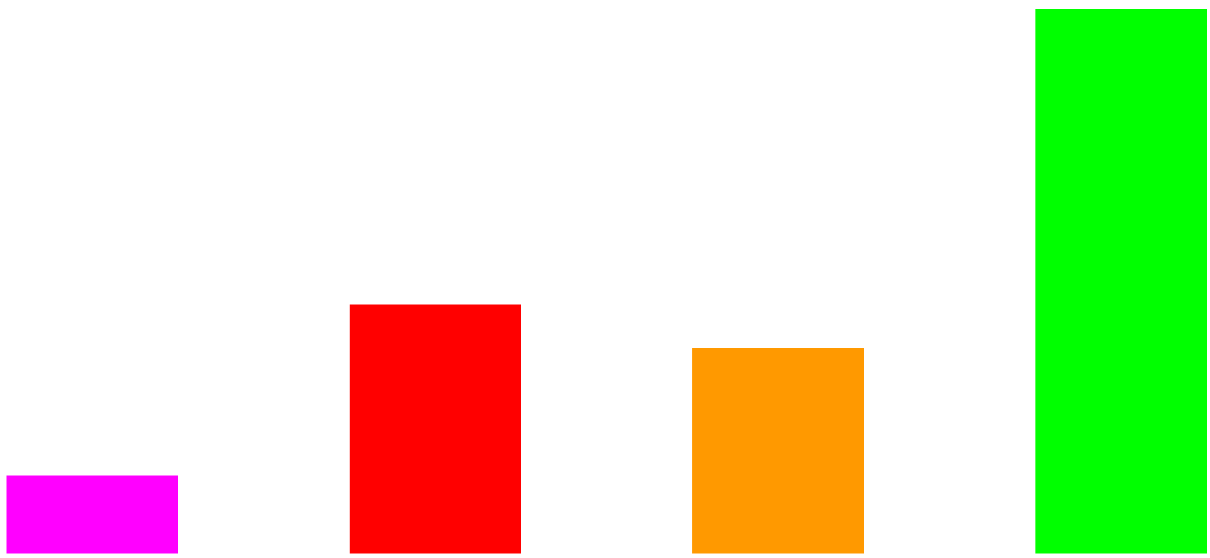
Extending the scientific foundation of the DNV GL Health Index methodology for electrical power equipment

Martijn S. Janssen, BSc

Delft University of Technology

Master of Science Thesis



Challenge the future

Martijn S. Janssen, BSc:
Extending the scientific foundation of the DNV GL Health Index methodology for electrical power equipment

For further information, contact the author: martijnj@gmail.com

Extending the scientific foundation of the DNV GL Health Index methodology for electrical power equipment

By

Martijn S. Janssen, BSc

in partial fulfilment of the requirements for the degree of

Master of Science

in

Electrical Sustainable Energy

at the Delft University of Technology,
to be defended on Thursday December 4, 2014 at 13:30 PM.

Thesis committee

Prof. dr. Johan J. Smit	TU Delft
Dr. Armando Rodrigo Mor	TU Delft
Dr. ir. Marjan Popov	TU Delft
Dr. ir. Dhiradj Djairam	TU Delft
Dr. ir. Jos M. Wetzer	DNV GL - Energy Advisory
Ir. Mischa E. Vermeer	DNV GL - Energy Advisory

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft Delft University of Technology

Summary

Western civilization is becoming dependent on electrical energy more than ever. This dependence sets high standards regarding reliability and security of supply. In order to meet the level of these standards, network operators consider grid planning and maintenance more frequently part of their core activities.

The present electricity network consists mainly of equipment put into service forty to fifty years ago. Similar figures appear from experience of network operators for the life expectancy of this equipment. Therefore, a large part of the electrical power equipment present in the current electricity network is approaching its life expectancy.

This causes the need for replacement of a large part of the electrical power equipment in the near future. The large amount of replacements is often referred to as the *replacement wave.* This increase in aged power equipment poses a potential threat to the security and reliability of supply and drives network operators to take adequate action.

During the past decade, new tools have become available that provide insight in the status of power equipment. These tools allow network operators to perform condition based maintenance, while taking safety and environmental issues into account.

The DNV GL Health Index (DNV GL HI) is one of these tools. The accuracy of this health index (HI) should be improved to maintain the value to its customers. In this thesis, the accuracy of the HI predictions is referred to as the *HI prediction quality.* This *HI prediction quality* shows the difference between the predicted and observed asset health. Furthermore, this *HI prediction quality* should meet a required level. Using this *required prediction quality level*, *validation* of the HI can be performed by comparing *the HI prediction quality* to the required level.

This thesis introduces systematic methods for *HI prediction quality quantification* and *validation* and shows the application of these methods to practical cases.

Chapter 2 reviews the previous work in the field of health indexing. It starts by describing the current state of art by comparing four of the present health indexing methodologies, including the methodology operated by DNV GL, according to five characteristics. For the characteristic covering the *HI foundation supporting methods*, limited literature was found on *HI prediction quality quantification* and *validation.* Furthermore, no practical case studies were found used for *HI prediction quality quantification* and *validation* of a complete HI, which indicated the need for research in this area using practical case studies. To conclude, this chapter has revealed that for such case studies, a *comparison between two datasets is necessary* and that *life data analysis* can be used for this comparison in such case studies.

Chapter 3 presents the research approach. This chapter starts explaining three datasets, predictions, observations and dummy data. Next, five methods for *HI prediction quality quantification* and *validation* are discussed. To conclude, the *life data analysis* is adopted for use in the analysis.

Chapter 4 explains the statistical methods and background used in this thesis. This chapter shows three types of probability functions. These functions are used to describe failure behaviour of assets using statistical distributions. Life data analysis is explained in this chapter in more detail. It is shown that for failure distribution fitting and parameter estimation, a parametric method using maximum likelihood estimation is most suitable to fit a failure distribution to the life data of the assets discussed in this thesis. To conclude, out of four methods to find the quality of the distribution's fit, the method using visual inspection was chosen.

Chapter 5 shows an overview of the collected and selected utility data. Furthermore, this chapter shows the preparation of this data for *prediction quality quantification* and *validation* of the DNV GL HI for *instrument transformers* and *distribution cables*.

Chapter 6 shows the results of the case study on *prediction quality quantification* and *validation* of the DNV GL instrument transformer Health Index. This case study using one *HI prediction quality quantification method* does not reject the hypothesis that it is possible to quantify *HI prediction quality* using utility data. No firm conclusion can be drawn regarding the finding of this case study on the DNV GL instrument transformer HI that the predictions were found to be *too pessimistic* and therefore *valid*.

Chapter 7 shows the results of the case study on *prediction quality quantification* and *validation* of the DNV GL distribution cable Health Index. This case study using one *HI prediction quality quantification method* does also not reject the hypothesis that it is possible to quantify *HI prediction quality* using utility data. No firm conclusion can be drawn regarding the finding of this case study on the DNV GL distribution cable HI that the predictions were found to be *too optimistic* and therefore *non-valid*. Furthermore, this practical case study reveals that for *HI prediction quality quantification* of cable circuits, besides a required data quality, the influence of the definition of a failure and changing reliability by changes in network topology should be incorporated.

Chapter 8 presents the conclusions and recommendations of this thesis.

Keywords

- Asset management
- Health index
- HI prediction quality quantification
- Validation
- Reliability
- Ageing assets

Acknowledgements

This thesis is a result of my studies in Delft. It would not have been completed without the help and support of many people.

To start, I want express my gratitude to dr. ir. J. M. Wetzer, first of all, for offering me this graduation project and in addition to that, for his guidance and support during the project. My second word of thanks goes to ir. M. E. Vermeer, for his guidance and support during the project. The meetings with both of you and your feedback always brought me new ideas and made me rethink concepts, which have undoubtedly contributed to the quality of this thesis.

Furthermore, I would like to thank Prof. dr. J. J. Smit for his support in performing this project. Many thanks go to my supervisor dr. ir. D. Djairam. The numerous chapters (and formats) I have sent you and the even greater number of corrections and suggestions you gave me have without any doubt improved this thesis and moreover, improved my writing skills. Lastly, my thanks go to dr. ir. M. Popov, for his time taking part in my graduation committee.

In addition, I am also thankful to ir. R. P. Y. Mehairjan and dr. L. A. Chmura for the valuable discussions in Delft on the research methods.

Next, a word of thanks is for ir. H. C. Wels, for his introduction in the reliability theory and his help on finding the right failure data. Furthermore, ir. E. J. de Haan deserves a word of thanks, for his valuable comments on my methods.

My thanks go out to all employees of the utilities to help me by their effort in assessing whether their utility data could be valuable for my research.

In particular, I am grateful to the two utilities and their employees who provided me with the necessary data. Firstly, to Piet Oosterlee, for providing me access to the necessary utility data. Secondly, to Jaap Moerland, for always being willing to collect and explain extra utility data when needed. Thirdly, to Erik Schenkel, for providing me access to the utility data and explanation on this data, vital for this research.

A great word of thanks goes to my parents Hans and Tineke, for their support, both emotional and financial. Furthermore, to my brothers Joris and Thomas, for their strong support during both the successful and less successful months.

Many thanks go to my friends in Delft, supporting me and providing me with a place to spend the night. I thank my roommates in Velp. Jos, Gijs and Thijs, you always achieved to keep me running after a hard day of work.

Finally, I would like to pay tribute to all pioneers in gliding and the members of my gliding club, who contributed to this sport, enabling me to see work from a different angle.

Many thanks to all of you. I wish you in particular, and of course every reader, a lot of pleasure in reading this thesis.

Contents

Summary v

Acknowledgements vii

Nomenclature xiii

1 Introduction 1
1.1 Asset management for electrical power equipment . . . 2
1.2 Assessment functions . . . 3
1.3 Research description . . . 3
1.4 Research problem . . . 4
1.5 Research objectives . . . 5
1.6 Research methodology . . . 6
1.7 Scientific challenges . . . 6
1.8 Scope of research . . . 7
1.9 Thesis outline . . . 8

2 Literature research 9
2.1 Existing HI SoTs . . . 10
2.1.1 DNV GL Energy HI SoT . . . 13
Characteristic 1: HI requirements . . . 14
Characteristic 2: Asset data types . . . 14
Characteristic 3: HI output representations . . . 16
Characteristic 4: Assessment functions . . . 19
Characteristic 5: HI foundation supporting methods . . . 21
2.1.2 EA Technology HI SoT . . . 25
Characteristic 1: HI requirements . . . 25
Characteristic 2: Asset data types . . . 25
Characteristic 3: HI output representations . . . 26
Characteristic 4: Assessment functions . . . 27
Characteristic 5: HI foundation supporting methods . . . 28
2.1.3 Kinectrics HI SoT . . . 30
Characteristic 1: HI requirements . . . 30
Characteristic 2: Asset data types . . . 30
Characteristic 3: HI output representations . . . 31
Characteristic 4: Assessment functions . . . 33
Characteristic 5: HI foundation supporting methods . . . 35
2.1.4 TU Delft HI SoT . . . 36
Characteristic 1: HI requirements . . . 36

Characteristic 2: Asset data types 37
Characteristic 3: HI output representations 38
Characteristic 4: Assessment functions 39
Characteristic 5: HI foundation supporting methods 40
2.2 Strengths/weaknesses analysis 41
2.3 Conclusions 43

3 Research approach **45**
3.1 HI prediction quality quantification and validation: methods overview 45
3.2 HI prediction quality quantification and validation: datasets 46
3.2.1 Predictions and observations datasets 46
3.2.2 Dummy dataset 49
3.3 Lifetime definition 49
3.4 Lifetime comparison 50
3.5 HI prediction quality quantification and validation: methods description 51
3.5.1 Method 1: New HI 51
3.5.2 Method 2: Utility criteria 52
3.5.3 Method 3: Binary classification 52
3.5.4 Method 4: Prediction error 54
3.5.5 Method 5: Life data analysis 57
3.6 Discussion of the methods 60
3.7 Adopted methods 63

4 Statistical methods and background **65**
4.1 Probability functions 65
4.2 Statistical distributions 66
4.2.1 Bathtub curve 67
4.2.2 Normal distribution 67
4.2.3 Weibull distribution 68
4.3 Life data analysis 69
4.4 Failure distribution fitting and parameter estimation 69
4.4.1 Visual inspection 70
4.5 Summary and conclusions 70

5 Utility data: collection and selection **71**
5.1 Data request procedure 71
5.1.1 Data requirements 71
5.1.2 Contacted organisations 73
5.2 Acquired data and data selection 73
5.2.1 Acquired data 73
5.2.2 Data selection 76
5.3 Data preparation: DNV GL HV instrument transformer Health Index 76
5.3.1 IT information 76
5.3.2 Utility data files summary 76
5.3.3 Data organization 77
5.3.4 HI asset data 78
Predictions dataset 78
Observations dataset 78
5.4 Data preparation: DNV GL distribution cable Health Index 78
5.4.1 Cable circuit information 78

5.4.2 Utility data files summary . . . 78
5.4.3 Data organization . . . 79
5.4.4 HI asset data . . . 82
Predictions datasets . . . 82
Observations dataset . . . 85

6 Case study: DNV GL instrument transformer Health Index 87
6.1 DNV GL IT HI . . . 87
6.1.1 Asset data . . . 87
6.1.2 HI remaining lifetime definition . . . 88
6.1.3 HI output representation used for the comparison . . . 88
6.1.4 Software and hardware . . . 88
6.2 Assumptions and hyphothesis . . . 88
6.3 Results . . . 89
6.3.1 Predictions . . . 89
6.3.2 Observations . . . 90
6.3.3 HI prediction quality quantification and validation by predictions and observations comparison . . . 90
6.4 Conclusions . . . 93
6.5 Recommendations . . . 93

7 Case study: DNV GL distribution cable Health Index 95
7.1 DNV GL distribution cable HI . . . 95
7.1.1 Asset data and settings . . . 95
7.1.2 HI remaining lifetime definition . . . 97
7.1.3 HI output representation used for the comparison . . . 97
7.1.4 Software and hardware . . . 97
7.2 Assumptions and hyphothesis . . . 97
7.3 Results . . . 98
7.3.1 Predictions . . . 98
Predictions ‘actual’ load . . . 98
Predictions ‘tweaked’ load . . . 99
7.3.2 Observations . . . 100
7.3.3 HI prediction quality quantification and validation by predictions and observations comparison . . . 101
Comparison predictions ‘actual’ load to observations . . . 101
Comparison predictions ‘tweaked’ load to observations . . . 103
7.4 Conclusions . . . 106
7.5 Recommendations . . . 106

8 Conclusions and recommendations 107
8.1 Conclusions . . . 107
8.2 Recommendations . . . 108
8.2.1 Recommendations for DNV GL . . . 108
8.2.2 Recommendations for further research . . . 109

Bibliography 111

Appendices 118

A Additional health index requirements, asset data and comparison 119
A.1 DNV GL HI SoT . . . 119
A.1.1 Additional requirements DNV GL HI SoT . . . 119
A.1.2 Asset data types DNV GL HI SoT . . . 120
A.1.3 Detection method classification DNV GL HI SoT . . . 123
A.2 EA Technology HI SoT . . . 124
A.2.1 Additional requirements EA Technology HI SoT . . . 124
A.2.2 Asset data types EA Technology HI SoT . . . 125
A.2.3 Detection method classification EA Technology HI SoT . . . 126
A.3 Kinectrics HI SoT . . . 127
A.3.1 Additional requirements Kinectrics HI SoT . . . 127
A.3.2 Asset data types Kinectrics HI SoT . . . 127
A.3.3 Detection method classification Kinectrics HI SoT . . . 129
A.4 TU Delft HI SoT . . . 130
A.4.1 Additional requirements TU Delft HI SoT . . . 130
A.4.2 Asset data types TU Delft HI SoT . . . 130
A.4.3 Detection method classification TU Delft HI SoT . . . 134

B Extended tables with asset data 135

C Extended tables with asset data 137

Nomenclature

AM	Asset Management
Assessment function	Any mathematical expression or formula within a HI that translates asset data into technical asset health
Asset data	Any type of data available about an asset
Asset health	Technical condition of an asset
Asset management decision support software	Software containing a HI, to aid asset managers in taking asset management decisions
CombiT	Combined transformer
Condition parameters	Any type of measured asset data that is linked to asset health
CT	Current transformer
DGA	Dissolved gas analysis
"EDPLK"	Paper-insulated lead covered (PE outer sheath)
Failure mode effects analysis (FMEA)	A step-by-step approach for identifying all possible causes of asset failure
Health index (HI)	The method used to translate asset data into asset health
Health index school of thought (HI SoT)	Group of one or multiple HIs with a similar approach, usually operated by institutions or companies
HI score	Score resulting from a health index that indicates asset health
IT	Instrument transformer
Online PD (SCG)	Online partial discharge measurements by DNV GLs Smart Cable Guard
Offline PD	Offline partial discharge measurements. In this thesis, they refer to the Oscillating Wave Test System (OWTS) or the 0.1 Hz Very Low Frequency (VLF) method.
PILC	Paper-insulated lead covered (steel armour outer sheath)
Mixed	Cable circuits that consist of combination of XLPE, PILC and "EDPLK"
RL	Remaining Lifetime
VT	Voltage transformer
XLPE	Cross-linked polyethylene

Chapter 1

Introduction

Western civilization is becoming dependent on electrical energy more than ever. This dependence sets high standards regarding reliability and security of supply. In order to meet the level of these standards, electrical network operators consider grid planning and maintenance more frequently part of their core activities.

The present electricity network consists mainly of equipment put into service forty to fifty years ago. Similar figures appear from experience of network operators for the life expectancy of this equipment. Therefore, a large part of the electrical power equipment present in the current electricity network is approaching its life expectancy.

This causes the need for replacement of a large part of the electrical power equipment in the near future. The large number of replacements is often referred to as the replacement wave.

A graphical example of the replacement wave is given by the chart in figure 1.1. This chart presents the number of installations and replacements of power equipment on a yearly basis. This chart confirms the increasing amount of aged power equipment.

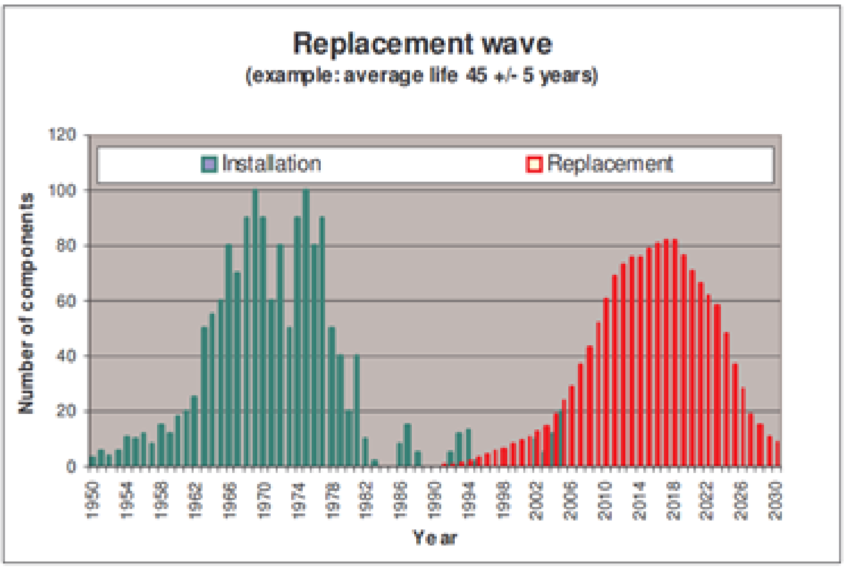


Figure 1.1: General representation of the replacement wave for electrical power equipment [1]

This figure introduces the replacement wave by means of a general representation. The number of yearly installations and replacements are represented by the green and red bars, respectively.

This increase in aged power equipment poses a potential threat to the security and reliability of supply and drives network operators to take adequate action.

Developments in software and computing power during the past decade have led to significant improvements in probability calculations using computer simulations. This caused new tools to become available that provide insight in the status of power equipment. These tools allow network operators to perform condition based maintenance, while taking safety and environmental issues into account. In the long run, this should lead to an increase in security and reliability of supply, better management of risks and reduction of costs.

The DNV GL Health Index (DNV GL HI) is one of these tools. The accuracy of this health index (HI) should be improved to maintain DNV GL's current position in the market of health indexing tools. In this thesis, the accuracy of the HI predictions is referred to as the *HI prediction quality*. This *HI prediction quality* shows the difference between the predicted and observed asset health. Furthermore, this *HI prediction quality* should meet a *required level* of prediction quality. Using this *required prediction quality level*, *validation* of the HI can be performed by comparing *the HI prediction quality* to the *required level*.

Section 1.1 gives an introduction to asset management for electrical power equipment. Next, section 1.2 gives the research description, followed by a general description of assessment functions in section 1.3. Thereafter, the research problem is presented in section 1.4, followed by the research objectives in section 1.5 and the research approach in section 1.6. Section 1.7 shows the scientific challenges, section 1.8 defines the scope of the research and finally in section 1.9, an outline of the thesis is given.

1.1 Asset management for electrical power equipment

The term *asset management* is used in multiple fields of study. This chapter explains *asset management* in the context of electrical power equipment.

In general, assets are defined as any items that have a distinct value to the organization. Electrical power equipment belongs to the physical assets owned by a network operator.

The British Standards Institution (BSI) defines asset management in its Publicly Available Standard (PAS55) [2] as follows:

> "Systematic and coordinated activities and practices through which an organization optimally and sustainably manages its assets and asset systems, their associated performance, risks and expenditures over their life cycles for the purpose of achieving its organizational strategic plan"

where an organizational strategic plan is defined as [2]:

> "Overall long-term plan for the organization that is derived from, and embodies, its vision, mission, values, business policies, stakeholder requirements, objectives and the management of its risks"

Figure 1.2 shows the concept of asset management for electrical power equipment. Asset managers of network operators perform the steps shown in this figure continuously according to the organizational plan. Table 1.1 explains the actions taken for each step in the figure.

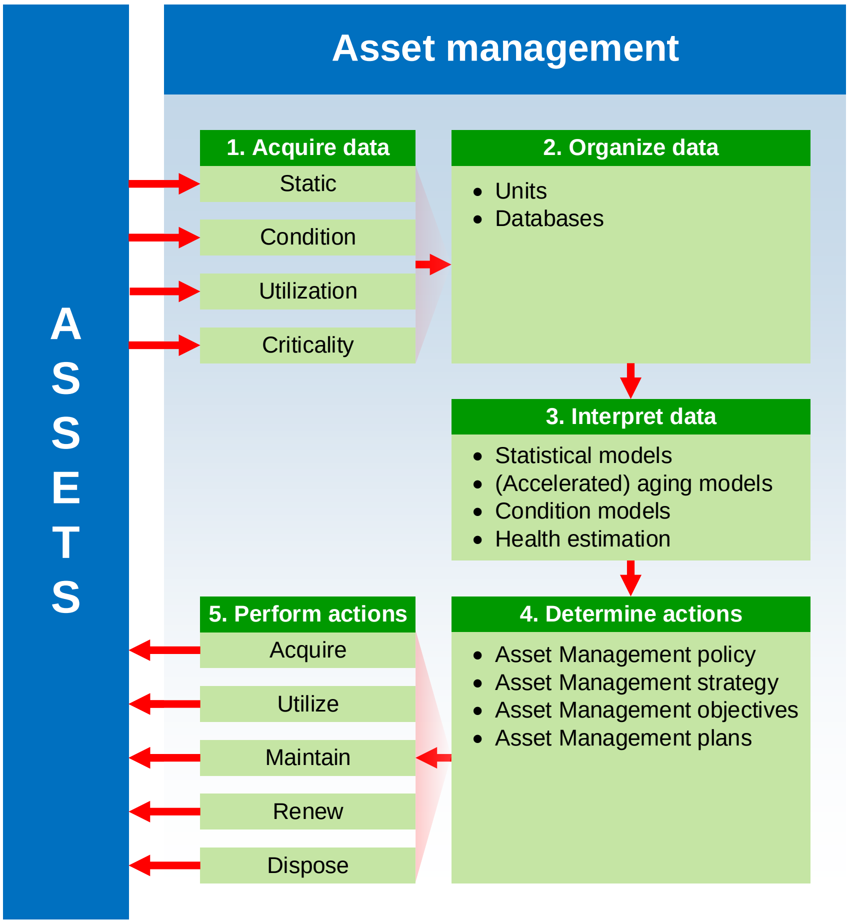


Figure 1.2: Asset management for electrical power equipment (step 4 and 5 based on [2]) *This figure shows the concept of asset management for electrical power equipment. The arrows indicate the direction of the process. The actions for each step are explained in table 1.1.*

1.2 Assessment functions

Assessment functions interpret the organized data. Part of this research is to develop a strategy for validation of these functions. Therefore, a general description of these functions is given in this section. They are described in more detail in chapter 2.

