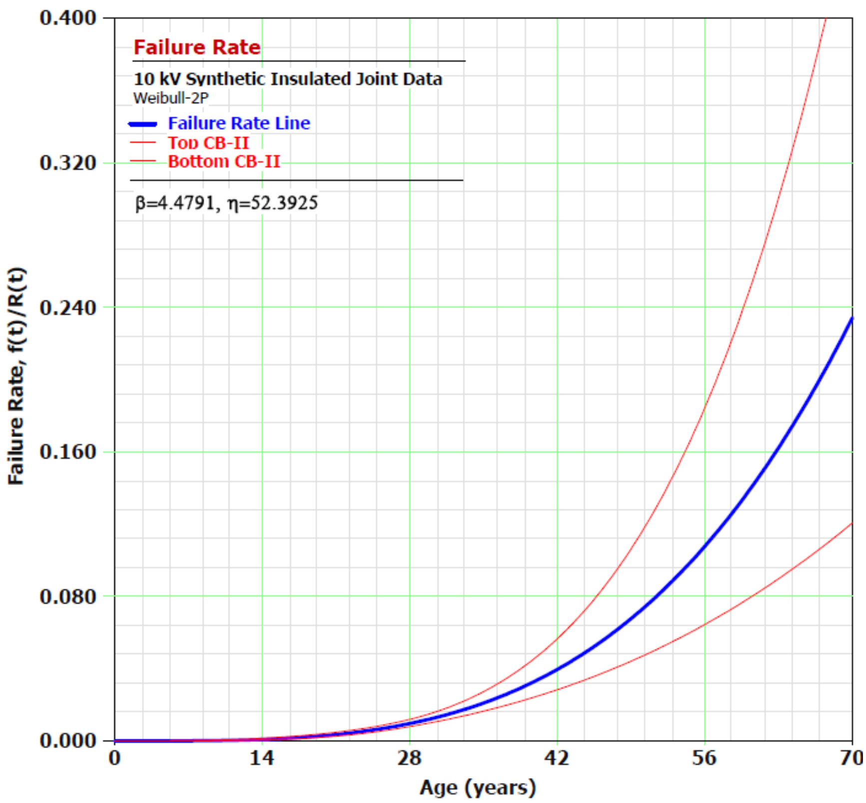


Figure 4.10: 2-parameter Weibull Failure Rate Function for the synthetic insulated cable joint population. The corresponding 90 % confidence bound are shown in red. From approximately 28 years the failure rate starts to increase steeper

The *cdf* plot shown in figure 4.8 refers to the probability that a synthetic insulated cable joint with a certain age will fail within a specific range. The statistical precision of the analysis is indicated in figure 4.8 by the corresponding confidence bounds. In figure 4.8, the 90 % confidence bounds are shown, which are relatively narrow because of the large amount of data.

It should be noted, that the data points shown in figure 4.8 do not seem to fit the related *cdf* curve. The reason for this is that when using MLE to draw a probability plot, the points are placed using a ranking method, while the line is drawn by using the estimated parameters (which are estimated with the MLE method). The MLE method is a numerical method, not a graphical method. This phenomenon will especially occur when the given data set contains many suspensions. Detailed explanations are outside the scope of this research, but for this report, it is important to note that when a line plotted using MLE does not seem to fit the ranked data points on the plot this does not necessarily imply a bad fit. Detailed information can be found in [21] [30].

From the *pdf* plot shown in figure 4.9, the expected average lifetime of the synthetic insulated cable joint population can be estimated. This expected average lifetime, also called mean life, is calculated and estimated to be 48 years with an upper confidence bound of 51 years and a lower confidence bound of 45 years.

Furthermore, the failure rate curve depicted in figure 4.10 indicates that the failure rates of synthetic insulated cable joints are increasing as the population gets older and eventually this failure rate starts to get steeper. This resembles the wear-out period of the bathtub curve. Despite the fact that the failure rate is increasing, it can be seen that the relative increase is slow.

4.3.2 Case 2: 10 kV Mass Insulated Cable Joint

Again, the parametric distribution fitting method is used for modelling the life data of the 10 kV mass insulated joints. The proposed procedure which is described in section 3.3.1 is followed. In this case study, the failure data and in-service data for mass insulated cable joints is employed.

4.3.2.1 Distribution Parameter Estimation & Goodness-of-Fit Test

From the life data analysis for the mass insulated cable joints, it is found that the lognormal distribution and the 2-parameter Weibull distribution are competing with each other. Taking into account the values of the statistical test (Goodness-of-fit test) together with engineering knowledge, the most suitable failure model is selected.

Table 4.3: Analysis results for two competing failure models based on the data set for mass insulated cable joints. With the K-S Test and the Normalized Correlation Coefficient, the higher the number, the worse is the fit. With the likelihood value, the closer the value to zero, the better the fit

	K-S Test	**Normalized Correlation Coefficient**	**Likelihood Value**
Lognormal Distribution	97.89	2.26	-1056.08
2-Parameter Weibull Distribution	80.85	1.66	-1110.16

From table 4.3, it can be seen that the 2-parameter Weibull distribution forms a better fit to the data set in compliance with the K-S test and the Normalized Correlation Coefficient. Resulting from this, the 2-parameter Weibull distribution is selected as failure distribution for the gives data set of the mass insulated cable joints.

4.3.2.2 Statistical Analysis

The selected failure distribution, namely the 2-parameter Weibull function, is used to describe the behaviour and characteristics of the 10 kV mass insulated cable joint population in terms of failure probability, failure rates and other related variables. The distribution functions for the 2-parameter Weibull distribution are used for calculating the related failure probabilities.

The fitted Weibull distribution results in a value for the shape parameter of $\beta=4.93$. The scale parameter (characteristic life) is $\eta=69.95$. The characteristic life, η, is the lifetime which on average 36.8% of the components under consideration would reach. This was discussed in chapter 3. The high value of the shape parameter, $\beta>1$, indicates that ageing is a predominating cause of failure. Additionally, a high shape parameter also indicates that failures are occurring close together, what simply implies a low variability in the failures. The corresponding probability plots for the 2-parameter Weibull distribution of the synthetic joints are shown in figures 4.11, 4.12 and 4.13 for the cumulative distribution function *(cdf)*, probability density function *(pdf)* and failure rate function, respectively.

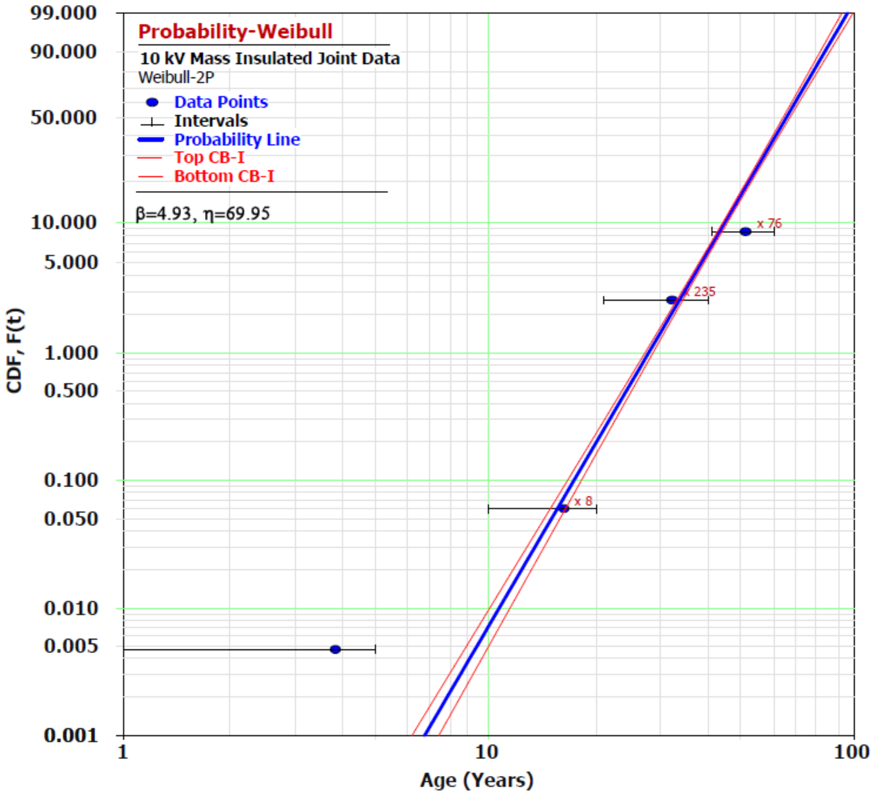


Figure 4.11: Cumulative distribution function (cdf) for the 10 kV mass insulated cable joint population. The 90 % confidence bounds are shown in red. It can be seen that the confidence bounds are narrow, because of the large amount of data

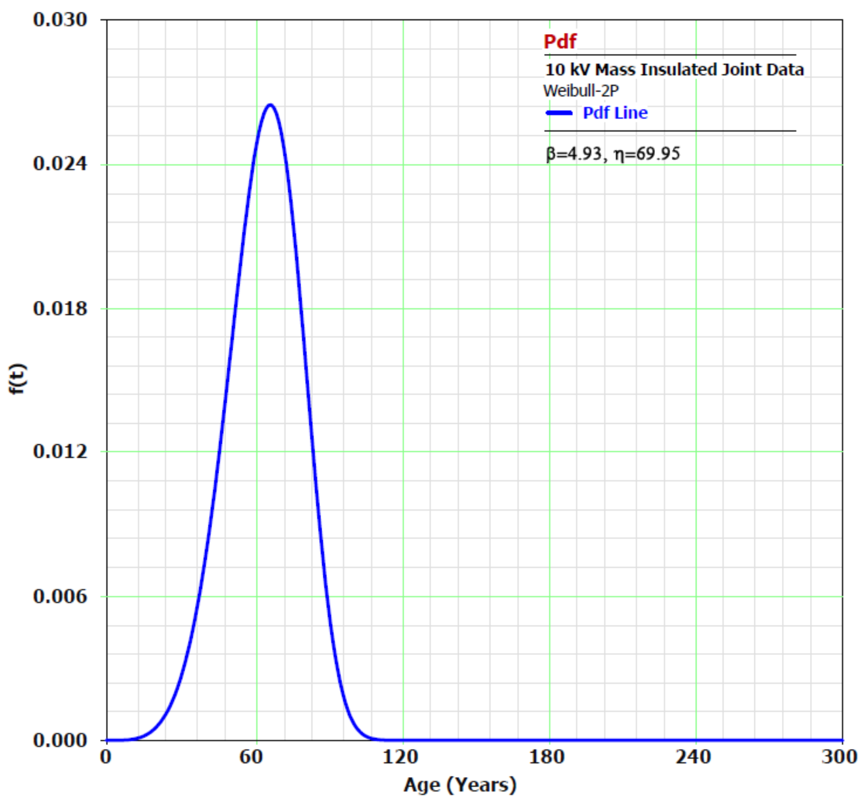


Figure 4.12: Probability density function (pdf) for the 2-parameter Weibull distribution. The probability that a mass insulated cable joint will fail within a certain range of age is indicated by the pdf plot

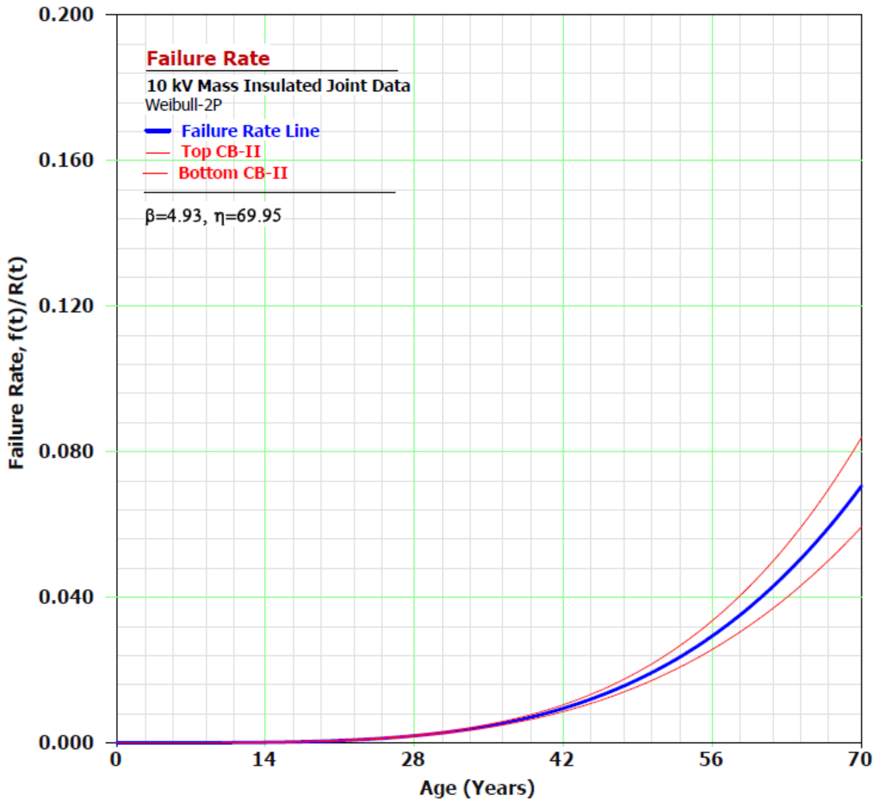


Figure 4.13: The failure rate curve for the 10 kV mass insulated cable joint population. The 90 % confidence bounds are relatively narrow. Starting at 28 years the failure starts to increase slowly

From the *pdf* plot depicted in figure 4.12, the expected average lifetime of the mass insulated cable joint population can be estimated. This expected average lifetime, also called mean life, is calculated and estimated to be 64 years with an upper confidence bound of 66 years and a lower confidence bound of 63 years. As can be seen from this, the characteristic life is not the same as the expected average lifetime. The characteristic life, η, is the lifetime which on average 36.8% of the components under consideration would reach, as mentioned earlier.