The organized data may consist of continuous or discrete values. The results of the assessment functions represent the technical condition of a part of the asset by the *time to required action*. *Time to required action* can either be the *remaining lifetime* (which is the time to replacement) or the *time to additional maintenance*, following from the calculations performed by the assessment functions. These calculations are based on international standards (IEC), literature (IEC, IEEE, Cigré) and expert knowledge.

1.3 Research description

The tool in focus during this research is the DNV GL HI. This tool is used by asset managers of utilities to gain insight in the health of their electrical power equipment. The tool aims to provide decision-support for medium to long term asset management planning. Currently, the tool is based on Microsoft Excel accompanied with the @RISK plugin for Monte Carlo

Table 1.1: Explanation of the steps in asset management for electrical power equipment
This table explains the actions for each step in figure 1.2.

Step	Explanation
1. Acquire data	This step includes the acquirement of all sorts of data per asset. Static data is data about the asset which does not change over time (e.g. nameplate data). Condition data is data from sensors in/around equipment and data from visual inspections. Utilization data contains information on regarding asset loading and asset loading patterns. Criticality data contains information per asset on the neighborhood in which the asset is situated and the importance of the asset in the electrical network.
2. Organize data	In this step, the data is put into a consistent format (i.e.: the units and database types are uniform and missing values are calculated).
3. Interpret data	This step uses ageing, condition and statistical models to estimate the technical condition (health) of the assets using the organized data of step 2. A major part of this step is performed in the DNV GL HI by algorithms called assessment functions. A general explanation of these algorithms is given in section 1.2.
4. Determine actions	During this step, the asset health is translated into actions aligned to the organizational plan of the network operator. In addition to its HI, DNV GL offers a separate risk module that provides decision-support for this step.
5. Perform actions	The final step performs the necessary actions resulting from step 4.

simulations. Chapter 2 explains the DNV GL HI in more detail.

Figure 1.3 shows the part of asset management covered by the DNV GL HI.
This research focuses on the part of the DNV GL HI responsible for the interpretation of the organized data. In the DNV GL HI, this part is largely covered by assessment functions.

1.4 Research problem

The algorithms of the DNV GL HI are currently based on both expert knowledge and existing standards. At this moment, the methods applied to validate its algorithms are limited.

The validation methods currently applied to the assessment functions are:

- Perform a check to ensure the input-output relationship of each assessment function is in line with its model(s) by using test data.
- Perform an overall check per asset type of the HI by looking at health assessments according to the client's data.

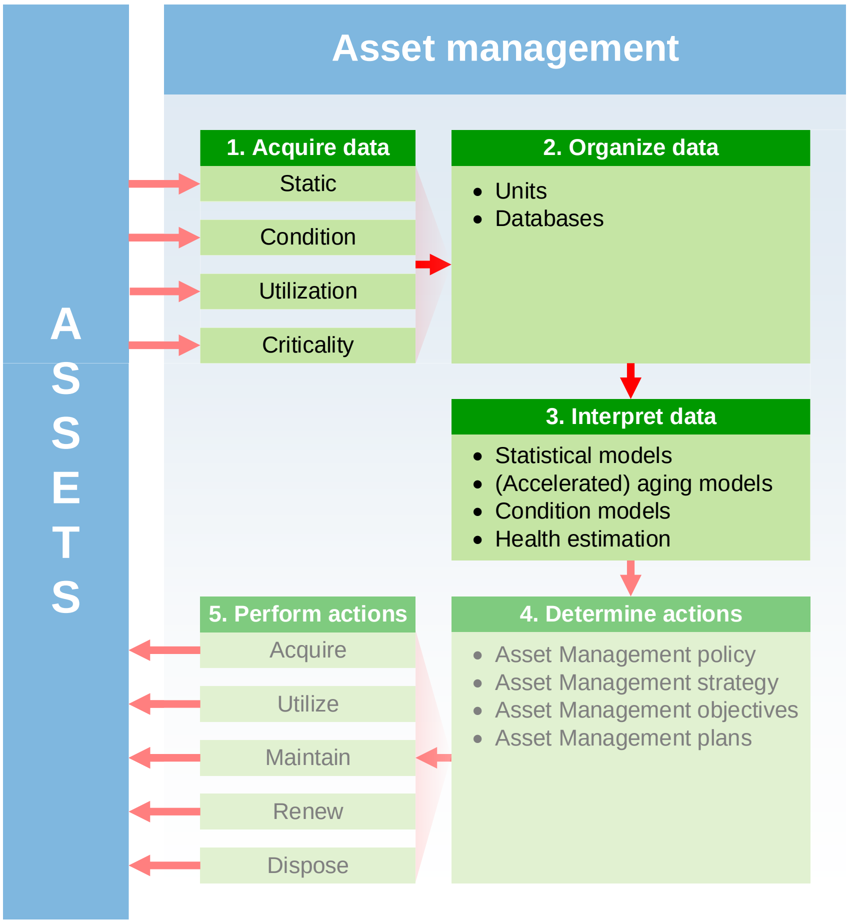


Figure 1.3: Part of asset management covered by the DNV GL HI

These validation methods ensure the accurate functioning of the assessment functions according to their respective models. However, they do not validate the models behind the assessment functions themselves. DNV GL aims for a more accurate and improved foundation for its HI. This leads to the main question addressed in this thesis:

How can the scientific foundation of the DNV GL HI be extended?

This question can be divided into three sub questions:

- In which way can *HI prediction quality* of the DNV GL HI be quantified?
- In which way can *HI prediction quality quantification* of the DNV GL HI be used to validate the DNV GL HI?
- Using additional data, which statements can be made on *HI prediction quality* and *validation* of the DNV GL HI?

1.5 Research objectives

The main objective of this research is to extend the scientific foundation of the DNV GL HI. This is performed by proposing methods for validation of its algorithms. The applied validation methods are described in more detail in section 1.6.

The research objectives are:

- Finding the current *state of art* in health indices and their *HI prediction quality quantification* and *validation methods* by means of a literature research
- Proposing and selecting methods to perform *HI prediction quality quantification* of the DNV GL HI
- Performing *validation* of the DNV GL HI using the *HI prediction quality quantification methods*
- Collecting and selecting data necessary for use in performing the *HI prediction quality quantification methods* and *validation* method
- Applying the selected *HI prediction quality quantification methods* and *validation method* using collected data to quantify the *HI prediction quality* and to show *validation* of the DNV GL HI

1.6 Research methodology

The research is divided into two main parts. These parts, a theoretical part and an analysis part, are described in this section.

The theoretical part explains the basics of asset management and provides an overview of the current state of art in health indexing and their *HI prediction quality quantification* and *validation* methods present by means of a literature research.

In the analysis part, methods are proposed to perform *HI prediction quality quantification* and *validation* of the DNV GL HI. Furthermore, data is collected and prepared for performing the methods for *HI prediction quality quantification* and *validation.* Finally, the developed methods for *HI prediction quality quantification* and *validation* are applied to the collected data to demonstrate *HI prediction quality quantification* and *validation* of the DNV GL HI.

1.7 Scientific challenges

In [3], the results of a study on optimizing the maintenance planning by looking at the condition indexing process are presented. In this study, a large part of asset management for electrical power equipment is covered.

As recommended by [3], more research should be done regarding the relationship between occurring failures and condition indicators. This is performed in this thesis by proposing methods for *HI prediction quality quantification* and *validation.* This section shows the scientific challenges in this thesis by presenting relevant questions.

Questions concerning the validation of the DNV GL HI are:

- Which requirements are present regarding the data necessary for use in performing the methods for *HI prediction quality quantification* and *validation*?
- Some assets have undergone maintenance which is either recorded or not recorded. How can historical maintenance data (or lack thereof) be included to improve the outcome of the validation result?

1.8 Scope of research

This research focuses on developing a strategy for *HI prediction quality quantification* and *validation* of the DNV GL HI. As it is required to finish the project within nine months, the scope of the research is limited.

The scope of research is summarized by the following points:

- As indicated in figure 1.4, this research focuses on *HI prediction quality quantification* and *validation* of the DNV GL HI. The research is limited to the DNV GL HIs that cover the following asset types:
  - HV Instrument Transformers
  - 10kV distribution cables
- Excluded from the research:
  - All other asset types
  - All other types of functions present in the DNV GL HI
  - The development of new assessment functions

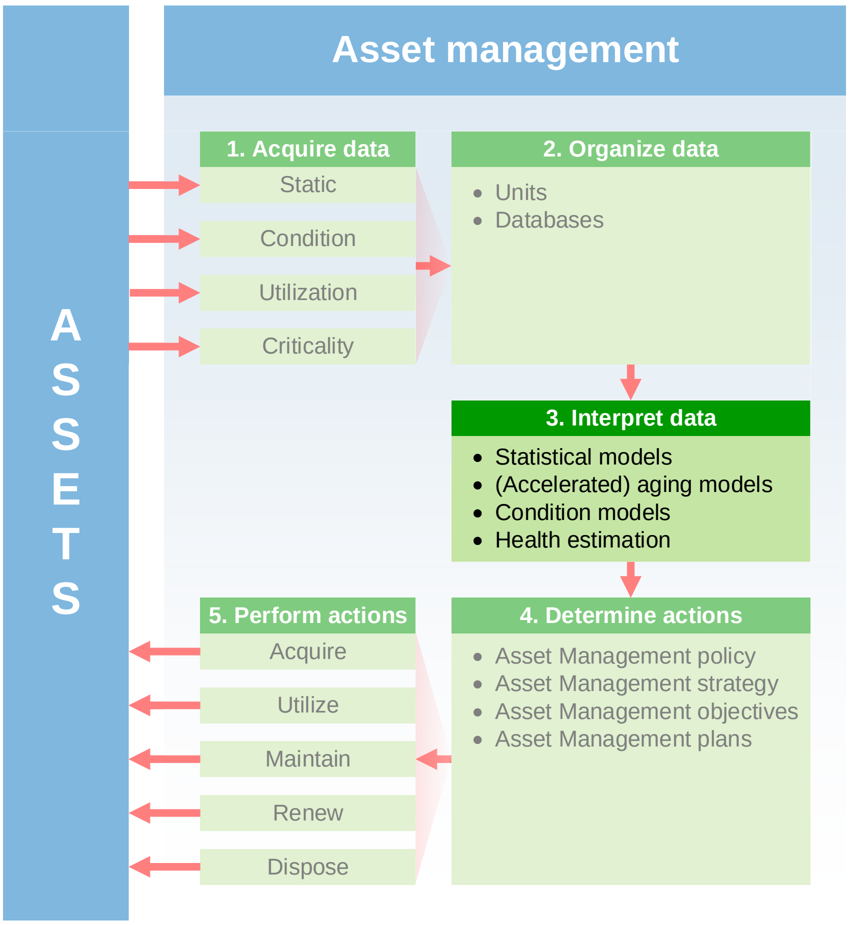


Figure 1.4: Part of asset management in focus during this MSc. research

1.9 Thesis outline

This thesis report consists of eight chapters.

Chapter 2 reviews the previous work in the field of health indexing. It starts by describing the current state of art by comparing present health indexing methodologies including the methodology operated by DNV GL. Then, it continues showing the present methods used to validate (parts of) health indices.

Chapter 3 presents the adopted methodology during the research. This chapter is started by describing the methodology adopted to validate (parts of) the health index.

Chapter 4 explains the statistical methods and background used in this thesis.

Chapter 5 gives an explanation of the data and software used for *HI prediction quality quantification* and *validation* of the DNV GL HI.

Chapter 6 shows the results of a case study which uses additional data for applying the *HI prediction quality quantification* and *validation* methods to data from a population of high voltage instrument transformers.

Chapter 7 shows the results of a case study which uses additional data for applying the *HI prediction quality quantification* and *validation* methods to data from a population of 10kV distribution cables.

Chapter 8 presents the conclusions and recommendations of this thesis.

Chapter 2

Literature research

The focus in this thesis will be on *prediction quality quantification* and *validation* of the technical aspects of a health index (HI). This technical focus is taken into account during the analysis of the existing HIs. HIs can be grouped according to their characteristics. These groups are referred to as *health index schools of thought*. For ease of reading, this text indicates them by HI SoTs.

This literature research will show an overview of presently known and used HI SoTs. Using this overview, the strengths and weaknesses of the DNV GL HI relative to these other HI SoTs are determined. Furthermore, it aims to find existing methods in *HI prediction quality quantification* and *validation*.

To achieve this, we start by describing different HI SoTs found in literature. Section 2.1 describes these SoTs by characteristics of their HIs and section 2.2 compares the HI SoTs by their characteristics. The sources used in this chapter are the libraries of IEEE, Cigré and repositories of various universities (e.g. Delft University of Technology).

Figure 2.1 shows the relation between the expressions and characteristics described in this chapter. The HI SoTs are at the top of the hierarchy. Underneath the HI SoTs, one or multiple HIs are found. Each HI is described according to five HI characteristics.

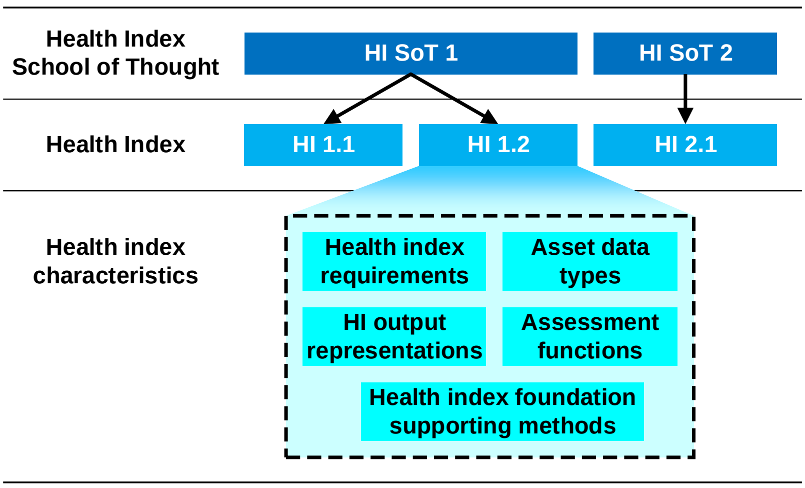


Figure 2.1: Hierarchical overview of the terminology used in this chapter

2.1 Existing HI SoTs

In literature, a wide variety of expressions is found for terminology within HI SoTs. For the ease of reading and prevention of misinterpretation, this chapter aims for uniformity in the terminology used for the concepts within HI SoTs. Table 2.1 explains these expressions as used in this chapter.

Table 2.1: Main HI expressions
This table contains the main HI expressions used in this chapter.

Expression	Explanation
Asset data	Any type of data available about an asset
Asset health	Technical condition of an asset
Assessment function	Any mathematical expression or formula within a HI that translates asset data into technical asset health
Asset management decision support software	Software containing a HI, to aid asset managers in taking asset management decisions
Condition parameters	Any type of measured asset data that is linked to asset health
Failure mode effects analysis (FMEA)	A step-by-step approach for identifying all possible causes of asset failure.
Health index (HI)	The method used to translate asset data into asset health
Health index school of thought (HI SoT)	Group of one or multiple HIs with a similar approach, usually operated by institutions or companies
HI score	Score resulting from a health index that indicates asset health

For each HI SoT, a description following from the literature research is given in the order indicated by table 2.2. Following this table, each section starts with characteristic #1, describing the requirements of a HI. Then, each section describes the HIs found by characteristics #2 through #4. To conclude, each section describes the methods found to support the foundation of their HIs by characteristic #5.

HIs are subject to continuous improvements. Therefore, there are several versions of most of the HIs. In most cases, the recent versions of these health indices lack published scientific literature and publicly available documentation of their institutions. Therefore, this chapter will not compare the most recent HI versions.

This chapter does not cover all HI SoTs. Table 2.3 gives an overview of the HI SoTs discovered during the literature research, including the number of publications associated with each HI SoT. As shown in this table, the majority of the literature found describes HI SoTs regarding electrical power equipment. The literature on integrated vehicle health management showed that for assets different from electrical power equipment, similar characteristics are considered.

Table 2.2: Overview of the HI characteristics described for each HI SoT

#	Characteristic	Description
1	HI requirements	Main requirements a HI needs to comply with.
2	Asset data types	Description of the asset data used by the HI SoT.
3	HI output representations	Representations for a single asset or a group of assets by the HI SoT using at least health index score(s).
4	Assessment functions	The mathematical expressions and formulas used to translate condition data into a health representation.
5	HI foundation supporting methods	Methods of HI SoTs to support the foundation of their HIs.

Table 2.3: Publications on HI SoTs discovered during this research

#	HI SoT	Type of assets in literature	Number of associated publications
1	DNV GL Energy	Electrical power equipment	15
2	EA Technology	Electrical power equipment	11
3	Kinectrics	Electrical power equipment	11
4	TU Delft	Electrical power equipment	31
5	Meridium	Electrical power equipment	6
6	IBM	Electrical power equipment	6
7	EDF	Electrical power equipment	6
8	Integrated vehicle health management	Cars, aeroplanes	3

As can be seen in table 2.3, significant differences exist in the number of publications found for each HI SoT. Moreover, the level of detail in describing the health indexing methods was found to vary significantly among the HI SoTs. The HI SoTs #5 through #8 lack a description of more than half of the characteristics discussed in this section. This chapter aims to make a comparison between the HI SoTs. In order to achieve a fair and thorough comparison, the HI SoTs #5 through #8 were disregarded.

The literature on these HI SoTs develops over time. To gain insight in developments of and dependencies between the researched HI SoTs over time, figure 2.2 shows an overview of the literature described in this chapter. In this figure, each arrow represents a reference between publications. For example: [4] refers to [5], [6] and [7].

This figure shows that in most cases, publications refer to its corresponding HI SoT. The arrows between the HI SoTs reveal interesting relationships. For example, two of the references in the Kinectrics HI SoT point to literature in the EA Technology HI SoT. These arrows indicate a relation between their HIs, which their description in literature indicates for this example.

In general, however, these links should be derived with caution. The goal of references pointing to previous literature might differ from further development of existing HIs. For example, a goal can be to criticize an existing method. In this case, not necessarily a dependency exists with the HI the reference points to. Sections 2.1.1 to 2.1.4 provide an overview of existing

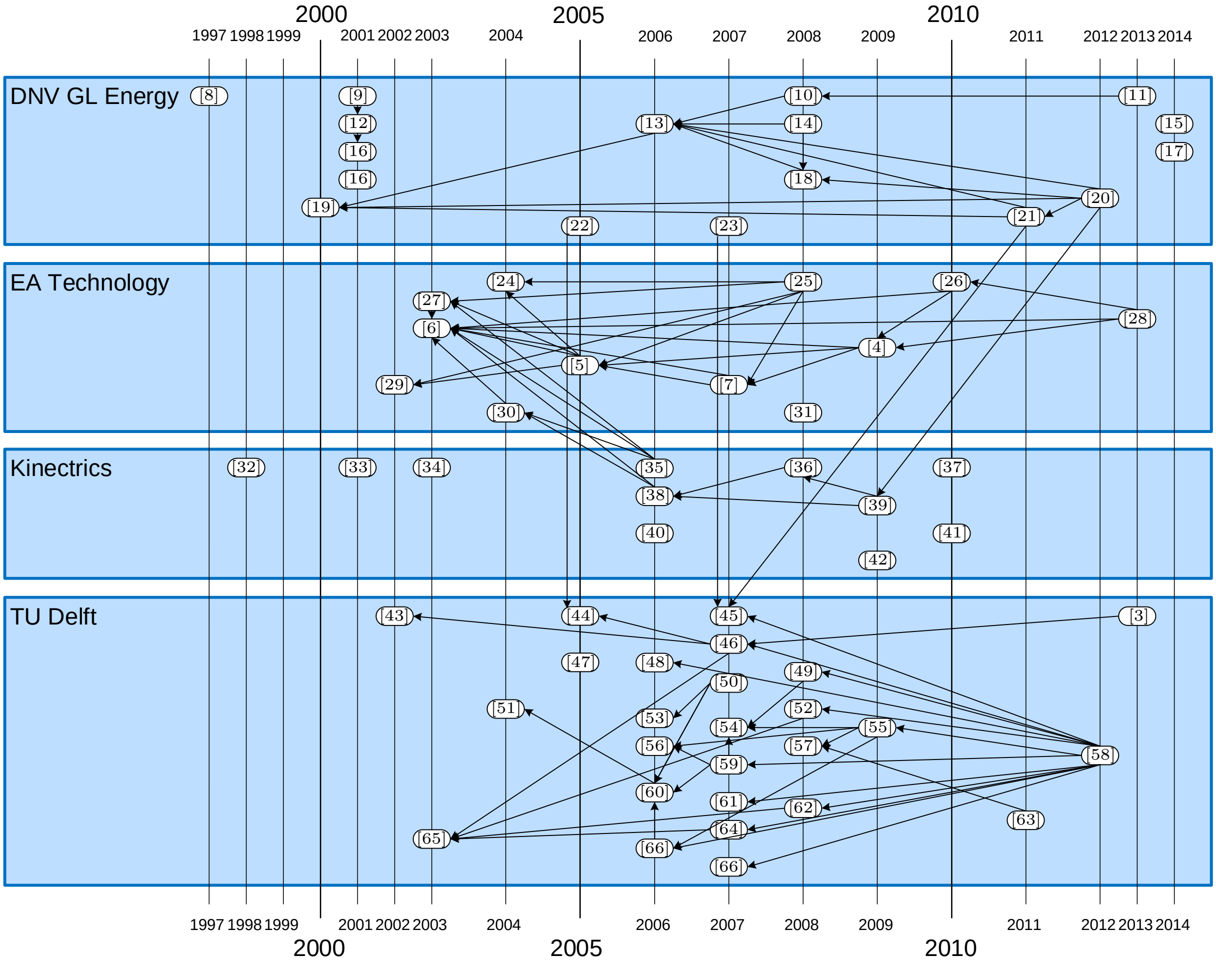


Figure 2.2: Overview of the literature in regarding HI SoTs
The references from the Kinectrics HI SoT to the EA Technology SoT indicate a relation between their HIs.

HI SoTs by a comparison of the HI characteristics discussed above.

2.1.1 DNV GL Energy HI SoT

In the beginning of this century, the Dutch energy market was liberalized [23]. This increased the competition between utilities. Furthermore, an upcoming replacement wave was expected due to an ageing asset base [8] (see chapter 1).

These two drivers forced Dutch utilities to take a closer look at the costs of their asset management activities. These activities shifted from optimizing for network performance avoiding risks towards taking controlled risks [12].

This shift in activities was accelerated by the trend of gradually shifting from corrective and time-based maintenance to condition-based maintenance (CBM) [19]. This shift led to the introduction of diagnostics in the asset management decision process. This, in turn, created the need for software tools to implement these rapid changes in maintenance processes.

DNV GL Energy has created several software tools for this purpose, two of which are MainMan [8] and the DNV GL HI [11].

Characteristic 1: HI requirements

For this characteristic, we describe the requirements found in literature for the DNV GL Energy HI SoT. The requirements are divided into categories. Below, the requirements regarding the characteristics of table 2.2 are presented.

Asset data types

- A HI should be able to cope with uncertainty in the input [11].
- HIs should incorporate the results of suitable diagnostic- and condition assessment tools to enable condition based maintenance (CBM) [19].

HI output representations

- HIs should provide a reliable forecast of future behaviour [10].

Assessment functions

- Assessment functions should be subjected to techniques that reduce the uncertainty of their predictions, two of which are [13]:
  - Coupling the assessment functions to externally measureable quantities.
  - Performing a sensitivity analysis on the assessment functions. The result of this analysis shows which (extra) input data most efficiently improves overall accuracy.
- From HIs for power transformers, it can be concluded that an uncertainty analysis should be performed when modelling ageing processes [21].
- Implementing CBM in a HI requires:
  - Good knowledge of failure and deterioration mechanisms including their criticalities. This is also known as Failure Modes Effects and Criticality Analysis (FMECA) [19, 9].
  - Presence of suitable indicators for the status of failure and degradation [19, 9].
  - Incorporation of reliable assessment tools to translate measurements into technical component health [19, 9].
- During incorporation of failure data analysis in an assessment function, the distinction should be made between equipment-parts and failure causes [9].

Besides the requirements regarding the characteristics, this HI SoT also describes requirements of additional characteristics. These requirements are listed in section A.1.1.

Characteristic 2: Asset data types

This characteristic shows the asset categories and number of asset data types found in literature for the DNV GL Energy HI SoT. The number of asset data types are presented in table 2.4. The table is divided into main categories by the first column. The second column shows shows the number of asset data types. The final column presents the reference to literature that indicated this asset data type. An extended version of this table is presented in section A.1.2.

Table 2.4: Number of asset data types as found in literature for the DNV GL Energy HI SoT

Category	Number of asset data types	Reference
Transformer	18	[18, 20, 9, 19]
Switchgear	19	[9, 19, 8, 67]
Power cables	7	[19, 16]
Generators	10	[19]

This table shows that, when compared to the other HI SoTs, the literature of the DNV GL Energy HI SoT describes a considerable number of asset data types.
Table 2.5 shows the result of classifying detection methods of asset data present within the DNV GL HI SoT for five different subjects. The first two columns in this table show the subject and for each subject, two classifications. The final column shows the number of detection methods belonging to each of the classifications.

Firstly, the assumptions taken for deriving the values of this table are explained.
For costs, a qualitative estimation of the costs of acquiring the asset data for *all assets* of a utility was performed. To enable this estimation, it was assumed that when the asset data should be acquired, it should be acquired for *all assets* of the utility.
A note regarding this assumption is necessary. Among the HI SoTs, different strategies exist in acquiring asset data. For example, in the DNV GL HI SoT, it is generally not required to acquire *all asset data types* for *all assets* of a utility. Furthermore, as most utilities use condition based maintenance, this data is readily available which significantly reduces the costs as the data should only be transferred into the format of the DNV GL HI. This comparison was performed using only the author's knowledge on the qualitative costs of (detection) methods.

For determining whether application of the (detection) methods requires the asset to be taken out of service, only the author's knowledge was used.
The maturity of the (detection) methods, the ability of (detection) methods to detect ageing and the accuracy of (detection) methods were determined qualitatively using the author's knowledge.

Table 2.5: Classification of the detection methods of the DNV GL Energy HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table summarizes the table in section A.1.3. This table contains a classification for the individual detection methods.

Subject	Subject classification	Number of detection methods for subject classification
Costs	Low (often already available)	5
	High	52
Out of service required	No	36
	Yes	21
Maturity	Mature	35
	Development phase	22
Ageing detection	Long term ageing	15
	Short term ageing	42
Accuracy	Good	49
	Bad	8

From table 2.5, the following conclusions can be drawn for the DNV GL HI SoT:

- The majority of the asset data mentioned in the DNV GL HI SoT is expensive to acquire for utilities for all assets if this data is not yet available for any of the assets.
- For acquiring most asset data, it is not required to take the equipment out of service.
- Most asset data types originate from mature (detection) methods.
- Most asset data types only cover short term ageing.
- Most (detection) methods achieve a good accuracy for the majority of the asset data.

Characteristic 3: HI output representations

This characteristic covers the *HI output representations* of the DNV GL Energy HI. This HI consists of predefined *HI output representations* and can include a limited number of additional representations on request by the cooperating utility [11].

This causes some different *HI output representations* to exist. Therefore, we start showing a common representation for a single asset and conclude with representations for multiple assets.

For a single asset, a colour scheme based on *required action* (additional maintenance/replacement) is used. The resulting colour depends on the component remaining lifetime (RL) or time to additional maintenance. These time frames are compared to two points in time:

1. The moment after the reference period.

2. The critical time.

The reference period is a period defined by the utility, typically 10 to 15 years. The critical time is defined by the time required for regular replacements, typically 3 to 5 years.

These time frames are used to define the HI scores, as indicated in figure 2.1.4 by the colours *green*, *orange*, *red* and *purple* [11]. These index scores indicate the required action, as explained in table 2.6.

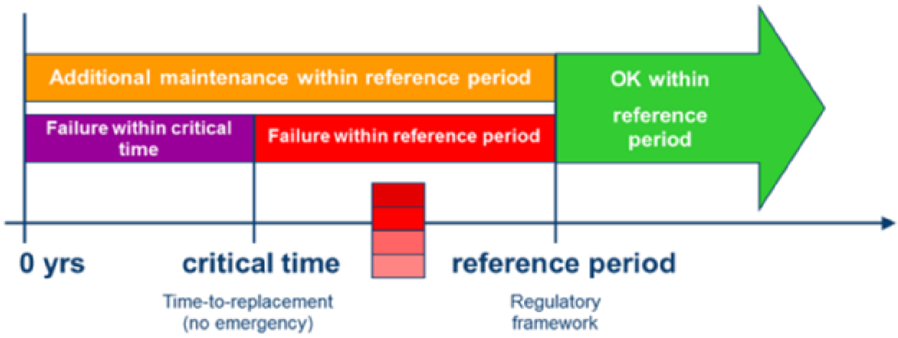


Figure 2.3: HI classification scheme [11]
This classification scheme defines the HI scores that indicate the required action. The required action is explained in table 2.6.

Table 2.6: Explanation of the colours assigned to a single asset (based on [11])
This table indicates and explains the required action for each colour of figure 2.3.

Colour	Required action	Explanation
Green	No action required (good condition)	Operation can be continued without any additional effort next to the standard scheduled maintenance within the reference period.
Orange	Additional maintenance required	Operation can be continued provided that additional maintenance or revision is carried out in addition to the standard scheduled maintenance within the reference period.
Red	Replacement within reference period	The end of life, or a significant failure probability, is expected within the reference period, additional maintenance is not sufficient to extend the life beyond the reference period.
Purple	Immediate replacement	The condition of the equipment is critical and requires immediate action.

The input parameters come with a given degree of uncertainty. When input parameters are missing, default values with higher degree of uncertainty are used. The DNV GL HI accounts for this uncertainty in the output by performing a sensitivity analysis using Monte Carlo simulations. The results from these simulations are used in defining the degree of uncertainty. This is graphically presented by the colour's intensity, as given in figure 2.4. Intensity increases with reliability of the result indicated in table 2.6 [11].

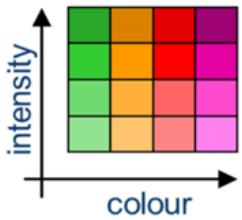


Figure 2.4: Colour codes and intensity [11]
This figure shows the four colours together with four intensities, indicating four levels of uncertainty.

A limited number of other results for multiple assets is possible on request by utilities. This includes *HI output representations* per asset type, region, voltage level or point-to-point connection. This is illustrated in figure 2.5.

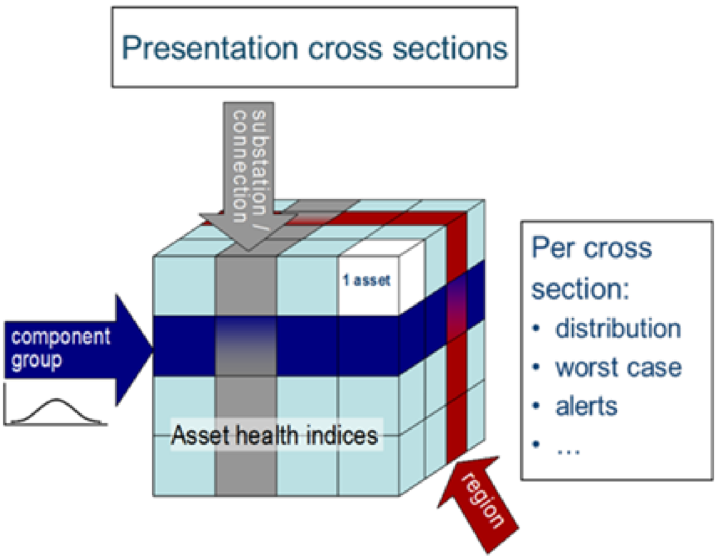


Figure 2.5: HI output representations for multiple assets based on various cross sections [11]

Figure 2.6 shows the HI scores for multiple assets based on the predicted required action (colour) and uncertainty (intensity). The highest bar in red colour and the highest intensity indicates that a large number of assets in this population require replacement within the reference period. Furthermore, the intensity of this bar indicates the high certainty of this result.

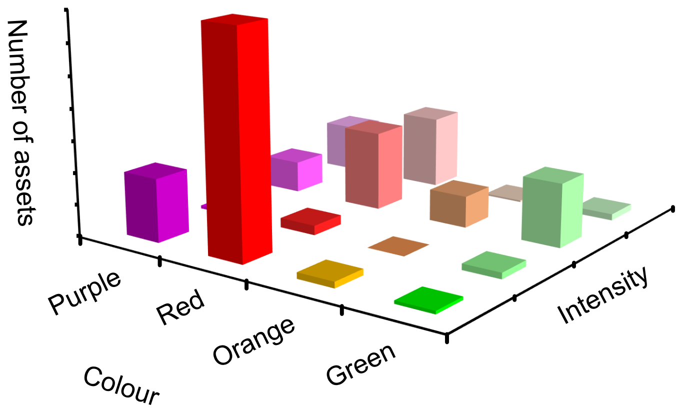


Figure 2.6: A 3D HI output representation for multiple assets including uncertainty of the results (based on [11])
The 3D chart in this figure uses predicted required action (colour) and uncertainty (intensity) to classify the assets. The highest bar indicates that a large number of assets in this population requires replacement within the reference period.

Furthermore, figure 2.7 shows a graphical representation of predicted replacement waves considering load growth, as presented in the HI of [11]. The installed number of transformers per year is indicated by the solid bars. For annual load growths of 1.5, 2 and 2.5%, a replacement wave is predicted. The solid line, indicating a predicted replacement wave for an annual load growth of 2.5%, shows a higher peak value for replacements than the dotted line, that indicates a load growth of 1.5%. Furthermore, for the solid line, the windows between the first and final replacements is smaller than for the dotted line.

Therefore, an increase in load not only causes the replacement wave to start earlier, it also decreases the time window during which the assets have to be replaced. This causes a higher peak in the replacement rate.

Characteristic 4: Assessment functions

This characteristic covers the part in literature on assessment functions which translate asset data into asset health. It starts with discussing the ideas behind the translations and concludes by indicating in which shapes or forms the assessment functions appear and indicates their internal relationships.

Three types of assessment functions are present in the DNV GL HI, as shown in figure 2.8 [11]:

- Statistical remaining life function
- Degradation remaining life function
- Condition remaining life function

The asset data on the left-hand side of figure 2.8 is translated into different remaining lifetimes (RLs) using these assessment functions. Thereafter, these RLs are weighed to find the HI score.

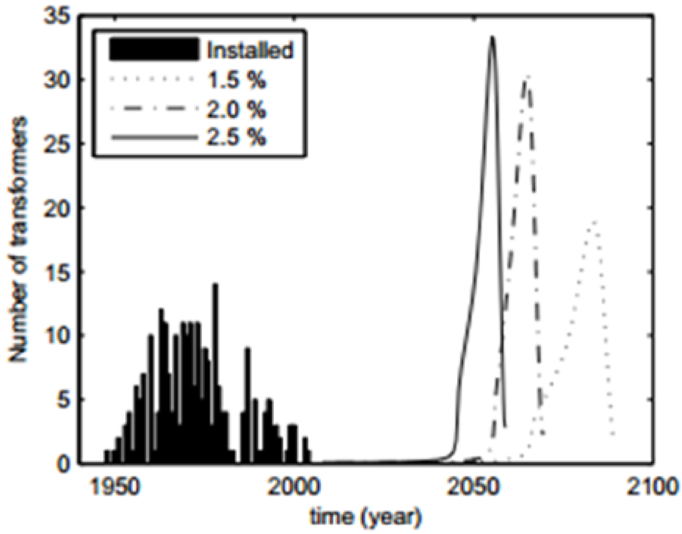


Figure 2.7: Predicted replacement waves example with annual load growths of 1.5%, 2% and 2.5% [10]
For an increase in load, the replacement wave starts earlier and shows a higher peak in the replacement rate.

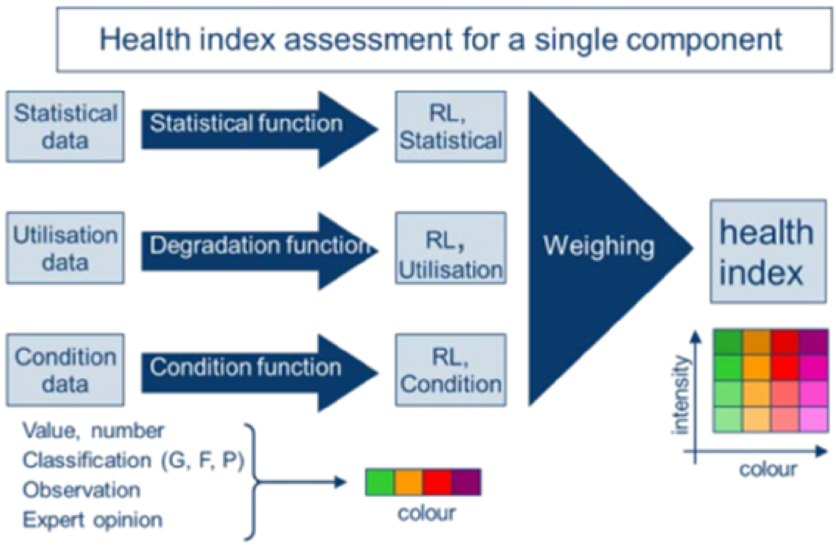


Figure 2.8: HI concept using three types of assessment functions [11]
The part on the left-hand side of this figure indicates the asset data types. Towards the right-hand side, the assessment functions, the remaining lifetimes they predict, their weighing and the HI score are shown.

The statistical function calculates a failure distribution. It arrives at this distribution by curve fitting on failure data or expert knowledge. If failure data is insufficient, FMEA analyses are used to estimate the distribution [11]. This failure distribution is combined with the present age of single component to calculate the components new distribution; this is shown in figure

2.9 and 2.10. The expected RL is calculated for each asset using [11]:

$$RL = \frac{x}{1 - F(t)} - t \tag{2.1}$$

In equation 2.1, t represents the current age of the asset. The value of x is calculated using [11]:

$$x = \sum_{n=t}^{\infty} f(n) * n \tag{2.2}$$

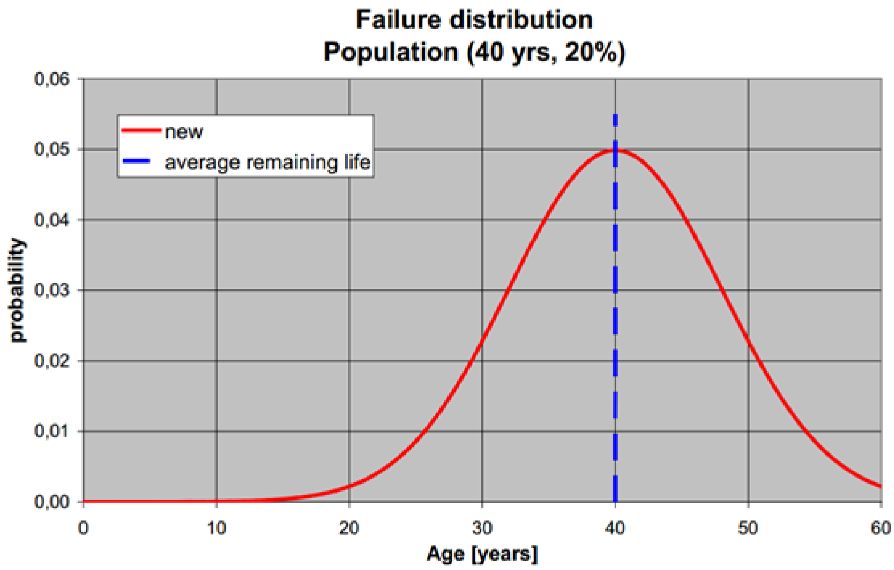


Figure 2.9: Example of a failure distribution (red) for an asset group (average lifetime: 40 years) with the average expected age of failure (the dotted blue line) [11]

Degradation remaining life functions predict the RL based on the utilization of the asset. These functions compare utilisation data to the design parameters of the asset. Examples of utilisation data mentioned in [11] are loading history for power transformers and number of operations for switchgear.

Condition remaining life functions determine the asset condition by using condition data. The asset condition can have three values: *good*, *poor* and *bad* [11]. Depending on the condition, the statistical RL is adjusted according to the scheme in figure 2.11.
In this HI, two techniques are used to manage missing data [11], deduction and statistical inference.

Construction of asset data by deduction is used when only a part of the required information is available. Construction of asset data by statistical inference is used when for a specific subset of the asset population, all required information is available. In this case, when taking a subset representative for the whole population, the missing information may be estimated by statistical inference.

Characteristic 5: HI foundation supporting methods

This characteristic covers the methods found in literature to support the HI foundation. In [18] and [68], a method is proposed for the validation of a degradation model. This model translates

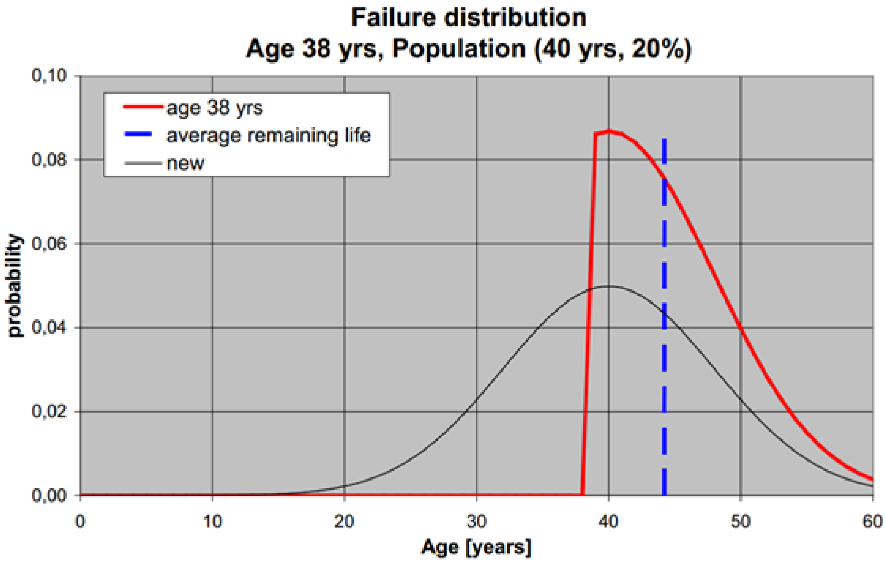


Figure 2.10: Example of an updated failure distribution (the red curve) for an asset of 38 years with the average expected age of failure (the dotted blue line) [11]
The statistical function uses the current asset age to update the failure distribution.

historical ambient temperature and transformer load into a *degree of polymerization* (DP). The DP value is a measure for the degradation of insulation paper in power transformers.

The validation is done by comparing a simulated DP values to DP values from measurements for the case of a machine transformer. Machine transformers are power or distribution transformers that are installed in a factory. In this case, it was installed in an aluminium plant. As a result, the loading pattern of this transformer was rather constant. This enabled extrapolation of the recorded pattern of two years to the entire lifetime of the transformer.

In figure 2.12, the simulated DP value is shown in time. The continuous line represents the simulated value in time with error margins indicated by dotted lines. The blue circle shows the measured value of polymerization as measured in 2007 where the blue line crossing the circle indicates the uncertainty induced by measurement error. The blue circle and its uncertainty are within the error margin of the simulation. This proves the validity of this specific degradation model for this transformer. This example shows validation of one of the degradation models present in the DNV GL HI.

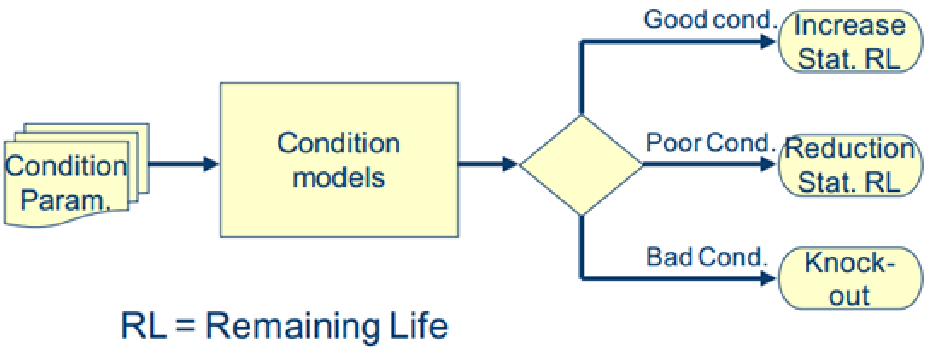


Figure 2.11: Graphical representation of the condition remaining life function [11].
Condition models adjust the statistical lifetime according to the condition.

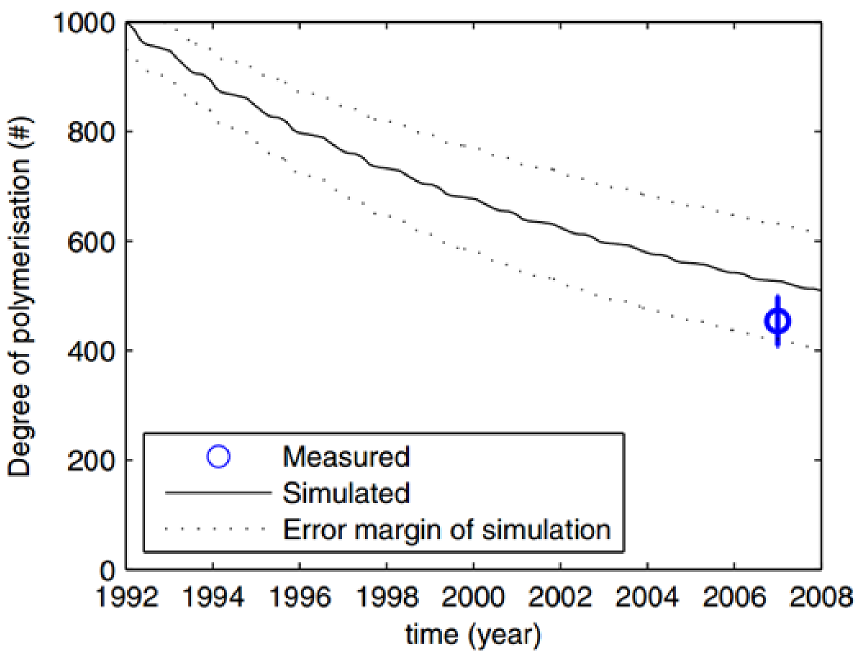


Figure 2.12: Comparison of the results from a simulated value for the degree of polymerisation to its measured value after transformer failure [18]
For this specific transformer, the analysis proves the validity of the degradation model for this case.