4.3.3 Case 3: 10 kV Oil Insulated Cable Joint

The parametric distribution fitting method is used for modelling the life data of the 10 kV oil insulated joints. The proposed procedure which is described in section 3.3.1 is followed. In this case study, the failure data and in-service data for oil insulated cable joints is employed.

4.3.3.1 Distribution Parameter Estimation & Goodness-of-fit Test

The life data for the 10 kV oil insulated cable joint population is fitted by means of statistical distributions. The results of the Goodness-of-fit Tests are used to determine the best fit. In this case, the normal distribution forms a better fit to the data set in compliance with the K-S test, the Normalized Correlation Coefficient and the Likelihood value. In table 4.4 a comparison between the Goodness-of-fit Test for the normal and 2-parameter Weibull distribution is made. Therefore, the normal distribution is selected as failure distribution for the gives data set of the oil insulated cable joints.

Table 4.4: Goodness-of-fit Test results for the normal distribution and the 2-parameter Weibull distribution based on the data set for oil insulated cable joints. With the K-S Test and the Normalized Correlation Coefficient, the higher the number, the worse is the fit. With the likelihood value, the closer the value to zero, the better the fit

	K-S Test	***Normalized Correlation Coefficient***	***Likelihood Value***
Normal Distribution	4.13E-8	1.5113	-449.4378
2-Parameter Weibull Distribution	0.605	4.3889	-450.1114

4.3.3.2 Statistical Analysis

The selected failure distribution, namely the normal function, is used to describe the behaviour and characteristics of the 10 kV oil insulated cable joint population in terms of failure probability, failure rates and other related variables. The distribution functions for the normal distribution are used for calculating the related failure probabilities.

The estimation of the parameters of the fitted normal distribution results in a mean value of $\mu= 54.03$ and standard deviation $\sigma= 15.02$. The corresponding probability plots for the 2-parameter Weibull distribution of the synthetic joints are shown in figures 4.14, 4.15 and 4.16 for the cumulative distribution function *(cdf),* probability density function *(pdf)* and failure rate function, respectively.

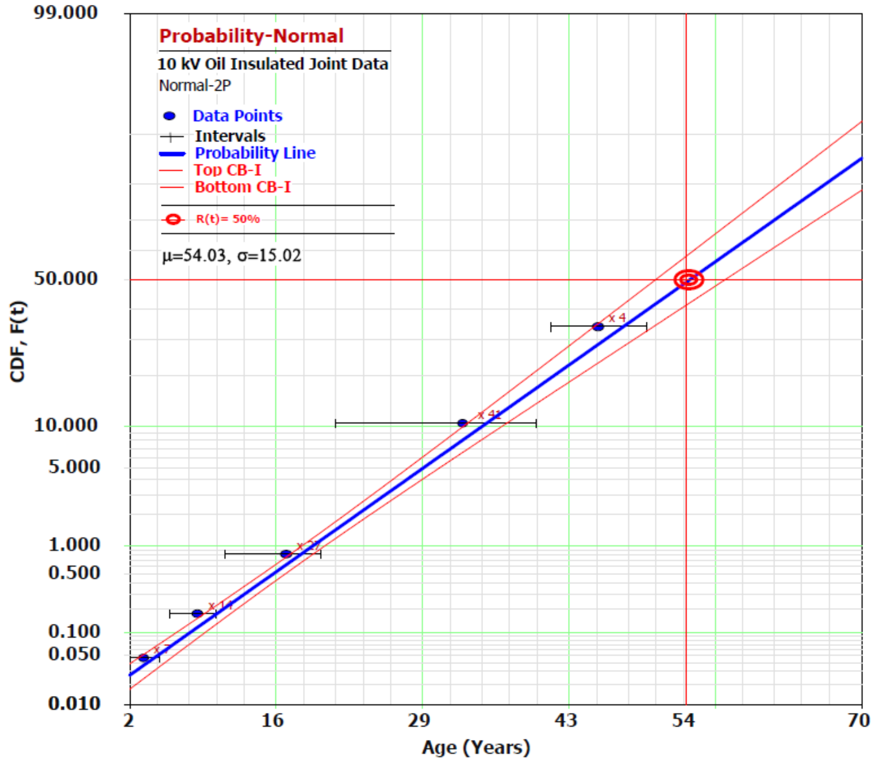


Figure 4.14: The cumulative distribution function (cdf) for the 10 kV oil insulated cable joint based on the normal distribution. The value of μ is said to be the point where R(t) = 50 %. This means that the estimate of μ can be read from the point where the plotted line crosses the 50 % cdf line

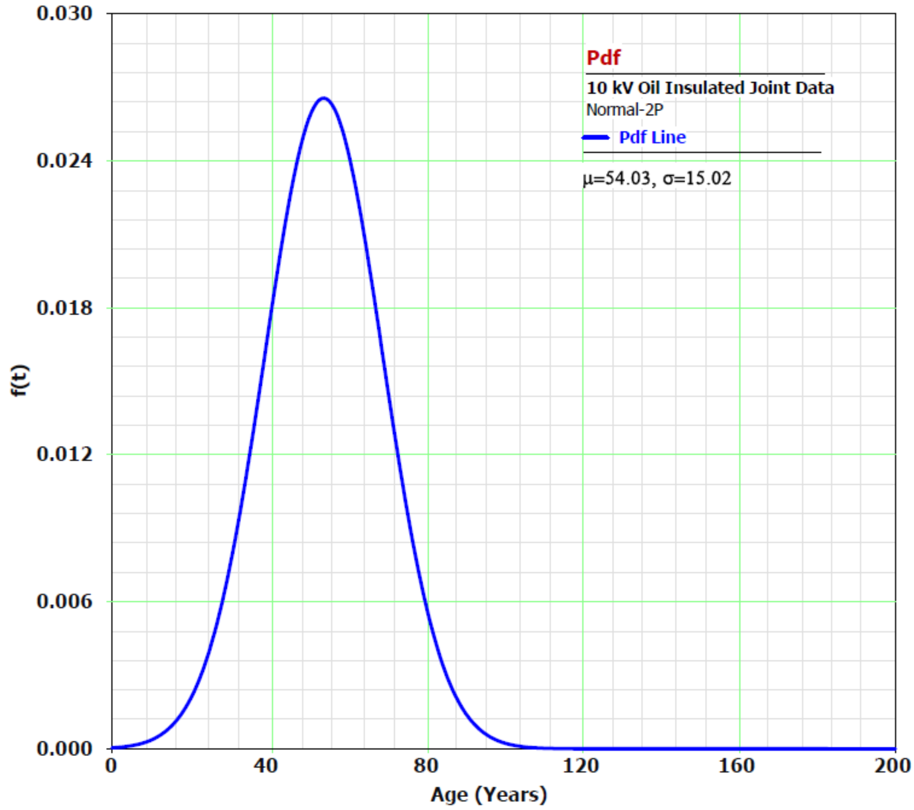


Figure 4.15: Probability density function (pdf) for the normal distribution. The probability that an oil insulated cable joint will fail within a certain range of age is indicated by the area under the pdf plot.

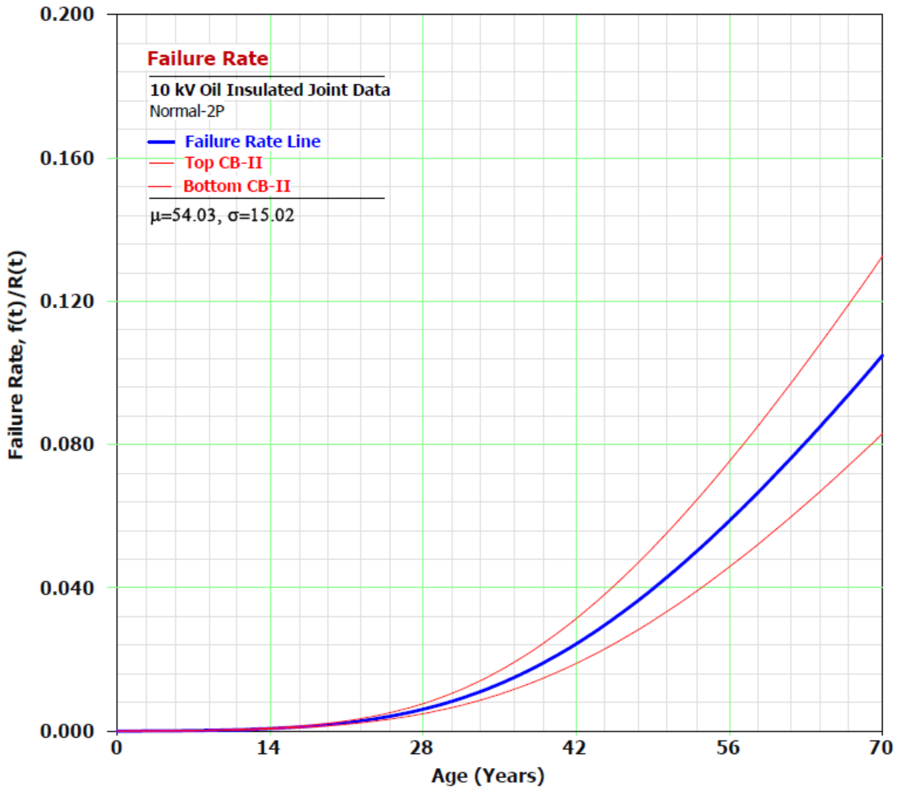


Figure 4.16: The instantaneous normal failure rate for the oil insulated cable joint population. The corresponding 90 % confidence bound are shown in red

From the normal *pdf* plot, depicted in figure 4.15, the expected average lifetime, also called mean life, μ, is 54 years. The upper limit for the mean life is 57 year, while the lower limit is 51 year. The mean life, μ, is also the location parameter of the normal *pdf*, as it located the *pdf* along the x-axis. The normal *pdf* has no shape parameter. This means that the normal *pdf* has only one shape, the bell shape. The standard deviation, σ, is the scale parameter of the normal *pdf*.

4.3.4 Comparison of the three types of cable joint populations

Generally, for electrical components, the failure rate and the *pdf* are the most important criteria besides the failure time [31]. The failure rate and *pdf* allow electrical components of different assets to be compared with each other. In this context, the failure rates and *pdf* of the three types of cable joint populations are compared with each other. In order to compare the *pdf* curves, the calculated results for the three different types of 10 kV cable joints are shown together in figure 4.17. Subsequently, the calculated failure rates of the different cable joints are shown in figure 4.18.

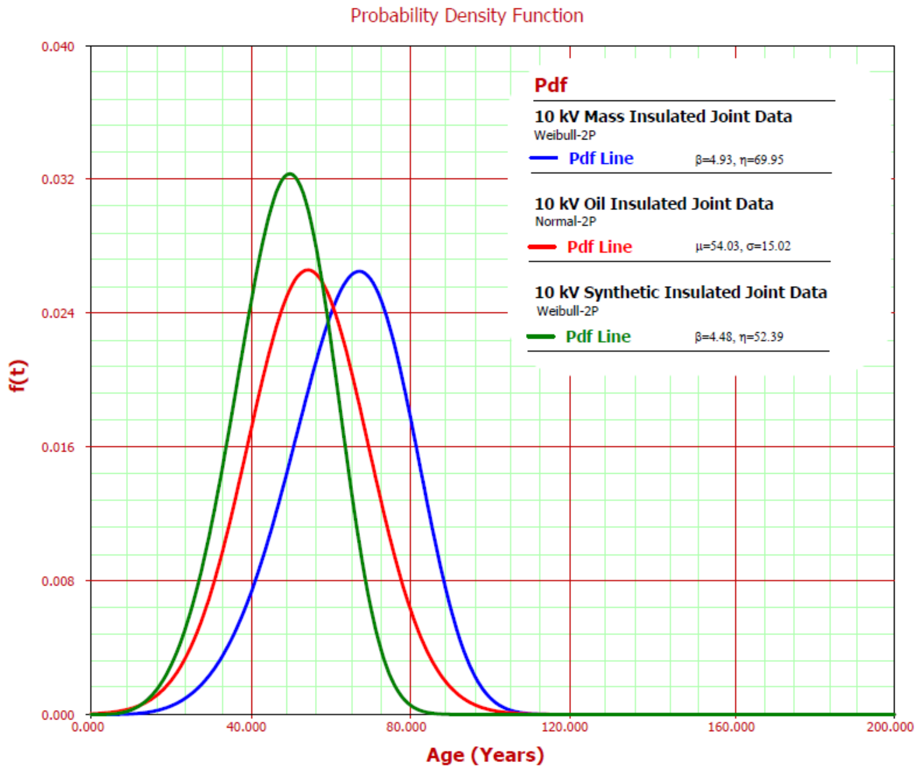


Figure 4.17: Probability density functions (pdf) for three different types of 10 kV cable joints. Green pdf line – Synthetic Insulated Cable joint; Red pdf line – Oil Insulated Cable joint; Blue pdf line – Mass Insulated Cable joints