In [21], two models are compared for this practical case. This comparison is shown in figure 2.13, where the simulated and measured DP values are shown. The continuous lines represent the simulated value in time with error margins indicated by dotted lines. Both lines use different degradation models. The black line results from a simulation using a degradation model developed by Lundgaard. The grey line results from a simulation using a degradation model developed by Emsley. The black circle shows the measured DP value , which was measured 454 in 2007 indicated by the circle, with 10% accuracy indicated by the line crossing the circle. The horizontal dashed lines between DP values of 200 and 300 indicate the DP threshold and its uncertainty margins.

Combined with the simulated DP-curve, this threshold could be used to predict the RL of the transformer.

Looking at the uncertainty margins of the simulated DP curves and the threshold line, the predicted RL that could be derived from these lines, still has a wide uncertainty margin. For example, when looking at the case of Emsley's degradation model, this uncertainty margin translates to almost 50 years.

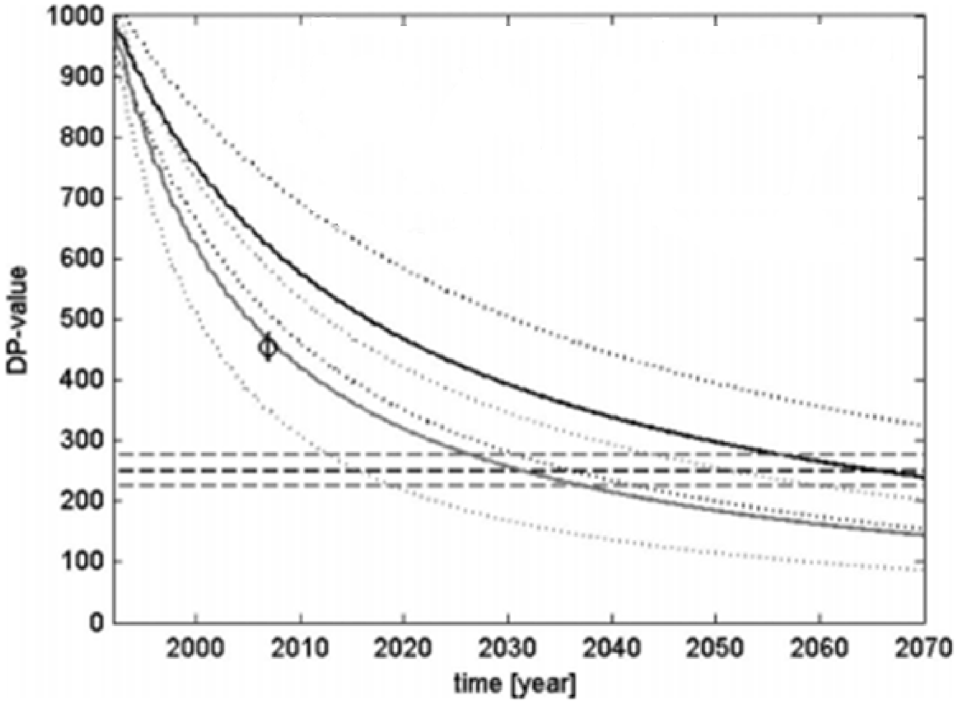


Figure 2.13: Comparison of the results from a simulated DP-value to the measured after failure (based on [21]).
The large width of the error margins induces large uncertainty in a possible predicted year of failure.

A different method is described in [16], where a cable circuit equipped with an online distributed temperature monitoring system is presented. This system is used in a cable circuit to compare actual temperature measurements to temperature values from simulations. These simulations are based on a thermal model using the loading of the cable system. It appeared that only a small number of measurements was required to validate the simulations.

2.1.2 EA Technology HI SoT

In the UK, the main driver for developing HIs was the sudden need for clear information on the present condition of assets. This sudden need was caused by two recent developments.

The first development was the upcoming replacement wave at the time, caused by the large number of equipment installed during the period 1950 - 1970.

The second development was a significant increase in economic and regulatory pressures in a short period put on electricity companies, where activities and decisions are usually undertaken with longer time frames in mind [6].

Characteristic 1: HI requirements

Asset data types

- The asset data types used to determine asset condition should be objective and verifiable. This should be achieved by the incorporation of knowledge on degradation and failure processes in assets [27].

HI output representations

- HI scores should be suitable to be quickly understood and interpreted by asset managers [27].

Besides the requirements regarding the characteristics, this HI SoT also describes more general characteristics. These characteristics are listed in section A.2.1.

Characteristic 2: Asset data types

In the EA Technology HI SoT, no specific asset data types are given. Therefore, no table to compare the asset data could be derived. However, the number of asset data types for each asset type is provided in literature. They are presented in table 2.7 as this still gives an indication of the completeness of this HI SoT regarding asset data types.

Table 2.7: Number of asset data types as found in literature for the EA technology HI SoT

Category	Number of asset data types	Reference
Transformer	24	[30]
Switchgear	169	[30]
Capacitors	9	[30]
Instrument transformer	42	[30]
Auxiliaries	127	[30]
Overhead lines	5	[27, 6]

Characteristic 3: HI output representations

This characteristic covers the output of the EA Technology HI. In general, it can be concluded that the *HI output representations* are defined together with the cooperating utility [6].

Many different *HI output representations* exist. Therefore, this characteristic starts with a common representation for just a single asset and concludes with representations for multiple assets. As described in [27], the general *HI output representation* is defined for single assets as a percentage or on a scale from 0 to 10. A *time to additional maintenance* is not defined in this HI, as this particular type of defects is expected to be detected during regular maintenance.

Furthermore, EA Technology links the HI to a probability of failure, as shown by figure 2.14. In this figure, the increase in probability of failure (PoF) becomes significant from a HI score of 7-8. This is done for the following main reasons:

- By defining this link, it becomes possible to derive and calibrate HIs [5].
- If this link is applied consistently [27]:
  - It enables a HI comparison between asset types.
  - It can provide insight in the future performance of the asset.

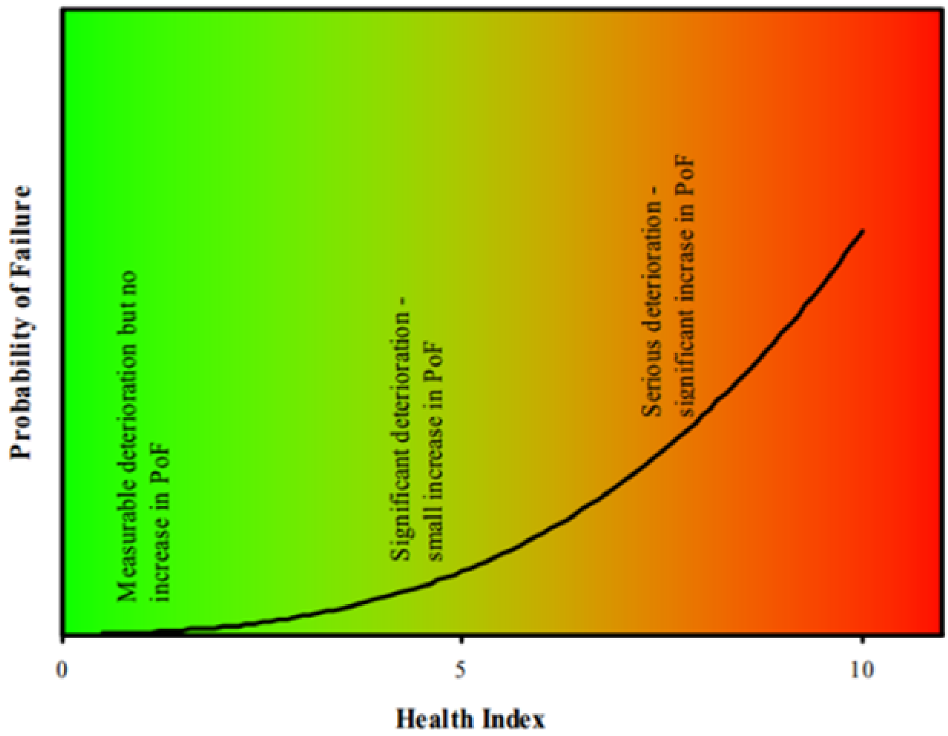


Figure 2.14: Example of the link between HI score and probability of failure (PoF) [25]
From a HI score of 7-8, the increase in probability of failure becomes significant.

According to utilities' wishes, different *HI output representations* can be derived. An example of a representation for a single asset type is given in figure 2.15. In this representation, the HI scores from *0* to *10* are translated into a score from *very poor* to *very good*. For the example assets in this figure, most assets are in a *good* to *very good* condition.

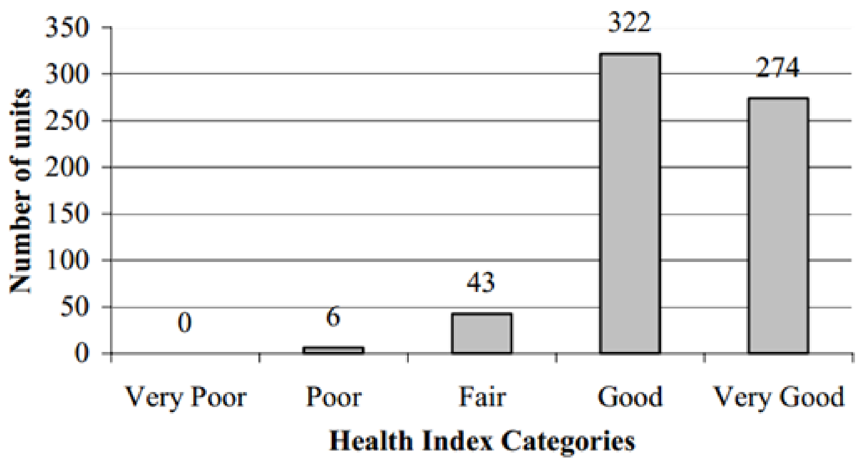


Figure 2.15: HI output representation for a single asset type [30]
The majority of the assets appear in good to very good condition.

Besides showing the HI, a *HI output representation* can also be based on risk and financial analyses. This representation can be in the form of a replacement profile (see chapter 1). An example of a replacement profile based on minimal overall costs is presented in figure 2.16. Keep in mind that due to the uncertainty in asset health prediction, the actual replacement profile can show large deviations from the presented profile.

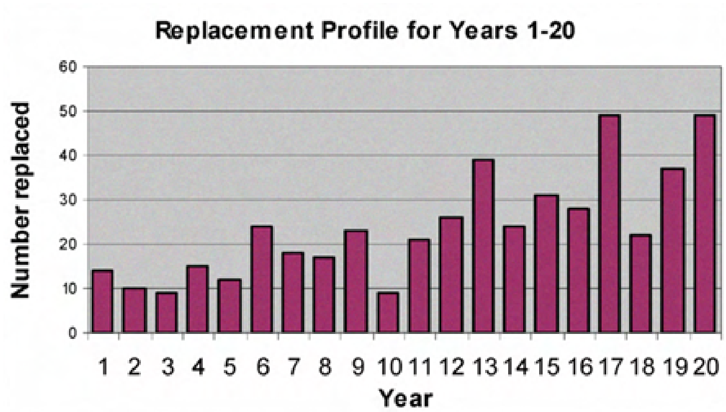


Figure 2.16: Optimized replacement programme [4]
The actual replacement profile can show large deviations from the presented replacement profile.

Characteristic 4: Assessment functions

As [6] shows with the example of dissolved gas analysis results interpretation, the focus of a HI is on providing the general overview of the asset's health. Therefore, the added value of performing separate condition assessments remains. Furthermore, EA technology incorporates loading of the assets in their risk index, as this allows for including risk calculations for the future based on predicted future load [28].

EA technology classifies the asset data types with respect to their individual impact on the

HI score into four categories. Table 2.16 explains these categories. The categories are used

Table 2.8: Four categories of asset data types [27, 30]

Relative degree of importance	Explanation
No impact	Indicator reflects defects or deterioration measures that have no impact on overall asset health.
Contributing factor	Indicator reflects defects or deterioration measures that range from low to high in importance, but typically in combination with other meaures as part of a formulation of generalized deterioration.
Combinatorial factor	Indicator reflects a measure which does not represent asset condition in isolation, but is a critical component in a complex logical and/or mathematical formulation of asset health.
Dominant factor	Indicator reflects the health of a dominant subsystem that makes up the asset, and end-of-life based on this single factor represents end-of-life for the entire asset.

to combine the subcomponent condition indicators into a single HI by means of importance weightings and formulas. The steps in this process are described by table 2.9.

After step 6, the asset overall health is converted into a discrete value ranging from "very poor" to "very good". To fine-tune the representation of the discrete values, the steps in table 2.9 may be repeated several times [27].

Characteristic 5: HI foundation supporting methods

This characteristic covers the actions EA Technology performs support the foundation of their HI. This process is embedded in the development of a HI.

In the initial stage, a HI should be developed based on pre-defined objectives and available information. Once the output representations of the HI can be used for asset management decisions, continuous improvement is achieved by introducing the feedback in the process.

The steps *Review and refine* and *Identify/Collect Additional Information* represent this feedback on the right-hand side of figure 2.17 [6].

Table 2.9: Steps performed in asset health calculation [27]

Step	Explanation
1	"Deterioration" assessments or scores are converted to health scores in a defined range from "perfect health" to "end-of-life".
2	Importance weighting is assigned to each factor in a range from "modest importance" to "very high importance".
3	General detoriation index is formulated by calculating the maximum possible score by summing the multiples of steps 1 and 2 for each factor.
4	The general deterioration index is normalized to a maximum score of 100 based on having a defined acceptable/minimum number of condition criteria.
5	Normalize the dominant factor to a maximum score of 100.
6	Calculation of the overall HI as the lesser of step 4 or 5, where 100% is excellent health and 0% is "poor" health.

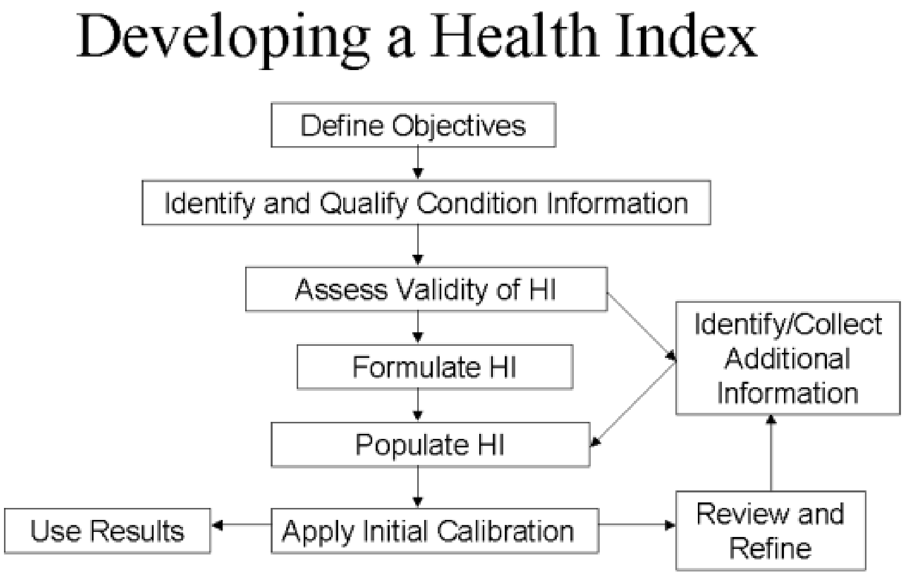


Figure 2.17: Steps in the development of a HI [6]
The feedback introduced by the two steps on the right-hand side of this figure ensure continuous improvement.

2.1.3 Kinectrics HI SoT

Utilities worldwide need to meet higher demands than ever regarding financial and technical performance [36, 39, 38]. To meet these needs, a balance between maintenance and associated risk needs to be found [36, 39, 38]. Furthermore, decisions regarding risk are mostly assigned to line management. To support this decision maker, analytical tools are required [33].

Kinectrics offers three of these tools, Asset Management Planner (AMP) [32], Risk-Based Asset Management (RiBAM) [33] and a HI [36, 39, 38].

Characteristic 1: HI requirements

Asset data types

- The asset data used by the HI should be relevant, up-to-date and reliable [35].
- The HI should use objective and verifiable observations of asset condition [36, 38].

HI output representations

- The HI score should be indicative of the suitability of the asset for continued service and representative of the overall asset health [36, 38].
- The HI score should be understandable and readily interpreted [36, 38].

Assessment functions

- As in most cases proper asset failure data is absent, expert knowledge needs to be included in the assessment functions for developing asset life expectancy and failure probability curves [35].

Besides the requirements regarding the characteristics, this HI SoT also describes more general characteristics, These characteristics are listed in section A.3.1.

Characteristic 2: Asset data types

This characteristic covers the number of asset data types found in literature by Kinectrics. They are presented in table 2.10.

Table 2.10: Number of asset data types as found in literature for the Kinectrics HI SoT

Category	Number of asset data types	Reference
Transformer	45	[36, 39, 35, 38]

Table 2.11 was constructed using the assumptions as explained in characteristic 2 of section 2.1.1. From this table, it appears that, for a large part similar as the other HI SoTs:

- Most asset data mentioned in the Kinectrics HI SoT is expensive to acquire for utilities for all assets if this data is not yet available for any of the assets.
- For acquiring most asset data, it is not required to put the equipment out of service.

Table 2.11: Classification of the detection methods of the Kinectrics HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table summarizes the table in section A.3.3. This table contains a classification for the individual detection methods.

Subject	Subject classification	Number of detection methods for subject classification
Costs	Low (often already available)	3
	High	42
Out of service required	No	36
	Yes	9
Maturity	Mature	43
	Development phase	2
Ageing detection	Long term ageing	21
	Short term ageing	24
Accuracy	Good	44
	Bad	1

- Most asset data types originate from mature (detection) methods.
- Most asset data covers both short and long term ageing.
- The majority of the (detection) methods responsible for the asset data achieve a good accuracy.

Characteristic 3: HI output representations

The general output of the Kinectrics HI is based on an asset health score ranging from *0* to *100*. The Kinectrics HI links this score to a condition, expected lifetime and associated requirements, as presented in table 2.12. From this table, the required action shifts according to the probability of failure from preventive to corrective maintenance.

Figure 2.18 shows the example of a *HI output representation* according to the rules presented in table 2.12. In this figure, a large number of assets is classified as *fair*. This might indicate a large number of medium-term replacements including the need for the asset manager to increase diagnostic testing.

This HI does not define a time to additional maintenance. Kinectrics considers types of defects associated with additional maintenance to be detected and handled within regular maintenance [39].

Figure 2.19 shows a *HI output representation* of risk for each of the asset locations. This representation gives asset managers a quick overview of critical locations. For this example, an

Table 2.12: HI scoring for Kinectrics (based on [36, 39, 38]) *This table provides a rough indication of the link between values used in various output representations.*

HI score	Condition	Probability of failure (PoF)	Equivalent status on life curve	Explanation
85-100	Very good	Low	First half of mean life expectancy	Normal maintenance
70-85	Good	Low but slightly increasing	Second one-third of mean life expectancy	Normal maintenance
50-70	Fair	Rapidly increasing but lower than PoF at mean age	Final one-third of mean life expectancy	Increase diagnostic testing, possible remedial work or replacement depending on criticality
30-50	Poor	Higher than PoF at mean age and increasing	First one-third after the mean life expectancy	Start planning process to replace or rebuild considering risk and consequences of failure
0-30	Very poor	Very high, more than double the PoF at mean age	Second one-third after the mean life expectancy	Immediately assess risk, replace or rebuild based on assessment

asset manager could give the assets indicated in yellow in the north-west the most attention. Other representations include top worst asset types or stations by actual dollars spent and total asset cost by age [37].

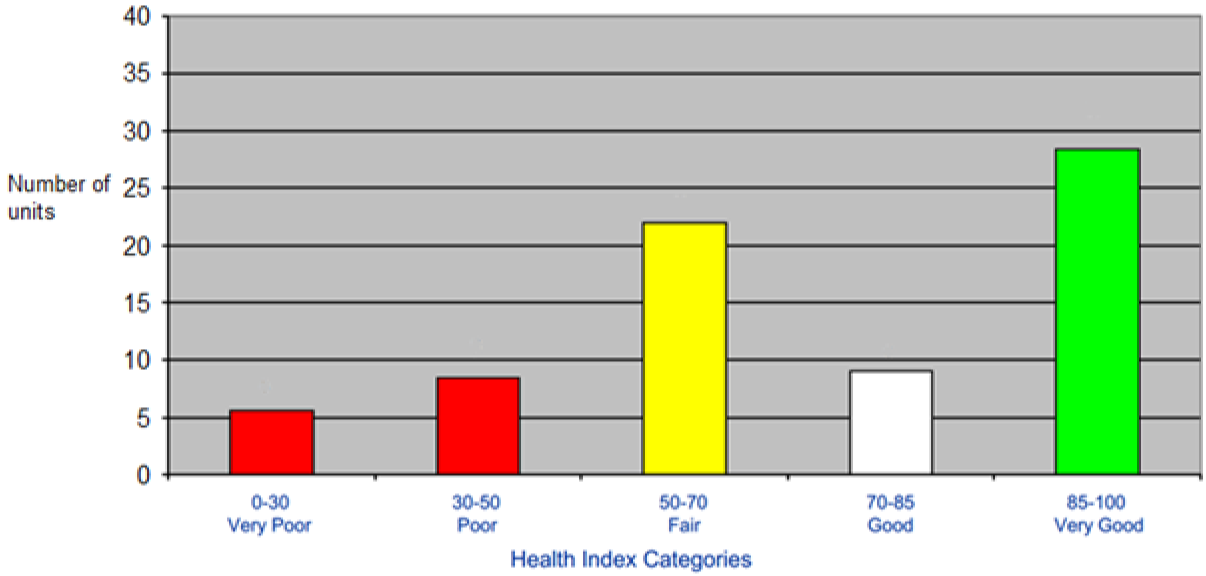


Figure 2.18: Kinectrics HI example output representation (based on [41])

A large number of assets is classified as fair, which can be a sign that a large number of medium-term replacements is expected together with a need to increase diagnostic testing.

Figure 2.19: Kinectrics HI output representation based on location [37]

This representation gives asset managers a quick overview of critical locations.

Characteristic 4: Assessment functions

The HI operated by Kinectrics uses HI sub-scores for each condition criterion. Two types of translations to these sub-scores are present; comparison to a fixed limit based on one asset data type and comparison to a limit based on multiple asset types. The resulting sub-scores can range from *A* to *E* and are converted into a factor between *4* and *0*. Furthermore, an asset data type specific factor is assigned as weighting factor for each HI sub-score. Kinectrics calculates the asset's HI score by taking the sum of the product of all HI sub-scores and weighting factors [36].

In [40], a sensitivity analysis is performed to find the robustness of HI scores when varying asset data. Despite this analysis, it is unclear whether a sensitivity analysis is integrated in the Kinectrics HI.

In [35], the method used to calculate the overall probability of failure is described. To start with, two probabilities of failure are calculated; a probability of failure *P1* based on the actual age and life curve and a probability of failure *P2* based on the condition and HI. *P2* is calculated linking the HI score to the probability of failure.

The graph in figure 2.20 shows a simplified version of this link.