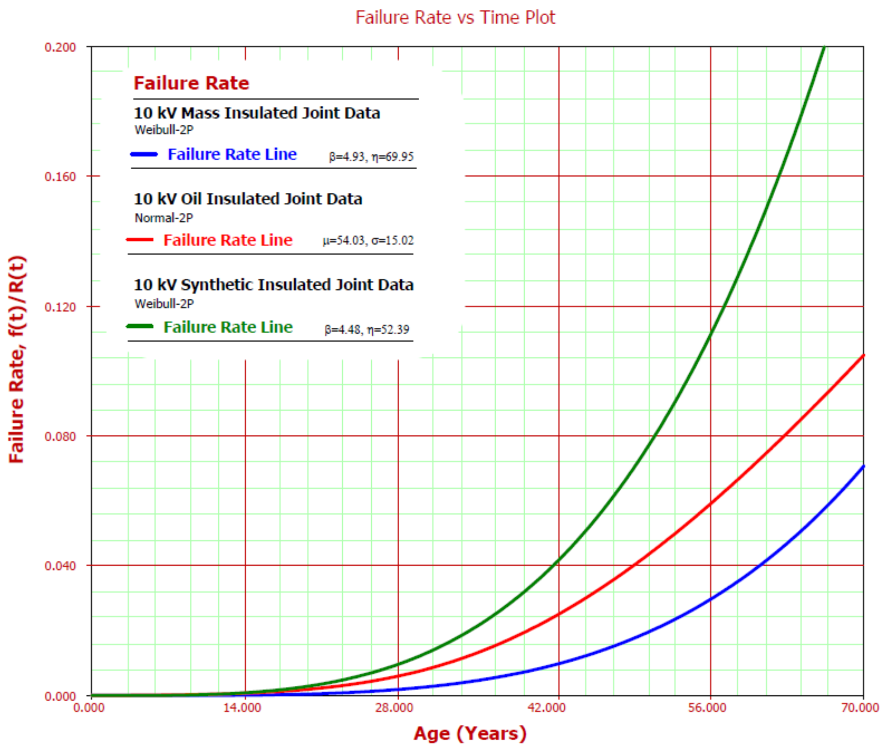


Figure 4.18: Failure rate curves for the three different types of cable joint populations. The failures can be compared with each other. For all three populations their failure rates rise over years according to the increasing right wing of the bathtub curve. However, each joint population shows different behaviour

From figure 4.17, the density of failure probability can be examined for the three types of cable joints. The peak value of the density of failure probability for the synthetic insulated cable joint is higher than the ones for the mass and oil insulated cable joints. Typically, this illustrates that synthetic insulated cable joints have a higher probability of failure when the components age is near the peak value. Furthermore, it can be seen from figure 4.17 that the probability of failure for the mass insulated cable joints is lower than for the synthetic and oil insulated cable joints. For instance, from figure 4.17, the failure probabilities of the cable joints for the first 10, 20 and 30 year of operation are listed in table 4.5.

Table 4.5: Gives the probability of failure for a specific mission time for three cable joint types

Type of cable joint	Failure prob. for 10 years	Failure prob. for 20 years	Failure prob. for 30 years
Synthetic insulated	*0.06 %*	*1.33 %*	*7.90 %*
Mass insulated	*0.0069 %*	*0.21 %*	*1.53 %*
Oil insulated	*0.17 %*	*1.17 %*	*5.47 %*

Figure 4.18 depicts the failure rate over time for three different cable joint types. It can be seen that the failure behaviour is different for each population. Additionally, it can be seen that the populations get older quite similarly; however, the rate of rise of the failure rate with equipment age differs slightly from each other. Furthermore, as shown in figure 4.18, the failure rates of the three cable joint populations are constantly low during the first several years of operation. In accordance with a typical wear-out failure mechanism, the failure rates will grow as the joints ages up to 20 years. Hence, the failure rate plots illustrate the increasing characteristic so that the distinct dependency on age is observable.

In the case of synthetic insulated joints, it can be seen that the failure rates are higher at earlier ages. This can be as result of improper joint installation and the influence of this on early failures. Improper assembly of cable joints can cause defect, and therefore may result in breakdown on the mid-long term.

Subsequently, the failure behaviours of oil insulated joints and of mass insulated joints differ from each other, even though they belong to the group of filled joints. An important reason why the failure rates for the oil insulated cable joints are higher can be the result of lower liquid levels in oil type joints. As mentioned in section 2.2.3.1, a lowered liquid level in joints filled with viscous material is often due to thermal heat cycles as result of the daily load cycles. Basically, a lowered liquid levels results in an impaired breakdown strength, which leads to partial discharges and finally to a breakdown.

Finally, it is very important to note that the failure probability plots shown in this section are based on calculation with incomplete failure data. The historical failure data used for the joints is a subset of the total number of failures. Historic failure data for 6 years have been used. At the same time, assumptions that are based on utility specific knowledge are made for the age distribution of the population which is still in service. Therefore, the failure rates obtained here are conservative values. In practice, the failure rates will tend to be higher than those implied by the models, since not all historical failures are reflected in this analysis.

4.5 Conclusions

This chapter has discussed the application of life data analysis for practical data. The analysis was carried out for three different types of 10 kV cable joints used in the MV distribution cable network of an area named "Region X". It was found that this area has the highest failure frequency for cable joints of this particular voltage level. Added to this, the quality of the available information, which is necessary for life data analysis, was found to be in relatively better condition when compared to the other areas.

Generally speaking, there are two types of data required for performing life data analysis, namely, failure data and suspensions (survived components). It was found that failure data for the three types of cable joints are completely available for a period of six years (2004-2009). Prior to that period, the databases have undergone many chances and many useful data was lost or incomplete. It can be concluded, that in most of the cases the exact age of a cable joint at the moment of failure is unknown to the repairman in the field. For that reason, estimations are made for the age by means of age intervals. Furthermore, it was found that information with reference to cable joints that are still in operation have discrepancies. Most of the time, the year of installation, hence age, of the cable joints is not specified. Especially, for mass insulated cable joints, a large portion (approximately 60 %) of the year of installation is not specified.

However, by using appropriate estimation techniques based on expert knowledge within the utility, much of the missing data can still be reasonably estimated to be used in the analysis. At the same time, it should be mentioned that Stedin has launched a data quality enhancement project in order to improve the missing data. This is necessary, because as a consequence of ageing workforce problems, it will be hard in the future to find experts with knowledge of installation records of old asset populations.

The collected failure data and survived unit data is used for performing the parametric distribution fitting analysis. From the analysis, it was found that the failure distribution with the best fit for:

- the synthetic insulated cable joint population was the 2-parameter Weibull distribution
- the mass insulated cable joint population was the 2-parameter Weibull distribution
- the oil insulated cable joint population was the normal distribution

When the probability plots for the three populations of 10 kV cable joints were compared with each other, it was found that the synthetic insulated joints have the highest failure rate, followed by the oil insulated joints and finally by the mass insulated joints. At the same time, we can conclude that the failure rates will grow as the joints ages up to 20 years in accordance with typical wear-out failure mechanisms. Before this time, the failure rates for the cable joints are constantly low. This suggests that the failure rates will tend to vary in different ways for each type of cable joint. When taking into account the probability density function *(pdf)*, the expected average lifetime of the joints can be estimated. For the synthetic insulated joints the expected average lifetime was estimated to be 48 years, for the mass insulated joints 64 years and for the oil insulated joints 54 years.

Fundamentally, the life data analysis performed in this chapter has illustrated that by applying appropriate statistical tools, large amounts of information, characteristics and behaviours of a certain distribution can be described by small numbers of parameters. Subsequently, the probability functions allow different assets to be compared with each other, and make reference to criteria like age, rate of rise of failure rates and mean life of components.

With the results of the failure rate function and the population of joints still in operation, the expected number of failures can be calculated. This will be covered in the next chapter. In this way, it is possible to estimate the failure behaviour of the joint populations and to verify the developed failure probability models.

5
AM Decision Support for Cable Joint Failures

In chapter 4, the application of parametric distribution fitting for three different types of 10 kV cable joint life data was described. Probability models were developed for each joint population. Based on the selected failure probability model, we found that it was possible to describe the behaviour and characteristics of the joint population in terms of failure probability, failure rates and other related variables. In this chapter, the developed failure models are used to facilitate the asset manager with tools in order to make sound decisions.

Firstly, section 5.1 starts with the background of possible tools which can be used to support the asset manager decision process, regarding the failure behaviour of a given population of components. Topics such as B(x)-lives, forecasting future failures and the failure count diagram will be described in this section.

Subsequently, in section 5.2, the B(x)-lives, forecasting future failures and the failure count diagrams are analyzed for all three types of 10 kV cable joint populations. In this section a subsequent sensitivity analysis is carried out for the synthetic insulated joints. From this sensitivity analysis the usefulness of the probability models to assess certain suspects groups of joints is shown. Furthermore, having performed these different failure assessments, more knowledge with regard to failure probability, ageing and failure frequency at a certain age will be created.

Finally, in section 5.3, conclusions considering the key findings of this chapter are described.

5.1 Background

5.1.1 Asset Management Support

Asset Management (AM) relies heavily on the use of information and data to facilitate the decision making process [2]. Even when databases are found to have missing or incomplete data, it is still possible to develop sensible failure probability models. Aspects, such as, failure probability, ageing and failure frequency are important [32]. Knowledge and information of these aspects can contribute in the decision process of AM. With the results of the statistical analysis from chapter 4, information regarding the failure probability and failure frequency at a certain age of asset groups in the near future of the three types of 10 kV cable joints can be extracted. As a verification, the actual failures that occurred over the past 6 years (period from 2004 to 2009 will be named verification period for the remainder of the text) can be used to assess whether the developed failure rate models are acceptable.

Therefore, the results from chapter 4 will be used to assess the reliability of the three cable joint populations by means of evaluating the level of reliability based on the age of the components. Afterwards, the developed failure rate models will be used to estimate the expected number of failure in the future. Furthermore, a failure count diagram will be proposed, which would give valuable information about the number of failures occurring at certain ages for each population.

5.1.2 B(x)-Lives

Statistics give the possibility of comparing populations of components with each other. The use of the percentile life, or B(x)-lives in engineering terminology, is encountered in almost every industry. The B(x)-lives indicates a certain level of reliability based on the age of the component, see [18] [21] [25] [26] [32].

In general, these parameters give the estimated time when the probability of failure will reach a specific point. For instance, if 10% of the cable joints are expected to fail by 15 years of operation, then, it can be stated that the B(10) life is 15 years. Hence, a B(x)-life of B(10) means that 10% of the total population will fail at a certain age and that 90% survives. The values of the B(x)-lives can assist the asset manager in anticipating which level of reliability is acceptable and at which age this level of reliability is reached.

Likewise, the asset manager can compare different groups of populations with each other. According to [32], different values can be used as an anticipated value for unreliability/reliability depending on the criticality of a failure. Furthermore, when the probability of failure becomes too high (unacceptable), components with ages higher than the accepted B(x)-life should be replaced or overhauled.

5.1.3 Predicting Future Failures

Ultimately, the asset manager is interested in anticipating how failure rates of certain assets will develop in the future. Predicting future performance is a very important objective for the asset manager. In this context, with the failure rate functions of the fitted distributions and with the population of components in operation (survived), a failure prediction for the near future (5 to 8 years) can be estimated.

At the same, this enables a methodology for the validation of the developed failure rate model. Simply, this means that with the developed failure rate models, the predicted failures for the verification period can be compared to the actual (real) occurred failures. If the predicted failures for the verification period and the actual failures are comparable, it means that the failure rate models are acceptable for predicting the future failures. According to [32], the expected number of failures $N_{f,e}$ is calculated as:

$$N_{f,e} = \sum_{i=0}^{Age\ oldest} \lambda(t_i) \cdot N_i \tag{5.1}$$

where $\lambda(t_i)$ is the failure rate at age i, and N_i is the number of units (cable joints) with age i in-service. Furthermore, the corresponding confidence bounds can be taken into account. By doing this, the variation in the number of expected failure can also be addressed.

5.1.4 Failure Count Diagram

From figure 4.4 (number of reported cable joint failures for the period 2005-2009), we found that the age at which a failure occurs is not exactly known in many cases. However, with the calculated failure rates, a failure count diagram can be developed. This diagram is basically a probability distribution of when a component can be expected to fail. This diagram gives, in relative term, how many components of an installed population of a particular age contribute to failures.

According to [17], the information from the failure count diagrams can be used to estimate the failure of components at a certain age, and additionally can be used for developing replacement or maintenance policies.