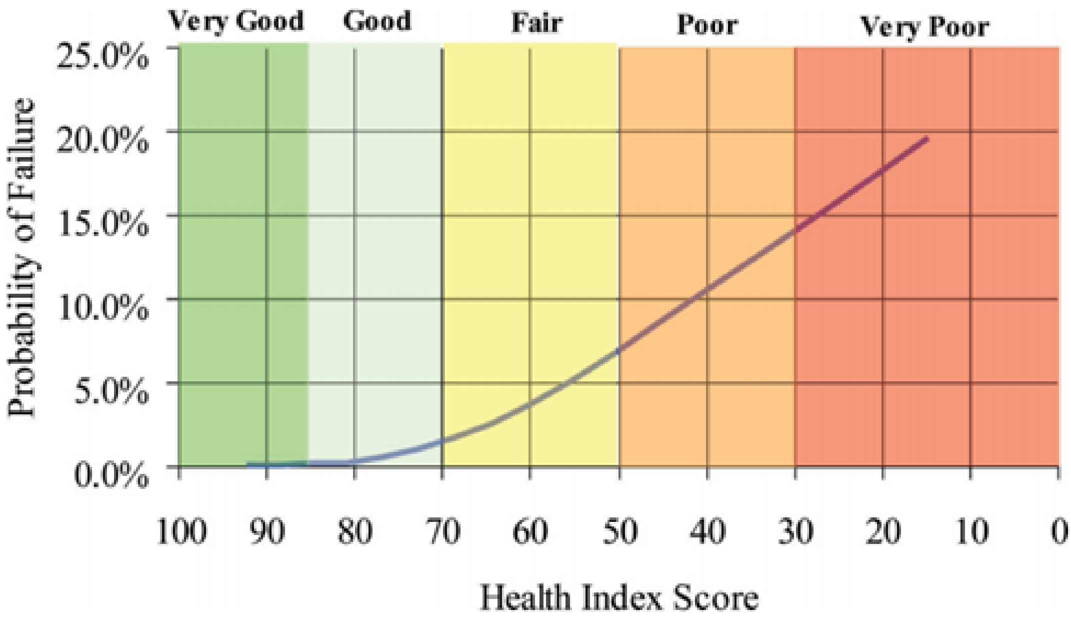


Figure 2.20: Simplified version of the link between probability of failure and HI score [39]

P1 and *P2* are combined to calculate the overall probability of failure. In this calculation, the importance of the HI is given a weight, as indicated by k in [35]:

$$P = P_1 + k(P_2 - P_1);\ 0 \leq k \leq 1 \tag{2.3}$$

Knowing the probability of failure for the asset, the effective asset age can be found using the asset life curve.

The example of figure 2.21 shows the steps involved in calculating the effective age from the HI. In this figure, both vertical axes indicate the probability of failure. In the left-hand graph, the horizontal axis shows decreasing asset health towards the right.

The right-hand graph shows the asset life curve. The horizontal axis in this graph indicates increasing effective age towards the right. Firstly, the HI score is translated into a probability of failure using the left curve. Subsequently, the effective age is determined using the curve on the right-hand side.

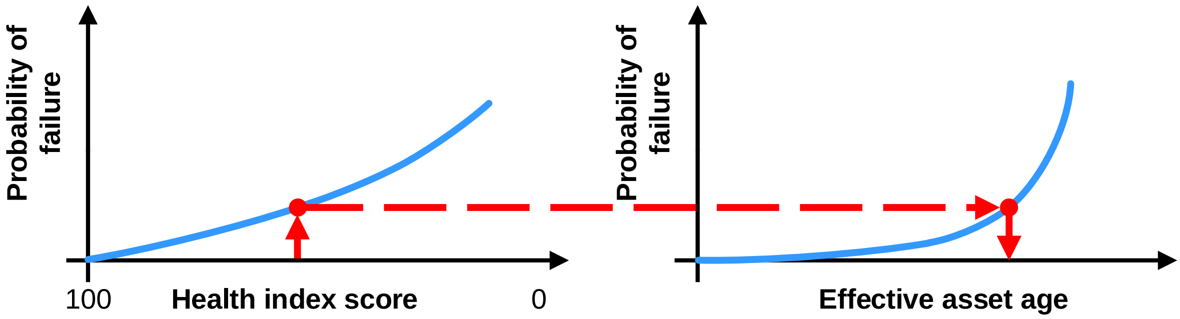


Figure 2.21: Example of extracting effective age from the probability of failure (based on [39])
The vertical axes in both graphs indicate probability of failure. The horizontal axis in the left graph shows decreasing asset health towards the right. The horizontal axis in the right graph shows increasing effective age towards the right.

Characteristic 5: HI foundation supporting methods

This characteristic covers the methods found in literature to support the HI foundation.

Firstly, a comparison between alarm levels was made during the formulation of the power transformer HI. Four recommendations taken from literature were considered for the choice of alarm levels for dissolved gases. For each gas concentration, the best health prediction was in most cases coupled to the worst case gas concentrations found in these recommendations [36]. In [39], a similar comparison is shown for the gases present in (on-line) tap changers.

Furthermore, [38] describes the development of the life expectancy curves. In this process, two phases can be distinguished. The initial curves are based on standards for the asset, accelerated ageing tests if available and industry's collective experience. Subsequently, the condition data was applied to a significant population and failure rates were determined over a period of time to fine-tune the curves if necessary.

2.1.4 TU Delft HI SoT

Besides companies, also universities are involved in research into (parts of) HIs.

Literature on HI SoTs of universities shows slight differences when compared to literature of HI SoTs of companies. For example, universities tend to describe findings on *parts of HIs*, while companies tend to describe *complete HIs*.

The importance of power transformers is emphasized by their value, repair times/efforts and outage costs, as stated in the work of Chmura [69]. The liberalization of electricity markets is mentioned to be the cause of several developments:

- A focus of utilities towards optimization of maintenance and investment costs (Gulski, Jongen, [60, 70, 59, 48, 64, 62, 55]).
- Requested higher reliability of the utilities by increasing cost of non-availability (Gulski, [60, 70]).
- In combination with increased power demand, higher flexibility and in addition to emergency repair, power cable failures may nowadays lead to loss of income or claims (Gulski, [57]).

Characteristic 1: HI requirements

Asset data types

- In addition to available failure data, condition data should be used to perform condition assessment (Smit, [50]).

HI output representations

- The first step in the development of a HI is to find what type of advice is needed by the asset owner (Gulski, [71]).

Assessment functions

- From the work of Gulski [50], the following basic steps in the condition indexing process are:
  - Gain general knowledge about the insulation defect.
  - Gain knowledge about ageing effects and the influence on diagnostic results.
  - Gather data with the aid of on-site diagnostic tools.
  - Perform statistical analysis on the collected data.
  - Generate norm values from the collected data and the statistical analysis.
  - Combine important diagnostic parameters to determine the condition index.
- End of life estimation can be performed by statistical analysis of life time data (Jongen, [58, 49, 62, 52, 72, 62]).
- When performing an end of life estimation, confidence bounds should be included [58].
- From the work of Smit [50], the following approach can be used to generate norm values:
  - Collect data for a particular high voltage component.

  - Use a distribution fitting tool to estimate the best fitting distribution for this data.
  - Based on the best fitting distribution calculate the norm values.

- For condition assessment, knowledge rules should be established (Smit, [50]).
- A HI should contain a reference for interpretation of the result from each diagnostic (Quak, [56]).
- For the assessment of the technical performance, the following key factors should be taken into account within a HI (Quak, [56]):
  - condition
  - load profile
  - asset structure
  - redundancy
  - environmental factors
- In supporting the decision between repair and replacement, knowledge from past experiences should be included in a HI (Jongen, [61, 66]).
- A HI should (support to) provide stakeholders with a clear overview of the effects by asset management decisions (Quak, [65]).
- For a predictive health model, the load type and allowed limits should be based on the preference of the utility (Bajracharya, [73, 74]).

Besides the requirements regarding the characteristics, this HI SoT also describes more general characteristics. The requirements regarding these characteristics are listed in section A.4.1.

Characteristic 2: Asset data types

Most descriptions found from the TU Delft make use of statistical analysis to predict future failure behaviour of single assets or asset populations. Complexity of assets is mentioned as the cause for the difficulty of interpreting the measurement results by physics. This complexity arises from the enormous variation in used materials, component ages, applied stresses and maintenance (Quak, [56]).

This characteristic covers the main asset categories and number of asset data types found in literature by the TU Delft. They are presented in table 2.13.

Table 2.13: Number of asset data types as found in literature for the TU Delft HI SoT

Category	Number of asset data types	Reference
Transformer	21	(Bajracharya, Chmura, Gulski, Jongen, [69, 63, 60, 45, 72, 15])
Gas insulated switchgear	4	(Chmura, [63])
Cable	42	(Gulski, [60, 70, 51])

Table 2.14: Classification of the detection methods of the TU Delft HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table summarizes the table in section A.4.3. This table contains a classification for the individual detection methods.

Subject	Subject classification	Number of detection methods for subject classification
Costs	Low (often already available)	6
	High	61
Out of service required	No	29
	Yes	28
Maturity	Mature	41
	Development phase	26
Ageing detection	Long term ageing	22
	Short term ageing	47
Accuracy	Good	41
	Bad	26

Table 2.14 was constructed using the assumptions as explained in characteristic 2 of section 2.1.1. From this table, it appears that, for a large part similar as the other HI SoTs:

- Most asset data mentioned in the TU Delft HI SoT is expensive to acquire for utilities for all assets if this data is not yet available for any of the assets.
- For acquiring asset data, it is in half of the (detection) methods required to take the equipment out of service.
- The majority of the asset data types originates from mature (detection) methods.
- The majority of the asset data covers short term ageing.
- The majority of the (detection) methods achieves a good accuracy for almost all asset data.

Characteristic 3: HI output representations

The HI description performed by the TU Delft in cooperation with network operators and a cable manufacturer results in a HI score. This HI score is determined from a category and reliability status of *1*, *6* or *9* [60].

This HI score is presented in table 2.15. This table provides a rough indication of the link between values used in various output representations. This table shows no clear distinction between time to additional maintenance and replacement (Smit, [50]).

Table 2.15: HI score for power cables (based on the work of Gulski [60])
This table provides a rough indication of the link between values used in various output representations.

Category	Reliability status	Condition status	Service life status	Condition index	Required action	Recommendation on maintenance
Normal	No problems	No defects or ageing symptoms observed	New or aged	9	No extra attention required, e.g. next inspection in 5 .. 10 years	No maintenance necessary
Defect initiation	No impact on reliability	Certain degree of insulation degradation observed; no harmful defects present	Strongly aged	6	Extra attention is needed, e.g. inspection within 1 year	Without any maintenance possible lifetime reduction
Defect	Asset can still be operated but the reliability is decreased					Maintenance is necessary
Failure	Asset can not be operated	Significant insulation degradation observed and serious defects are present	Nearby end of lifetime	1	Maintenance is necessary, e.g. repair or replacement	Based on economics, repair or replacement maintenance is necessary

Furthermore, statistical analysis can be utilized to find representations for power transformers and power cables (Chmura, Jongen, Smit, [69, 58, 50, 54, 66, 45, 49, 62, 52]). An example of such representation for mass insulated cables is given in figure 2.22.

This figure shows the expected failures on a yearly basis including the 90% confidence bounds of this value. These confidence bounds show that the number of yearly failures is expected to stay within the interval of the two bounds with a confidence of 90%.

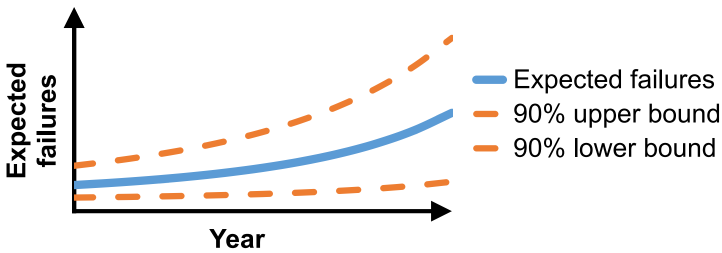


Figure 2.22: Example failure prediction for a population of mass insulated cables (based on the work of Jongen [52]).

The predictive health method as described in the work of Bajracharya [15] combines asset data together with ambient temperature to predict the maximum loading of power transformers.

Characteristic 4: Assessment functions

The study of Chmura [63] mentions a method to calculate cable degradation based on loading. This method is based on the quadratic relation between current and conductor temperature.

Subsequently, the conductor temperature can be translated into degradation using a relation derived by experiments.

Next, the study described by Chmura [63] reports a method for estimating insulation life consumption (ILC) based two asset data types on testing voltages and relative $tan\,\delta$. This relation was also derived by experiments.

Subsequently, the study described in the work of Smit [50] proposes a method to rank insulation condition of power cables based on partial discharge measurement results. This method uses statistical methods to describe an assessment function based on several partial discharge ratings.

Finally, the work of Jongen [66] shows the use of statistics to describe future failure behaviour of assets. The process of describing future behaviour using statistical analysis consists of multiple steps. The main steps in the process of describing present failure behaviour include distribution fitting to the data, goodness-of-fit tests and finding confidence bounds. Subsequently, an additional analysis describes the future failure behaviour making use of the present failure behaviour.

To conclude, the studies of Bajracharya [15, 74] describe a predictive health method. This method is different from the aforementioned HIs regarding its goals and working principle. However, as this method could play an important role in the future of health indexing, it is included in this text. This method is part of a framework that prevents (excessive) asset ageing. To illustrate its working principle, this framework has been applied to the example of power transformers. Each transformer in a network predicts the values of its maximum load based on the predicted transformer oil temperature. The transformers interact and take actions according to these values. For this example, re-routing power is a possible action in the case of predicted overloading. These actions lead to an overall decrease in ageing.

This method uses thermal models in finding the instantaneous maximum load. These models are comparable to assessment functions, the difference is that they translate asset data into instantaneous asset operating limits instead of asset health. The maximum load for the example of a power transformer is determined by the maximum winding temperature at a certain location in the transformer, the so-called *hotspot* temperature. In the studies of Bajracharya [15, 74], a discrete-time state-space model is used to simulate the dynamics of this and related relevant temperatures.

Characteristic 5: HI foundation supporting methods

Most TU Delft literature focuses on using statistics to describe failure rates using failure data. However, one reference showed a method to minimize the difference between predicted and observed asset health. This method is explained in this characteristic.

The study of Quak [46] describes a condition index selection process for cables. The main steps are to collect partial discharge measurements, use knowledge rules to set boundary values and based on these boundaries assign a condition index. The proposed future actions are stored in the same database as the discharge measurements. This feedback enables a continuous improvement of the condition index selection process by comparing proposed actions to measured condition over time.

2.2 Strengths/weaknesses analysis

In this chapter, the described HI SoTs were analysed with respect to their strengths/weaknesses. This has provided insight in how they compare to each other regarding the relevant aspects for the remainder of this thesis. Table 2.16 presents this analysis.

From this table, the following conclusions on strengths/weaknesses can be drawn:

HI requirements

- The DNV GL Energy and TU Delft HI SoTs describe the most requirements regarding HIs. Furthermore, only these HI SoTs make statements regarding the process of developing a new HI.

Asset data types

- Among all HI SoTs, only the Kinectrics HI SoT describes the HI for one asset type. This suggests that the scope of the Kinectrics HI SoT is limited.

HI output representations

- Only the DNV GL Energy HI SoT includes both time to additional maintenance and remaining lifetime in its HI output representations.
- Many types of health representations exist among the discussed HI SoTs, only the representation of the TU Delft lacks the option of health estimation for a single asset.

Assessment functions

- The method in the assessment functions of the TU Delft HI SoT uses the lowest number of asset data types.
- Only the assessment functions of the DNV GL Energy and TU Delft HI SoT include uncertainty.

HI foundation supporting methods

- All HI SoTs describe methods to incorporate continuous improvements based on feedback in the development of HIs or the fine-tuning of the assessment, except for the DNV GL HI. This HI describes methods support its foundation by practical cases.

Table 2.16: Strengths [+] / weaknesses [-] analysis of the discussed HI SoTs

Characteristic	Subject	DNV GL Energy	EA Technology	Kinectrics	TU Delft
HI requirements	General	[+] Many requirements described: • Asset data types • HI output representations • Assessment functions • Compatibility • Implementation • Content /information handling • Capabilities	[-] Few requirements described: • Asset data types • HI output representations • Capabilities	[-] Few requirements described: • Asset data types • HI output representations • Assessment functions • Capabilities	[+] Many requirements described: • Asset data types • HI output representations • Assessment functions • Content /information handling • Capabilities
Asset data types	Asset types	[+] Many asset types and very general: • Power transformers • Switchgear • Power cables • Generators	[+] Many asset types and very specific: • Multiple types of transformer • Multiple types of switchgear • Capacitors • Multiple types of instrument transformer • Multiple auxiliaries • Overhead lines	[-] One asset type and very specific: • Power transformers • Distribution transformers	[+] Many asset types and very specific: • Multiple transformers • General GIS technologies • Multiple types of cables • Cable accessories
HI output representations	General	[+] Single/multiple assets	[+] Single/multiple assets	[+] Single/multiple assets	[-] Only multiple assets
Assessment functions	Method	[+] Lowest remaining lifetime by using most asset data types	[+] Sumproduct by using most asset data types	[+] Sumproduct by using most asset data types	[-] Statistical analysis using small number of asset data types
	Uncertainty	[+] Included in Monte Carlo simulation	[-] Not specified	[-] Unclear whether included in HI	[+] Included in statistical analysis
HI foundation supporting methods	General	[+] Methods based on practical cases	[+] Methods for continuous improvements	[+] Methods for continuous improvements	[+] Methods for continuous improvements

2.3 Conclusions

The main focus of this chapter has been on providing an overview of the four main HI SoTs and to find present methods in *HI prediction quality quantification* and *validation*. As mentioned in section 2.1, this literature research does not cover all existing HI SoTs. In addition to this, not all HI SoTs are described (in detail) in published scientific literature and publicly available documentation of their institutions.

Regarding this overview, the following conclusions are drawn:

HI requirements

- The DNV GL Energy and TU Delft HI SoTs describe the most requirements regarding HIs.

Asset data types

- Among all HI SoTs, only the Kinectrics HI SoT describes the HI for one asset type. This suggests that the scope of the Kinectrics HI SoT is limited.

HI output representations

- Only the DNV GL Energy HI SoT includes both time to additional maintenance and remaining lifetime in its HI output representations.
- Many types of health representations exist among the discussed HI SoTs, only the representation of the TU Delft lacks the option of health estimation for a single asset.

Assessment functions

- The method used to translate data into asset health differs among the HI SoTs.
- Only the assessment functions of the DNV GL Energy and TU Delft HI SoT include uncertainty.

HI foundation supporting methods

- All HI SoTs describe methods to incorporate continuous improvements based on feedback in the development of HIs or the fine-tuning of the assessment, except for the DNV GL HI. This HI describes methods support its foundation by practical cases.
- In general, the HI SoT in scope presents only limited literature on *HI prediction quality quantification* and *validation*. Furthermore, no practical case studies were found used for *HI prediction quality quantification* and *validation* of a complete HI. This suggests that research using practical case studies for *HI prediction quality quantification* and *validation* of a HI are necessary.

Besides the conclusions on the HI characteristics, this literature research has also revealed conclusions regarding the methods to be used for quantifying HI prediction quality:

- The HI foundation supporting methods, as described by the DNV GL HI SoT, lead to the insight that for quantifying HI prediction quality and, as the next step, validation, a *comparison between two datasets* is necessary. Chapter 3 explains this in more detail.

- *Life data analysis* (LDA) using distribution fitting, as described by the TU Delft HI SoT, can be applied in order to find future failure behaviour of assets using the development of present failure behaviour. LDA is possible on both of the two aforementioned datasets and is explained in further detail in chapters 3 and 4.

Chapter 3

Research approach

This chapter presents the approach used in this thesis to extend the scientific foundation of the DNV GL Health Index (HI).

The methods as described in this approach aim for *HI prediction quality quantification* and *validation* in order to find future improvements of the DNV GL HI. This approach considers five methods for *HI prediction quality quantification* and *validation*. The methods described in this chapter are the result of a combination of methods described in literature and discussions at the Delft University of Technology and at DNV GL in Arnhem.

Section 3.1 provides an overview of the methods. Section 3.2 provides an overview of the datasets distinguished in the methods of this chapter. Section 3.3 explains the definition of predicted and observed lifetime. Section 3.4 explains how the predicted and observed lifetimes can be compared. Section 3.5 discusses the five methods in further detail. Section 3.6 discusses the methods by strengths/weaknesses and section 3.7 shows the methods adopted during the research.

3.1 HI prediction quality quantification and validation: methods overview

The predictions of HIs consist of the statements HIs make on asset health (see section 2.1). In order to quantify the quality of these statements, this section will use the statements on (asset) remaining lifetime (RL). This RL is the period the asset is expected to survive, starting from the moment of the HI prediction. The methods described in this section aim to quantify the quality of HI predictions. This so-called *HI prediction quality* is given by the difference between predicted and observed asset health.

Figure 3.1 provides an overview of the five *HI prediction quality quantification* methods described in this section. Table 3.1 explains these methods in more detail. The methods presented in figure 3.1 are presented in groups by the dataset types they use. Each method can use multiple datasets. This chapter makes a distinction between three dataset types: *predictions*, *observations* and *dummy data*. These three dataset types are explained in more detail in section 3.2.

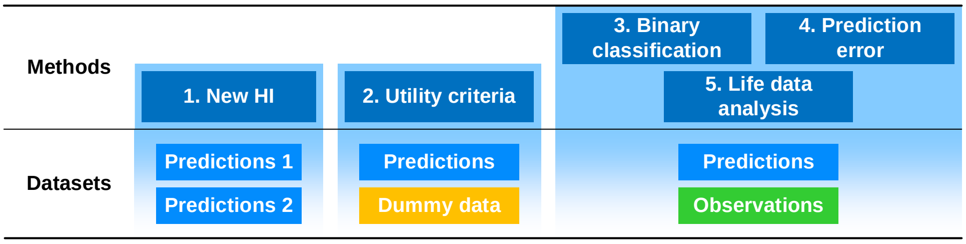


Figure 3.1: Overview of the described HI prediction quantification methods
Each method uses at least one dataset of the type predictions.

Table 3.1: Explanation of the methods for *HI prediction quality quantification* and *validation*

Section	Method	Datasets	Explanation
3.5.1	1. New HI	Predictions 1 Predictions 2	Comparison between a *new HI* and the DNV GL HI.
3.5.2	2. Utility criteria	Predictions Dummy data	Comparison of the assessment functions to *utility criteria*, which results in differences and similarities between the assessment functions and *utility criteria*.
3.5.3	3. Binary classification	Predictions Observations	Comparison of predictions to observed health by using *binary classification* based on the predicted health and observed health which results in a percentage of misqualification.
3.5.4	4. Prediction error	Predictions Observations	Comparison of predictions to observed health by counting the number of occurrences of the difference between the predicted lifetime and observed lifetime per asset which results in an indication for the *prediction error*.
3.5.5	5. Life data analysis	Predictions Observations	Comparison of predicted lifetimes to observed lifetimes using *life data analysis*.

3.2 HI prediction quality quantification and validation: datasets

3.2.1 Predictions and observations datasets

The *predictions dataset* contains for each asset a *predicted lifetime*. The *observations dataset* contains for each asset a*n observed lifetime* or an *observed operating time*.

The HI predictions are based on a large number of asset data types. Some events cannot be predicted, even though the available asset data originates from state-of-the-art detection techniques that include all reasonably known failure modes. An example of such an event is a failure due to excavation work. The assets that fail due to these events are therefore excluded from both datasets.

The HI translates the asset data into RLs. Combined with the commissioning year and the year of the HI predictions, these RLs are subsequently translated into predicted lifetimes. Depending on the method, these lifetimes are translated into the predictions dataset.

When the method uses only the exact moment of failure, the predictions dataset originates directly from the predicted lifetimes. When the method includes the moment of predicted failures relative to a reference moment, this information is included by translation of predicted lifetimes to so-called *suspensions*, which are explained in the end of this section.

The blue boxes in figure 3.2 show the steps from the asset data to the prediction dataset. The dataset of observations originates from a *failure database* combined with a database that contains the *entire asset population*. A *failure database* consists of a chronological list of failures and associated information. The *entire asset population* includes all assets in service at the point in time the HI predicted the RLs. The green boxes in figure 3.2 show this combination.