The mentioned tools in sections 5.1.2, 5.1.3 and this section, when developed for a specific case, can form the foundation for studies of proposed maintenance actions, replacement policies as well as other AM strategies. The most important aspect in the following analytical results is the fact that such analysis of age and failures can provide the type of tool required to support AM decision making.

5.2 10 kV Cable Joint Failure Assessment

5.2.1 B(x)-Lives

As already explained, the B(x)-lives can be used to obtain the age of a component where a certain level of reliability is obtained. With this information the asset manager can decide which level of reliability he or she is willing to accept, and what age the components need to reach in order to fulfil such levels of reliability. The B(x)-lives for the three different populations of 10 kV cable joints are shown in table 5.1. The following B(x)-lives are listed: B(1), B(10), B(25), and B(50). Note that B(50) is the same as the mean life of the populations under consideration.

Table 5.1: B(x)-lives of synthetic, mass and oil insulated 10 kV cable joints. The corresponding upper and lower 90 % confidence bounds are also listed for all three populations of cable joints

	Synthetic Insulated Cable Joint Population			Mass Insulated Cable Joint Population			Oil Insulated Cable Joint Population		
	Component Age (year)			Component Age (year)			Component Age (year)		
	90 % Bound	*B-life*	*90 % Bound*	*90 % Bound*	*B-Life*	*90 % Bound*	*90 % Bound*	*B-life*	*90 % Bound*
B(1)-life	17	**19**	21	26	**27**	28	18	**19**	20
B(10)-life	30	**31**	33	43	**44**	45	33	**35**	37
B(25)-life	38	**40**	42	53	**54**	56	42	**44**	46
B(50)-life (mean life)	45	**48**	52	67	**65**	63	51	**54**	57

From table 5.1, the asset manager can assess and compare the reliability of the three cable joint population with each other. For instance, it can be seen that 10% (B(10)-life) of the total synthetic insulated joint population will fail at an age of roughly 31 years. The oil insulated cable joints reach comparable ages (35 years) for the same level of reliability, while the mass insulated joints reach a higher age of roughly 44 years.

Additionally, the B(x)-lives can be used to assess how many cable joints are actually older than a certain chosen B(x) level. The level of B(x)-life which the asset manager can choose for a certain population of components, depends on the network type, component, impact of failure etc. If, for example, the asset manager is interested in getting to know how many cable joints of each population are older than the B(10)-life, then the calculated B(10)-life together with the in-service cable joints can be used for this. In table 5.2, the amount of cable joints older than the B(1) and B(10)-life is listed for each type of 10 kV cable joint.

Table 5.2: The level of reliability for a given population together with the % of cable joints which are older than the related B(x)-life. The asset manager can decide which level of reliability criteria to set for each population

	% of population older for synthetic joints	% of population older for mass joints	% of population older for oil joints
B(1)-Life	35%	46%	23%
B(10)-Life	21%	5%	2%

From table 5.2 it is found that 35% (680 joints) of the synthetic insulated cable joint population is older than the B(1)-life, while 21% (410 joints) of the same population is older than the B(10)-life. In case of the mass insulated joints, 46% (6530 joints) of the populations is older than the B(1)-life, while 5% (640 joints) is older than the B(10)-life. For the oil insulated joints the values are 23% (3160 joints) and 2% (125 joints) older than, respectively, the B(1)-life and B(10)-life. The total population for the synthetic, mass and oil insulated joints are approximately, 1950, 14460 and 15340, respectively.

Based on these analytical results, the asset manager can decide which portion of the population is in the end of life given the selected level of reliability. From these analytical results, more knowledge is created on the reliability of the total population. In the way forward, the asset manager can prepare maintenance or replacement policies in order to deal with the portion of the population which has exceeded the required reliability criteria.

5.2.1.1 Sensitivity Analysis

In this section, a sensitivity analysis is performed for the synthetic insulated cable joint population. In section 4.2, the collection of available 10 kV cable joint data was described. Both, failure data and in-service data for all three types of 10 kV cable joints was discussed and presented in section 4.2. The total number of reported failures for the period 2004 and 2009 for the 10 kV synthetic joints is shown in figure 5.1.

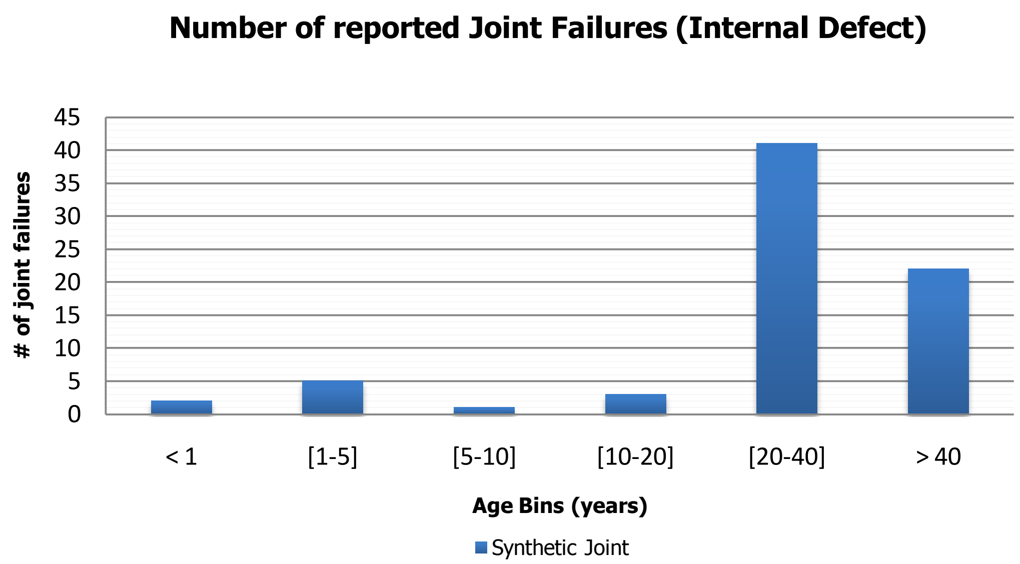


Figure 5.1: In this figure the internal defect failures for 10 kV synthetic insulated joints is shown for the period 2004-2009.

Stedin, indicated that the failures which are reported in the age intervals [20-40] and [>40] years are probably failures of 10 kV resin joints that were applied in the 1970s. Usually, these types of joints were registered as synthetic insulated joints. These types of joints were used in 10 kV three-phase paper-oil insulated MV cables. These joints are known as "Nekaldiet" joints.

These joints have contributed significantly to outages because of breakdown of the resin insulation; however, they are not applied anymore. Therefore, Stedin was interested to assess whether the population of synthetic cable joints without the suspect "Nekaldiet" failures had a higher reliability or not. In order to analyse this it was required to exclude the recorded "Nekaldiet" failures in the age intervals [20-40] and [>40] years from the statistical analysis.

By using expert knowledge from the utility, we decided to exclude all the failures that were recorded in the age bin [>40] years from the analysis, because we had the impression that the "Nekaldiet" joints are in operation long enough to have reached ages higher than 40 years. Furthermore, a number of failures from the age interval [20-40] years were also excluded from the statistical analysis. We had the impression that, not only "Nekaldiet" joints, but also other types of synthetic insulated joints could have been in operation longer than approximately 20 years. In this context, two scenarios were used to exclude recorded failures from the age interval [20-40] years. In the first scenario 10 failures were excluded, while in the second scenario 20 failures were excluded.

A sensitivity analysis is performed in this section for two scenarios. The parametric distribution fitting method was used for calculating the probability distributions and parameters.

Besides the recorded failure data, the in-service population data for synthetic joints was also adjusted for this sensitivity analysis. Together with the experts we decided to exclude all installed synthetic joints which were older than 40 years. We had the impression that it was very likely that this group of synthetic joints belonged to the "Nekaldiet" type of cable joint, since; these joints were installed a few decades ago.

Two scenarios were analyzed. Scenario 1 excludes all failures from the age bin [>40] years and 10 failures from the age bin [20-40] years. Scenario 2 excludes all failures from the age bin [>40] years and 20 failures from the age bin [20-40] years. The corresponding B(x)-lives of these two scenarios were compared with the B(x)-lives of the original case (described in section 5.2). The results are shown table 5.3.

Table 5.3: B(x)-lives of synthetic insulated 10 kV cable joints for the original case and two other scenarios.

	Synthetic Insulated Cable Joint Population Original			Synthetic Insulated Cable Joint Population Scenario 1			Synthetic Insulated Cable Joint Population Scenario 2		
	Component Age (year)			Component Age (year)			Component Age (year)		
	90 % Bound	*B-life*	*90 % Bound*	*90 % Bound*	*B-Life*	*90 % Bound*	*90 % Bound*	*B-life*	*90 % Bound*
B(1)-life	17	**19**	21	12	**15**	17	11	**14**	17
B(10)-life	30	**31**	33	31	**33**	37	34	**39**	46
B(25)-life	38	**40**	42	42	**49**	56	49	**62**	78
B(50)-life (mean life)	45	**48**	52	55	**67**	82	67	**92**	125

From the analytical results shown in table 5.3, we can conclude that the reliability (B(x)-life) of the 10 kV synthetic insulated joint populations is higher when the suspect "Nekaldiet" joints are excluded from the analysis.

In figure 5.2 the failure rate plots for the original data and the two mentioned scenarios are shown for the synthetic insulated cable joints. Resulting from figure 5.2, it can be found that the failure rates are considerably lower for the synthetic joints when the suspect "Nekaldiet" failure records are excluded from the statistical analysis. Thus, the "Nekaldiet" joints negatively impact the overall reliability of the synthetic insulated joint population.

More specifically, the asset manager should consider removing all "Nekaldiet" cable joints which are still in operation in the area "Region X" and probably other areas as well.

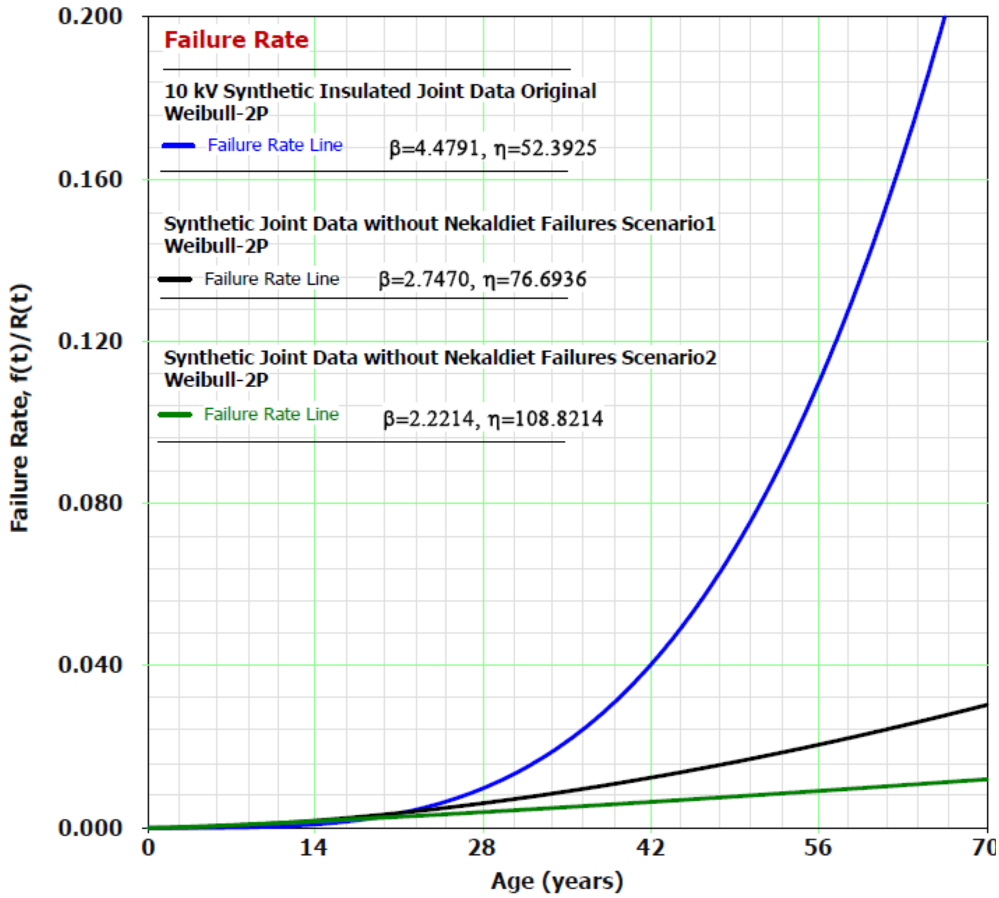


Figure 5.2: This figure shows the failure rate plots for three subsets of life data for synthetic insulated cable joints. The blue failure plot represents the original data record, while the black failure plot represents scenario 1 and the green failure plot represents scenario 2

5.2.2 Predicting Future Cable Joint Failures

With the developed failure rate models for each population and the number of components in operation, the asset manager can anticipate the development of future cable joint failures. This is illustrated in this section for the three cable joint populations. Before proceeding with the analysis, some basic aspects need to be mentioned, which are:

- Each time one cable joint fails, it is repaired by two new cable joints. Therefore when predicting the number of failure for the next year, this should be addressed. However, mass insulated joints that failed are not replaced by two new mass insulated joint. Instead of this, it is the policy of Stedin to repair all mass insulated joint failures with two new oil insulated joints. This is taken into account in the analysis.