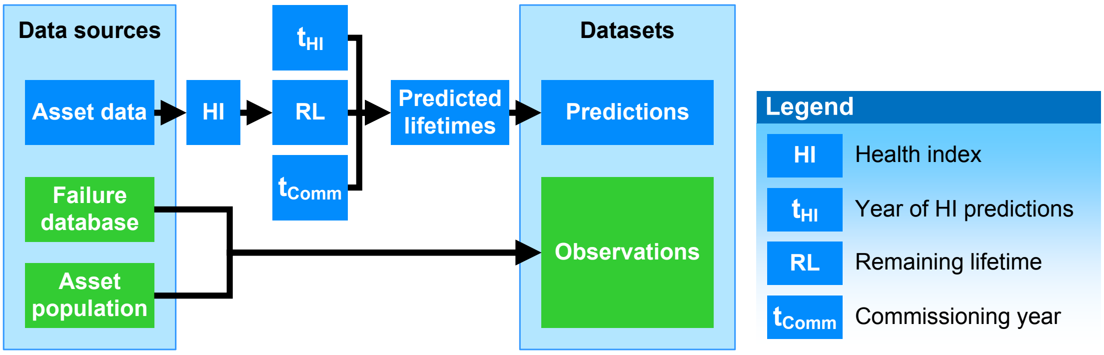


Figure 3.2: Origin of the predictions and observations datasets

As indicated in figure 3.2, each asset is assigned a prediction of the HI in terms of RL, starting from the year of the HI prediction (t_{HI}). The definition of RL is given in section 3.3. Figure 3.3a indicates the latter as the prediction moment, indicated in red colour. Furthermore, the commissioning year (t_{Comm}) is extracted from the asset data. This information enables the finding of the predicted lifetime according to:

$$Predicted\ lifetime = (t_{HI} - t_{Comm}) + RL \tag{3.1}$$

The moment of the most recent failure record is called the reference time (t_{Ref}). This thesis indicates the period between t_{HI} to t_{Ref} by the *time window*. In the majority of the cases, the asset is predicted to fail after t_{Ref}. These cases are called *predicted suspensions*. Figure 3.3b shows an example of this case.

Besides the predictions dataset, the dataset of observations originates from a failure database combined with a database of the entire asset population. This dataset is based on the observation moment at t_{Ref}, indicated in figure 3.4a and 3.4b.

The fact that assets are present in the failure database is shown in figure 3.4a by a failure within the time window. The assets that are missing in the failure database but are present in the database of the entire asset population are assumed to be still operational at t_{Ref}, as shown by the *observed suspension* in figure 3.4b.

In particular cases, assets are taken out of service during the time window due to a high failure expectancy. These replacements in these cases are called *preventive replacements*. For the assets in these cases, no conclusion can be drawn from the difference between predicted and observed lifetimes. Therefore, all assets of this group are excluded from both datasets.

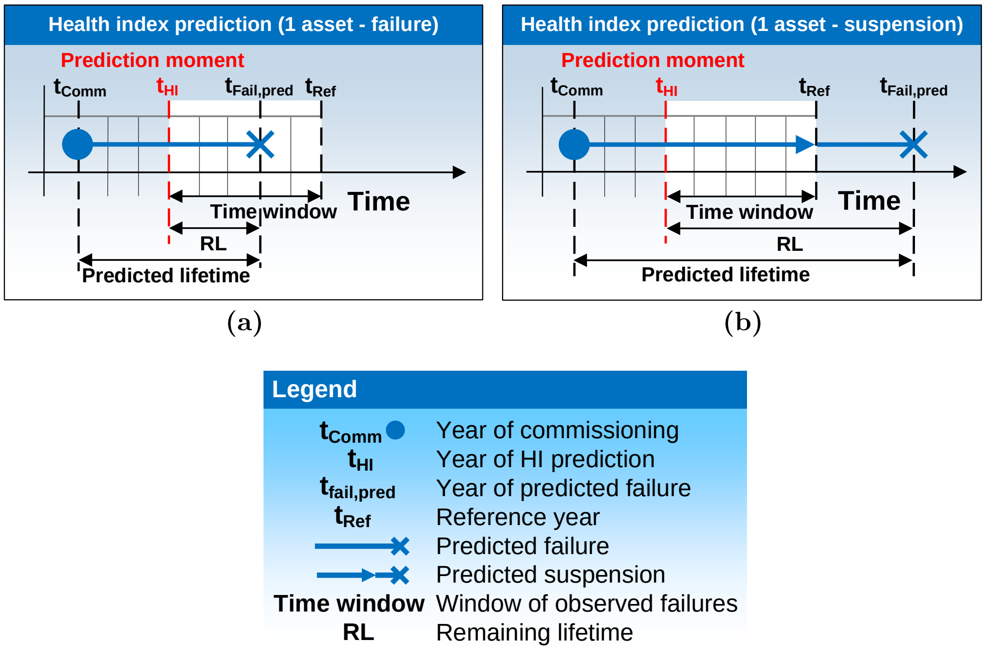


Figure 3.3: Explanation of: (a) a predicted failure; (b) a predicted suspension
t_{HI} is the prediction moment.

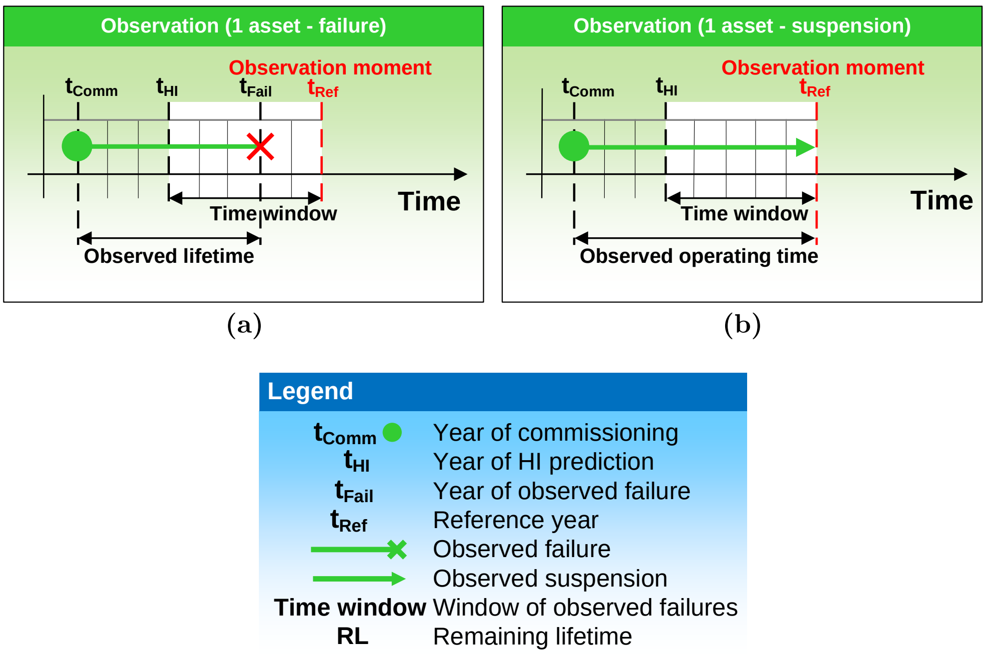


Figure 3.4: Explanation of: (a) an observed failure; (b) an observed suspension
t_{Ref} is the observation moment.

3.2.2 Dummy dataset

As an illustration of a dummy dataset, consider the following example, where fictional alarm levels for distribution cables are used. These alarm levels are shown in table 3.2. The alarm

Table 3.2: Example of alarm levels of three asset data types for distribution cables

Asset data type	Alarm level
Partial discharge magnitude [pC]	5000
Tangens delta [%]	1
Historical peak load [% S_{rated}]	120

levels in this table are used to generate *dummy datasets* containing values for the combinations of asset data types. These dummy datasets contain fictional assets with values that cover all alarm level combinations. These combinations enable a comparison between the utility's alarm levels and the alarm levels present in the DNV GL HI. To test all combinations of N asset data types, 2^N criteria combinations are generated. The example with three asset data types presented in table 3.2 results in the eight dummy datasets presented in table 3.3.

Table 3.3: Example of alarm levels of three asset data types for distribution cables

Asset [#]	Partial discharge [pC]	Tangens delta [%]	Historical peak load [% S_{rated}]
1	0	0	0
2	0	0	120
3	0	1	0
4	0	1	120
5	5000	0	0
6	5000	0	120
7	5000	1	0
8	5000	1	120

3.3 Lifetime definition

To compare the predictions and observations datasets, the definition of lifetime should be clear for each dataset. Ideally, exactly the same definition for lifetimes would apply to each of the datasets that are used by the methods for *HI prediction quality quantification* and *validation*. However, this appears not to be the case, as will be shown in the case studies of chapter 6 and 7.

For each version of the DNV GL HI, the RL of an asset is defined in the HI as the period during which *at most* $X\%$ of the assets with the same values for asset data is expected to fail. Therefore, the *predicted lifetimes* in the *predictions dataset* are based on the period from commissioning to the moment when *at most* $X\%$ of the assets is expected to have failed with the same values for asset data. This moment is equal to the moment the failure probability is at most X%, given by:

$$Predicted\ lifetime = t_{Pred:max.\ X\%\ failed} - t_{comm} \tag{3.2}$$

The observed lifetimes are defined differently from the predicted lifetimes. The *observed lifetimes* are defined as the period starting from commissioning to the moment of the first observed failure. This is given by:

$$Observed\ lifetime = t_{Obs:failed} - t_{comm} \tag{3.3}$$

3.4 Lifetime comparison

The previous section has shown that the predicted and the observed lifetimes are defined differently. This section explains how these differently defined lifetimes can be compared. To explain this comparison, the difference between predicted and observed health is taken to be zero, i.e. the *HI prediction quality* has the maximum value. In this thesis, this maximum value is taken *1*.

For this case, the predicted lifetime equals the lifetime during which a maximum of X% of the assets that have the same values for asset data are observed to fail. A prediction holding a maximum *HI prediction quality* for an asset which is expected to fail equals the lifetime when the failure probability for that asset equals 50%. Figure 3.5 shows this graphically. The blue line and dotted lines show that the predicted lifetime is the period from commissioning until X% of the assets with the same values for asset data have failed. The green line and dotted lines show that the observed lifetime is the period from commissioning until the failure probability equals 50% for assets with the same values for asset data.

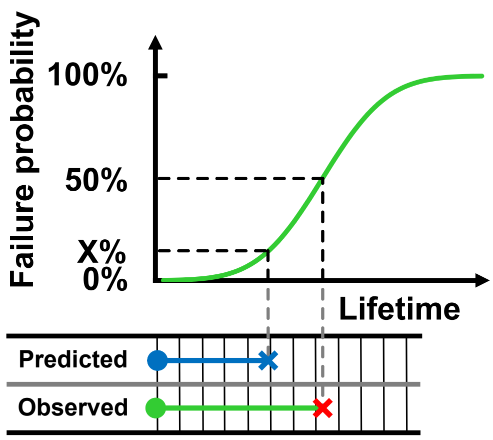


Figure 3.5: Explanation of the difference in definition between the predicted and observed lifetime for zero difference between predicted and observed health

Using this definition, it is not possible to directly compare the predicted asset health to the observed asset health for *single* assets. To do this comparison, it would require to adjust the predicted RL for a single asset to the period after which *50%* of the assets that have the same values for asset data have failed. To do so, the following two requirements should be met:

1. The HI needs to be adjusted for this new RL definition. This requires all assessment functions to be re-written using expert knowledge, to make predictions that are in line with this new definition.
2. It should be proven, that when the predictions of the *adjusted HI* are valid, the predictions of the *original HI* are also valid.

To fulfil these requirements, adjusting the present HI to meet the above-mentioned requirements using DNV GL's experts would require these experts to be consulted more frequently than feasible for this graduation project. Therefore, it can be concluded that within this thesis, an adjusted HI following a different predicted lifetime definition can neither be developed nor validated.

However, the lifetime definition does allow comparison for a *group* of assets. When predictions are made for a *group of assets*, the lifetime for which the failure probability of the predicted lifetimes equals 50% is equal to the lifetime for which *at most* X% of the assets have failed. At most X% failed assets equals a maximum failure probability of X%. This is explained in further detail in section 3.5.5.

As can be concluded from this section, in this thesis, only a comparison for a *group* of assets is possible. However, in the future, also a comparison for a *single* assets could become possible. Therefore, this chapter also describes methods using comparison for *single* assets.

3.5 HI prediction quality quantification and validation: methods description

3.5.1 Method 1: New HI

The first step in this method is to develop a new HI in addition to the existing DNV GL HI.

Next, the existing DNV GL HI and the new HI are provided with asset data (see section 3.1). This asset data consists of actual data from utilities or fictive data.

This method uses three steps:

1. The new HI and the DNV GL HI translate the asset data into RLs. According to the translation described in section 3.1, these RLs are translated into predicted lifetimes.
2. The predicted lifetimes of the two HIs are compared. Using this comparison, the *HI prediction quality* is derived. Methods to compare predicted lifetimes are discussed in sections 3.5.3-3.5.5.
3. For this comparison, a maximum difference between the predictions of the two HIs is defined using the methods discussed in 3.5.3-3.5.5 to enable *validation* of the DNV GL HI.

Figure 3.6 shows the abovementioned steps of this method.

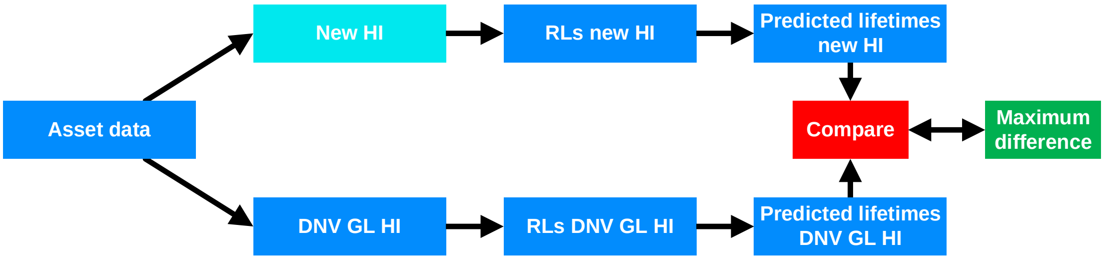


Figure 3.6: Comparing remaining lifetimes (RLs) of a new HI to the DNV GL HI

3.5.2 Method 2: Utility criteria

Most utilities base the decision to take equipment out of service on certain criteria. In this section, these criteria are referred to as *utility out of service criteria*. Using these criteria, dummy datasets containing values for several asset data types are constructed to compare the *HI prediction quality* of the DNV GL HI and the *utility out of service criteria*.

The DNV GL HI translates the asset data of dummy datasets into RLs. Based on a reference RL, RL_X, the RLs are translated into a *status prediction*. This *status prediction* can adopt two values: *good* or *critical*. The final step in this method is to compare the *status predictions* to the *utility out of service criteria*. In this comparison, the number of cases for which the status predictions match the *utility out of service criteria* can be used as an indicator for *HI prediction quality*. However, this method can not be used for a *validation* of the DNV GL HI, as the predictive value of *utility out of service criteria* is unclear. Figure 3.7 shows an overview of this method.

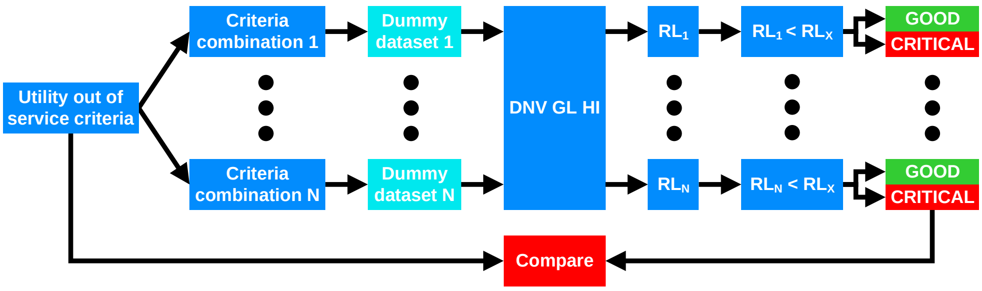


Figure 3.7: Comparing the results of dummy datasets based on out of service criteria used by utilities
The number of cases in which the status predictions match the utility out of service criteria can be used as an indicator for HI prediction quality.

3.5.3 Method 3: Binary classification

This method classifies the outcome of entries in the predictions dataset of the HI for each single asset to four cases. This classification is applied to multiple assets to find *HI prediction quality* by looking at parameters resulting from the binary classification. These parameters are explained in the end of this section.

Both the predictions and observations datasets are used to perform this classification. To explain this classification, the *time window* between two points in time is used: t_{HI} and t_{Ref} (see section 3.2.1).

The classification in this method distinguishes four cases:

I. The HI predicts the asset to fail *within* the time window and it failed *during* the time window.

II. The HI predicts the asset to fail *within* the time window but it failed *after* the time window (an *observed suspension*).

III. The HI predicts the asset to fail *after* the time window (a *predicted suspension*) but it failed *during* the time window.

IV. The HI predicts the asset to fail *after* the time window (a *predicted suspension*) and it failed *after* the time window (an *observed suspension*).

In figure 3.8, these cases are further illustrated. Case I and IV show that the prediction matches the observation. Therefore, these cases are coloured green in this figure. Case II and III show that the prediction differs from the observation. Therefore, these cases are coloured red in this figure.

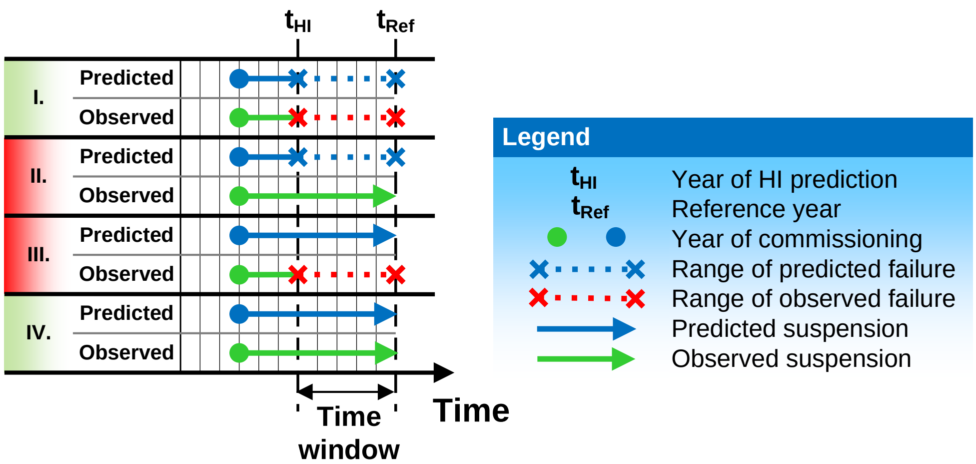


Figure 3.8: Graphical representation of the four cases, explained for a single asset

This method is applied to assets of a utility. The four cases form the basis for five *HI prediction quality parameters*. Figure 3.9 shows the *HI prediction quality parameters*:

1. **Positive predictive value.** Given a predicted failure, the probability that the asset failed.

2. **Negative predictive value.** Given a predicted suspension, the probability that the asset failed.

3. **True positive rate.** Given an observed failure, the probability that asset failure was predicted.

4. **True negative rate.** Given a suspension, the probability that a suspension was predicted.

5. **Accuracy.** Given an outcome, the probability the outcome was predicted.

For these *HI prediction quality parameters*, *threshold values* should be chosen. *Validation* is considered successful when all *HI prediction quality parameters* meet their *threshold values*.

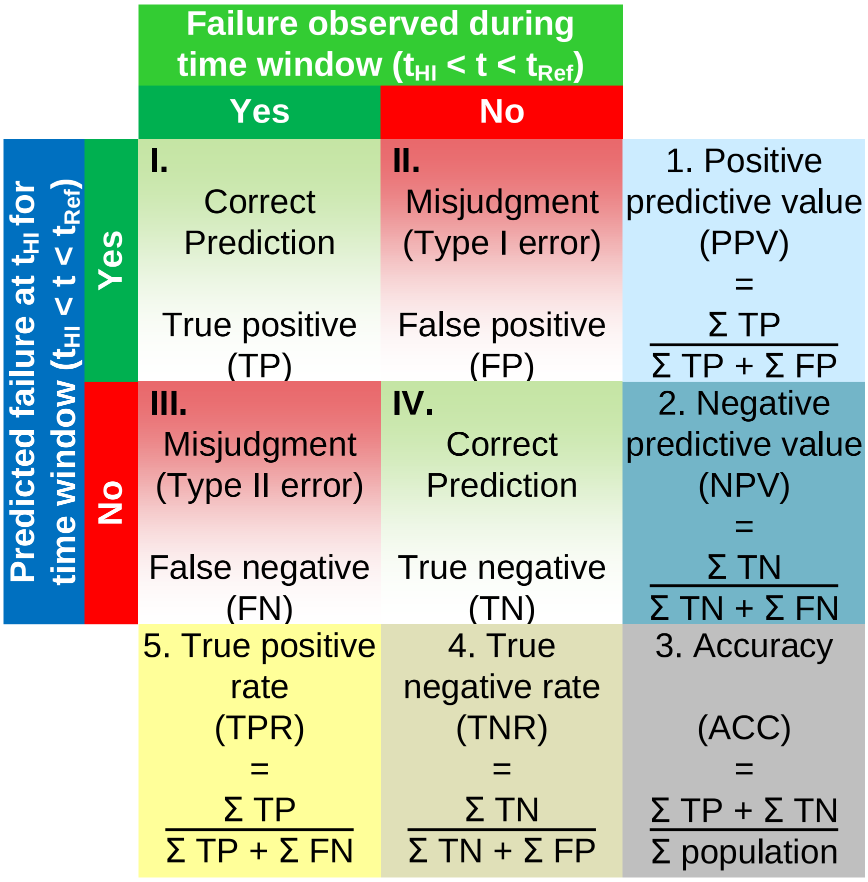


Figure 3.9: Comparing the results of predictions to observations

3.5.4 Method 4: Prediction error

The method introduced in this section compares the predictions and observations for individual assets. For each asset, the time difference between predicted and observed moment of failure is determined. The method from the previous section disregards the information of this difference by only looking at the *interval* during which failures were predicted or observed. Using this time difference, the *prediction error* is defined for *HI prediction quality quantification*. Subsequently, two *prediction error* representations are discussed in this section and used for HI *validation*.

In this method, six cases are distinguished:

X_1: The asset fails *within* the time window *exactly in* the year it was predicted to fail.

X_2: The asset fails *within* the time window *before* the year it was predicted to fail.

X_3: The asset fails *within* the time window *after* the year it was predicted to fail.

X_4: The asset fails *after* the time window but was predicted to fail *within* the time window.

X_5: The asset fails *within* the time window but was predicted to fail *after* the time window.

X_6: The asset fails *after* the time window and was predicted to fail *after* the time window.

These cases are shown in figure 3.10a. This method bases *HI prediction quality* on the difference between the predicted and observed year of failure.

For cases X_4 and X_6, the observed failure occurs after the time window at an unknown moment. The fact that this moment is unknown adds uncertainty to assessment of the *prediction quality.*

For case X_4, only the minimum time the prediction differs from the observed failure (the period from the predicted failure to the reference time) is known. In this case, this minimum difference is taken.

For case X_6, no conclusion can be drawn based on the difference between the moment of predicted failure and the observed failure. In this case, the asset is excluded from the analysis. Furthermore, this case is added to the number of occasions where difference between prediction and observation is unclear.

The values for examples of the abovementioned cases are presented in figure 3.10b. For example, the difference between predicted and observed moment of failure for case X_3 is 3 years. The minus sign is added to show that the HI prediction was too pessimistic for this case.