- Every failed cable joint is removed from the population and the failure predictions are estimated by taking into account the remaining population of cable joints. The numbers of failed joints in a particular year are replaced by new joint for the next year of interest.

- When predicting the future failures, it is required to increase the age of the remaining population for every next year. Therefore, the population of cable joints is made one year older for every subsequent year to come.

By using equation 5.1, the expected failures are estimated. This is estimated for the coming 6 years for all three types of cable joints. Besides predicting the future number of failures, an additional analysis is performed in order to assess whether the developed failure rate model is in agreement with the actual historic failures.

5.2.2.1 *Failure Prediction for Synthetic Insulated Joints*

The calculated failure rate curve shown in figure 4.10 for the synthetic insulated cable joint data together with the population of synthetic joints still in operation is used for predicting the number of expected failures for the coming 6 years. This is illustrated in figure 5.3.

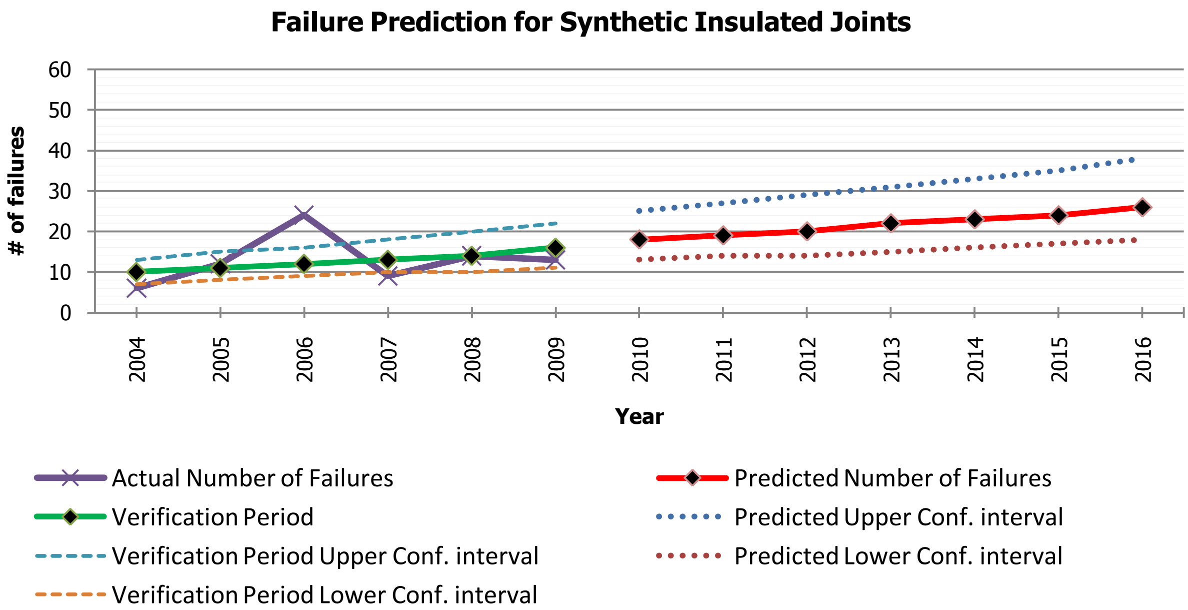


Figure 5.3: Estimation of the number of total expected failures for the coming six years for 10 kV synthetic insulated joints. The red line gives the number of predicted failure starting at 2010 until 2016. The corresponding 90% confidence bounds are also shown. In the period 2004-2009 the actual number of failure (purple line) is compared to the estimated number of failures for that period

From figure 5.3, it can be seen that the estimated number of failure (green line) based on the analysis are comparable with the actual number of failures in the period 2004-2009. As result of this, it can be concluded that the developed failure rate model reasonably describes the failure behaviour of the considered population.

The developed failure rate model for the synthetic joints is used together with the population of 2010 to predict the number of failures in that particular year. For 2010, a number of 18 failures are predicted with a variation between 13 and 25 when taking into account the respective 90% lower and upper confidence bounds.

From this point on, for every next year, the ages of the remaining population of joints are made one year older. At the same time, the estimated failures from the previous year are subtracted from the population. It is also taken in account that every joint failure introduces two new joints. With this information, the asset manager can determine whether the expected numbers of future failures are acceptable, or, whether structural replacement is necessary in the coming years.

5.2.2.2 *Failure Prediction for Mass Insulated Joints*

The calculated failure rate curve, shown in figure 4.13, for the mass insulated cable joint data together with the population of mass joints still in operation is used for predicting the number of expected failures for the coming 6 years. This is illustrated in figure 5.4.

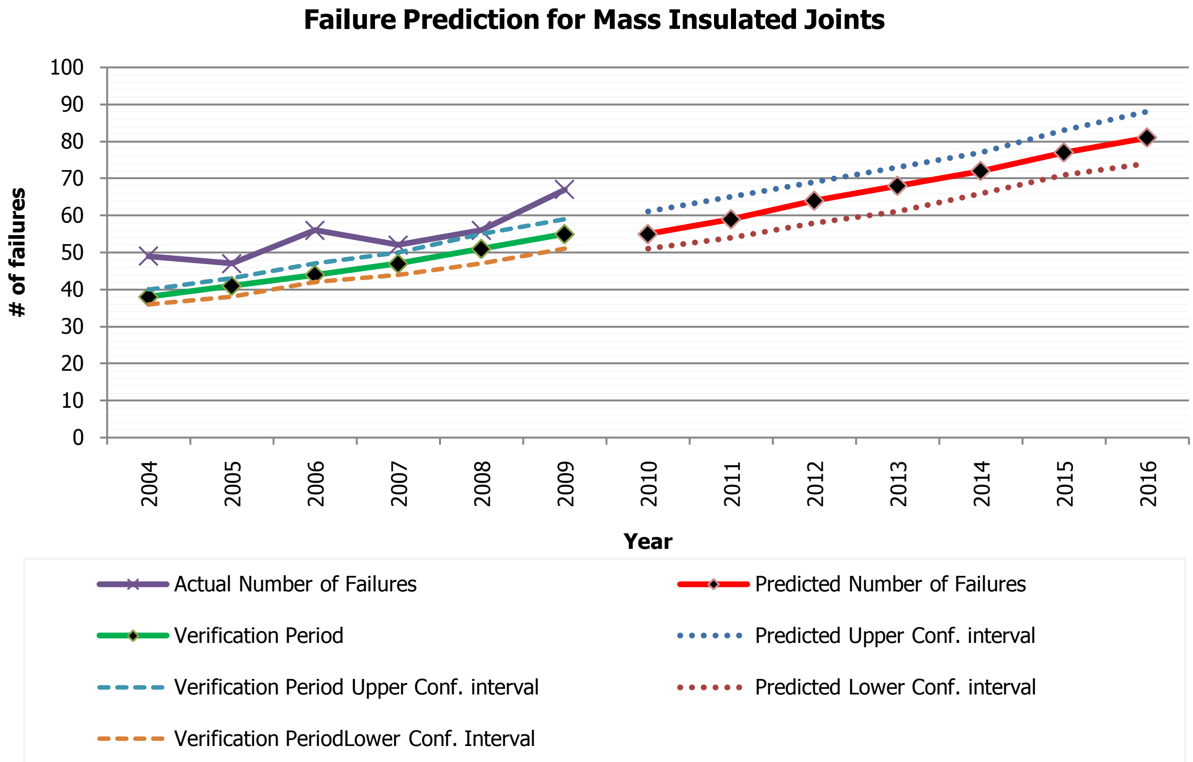


Figure 5.4: Failure prediction for the mass insulated joints together with the corresponding 90% confidence bounds. From these analytical results it can be seen that actual failures are higher than the estimated failure for the period 2004-2009. However, the number of failures is rising for the coming 6 years.

From figure 5.4, it can be seen that, for the period 2004-2009, the estimated failures are lower than the actual occurred failures. This might indicate that the developed failure rate model does not properly describe the failure behaviour of the mass joint population. However, it might also mean that the estimations that were made for the part of the population with unknown age are not in accordance to the actual situation.

It should be noted, that for almost 60% of the mass joint population no exact age was specified in the database. This 60% of the population corresponds to roughly 5300 mass joints. These joints were divided proportionally to the joint with age. In order to assess whether this estimation, regarding the 5300 joints, might be an improper estimation, the 5300 joints are not divided proportionally but according to a certain age interval. The new estimation for the in-service joints is based on engineering knowledge from the utility.

From this expert knowledge, it was found that mass joints were mostly used a few decades ago. Thus, it is assumed that the 60% of mass joints without specified age might be between 25 and 50 years old. Based on the new assumption, the parametric distribution fitting procedure is performed again and the resulting failure rate function together with the population of mass joints in-service (including the new estimation) is used to determined the expected future failures. In figure 5.5, the expected number of failures based on the new population estimations is shown.

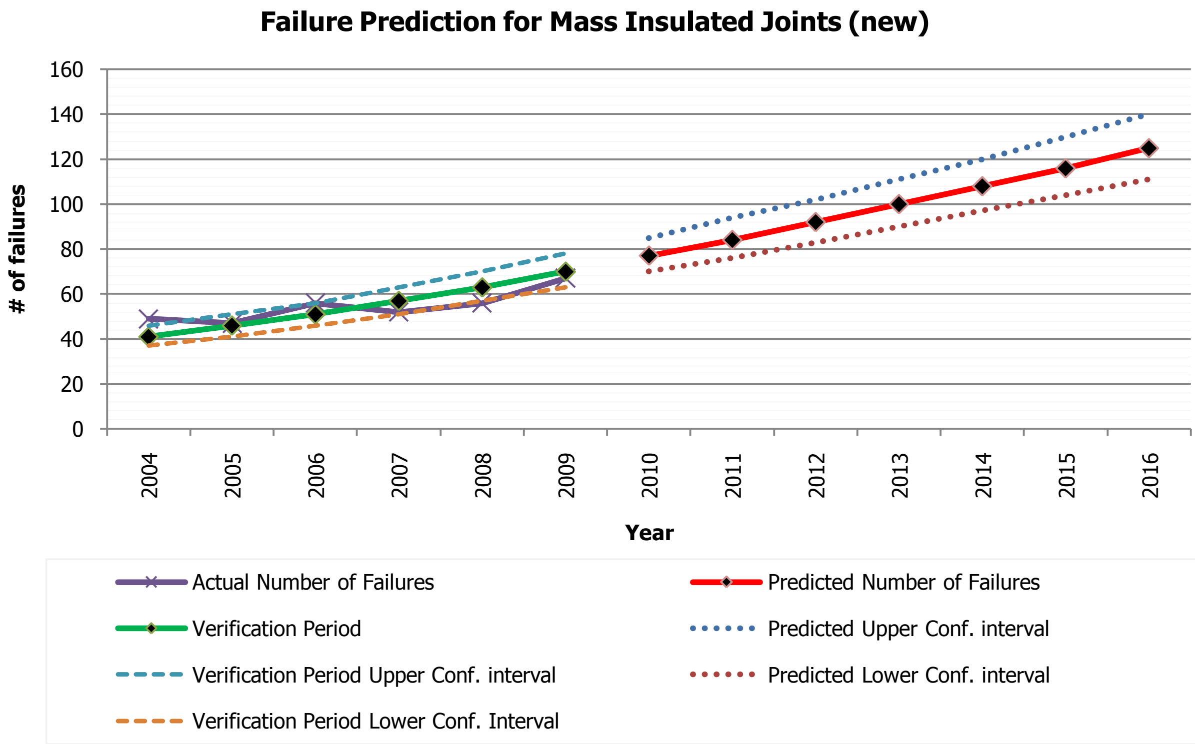


Figure 5.5: Second failure prediction for mass insulated joints based on new assumption for the portion of the joint population without specified age. In this case, it is assumed that the joints without specified age are 25-50 years old. With this information a new failure prediction is made. For the period 2004-2009 it is found that the total number of actual failures is comparable with the total number of estimated failure for that period

From figure 5.5 it can be seen that with the developed failure rate model and the second attempt of estimating the 60% of unknown population, the number of calculated expected failures is more is agreement with the actual occurred failures in the period 2004-2009. Under these circumstances, it can be concluded that based on the analytical results, it seems probable that the population of mass joints without recorded age (60% of the population) is older than 25 years.

However, it should be noted, that these assumptions are based on the available data at the moment of the present study. Another way of arguing might reveal that there have been more failures of mass joints in the past, of which the records are missing, and therefore the failure rates obtained here are conservative values. Whether the mass joint population is older or the number of failures in the past is higher, in either case, the asset manager now has more knowledge on the failure behaviour of the mass insulated joints.

This case illustrates that, by choosing appropriate statistical models and engineering reasoning it is still possible to create valuable information on the failure behaviour of population, even in case of uncertain or missing data.

Subsequently, if it is assumed that the expected failure shown in figure 5.5 is representative for the mass joint population, then it can be found that the number of failure will be increasing every year. In 2010, roughly 77 failures are expected and this number will grow to approximately 125 failures in 2016. Even though failed mass joints are replaced by oil insulated joints, the number of failures in the coming years will gradually increase. Finally, this information can form the foundation for the asset manager to determine if the expected numbers of future failures are acceptable, or, whether structured replacement is necessary in the coming years.

5.2.2.3 Failure Prediction for Oil Insulated Joints

The calculated failure rate curve, shown in figure 4.16, for the oil insulated cable joint data together with the population of oil joints still in operation is used to predict the number of expected failures for the coming six years. This is illustrated in figure 5.6.

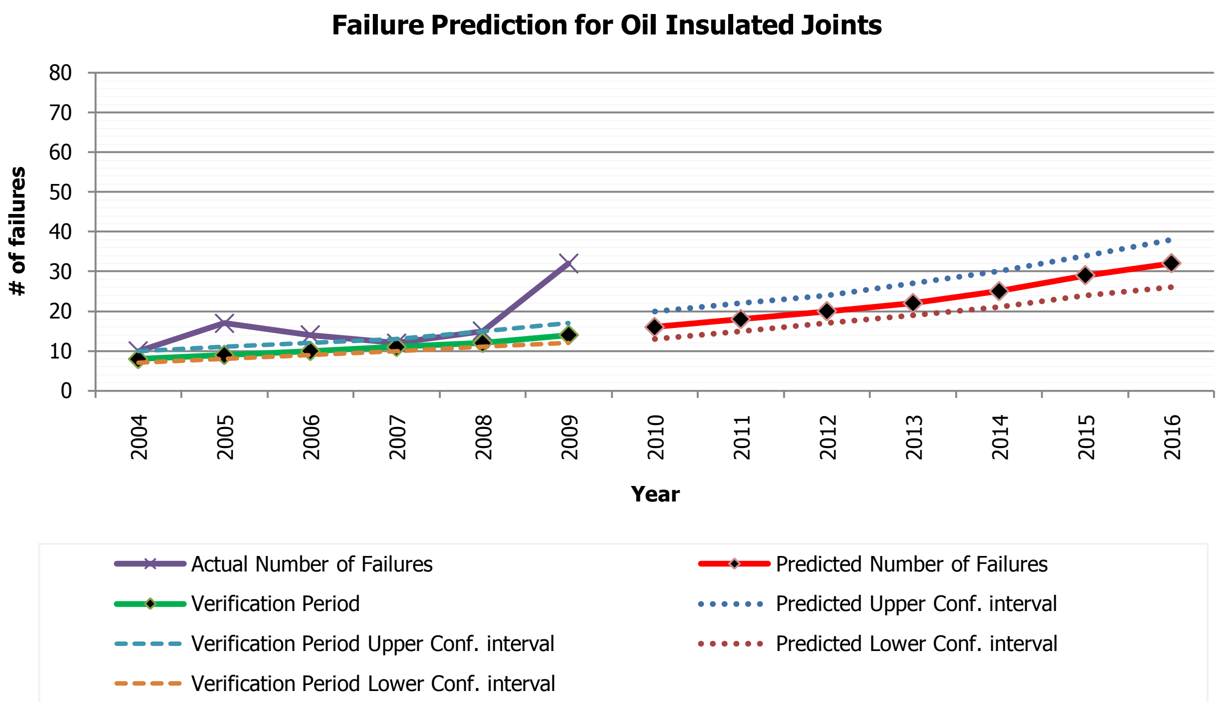


Figure 5.6: The number of calculated expected failures for the oil insulated cable joints. Also shown are the corresponding 90 % confidence lower and upper bounds. For the period 2004-2009, the number of calculated expected failures for the oil joints (green line) is reasonably comparable with the number of actual occurred failures (purple line). Furthermore, the failure prediction for six year from now illustrates that the number of failures are increasing gradually every next year

From figure 5.6, it is found that the number of actually occurred failures for the period 2004-2009 is reasonably comparable to the calculated number of failures, except for the year 2009. In 2009 Stedin encountered higher number of failures then in the years prior to 2009. Ground conditions (ground sagging mostly due to peaty soil) are recorded as the main cause of this higher number of failures in 2009.

In general, it can be concluded that the developed failure rate model reasonably describes the failure behaviour of the considered population. Subsequently, the failure rate model is used for predicting the number of failure in the future. Resulting from this, it is found that the number of failures is increasing for every following year. From this point on, for each following year, the population of joints in service are made one year older. Concurrently, the estimated failures from the previous year are subtracted from the population. It is also taken into account that every joint failure introduces two new joints.

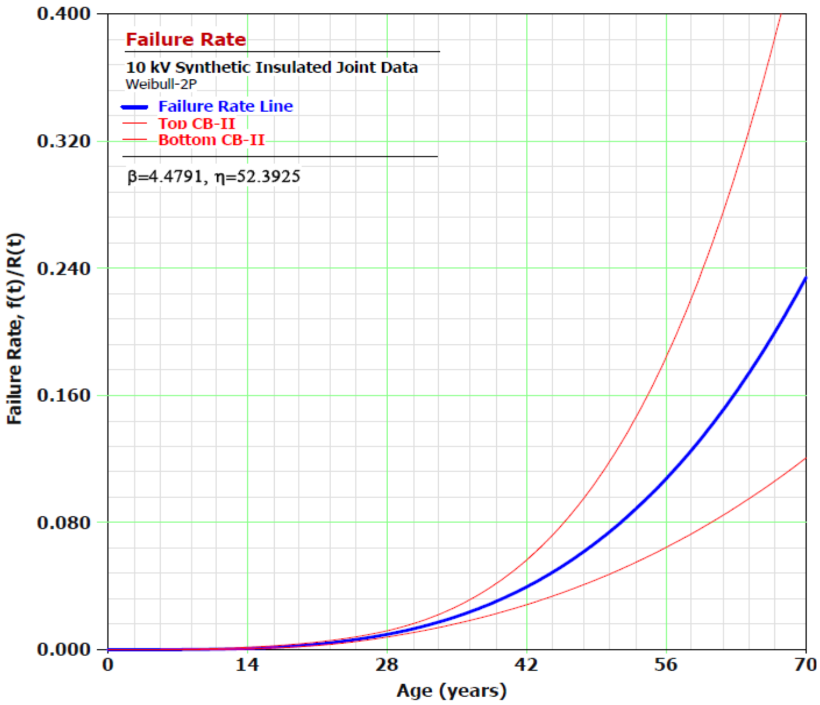
With this information, the asset manager can determine if the expected numbers of future failures are acceptable or whether structured replacement is necessary in the coming years.

5.2.3 Failure Count Diagram

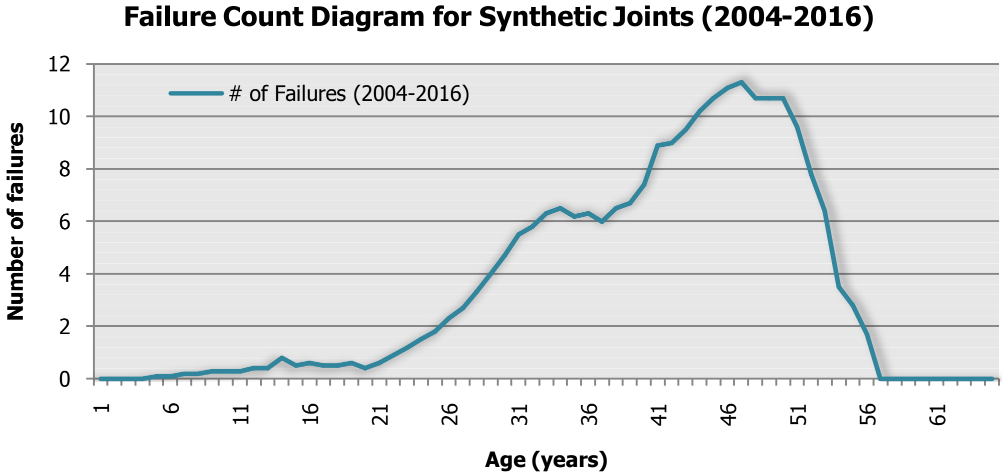
With the failure count diagram, it will be possible to assess the impact that typical increasing failure rates have on the installed equipment base. This diagram gives, in relative terms, how many components of a particular age contribute to the total number of failures of the installed population.

5.2.3.1 Failure Count Diagram for Synthetic Insulated Joints

As figure 5.7a shows, every year, as the cable joints grow older, their failure rate increases.



(a)



(b)

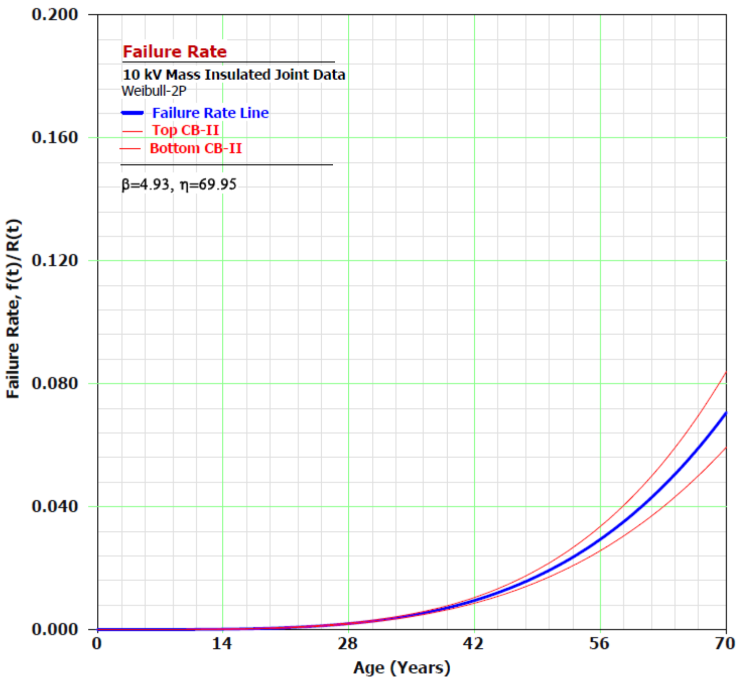
Figure 5.7: Shows the failure rate plot and failure count diagram for the synthetic insulated joint population. (a) Shows the failure rates as function of age. (b) Shows the number of failures occurring each year. The maximum is reached in year 47, when the combination of escalating failure rate and number of remaining units peaks

However, every year, because many joints have already failed in the previous years, there are fewer units remaining to potentially fail in the next year. In this case, the number of units that can be expected to fail in any year is the failure rate for that age multiplied with the number of remaining in that year. It is interesting for the asset manager to anticipate when this value reaches a maximum. Figure 5.7b illustrates this for the 10 kV synthetic insulated cable joint population. Resulting from figure 5.7b, it is found that the maximum number of failures is reached in year 47 for synthetic insulated joints.

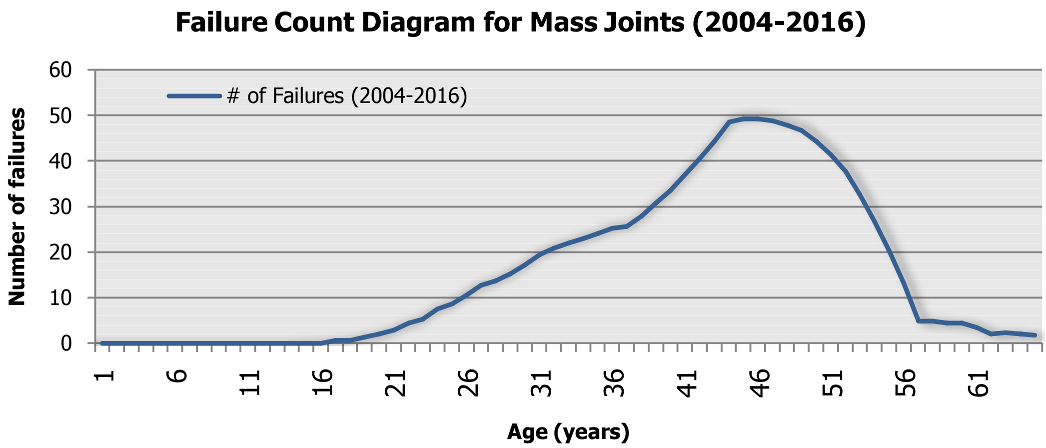
This maximum is reached when a combination of rapidly increasing failure rate and high number of remaining units peaks. More than half of the failures occur in the range of 21 and 51 years. Despite the higher failure rate for joints older than 51 years, (see figure 5.7a) it can nevertheless be concluded that failures in intermediate years are the real cause of the system reliability problem. It is still those synthetic joints that have reached 47 years of service that contribute most to the systems problems. Perhaps, cable joints older than that fail with higher likelihood but there are most of the time too few to generate a high total failure count. The failure count diagram is a representation of the relative contribution to failures of synthetic joints as function of their age. This diagram can be seen as an important tool in managing reliability and replacement policies.

5.2.3.2 Failure Count Diagram for Mass Insulated Joints

Similar failure count analysis is performed for the mass insulated joint population assuming, that for the 60% of the population with unknown age, the age is between 25 and 50 years. In figure 5.8, the corresponding failure rate plot and failure count diagram are depicted.