Figure 3.10c shows the ordering of the values based on whether they represent too optimistic, or too pessimistic values. The top of figure 3.10d shows the result representation by a histogram counting the number of occurrences for each prediction difference interval. The bottom of figure 3.10d shows the parameters that are derived using figure 3.10c. Both representations quantify the *HI prediction quality.*

For the histogram in top of figure 3.10d, reference values are established. *Validation* of the DNV GL HI is successful when the values in this histogram comply with the reference values.

For the parameters in the bottom of figure 3.10d, reference values are established. *Validation* of the DNV GL HI is successful when the values of these parameters comply with the reference values.

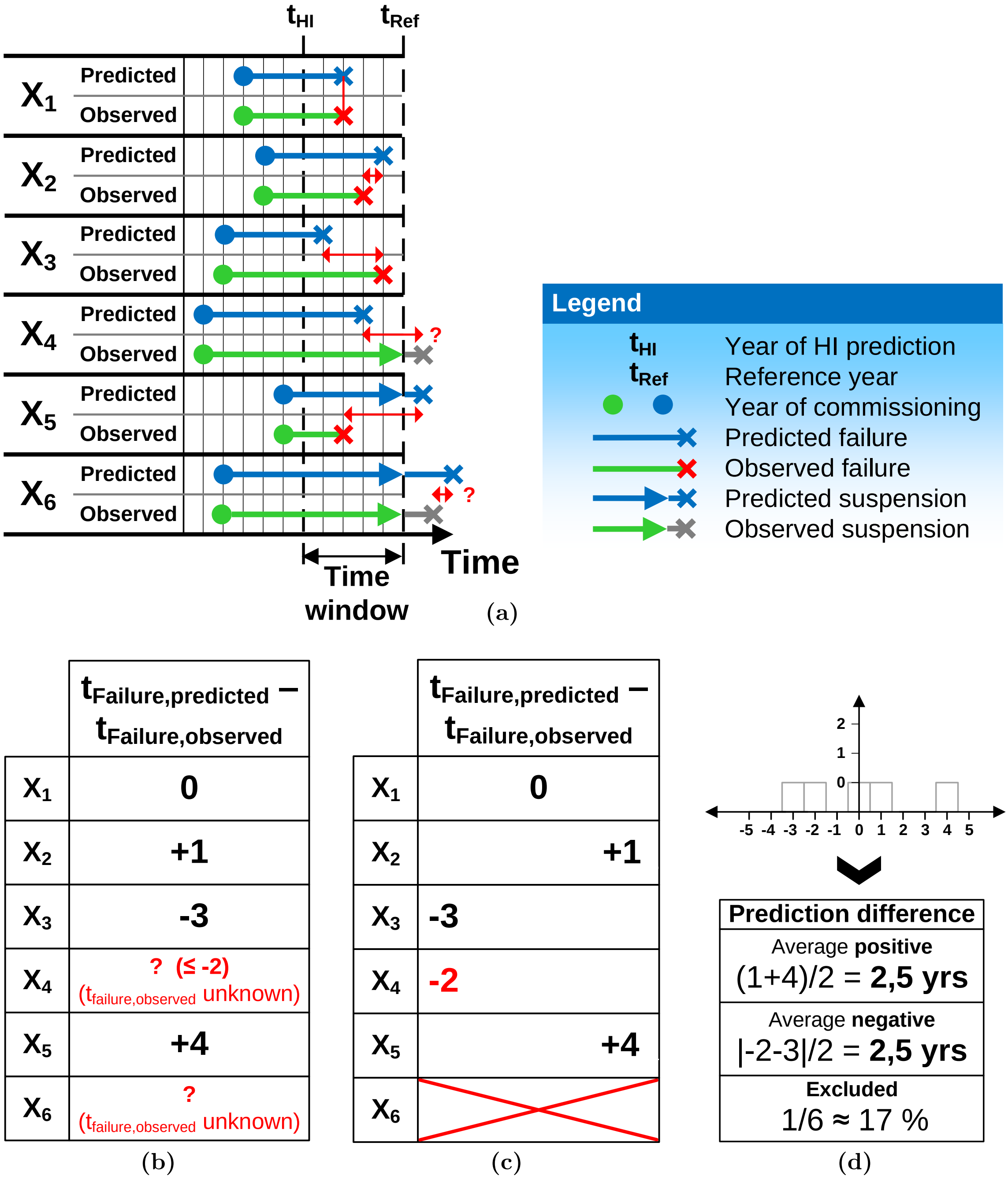


Figure 3.10: Explanation of the prediction error method: (a) Graphical representation of the six cases distinguished in this method; (b) Prediction error quantification for the six cases; (c) Positive and negative prediction error quantification; (d) Representations for prediction error quantification

This method results in representations that take into account most of the available information.

3.5.5 Method 5: Life data analysis

The method introduced in this section compares predictions and observations for a *group* of assets. The *HI prediction quality* is quantified by the difference between predictions and observations on *group level*. *Validation* is done by comparing this difference to a reference value.

In this method, the following four cases are distinguished.

For the predictions dataset:

1. The asset was predicted to fail within the time window.
2. The asset was predicted to fail after t_{Ref}.

For the observations dataset:

3. The asset was observed to fail within the time window.
4. The asset was observed to fail after t_{Ref}.

For each case, the RL of the asset is translated into the corresponding predicted lifetime. This translation is explained in section 3.2.1.

In figure 3.11a, the predictions for assets X_1 - X_3 illustrate case 1. For this case, the *failure* is predicted within the time window. Therefore, the asset in this case is in the failed ('F') state.

For this case, the period starting from the commissioning year and ending at year of predicted failure is called state total time.

The predictions for assets X_4 and X_5 illustrate case 2. For this case, the moment of failure is predicted after the time window. Therefore, this case is in the suspended ('S') state, i.e. a *suspension* is predicted. For this case, the period starting from the commissioning year and ending at t_{Ref} is called *state total time.*

The observations dataset is translated in the same manner as the predictions dataset.

In figure 3.11b, the observations for assets X_1, X_2 and X_4 illustrate case 3. This case is in the failed ('F') state. For this case, the period starting from commissioning year and ending at the moment of failure is called *state total time.* The observations for assets X_3 and X_5 illustrate case 4. This case is in the suspended ('S') state. For this case, the period starting from the commissioning year and ending at t_{Ref} is called *state total time.*

Life data analysis is used to compare both datasets. (Reliability) life data analysis is the study and modelling of lifetimes [75] and is explained in further detail in section 4.3.

The steps necessary for life data analysis are:

1. Gathering the life data (this is described in section 3.1).
2. Selecting a lifetime distribution that fits the data and models the lifetimes.
3. Estimation of the parameters that fit the distribution to the data.
4. Generation of curves and results to estimate life characteristics.

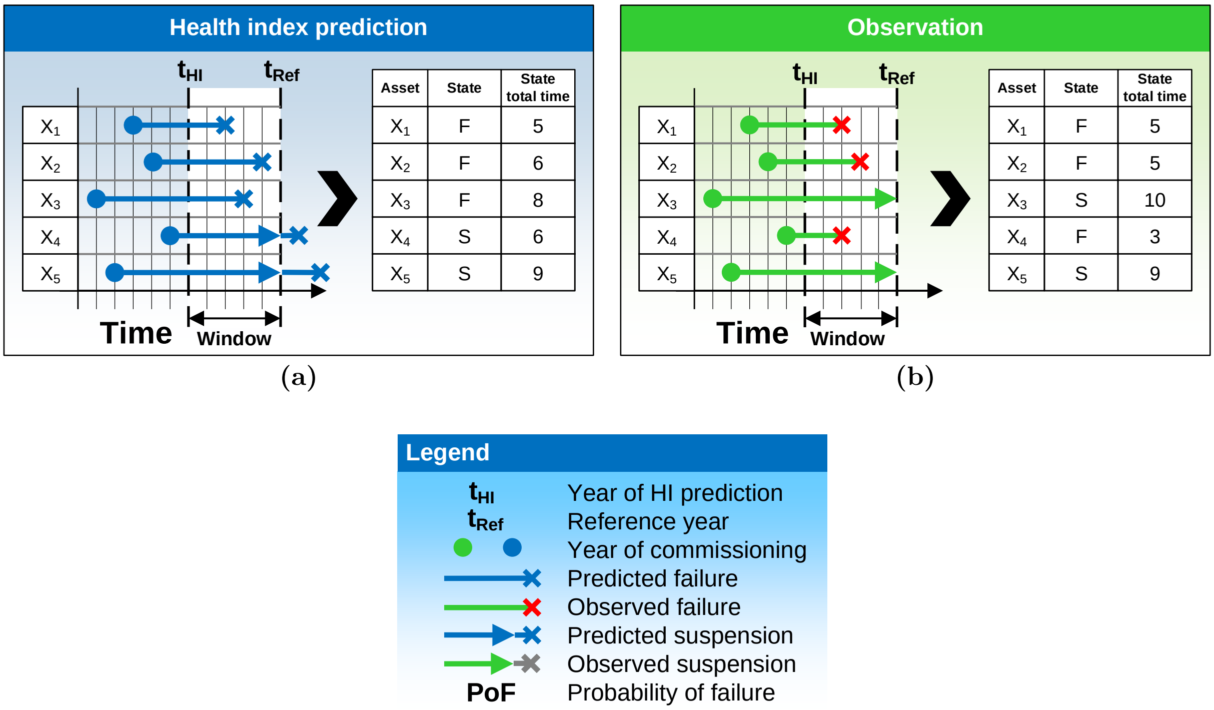


Figure 3.11: Preparation of datasets containing: (a) prediction life data; (b) observation life data

This method uses the curves of step 4 to compare the life characteristics of the predictions and observations datasets. This method uses the lifetime definition as described in section 3.3.

Figure 3.12 shows the lifetime curves for the predictions and observations datasets. In figure 3.12a, the predicted health equals the observed health. In this figure, the *threshold value* between a valid and a non-valid *observed* failure probability, for the average predicted lifetime, equals X%.

1. When the average predicted lifetime, i.e. the lifetime for which a failure probability of 50% is predicted, shows an observed failure probability below this *threshold value*, the prediction is considered *not valid*. This is shown by the *red area* in the figure.

2. When the average predicted lifetime, i.e. the lifetime for which a failure probability of 50% is predicted, shows an observed failure probability below this *threshold value*, the prediction is considered *valid*. This is shown by the *green area* in the figure.

Figure 3.12b shows a case for which the average observed failure probability for the average predicted lifetime is above the defined average predicted failure probability as given by the definition of section 3.4. For this case, the HI predictions are *too optimistic*. In this case, the HI prediction is *not valid* and the the lowest possible HI prediction quality is achieved.

Figure 3.12c shows a case for which the average predicted health is below the average observed health. For this case, the HI predictions are *too pessimistic*. In this case, the HI prediction is *valid*.

The orange arrow in the right of this figure shows the distance from the predicted failure probability to the defined predicted failure probability. This *distance* (in percent) is inversely proportional to the *HI prediction quality* and is defined by:

$$HI\ prediction\ quality = 1 - \frac{distance\ [\%]}{X\%} \tag{3.4}$$

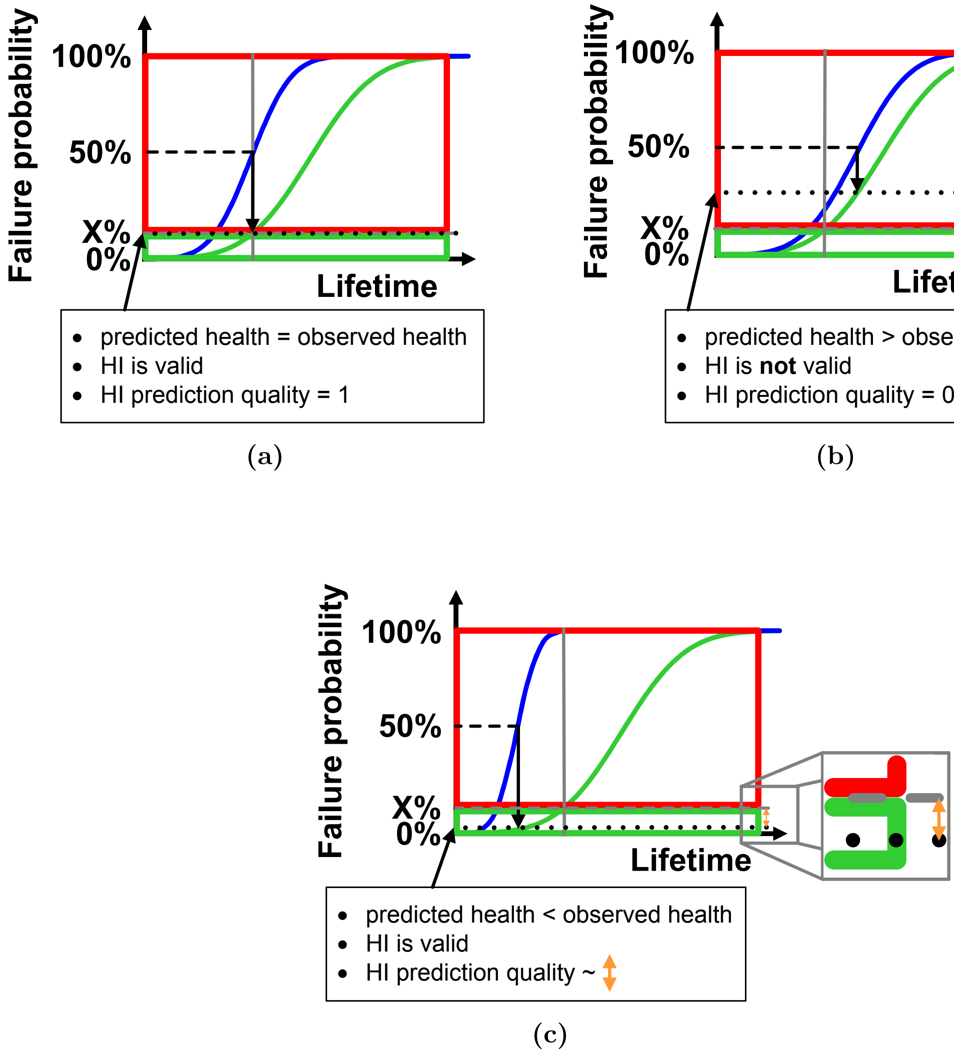


Figure 3.12: Prediction quality and validity by lifetime for: (a) predictions that match observations; (b) too optimistic predictions; (c) too pessimistic predictions
The orange arrow in figure 3.12c shows the distance of this prediction from the best HI prediction quality, which is shown in figure 3.12a.

3.6 Discussion of the methods

The five methods described in section 3.1 show that, using these methods, it is possible to define *HI prediction quality* and to perform *validation* of the DNV GL HI. This section discusses the methods regarding their strengths and weaknesses. Table 3.4 summarizes the strengths/weaknesses of the methods discussed in this section.

The first method is to develop a new HI. This method has the advantage that developing a new HI is a good exercise to gain knowledge on the working principles behind the DNV GL HI. Furthermore, this method might reveal weaknesses in the DNV GL HI. The new HI is developed using only the experience of the author. The DNV GL HI is developed in multiple years using expert knowledge.

Therefore, the drawback of this method is that the new HI will only reveal a very limited number of weaknesses of the DNV GL HI. Besides, the new HI is also likely to contain weaknesses, which further limits the value of this method.

The second method is to compare experience of utilities to the expert knowledge contained in the DNV GL HI. The advantage of this method is that the comparison can reveal weak points of the DNV GL HI by differences between experience of utilities and the expert knowledge contained in the DNV GL HI. This then enables possible improvement of the DNV GL HI.

However, the predictive value of these *out of service criteria* is unclear and requires further research. In addition to this, this method can not be used for a *validation* of the DNV GL HI. Furthermore, most of the *out of service criteria* are based on expert knowledge and values provided by manufacturers. The expert knowledge put in these values could be subjective. Manufacturers do not make these values publicly available. Besides, utilities need to maintain a good relationship with manufacturers.

Therefore, these values as provided by manufacturers are hard to obtain.

The third method is to classify the outcome of entries in the predictions dataset of the HI for each single asset. This method has the advantage that it is possible to apply it (partly) to the acquired data. Furthermore, it is applied in several fields of study. The disadvantage of this method is that each of the quality parameters only represents the *HI prediction quality* in a single number.

Therefore, it is complicated to do statements on *HI prediction quality* based on the values of these parameters.

Another disadvantage arises from the information lost when translating predictions and observations in the four cases. The lost information consists of the observed moment of failure and the predicted moment of failure.

The fourth method is to compare predictions and observations of *individual* assets. The advantage of this method is that a minimum amount of information is lost when finding the prediction error. The disadvantage of this method is that the reference values for *validation* should still be defined properly.

The fifth method is to compare predictions and observations of a *group* of assets. The advantage of this method is that it enables comparison between predictions and observations on group level. This enables to quantify *HI prediction quality*. The disadvantage of this method is that no detection is possible of individual differences between predictions and observations. This method can compare the two datasets in meaningful ways to find points of improvement for the DNV GL HI. Despite this, this method does not take into account the differences between predictions and observations of individual assets.

Figure 3.13 explains this by showing the data points of predicted and observed failures. In this figure, the predicted failure of data point 5 has the largest difference in failure probability compared to the observed failure of data point 5. The other data points show as well a (large) difference in failure probability (execpt for data points 3). However, the *HI prediction quality* has the maximum value, *1*.

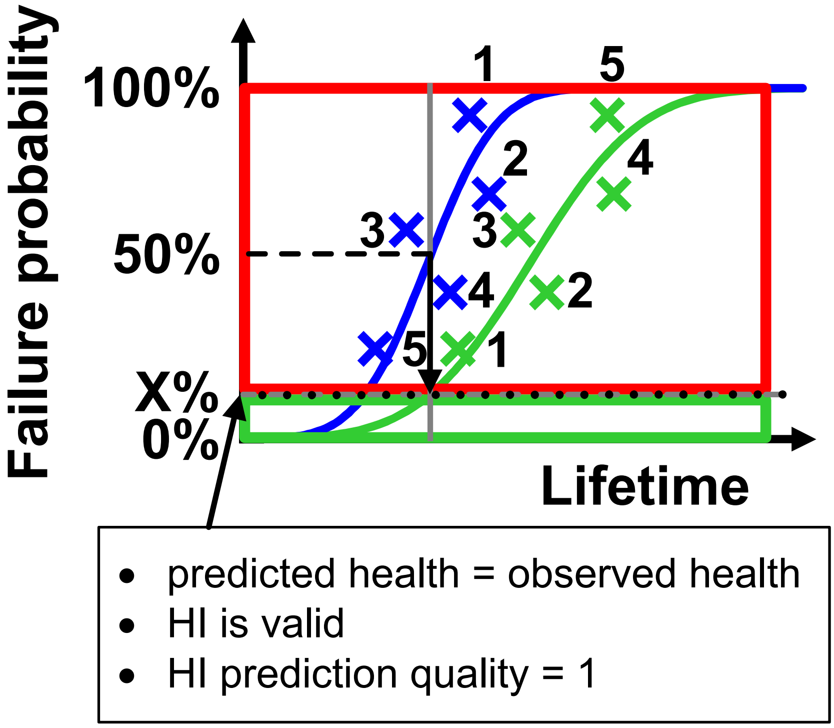


Figure 3.13: Copy of figure 3.12c, showing an example of individual differences in predictions and observations datasets, which still result in the maximum value for HI prediction quality

Table 3.4: Strengths/weaknesses analysis of the methods for *HI prediction quality quantification* and *validation*

Section	Method	Strengths	Weaknesses
3.5.1	1. New HI	• Better understanding DNV GL HI	• Reveals limited number of weaknesses in DNV GL HI • *New HI* also likely to contain weaknesses
3.5.2	2. Utility criteria	• Reveals weaknesses in DNV GL HI	• Predictive value *out of service criteria* unclear • Criteria hard to obtain
3.5.3	3. Binary classification	• Can be applied (partly) to data • Applied in several fields of study	• Hard to do statements on *HI prediction quality* • Large amount of information lost in representation
3.5.4	4. Prediction error	• Low amount of information lost in representation	• Room for improvement in *HI prediction quality* definition
3.5.5	5. Life data analysis	• Comparison on group level	• Individual differences between predictions and observations neglected

3.7 Adopted methods

The main argument for adopting the methods is their value for both the research described in this thesis and future research. This value depends on two properties of the method:

1. The extent to which the method is suitable for *HI prediction quality quantification* and *validation*.
2. The extent to which the method is suitable for enhancing the *HI prediction quality* for *validation* of the DNV GL HI by future research.

Given the disparity between the level of expertise of the author and the experts that developed the DNV GL HI, it is not realistic to expect that a *newly developed HI* by the author will outperform the DNV GL HI. Therefore, developing a new HI does not yield reliable statements on the *HI prediction quality*. This led to the decision to exclude this method from the research.

A comparison using utility criteria can only yield differences between the expert knowledge in those criteria and the expert knowledge in the DNV GL HI. These differences could only yield very specific points of improvement. Therefore, no quantification of the overall *HI prediction quality* is possible using this method. As this thesis aims to extend the overall foundation of the DNV GL HI by quantification of the *HI prediction quality*, this method was excluded as well.

Binary classification of the outcomes for each single asset is possible. However, using binary classification, a large amount of information is disregarded during the classification. This significantly reduces the usability to enhance the *HI prediction quality* of the DNV GL HI in future research.

The *prediction error method* is able to provide several statements on the *HI prediction quality*. Besides, this method can also be used for enhancing the *HI prediction quality* of the DNV GL HI.

The *life data analysis method* also enables statements on the prediction quality. Besides, its results can be used to improve the *HI prediction quality* of the DNV GL HI. However, this method can not be used for comparison on *group level* due to the definition of predicted lifetime.

To conclude, the *prediction error* and *life data analysis* methods complement each other regarding their weaknesses. As the *prediction error* method can not be used for comparison on group level, this method was excluded as well.

Therefore, only the *life data analysis* is adopted in this thesis to find the *HI prediction quality* and for *validation* of the DNV GL HI.

Chapter 4

Statistical methods and background

This chapter presents and discusses the statistical methods used in this thesis for *HI prediction quality quantification* and *validation.*

Section 4.1 presents an overview of common probability functions. Section 4.2 uses these probability functions to describe three types of statistical distributions. Section 4.3 describes the concept of life data analysis, as introduced in chapter 3, in more detail. Section 4.4 discusses the assumptions made for failure distribution fitting and parameter estimation. Finally, section 4.5 summarizes the main conclusions of this chapter.

4.1 Probability functions

Probability functions describe the distribution of probabilities. In reliability engineering, they are mostly used to describe *failure behaviour* of assets using *statistical distributions. Statistical distributions* are explained in section 4.2.

Failure data of assets consists of the time to failure, representing the time span until failure, occurs [58]. For this so-called *time-to-failure* data, the random variable *X*, the age of the failed component, can have a value ranging from zero to infinity.

The *probability density function* (PDF) is a function $f(t)$ that describes the likelihood of random variables to take on a certain value. In reliability engineering, this function describes the distribution of failure probability. An example of a PDF is given in figure 4.1.

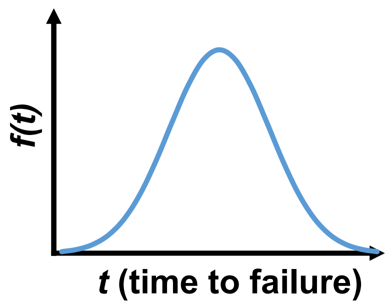


Figure 4.1: Example of a probability density function

Taking the integral of the PDF results in the *cumulative distribution function* (CDF). This function, like the PDF, describes the likelihood of random variables to take on a certain value *for a specified interval.* In reliability engineering, this function describes the probability of failure for the interval starting from zero to t, and is denoted by *F(t).*

The CDF relates to the PDF as:

$$F(t) = \int_{-\infty}^{t} f(t)dt \tag{4.1}$$

An example of a CDF is given in figure 4.2.