(a)



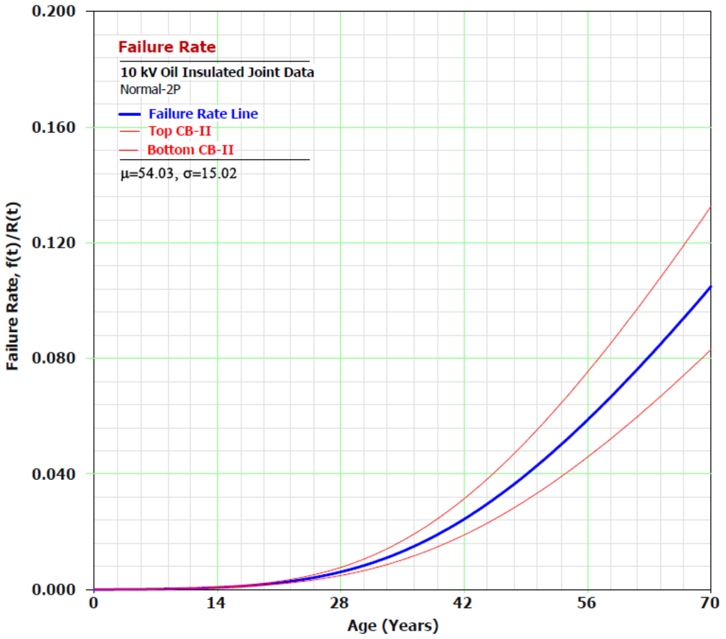
(b)

Figure 5.8: Shows the failure rate plot (a) and failure count diagram (b) for the 10 kV mass insulated joint population. From figure (b) is can be seen that no failures occur for mass joints younger than 16 years. Due to the fact that mass joint are not applied anymore, and resulting from this, every failed mass joint is replaced by oil joints. The failure count diagram peaks at 46 years

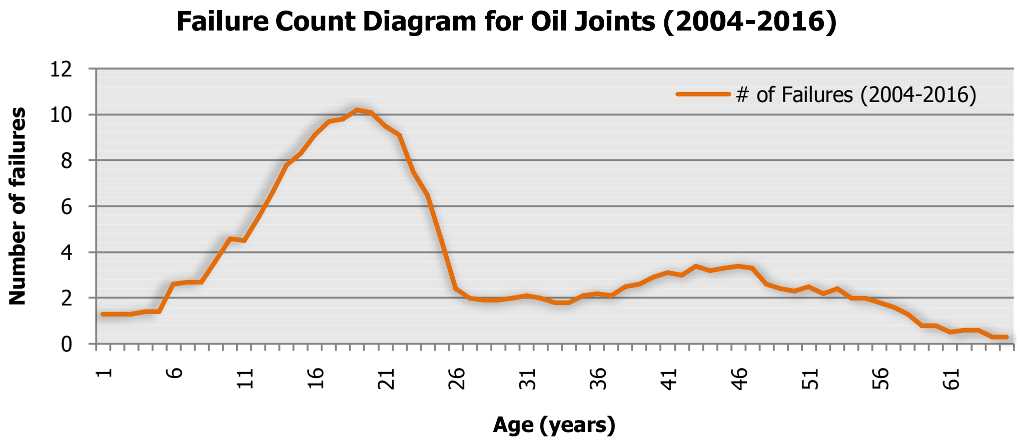
As figure 5.8a shows, every year, as the mass cable joints grow older, their failure rate increases. From figure 5.8b, it is found that the maximum number of failures is reached in year 46 for mass insulated joints. Again, this maximum is reached when a combination of escalating failure rate and high number of remaining units peaks. Although, the failure rate for this population increases after 56 years (see figure 5.8a) it is found from the failure count diagram that few mass joints survive to encounter such high failure rates and to impact the reliability. Instead, as was the case for the synthetic joints, the high impact failure level that is the main source of failures for the utility are found to be caused by mass joint of intermediate age (36 years until 51 years).

5.2.3.3 *Failure Count Diagram for Oil Insulated Joints*

The failure rate plot and the failure count diagram for the oil insulated joints are depicted in figure 5.9.



(a)



(b)

Figure 5.9: The failure rate function and failure count diagram for the 10 kV oil insulated cable joint population are shown in this figure. From figure (b) is can be seen that the maximum value is reached at 20 years, while after 36 years the number of failure slightly increase again. The first peak at 20 years is due to high number of oil joints applied in last two decades

The failure count diagram for the oil insulated cable joint population looks different than the ones for the synthetic and mass insulated joints. In figure 5.9b, it can be seen that the maximum value for failures is reached at an earlier time, at year 20. Because the policy of Stedin is to replace all failed mass joints with oil joints, the population of oil joints for the past two decades has grown. Resulting from this, there are many units with ages between 11 and 26 years and therefore, the contribution to failures is high.

From here on, the number of failures decreases until age 36 and then gradually increases again. This peak in failures is reached when a combination of escalating failure rate and high number of remaining units peaks again. Thereafter, even though the failure rate keeps increasing every year, the number of failures actually occurring decreases, because there are fewer and fewer oil joint left each year and therefore the net number of failures decreases.

5.3 Conclusions

In this chapter, the results of the statistical analysis are used as a fundament to develop a number of tools for evaluating the failure behaviour of the three types of investigated cable joint populations. The information from these evaluations can be used to assist the asset manager in his or her decision, regarding the failure behaviour of the three types of investigated cable joint populations. In general, the most important conclusion that can be drawn from the foregoing analytical results of this chapter is that the analysis of age, expected future failures and total number of failures can provide the type of tool needed to support the AM decision making processes.

It is found that with the results of the B(x)-lives the asset manager can anticipate which level of reliability is acceptable. With the selected B(x)-life, the asset manager can assess at what age this level of reliability is reached. Furthermore, when the probability of failure exceeds a certain value, components with ages higher than the selected B(x)-life can be replaced or overhauled. When the B(1)-lives for the three joint population are compared with each other, it is found that 35%, 46% and 23% of the population for respectively the synthetic, mass and oil joints are older than the B(1)-life. The B(10)-lives are also compared, and from this, it is found that 21%, 5% and 2% of the population for respectively the synthetic, mass and oil insulated joints are older than the B(10)-life. In the end, the level of B(x)-life which the asset manager can choose for a certain population of components, depends on the network type, component type, impact of failure etc.

Additionally, we found from the sensitivity analysis that with the failure probability models technical reliability assessment can be carried out for suspect group of assets within a population. From the sensitivity analysis for the "Nekaldiet" joints, we found that the failure records of these joints had a negative impact on the overall reliability of the synthetic joint population.

Forecasting asset performance is one of the main responsibilities of the asset manager. With the developed failure rate models for each population and the number of components in operation, the asset manager can anticipate the development of future cable joint failures. From the analysis of the synthetic joints, it can be concluded that the number of failure will increase in the years to come. Furthermore, it is found that the backward estimated number of failures, based on the analysis, is comparable with the actual number of failures in the period 2004-2009. As result of this, it can be concluded that the developed failure rate model reasonably describes the failure behaviour of this population.

A similar conclusion can be drawn for the oil insulated cable joint population.

In case of the mass joints the situation was found to be different. For the mass joints the actual failures occurred in the period 2004-2009 where higher than the calculated number of failures. Therefore, a second recalculation was performed for analyzing this population. In the second attempt the 60% of mass joints without specified age where not divided proportionally across the existing population, but divided according to a certain age interval. Based on utility knowledge, the 60% of mass joints without known age were assumed to be between 25 and 50 years old. Afterwards, the failure rate model and the new age distribution for the mass joint population were used to calculate the number of failures. From this, it is found that the calculated number of failures is comparable with actual occurred failure for the period 2004-2009. Thus, it can be concluded that the mass joint population might be older than previously estimated. Another way of arguing might reveal that there have been more failures of mass joints, of which the records are missing, and therefore the failure rates obtained here are conservative values.

The failure count diagram makes it possible to determine when a component can be expected to fail. From the failure count diagram analysis, it can be concluded that the maximum number of failures is reached when a combination of escalating failure rate and high number of remaining units peaks. This maximum value is found to be different for each cable joint population. A general conclusion, which holds for all three joint populations, is that failures in intermediate years are the real cause of system reliability problems.

Accordingly, it can be concluded that the very high failure rates that develop after 50 to 60 years of operation have little impact on the utilities quality of service, because few component survive to see such high failure rates. From the analysis, it arises that the synthetic joint population reaches the maximum number of failures for components with ages of roughly 47 years. From the failure count diagram for the mass joints, it can be concluded that the maximum number of failures for mass joints can be expected to occur with an age of roughly 46 years. Finally, for the oil insulated joints, it can be concluded that the expected number of failures peaks for components with an age of around 20 years. Added to this, another peak is found for component with an age of roughly 46 year.

With the information, which has now been developed for each joint population, the asset manager can determine which level of reliability is acceptable and which part of the population has a high failure probability and thus deteriorates the utilities quality of service. The maximum age of certain components in relation to the requested reliability and the failure expectation gives valuable information to the asset manager. Basically, resulting from the analytical results, it can be concluded that even though the data was either missing or incomplete, it is still possible to develop sensible probability methods in order to provide the asset manager with useful information for supporting AM decision processes.

6
Conclusions & Recommendations

In this chapter, the conclusions of this study are presented and recommendations on further improvement of this study are made.

Firstly, in section 6.1, the conclusions are presented in two parts. In the first part, conclusions are presented with regard to the failure related defects in MV underground distribution networks and the analyzed historic failure statistics for 10 kV MV underground networks. In the second part, conclusions are presented with regard to the applied statistical life data analyses method. Some general conclusions are drawn with regard to the statistical life data analyses. Thereafter, some conclusions are drawn with regard to the investigated 10 kV cable joint populations.

Lastly, in section 6.2 recommendations will be made on further improvement of this study. Data availability will form an important aspect, as well as challenge, in further improvements and thus adoption of probabilistic models in AM decision-making processes.

6.1 Conclusions

Historic Failure Statistics

- From the analyzed historic failure statistics, we can conclude that more than 80% of power-delivery outage related failures in 10 kV MV networks are caused by failures in cable systems.

- From a component level failure pattern analysis of cable systems, we can conclude that, for area "Region X", approximately 65% of breakdowns are caused by internal component related defects. The remaining 35% of failures are caused by external defects such as excavator digging.

- We also concluded that out of the 65 % of internal component related defects, the majority of failures (44%) occur in 10 kV cable joints.

- From more in-depth historic analysis, we can conclude that in data records three types of 10 kV cable joints can be distinguished and their share in the overall 44 % of joint failures are:
    - Mass Insulated Joints -> 57%
    - Oil Insulated Joints -> 25%
    - Synthetic Insulated Joints -> 18%

Statistical Life Data Analysis (LDA)

- It was found that, when sufficient life data and appropriate statistical tools (which take into account interval-data and censored-data) are available, then, statistical life data analysis can be applied to assess the failure probability of a certain population.

- We can conclude that, with life data analysis large amounts of information, characteristics and behaviours can be described by small numbers of parameters. Subsequently, the probability functions allow different assets to be compared with each other, and make reference to criteria like age, rate of rise of failure rates and mean life of components.

- We can conclude that the failure probability models of a certain population of components can give valuable information with regard to failure behaviour, which can facilitate sound AM decision processes in a mid and long term basis.

Application of LDA with regard to three types of 10 kV cable joint populations

- From the LDA we found that the failure distribution with the best fit for:
    - the synthetic insulated joint populations was the 2-parameter Weibull distribution
    - the mass insulated joint population was the 2-parameter Weibull distribution
    - the oil insulated joint population was the normal distribution

- When the failure rate plots of the three populations are compared with each other, we can conclude that the synthetic insulated joints have the highest failure rate, followed by the oil insulated joints and finally by the mass insulated joints.

- We can conclude that, during the first 20 years of operation the failure rates for all three populations of joints are constantly low, however, after this period the failure rates grow as the joints ages. Each cable joint population exhibits a different rate of rise of the failure rate over time.

- Furthermore, we found that the average expected lifetime for the synthetic insulated joint population was estimated to be 48 years, for the mass insulated joint population 64 years and for the oil insulated joint population 54 years.

The failure probability models are applied as a fundament to develop a number of tools for evaluating the failure behaviour of the three types of 10 kV cable joint populations.

- We can conclude that, with the B(x)-life concept the asset manager can assess which level of reliability is acceptable. If the probability of failure exceeds a certain accepted level of reliability, then the asset manager can consider structured replacement/overhauling or maintenance.

- Furthermore, we found that the developed failure rate models can be used to statistically predict the occurrence of joint failures in the future. These analytical results can facilitate the asset manager to determine if the forecasted joint failures are acceptable or if replacement is necessary in the coming years.

- With the developed failure count diagrams, it is possible to determine when a component can be expected to fail based on the installation records and failure rates.

- From the failure count diagram, we can conclude that the maximum number of failures for a population of cable joints is reached when the rapidly increasing failure rate and high number of remaining units coincide.

- We can conclude that the failures in intermediate years are the real cause of system reliability problems. For instance, we found for the synthetic and mass insulated joint population that the failures in intermediate years are around 20-50 years. However, for oil insulated joints the numbers of failures are found to peak around 10-25 years and a small peak after 36 years.

- Accordingly, we can conclude that the very high failure rates that develop after 50-60 years of operation make little impact on the utilities quality of service, because few components survive to see such high failure rates.

On the whole, we can conclude that, even though the data was either missing or incomplete, the analytical results described in this thesis have shown that it is possible to develop sensible probability models to facilitate the asset manager in typical AM decision processes.

As underground cable systems age in service, these systems will be characterized by increasing failure rates and result in higher costs for repair and consumer costs for degraded system performance. Presently, most approaches are reactive and are not able to strictly address the fact that cable system failure rates will rise in future years. Probabilistic failure rate modelling can help utilities to better predict failures and asset managers in their decision support regarding replacement, maintenance and cost-effective budget plans.

6.2 Recommendations

- To further adopt statistical failure probability modelling into the AM decision-making framework, data collection will play an important role in the usefulness of these models. Therefore, it is recommended that power distribution utilities keep complete and accurate records with regards to cable system installation, replacement, repair and outage data. Understanding and estimating the costs and benefits of such database systems is an important aspect for companies, as database systems are costly for an organisation. Therefore, the utilities are recommended to only record data that is required for certain decision-making goals. Based on the goal and strategy of the asset manager, the level of detail of recording data can be better focused.

- The developed annual failure rate models can be applied to determine the economical optimum replacement time and cost. Therefore, further economic evaluations considering annual failure probability are recommended.

- In this thesis, the effect of component age on failure probability was extensively discussed, however in literature it is found that the electro-thermo-mechanical stresses should also be addressed in representative life models. For instance, a thorough electro-thermo-mechanical life model of electrical components can be established by the Arrhenius Model. Most of the time utilities do not have detailed electro-thermo-mechanical information; however, with the shift towards more condition based maintenance in distribution networks, this information will become more available in the near future. Further research in this context is recommended.

- Finally, the presented parametric distribution fitting method has been applied to cable joint data, however, it should also be applicable to other asset groups. Based on this, it is recommended to study the application of this method for other asset groups in order to gain more knowledge of statistical methods for different component groups. Preferably, the analysis can be beneficial for transformer, underground cable, circuit breakers, voltage and current transformers etc, however, it can be applied to non-electrical components such as gas and heat pipe infrastructures.

References

1. **Brown, Richard E.** *Business Essentials For Utility Engineers.* s.l. : CRC Press Taylor & Francis Group, 2010. 978-1-4398-1196-2.

2. **Cigre, WG D1.17.** *Generic Guidelines for Life Time Condition Assessment of HV Assets and Related Knowledge Rules.* s.l. : Cigre, June 2010. 978-2-85873-107-7.

3. —. *Fundamental Aspects of Data Quality for HV Asset Condition Assessment.* s.l. : Cigre, October 2006. Vol. Electra.

4. **Grotenhuis, B. J., Kerstens, A. and Riet van, M. J.M.** *Dutch distribution cable: improvements in construction, reliability, environment and efficiency.* s.l. : CIRED, June 1997.

5. **Wester, Frank.** *Condition Assessment of Power Cables using Partial Discharge Diagnosis at Damped AC Voltages.* Rotterdam, The Netherlands : Optima Grafische Communicatie , December 2004. 90-8559-019-1.

6. **Zhang, Xiang, et al.** Estimation of the Lifetime of the Electrical Components in Distribution Networks. *IEEE Transactions on Power Delivery.* January 2007, Vol. 22, 1.

7. **Buys, P. M.** *Storingen- en trendanalyse Elektriciteitsnetten Stedin.* Rotterdam : Asset Management Stedin, May 2010 .

8. **Gulski, E., et al.** *PD Knowledge Rules for Insulation Condition Assessment of Distribution Power Cables.* 5th Volta Colloquium on PD Measurement.

9. **Mashikian, M. S. and Szatkowski, A.** *Medium Voltage Cable Defects Revealed by Off-Line Partial Discharge Testing at Power Frequency.* s.l. : IEEE Electrical Insulation Magazine, Juli/August 2006.

10. **Cigre.** *Knowledge Rules for Partial Discharge Diagnosis in Service.* s.l. : Cigre, April 2003.

11. **Kreuger, F. H.** *Industrial High Voltage, Volume 1,2,3.* Delft : Delft University Press, 1991. 90-6275-561-5.

12. **Slootweg, J. G., Postma, A. and Steennis, E. F.** *A Practical Approach Towards Estimating Remaining MV Cable Life.* Vienna, 19th International Conference on Electricity Distribution : CIRED, May 2007. Paper 0221.

13. **Hirji, Ali and Patrawala, Dinesh.** *Reliability Aspects of Power Cables and Accessories.* s.l. : Raychem RPG LTD.

14. **Cuppen, Andre N., Steennis, E. Fred and Wielen van der, Peter C.J.M.** *Partial Discharge Trends in Medium Voltage Cables measured while in -service with PDOL.* s.l. : IEEE, 2010. 978-1-4244-6547-7.

15. **Wielen van der, P. C.J.M.** *On-line Detection and Location of Partial Discharges in Medium-Voltage Power Cables.* Eindhoven, the Netherlands : Eindhoven University Press, 2005. 90-386-1683-X.

16. **Brown, Richard E.** *Electric Power Distribution Reliability, second edition.* s.l. : CRC Press Taylor & Francis Group, LLC, 2009. 978-0-8493-7567-5.

17. **Willis, H. Lee.** *Power Distribution Planning Reference Book, Second Edition, Revised and Expanded.* New York : Marcel Dekker, Inc, 2004. 0-8247-4875-1.

18. **Jongen, R. A., et al.** *Application of Statistical Methods for Making Maintenance Decisions within Power Utilities.* s.l. : Deis, IEEE Electrical Insulation Magazine, Nov/Dec 2006.

19. **Klaassen, Klaas B. and Peppen van, Jack C.L.** *System reliability, Concepts and applications.* Delft, The Netherlands : VSSD, 1989-2006. 90-71301-68-0.

20. **Ghiani, Emilio, et al.** *Renewal Planning of Aged Distribution Networks.* s.l. : IEEE, 2007-2009.

21. **Corporation, Reliasoft.** *Life Data Analysis (Weibull Analysis) Reference Book.* s.l. : Reliasoft .

22. **Mann, Schafer and Singpurwalla.** *Methods for Statistical Analysis of Reliability and Life Data.* s.l. : Wiley, 1974.

23. **Moon, Jong-Fil, et al.** *Time-varying failure rate extraction in electric power distribution equipment.* Stockholm, Sweden : 9th international Conference on Probabilistic Methods Applied to Power Systems, June 2006.

24. **Altman, Douglas G. and Bland, J. Martin.** *Parametric vs. Non-parametric methods for data analysis.* s.l. : BMJ, July 2009. Vol. 339.

25. **Jongen, Rogier, Gulski, Edward and Smit, Johan.** *Statistical Analysis of Diagnostic- and Life Time Data of High Voltage Components .* s.l. : IEEE, 2007. 978-981-05-9423-7.

26. **Quak, Benjamin.** *Information Strategy for Decision Support in Maintaining High Voltage Infrastructures .* s.l. : TU Delft (PHD Thesis), 2007. 978-90-8559-334-8.

27. **methods, NIST/SEMATECH e-Handbook of statistical.** *Engineering Statistics Handbook.* s.l. : NIST (agency of the U.S. Commerce Deparment), 2003.

28. **Nelson and Wayne.** *Applied Life Data Analysis.* New York : John Wiley & Sons, 1982.

29. **Rinne, Horst.** *The Weibull Distribution, A Handbook.* s.l. : CRC Press Taylor & Francis Group, 2009. 978-1-4200-8743-7.

30. **Reliasoft.** Maximum Likelihood Estimation. *Reliability HotWire e-magazine.* Nov 2001, 9.

31. **Zhang, Xiang and Gockenbach, Ernst.** Assessment of the Actual Condition of the Electrical Components in Medium-Voltage Networks. *IEEE Transaction on Reliability.* June 2006, Vol. 55, 2.

32. **Jongen, R. A., Smit, J. J. and Janssen, A. L.J.** *Application of Statistical Analysis in the Asset Management Decision Process.* Beijing, China : International Conference on Condition Monitoring and Diagnosis (IEEE), April 2008.

33. **Reliasoft.** Comparison of MLE and Rank Regression Analysis when the Data Set Contains Suspensions. *Reliability HotWire .* June 2002, 16.

Nomenclature

List of Abbreviations

AM	Asset Management
LV	Low Voltage
MV	Medium Voltage
HV	High Voltage
LDA	Life Data Analysis
MLE	Maximum Likelihood Estimation
LSE	Least Squares Estimation
DSO	Distribution System Operator
O&M	Operation & Maintenance
POF	Probability of Operational Failure
PSCFR	Permanent Short Circuit Failure Rate
MTTS	Mean Time To Switch
pdf	Probability Density Function
cdf	Cumulative Density Function
TKV	Ten Kilo Volt System (Asset Database Utrecht)
K-S Test	Kolmogorov-Smirnov Test
LKV	Likelihood Value

List of Symbols

kV	Kilo Volt (kV)
$R(t)$	Reliability Function
$\lambda(t)$	Failure Rate Function
$\mu(t)$	Mean Life
σ	Standard Deviation
β	Shape Parameter (Weibull Distribution)
η	Scale Parameter (Weibull Distribution)
γ	Location Parameter (Weibull Distribution)
ρ	Rho (Correlation Coefficient)

List of Figures & Tables

Appendix A
MLE vs. Rank Regression

In this appendix, we will discuss why Maximum Likelihood Estimation (MLE) is suggested to be used for parameter estimation when the data set contains a large number of suspensions. This will be explained with an example [33] of a comparison of MLE and Rank Regression analysis when the data set contains suspensions.

Rank regression method is the most widely used method for performing suspended items analysis; however, there is a shortcoming to the method that is important to understand. When using rank regression methods to take into account that some components did not fail, only the position (time-to-failure) where the failure occurred is taken into account, and not the exact time-to-suspension. The following example illustrates this for two cases of life data.

Table A1: Two data sets to be analyzed with rank regression and MLE methods for comparison purposes **[33]**

Case 1			**Case 2**		
Number	***State*, F or S***	***Life of item, hr***	***Number***	***State*, F or S***	***Life of item, hr***
1	F_1	1000	1	F_1	1000
2	S_1	1100	2	S_1	9700
3	S_2	1200	3	S_2	9800
4	S_3	1300	4	S_3	9900
5	F_2	10000	5	F_2	10000
** F-> Failed, S-> Suspended*					

Usually, the shortfall mentioned above for rank regression is significant when the number of failures is small and the number of suspensions is large and not spread uniformly between failures. This is the case for the example data in table A1. In cases like this it is usually recommended to use the MLE instead of rank regression, since MLE does not look at the ranks or plotting positions, but rather considers each time-to failure or suspension.

The estimated parameters for the Weibull 3-parameter distribution using rank regression method are for both cases (case 1 and 2):

- β=0.81 and η= 11,383 hr

However, the MLE results for case 1 are:

- β=1.32 and η=6883 hr

And the MLE results for case 2 are:

- β=0.93 and η=21,447 hr

This example has shown that there is a difference in the results of the two data sets calculated using MLE and the results using rank regression. The results for both cases are the same when the rank regression estimation technique is used. This is due to the fact that rank regression considers only the positions of the suspensions. However, the MKE results are quite different for the two cases. For case 2 the value of η is much higher, which is due to the higher values of suspension times in case 2. This is because MLE technique, unlike rank regression, considers the value of the suspensions when estimating the parameters.

Appendix B
Cable Joint Type Overview

In this appendix, we will briefly indicate the type of cable joints that are categorized under synthetic, mass and oil insulated joints. Most of the time the type of cable joint that belongs to a certain category is based on the insulation material used for the cable joint and the brand. In table B1 an overview of which type of joint belongs to which category is shown. The categorization is based on information from the utility itself.

Table B1: In this table an overview of the different types and categories of 10 kV cable joints as recorded in the TKV database is shown. The information in this table has been made available by Stedin

Date		**9/4/2009**
Code	**Description from TKV**	**Type of Insulation**
0	Lovink KGX	Oil
1	Lovink	Oil
2	Massa	Mass
3	Nekaldiet	Synthetic
4	Olie	Oil
5	Raychem	Synthetic
6	DON SO127	Unknown
7	Schutmof	Mass
8	Onbekend	Unknown
9	Cellpack	Synthetic
A	Tece	Synthetic
B	Euromold	Synthetic
C	Kabeldon SMTD	Oil
D	Elastimold	Synthetic
E	Tyco	Synthetic
F	Kabeldon ABB	Oil

From table B1 it can be seen that for a number of cable joints the utility was unable to specify the type of insulation material. Furthermore, the names mentioned in the second column of the table are vendor specific names. For instance, a Kabeldon cable joint is a product of ABB. The information from this table is used to categorize the three different cable joints population, namely, synthetic insulated, mass insulated and oil insulated 10 kV cable joints.

High-Voltage Technology and Asset Management
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Address
Mekelweg 4
2628 CD Delft

Postal address
P.O. box 5031
2600 GA Delft

Telephone: +31 15 278 2870
Fax: +31 15 278 8382
Website: http://htm.ewi.tudelft.nl