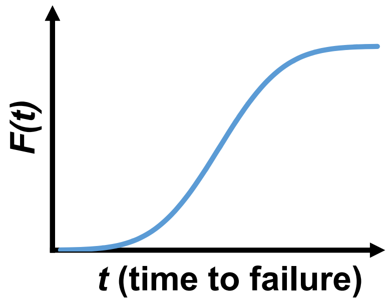


Figure 4.2: Example of a cumulative distribution function (CDF), based on the PDF of figure 4.1

Next to the PDF and CDF, reliability engineering often uses the *failure rate function*, $\lambda(t)$, to represent reliability (also known as the *hazard rate function*, *h(t)*). This function describes the failure rate (e.g. yearly number of failures) for a moment, indicated by t. The failure rate function relates to the PDF and CDF as:

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \tag{4.2}$$

An example of a failure rate function is given in figure 4.3.

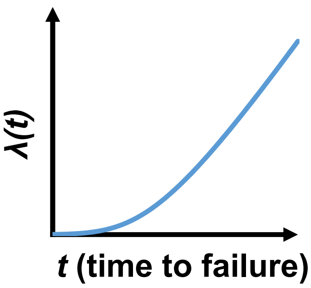


Figure 4.3: Example of a failure rate function, based on the PDF and CDF of figures 4.1 and 4.2

4.2 Statistical distributions

Several types of distributions exist. In this thesis, the distributions are used for modelling the failure behaviour of assets. In failure behaviour of assets, multiple stages can be distinguished. These stages can be described with a single curve, of which the shape is similar to a bathtub. Section 4.2.1 explains this so-called *bathtub curve*. Sections 4.2.2 and 4.2.3 explain the two distribution types present in this thesis.

4.2.1 Bathtub curve

The failure behaviour of assets is generally characterized by three stages:

Stage 1: The stage of *infant mortality*. During this stage, most failures occur shortly after installation and a decreasing failure rate $\lambda(t)$ is observed. This stage starts at the assets commissioning year.

Stage 2: The stage of *normal life*. During this stage, failure behaviour is random and a constant failure rate $\lambda(t)$ is observed. This stage is between the infant mortality and wear-out stage.

Stage 3: The *wear-out* stage. During this stage, most observed failures are due to ageing and an increasing failure rate $\lambda(t)$ is observed.

Figure 4.4 shows an example of a bathtub curve as function of time for each of the three stages. In the figure, the horizontal axis shows the time to failure t. The vertical axis shows the failure rate $\lambda(t)$. The failure behaviour is illustrated by this figure for each stage.

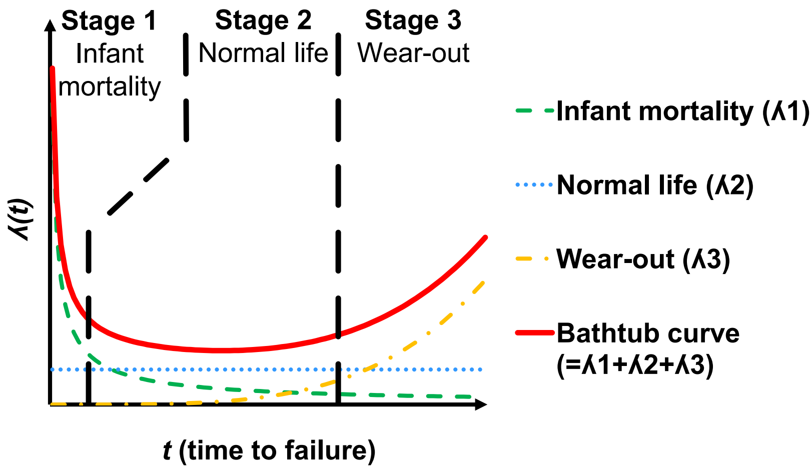


Figure 4.4: Example of a bathtub curve including the three stages

4.2.2 Normal distribution

The normal distribution is characterized by two parameters:

1. μ, representing the *mean value* of the distribution
2. σ, representing the *standard deviation* of the distribution

The normal distribution is described by the PDF:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2} \tag{4.3}$$

Figure 4.5 shows a plot of this PDF with the two parameters, μ and σ. For expert judgement these parameters are not only well known, but also best interpretable. Therefore, failure behaviour is described in the DNV GL HI using the parameters of the normal distribution whenever expert judgement is required.

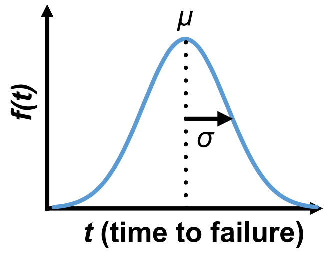


Figure 4.5: Example of a normal probability density function

4.2.3 Weibull distribution

For lifetime modelling, usually the Weibull distribution is used. Its strength lies is its flexibility, which enables to describe the different life stages (see section 4.2.1) of a population of assets [76].

The Weibull distribution is characterized by three parameters:

1. η, the *scale* parameter
2. β, the *shape* parameter
3. γ, the *location* parameter

The *location* parameter determines the moment when the first failures occur. For electrical power equipment, the first failures can already occur from the *moment of commissioning*. Therefore, the value of this parameter is set to zero in the majority of the reliability studies on electrical power equipment. This results in a Weibull PDF characterized by the two parameters η and β, also known as the *2P Weibull PDF*:

$$f(t) = \left(\frac{\beta}{\eta}\right)\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \mathit{with}\ \beta, \eta, t > 0 \tag{4.4}$$

Figure 4.6a shows an example of the *2P Weibull PDF* for three different values of η. This example shows the effect of η on the *scale* of the PDF. Figure 4.6b shows an example of the *2P Weibull PDF* for three different values of β. This example shows the effect of β on the *shape* of the PDF.

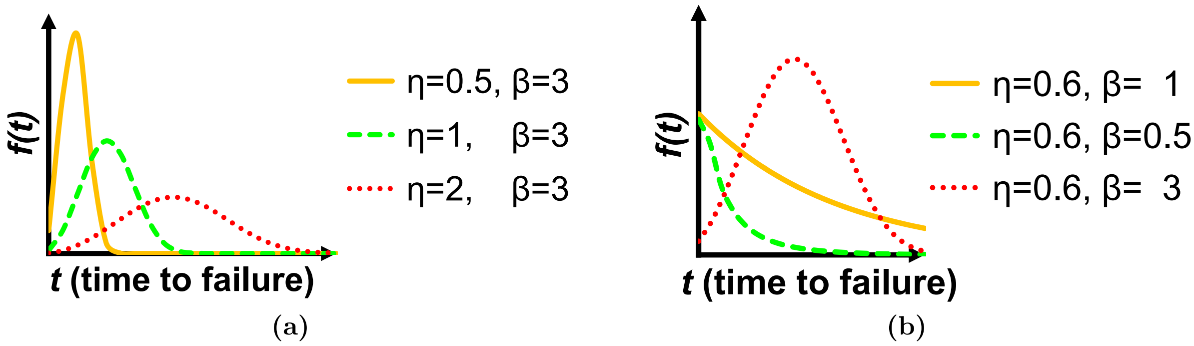


Figure 4.6: *2P Weibull PDF* for three different values of: (a) the *scale* paramater η; (b) the *shape* parameter β

4.3 Life data analysis

A very brief introduction into *life data analysis* (LDA) has already been given in chapter 3 by the application in a method for *HI prediction quality quantification* and *validation*. This section explains this method for analysing life data in more detail. As shown in the work of Mehairjan [76], two types of methods exist for analysing life data: *parametric* and *non-parametric*.

Of these two methods, the work of Mehairjan [76] shows that *parametric* methods are most suitable for describing failure behaviour of a population, as these methods are most appropriate for large data samples.

As LDA is used in this thesis to compare the failure behaviour of asset populations, the parametric method is chosen. As mentioned in chapter 3, the four steps for life data analysis are:

1. Gathering the life data (this is already described in chapter 3).
2. Selecting a lifetime distribution that fits the data and models the lifetimes.
3. Estimation of the parameters that fit the distribution to the data.
4. Generation of curves and results to estimate life characteristics.

The assumptions for steps 2 and 3 will be explained in section 4.4.

4.4 Failure distribution fitting and parameter estimation

Failure distribution fitting is the step during which a lifetime curve is fitted to the life data.

The work of Mehairjan [76] presents three methods for failure distribution fitting:

1. Probability plotting
2. Rank regression analysis (Least Squares Estimation, LSE)
3. Maximum likelihood estimation (MLE)

From these methods, only MLE can take into account large numbers of suspensions. The datasets used in this text contain a large number of suspensions. Therefore, this thesis takes MLE into account as the method for distribution fitting. Using MLE, the parameters of the lifetime distribution are estimated.

The work of Mehairjan [76] describes the main methods to quantify the quality of the failure distribution fit to the data. In summary, these methods include:

- *Visual inspection*, by visually looking how well the curve fitted to the data represents the data points
- *Correlation coefficient*, which shows how well the probability line fits the data.
- *Likelihood value*, which can be used to in addition to visual inspection to assess the fit of the distribution to the dataset.
- Other methods, Kolmogorov-Smirnov test, Anderson-Darling test, etc.

From these methods *visual inspection* will be used to quantify quality of the failure distribution fits to the data and is explained in the next section.

4.4.1 Visual inspection

Visual inspection consists of visually looking how well the curve fitted to the data represents the data points. The 90% confidence bounds show that the failure probability F(t) is expected to stay within the interval of the two bounds with a confidence of 90%. In the work of de Haan [77], it is shown that during the distribution fitting, data points can be outside the confidence bounds of the distribution. These data points are called *outliers*. Figure 4.7 shows an example of a fit with two outliers (coloured grey). Furthermore, the work of de Haan [77] states that when the number of outliers is very large, the data quality appears to be low. Low data quality causes the data to be unrepresentative for modelling failure behaviour of assets.

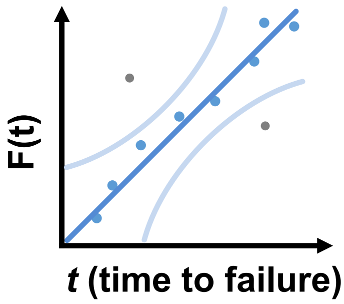


Figure 4.7: Failure probability plot with outliers

4.5 Summary and conclusions

This chapter has presented the statistical methods and background used for *HI prediction quality quantification* and *validation*.

This chapter has shown three types of probability functions, the *probability density function*, the *cumulative probability function* and the *failure rate function*. These functions are used to describe failure behaviour of assets using *statistical distributions*. Three statistical distributions have been presented in this chapter, the *bathtub curve*, the *normal distribution* and the *Weibull distribution*.

Life data analysis, as introduced in chapter 3 has been explained in this chapter in more detail. It was shown that for failure distribution fitting and parameter estimation, a parametric method using MLE is most suitable to fit a failure distribution to the life data of the assets discussed in this thesis. To conclude, from four methods to find the quality of the distribution's fit, the *visual inspection* method was chosen.

Chapter 5

Utility data: collection and selection

The adopted research methods introduced in chapter 3 rely on utility data to enable *HI prediction quality quantification* and *validation*. This utility data includes operational and failed assets. Chapter 3 showed that these methods enable prediction quality quantification and that, with future research, these methods could be used to enhance the prediction quality of the DNV GL HI.

This chapter describes the collection and selection procedure of this failure, maintenance and operation data.

Section 5.1 starts with the explanation of the data request process. In this request, data requirements are formulated and used in a data request towards utilities. Section 5.2 provides an overview of the acquired data and the data selected to be used as input for the methods adopted in chapter 3.

5.1 Data request procedure

This section explains the steps taken during the data request. The first step is to define the data requirements stated in the request, which will be described in section 5.1.1. The second step is to request the data according to these data requirements. Section 5.1.2 presents the organizations that were contacted regarding this data request.

5.1.1 Data requirements

The adopted methods of chapter 3 set data requirements for the utility data. This section starts with an overview of the data requirements the data should at least meet to enable application of the *life data analysis* method (see section 3.5.5). The data requirements are described per data type.

All data types should at least meet the following *minimum* requirements:

- The data should be compliant with a version of the DNV GL HI. Therefore:
  - a version of the DNV GL HI should cover the asset type of the data.

- Sufficient data should be present to create a predictions dataset using the DNV GL HI. This data should at least:
  - include the ages of in-service components to enable statistical analysis [58].

  - uniquely identify individual assets to ensure that exactly the same assets are used in the predictions and observations datasets.

- The failure data should include a failure database:
  - containing at least the date and cause of failure.
  - containing a sufficient number of failures. For insulation breakdown data, data containing at least ten failures should be used if possible [78].

When the data fulfils the above-mentioned minimum requirements, it is possible to quantify the prediction quality of statistical remaining life functions of the DNV GL HI (see chapter 2).

To include *quantification* of the *HI prediction quality* of the utilisation remaining lifetime functions, the data should, in addition to the above-mentioned minimum requirements, meet the following requirements:

- The data should contain utilisation (loading) information which suits the format as required by the DNV GL HI.
- The data should show a significant relation between the measured values in the asset data and failure modes.
- The data should include the unit type and its operating limits [46].

To include *quantification of the HI prediction quality* of the condition remaining lifetime functions, the data should, in addition to the above-mentioned minimum requirements, meet the following requirements:

- The data should contain diagnostics (condition data) which suits the format as required by the DNV GL HI.
- The data should show a significant relation between the measured values in the asset data and failure modes.
- The data should include threshold levels allowing interpretation of the data according to the units and threshold levels of the DNV GL HI.

Quantification of the prediction quality is possible for all types of assessment function, when the data meets all abovementioned requirements.

In addition to this, requirements for *strengthening* the methods for quantification of the HI *prediction quality* are:

- The data should be available in a reliable and consistent dataset [58].
- The data should preferably include a maintenance database containing each maintenance activity together with its date.

To conclude, a requirement to *enhance the speed of data analysis*, which allows for a study using larger datasets, is:

- The data should be available in a format suitable for quick data analysis (e.g. available in a database rather than separate paper reports).

5.1.2 Contacted organisations

To find the data required for the research, the following persons and organisations were contacted:

1. The supervisors of both the TU Delft and DNV GL.
2. Within DNV GL, Henk Wels, risk and reliability engineer. He provided access to a significant number of reports on failure data. Besides, he had access to condition data for a number of power transformers. The number of transformers covered by this data appeared insufficient for application of the adopted method. Furthermore, he could not provide the load profiles and other asset data essential for use in the adopted methods.
3. Maurice Roovers, manager Energy Infrastructure at the umbrella organization for Dutch utilities, *Netbeheer Nederland*.
4. All Dutch utilities. The result of this data request is shown in table 5.1.

Table 5.1: Contacted utilities regarding data
Four utilities were willing to cooperate and did provide data.

Utility	Data provided	Comment
'A'	No	No agreement established to provide data
'B'	No	Number of assets insufficient for a case study
'C'	No	No agreement established to provide data
'D'	No	Failure data does not fulfil the requirements (individual assets are not indicated)
'E'	Yes	Willing to cooperate
'F'	Yes	Willing to cooperate
'G'	Yes	Willing to cooperate
'H'	Yes	Willing to cooperate

Four out of the eight contacted utilities were willing to cooperate and did provide data. The data acquired from these utilities and the data selection are described in the next section.

5.2 Acquired data and data selection

5.2.1 Acquired data

As mentioned in section 5.1.2, a total of four utilities provided data. The level of detail of the acquired data depended on two main properties of this data:

1. The number of asset owned by the utility that failed during the time window, should be sufficient to apply the adopted methods.
   If this criterion was not met, no additional/more detailed data was requested/acquired.

2. The time windows of data for the predictions and observations datasets should be sufficiently wide to enable application of the adopted methods.
   If this criterion was not met, no additional data/more detailed was requested/acquired.

A wide variety of expressions is found in terminology of the utility data. Table 5.2 explains these expressions as used in this chapter.

Table 5.2: Main utility data expressions
This table contains the main utility data expressions used in this chapter.

Expression	Explanation
DGA	Dissolved gas analysis
IT	Instrument transformer
CT	Current transformer
VT	Voltage transformer
CombiT	Combined transformer
Online PD (SCG)	Online partial discharge measurements by DNV GLs Smart Cable Guard.
Offline PD	Offline partial discharge measurements. In this thesis, they refer to the Oscillating Wave Test System (OWTS) or the 0.1 Hz Very Low Frequency (VLF) method.
XLPE	Cross-linked polyethylene
PILC	Paper-insulated lead covered (steel armour outer sheath)
“EDPLK”	Paper-insulated lead covered (PE outer sheath)
Mixed	Cable circuits that consist of combination of XLPE, PILC and “EDPLK”

Table 5.3 summarizes the acquired data. This table shows the available asset data for each utility. This asset data is presented by the number of assets for which data is provided and the range of this data. For example, the commissioning year of PILC distribution cables for utility ‘H’ ranges from 1930 to 2013.

The reports of off-line partial discharge measurements for the same distribution cables show measurements from 1996 to 2013. Data, for which the range is unclear, is indicated by ‘n.a.’. The number of assets at SCG represents the number of assets on which a SCG system has ever been installed. For each utility, less SCG systems than this number are concurrently operational.

Visual inspection parameters consist of corrosion, sweating, leaking and dripping. This table shows that the largest number of failure records are available for distribution cables.

Table 5.3: Available asset data
This table shows the available asset data for the four utilities.

			Utility 'E'		Utility 'F'		Utility 'G'		Utility 'H'	
Asset type	Data type	Type	Assets	Range	Assets	Range	Assets	Range	Assets	Range
Transformers (MV&HV)	Asset data	Commissioning year	-	-	1958-2012	151	1966-2011	89	-	-
		DGA	-	-	89	n.a.	151	n.a.	-	-
	Failure data	Failures	-	-	-	-	-	-	-	-
Instrument trans-form-ers (HV)	Asset data	Commissioning year (HI 2007)	-	-	-	-	11643	1950-2012	-	-
		Visual inspection	-	-	-	-	14064	2013	-	-
		DGA	-	-	-	-	1719	>2007	-	-
		CT	-	-	-	-	840	>2007	-	-
		VT	-	-	-	-	706	>2007	-	-
		CombiT	-	-	-	-	173	>2007	-	-
	Failure data	Failures	-	-	-	-	12	2005-2013	-	-
		CT	-	-	-	-	7	2005-2012	-	-
		VT	-	-	-	-	5	2011-2013	-	-
Distribution cables	Asset data	Commissioning year	-	-	-	-	-	-	1008	1930-2013
		PILC	-	-	-	-	-	-	898	1930-2013
		XLPE	-	-	-	-	-	-	44	1997-2013
		"EDPLK"	-	-	-	-	-	-	15	1970-2004
		Mixed	-	-	-	-	-	-	51	1931-2012
		Online PD (SCG)	6	2008-2014	18	2007-2014	-	-	18	2011-2014
		Offline PD reports	-	-	-	-	-	-	76	1996-2013
		Maximum load	-	-	-	-	-	-	1296	2013
		Circuit diagrams	-	-	-	-	-	-	n.a.	n.a.
	Failure data	Interruptions	-	-	-	-	-	-	97	2006-2014
		PILC	-	-	-	-	-	-	83	2006-2014
		XLPE	-	-	-	-	-	-	14	2006-2014

5.2.2 Data selection

As a result of the requirements mentioned in section 5.1.1, only a part of the data was used for *HI prediction quality quantification* of the DNV GL HI. This section describes the data selection.

From table 5.3, it appears that for the distribution cables of utilities 'E' and 'F', distribution cable diagnostics of only 6 and 18 distribution cables were available. Due to this limited amount of data, this data was not used during the research.

For utility 'F', diagnostics of 89 transformers were available. As no transformer failures were reported, this data was not used during the research.

The data on the instrument transformers of utility 'G' included inspection and failure data. This failure data was, however, limited to only seven and five failures, for current and voltage transformers, respectively. Furthermore, only inspection data from 2013 was available. Despite the limited amount of failure and inspection data, the data was already used in a version of the health index. Therefore, this data helped in the practical aspects of handling the large amount of asset data. The number of failures in this data was sufficient to show application of the *life data analysis* method (see section 3.5.5). Section 5.3 shows the process of preparing this data for the analysis in the case study of chapter 6.

The data of the 10kV distribution cables of utility 'H' included diagnostics, utilization and failure data. Furthermore, the number of failures (97) present in this data was sufficient to apply the *life data analysis method* (see section 3.5.5). Section 5.4 shows the process of preparing this data for the analysis in the case study of chapter 7.

5.3 Data preparation: DNV GL HV instrument transformer Health Index

This section provides a detailed overview of the available utility data on HV instrument transformers (ITs), as presented in table 5.3 and shows how the data is prepared for the HI assessment. This includes the assumptions on data that was not available from the utility and originated from other sources.

5.3.1 IT information

The data used in this case study covers over 13000 ITs of a HV transmission network. The data shows three types of ITs: current transformers, voltage transformers and combined transformers. This HI version was developed and used for HI predictions in 2009. The failure database contained failures over the period 2005-2013. Therefore, the time window (see section 3.2.1) was set to the period 2009-2013. From the dataset of ITs, only the ones commissioned before 2009 were selected for this analysis.

5.3.2 Utility data files summary

This section provides an overview of the data files available for the case study. Table 5.4 shows the asset data types available for each data file.

Table 5.4: Summary of the utility's data files

#	Data file	Format	Description	Unit
1	Failure database 20140605	MS Excel	Date of failure	[dd-mm-yyyy]
			Rated voltage	[kV]
			IT type	{CT, VT, CombiT}
			Location	[-]
			Brand/type	[-]
			Manufacturing year	[yyyy]
			Event	[-]
2	Healthindex 2009 v4.2 - ITs	MS Excel	Substation	[-]
			Section	[-]
			Component identification number	[-]
			Manufacturer	[-]
			Type	[-]
			Commissioning year	[-]
			Rated voltage	[kV]
			Location	{Inside, Outside}
			Gas insulated switchgear	{Yes, No}
3	Rekenblad 20131007 Overzicht STRMTRF	MS Excel	Substation	[-]
			Object description	[-]
			Substation type	[-]
			Section	[-]
			Type of section	[-]
			Component identification number	[-]
			Status	{In service, Removed, Out of service}
			Type of IT	{VT, CT, CombiT}
			Manufacturer	[-]
			Manufacturing year	[-]
			Corrosion	{Yes, No}
			Sweating	{Yes, No}
			Leakage	{Yes, No}
			Dripping	{Yes, No}

5.3.3 Data organization

Data file #3 in table 5.4, which is from 2013, includes information on corrosion, sweating, leakage and dripping. These parameters are called *condition indicators* (CIs) and are obtained by visual inspection of the asset. Data file #1 does not include these CIs. It was considered to include the CIs of data file #3 for finding the predictions dataset, assuming that the values for these CIs do not improve in time. However, no maintenance records were available for this utility. Therefore, it was not possible to check the which assets should be excluded due to maintenance during the time window. Consequently, only the asset data from data files #1 and #2 of table 5.4 were used.

For the data of this utility, it was not possible to find the corresponding asset data for the failed assets directly, as the failure database lacked a component identification number. However, based on the data on the failed asset, as available in the failure database, it was possible for each failed asset to derive the corresponding predicted RL.

5.3.4 HI asset data

For this utility, the data was already put in the HI format and results for RL were found in 2009. Therefore, the HI asset data equals the data given in data file #2. This data includes only the assets with a known commissioning year before 2009. For the ITs, the manufacturing year was assumed to be equal to the commissioning year.

Predictions dataset

The predictions dataset originated from the predicted RLs of the HI in 2009. For these assets, at t_{HI}, only the commissioning year was available.

Observations dataset

Failures in this dataset originate from the assets present in the failure database that failed during time window, for this data the period 2009-2013. Suspensions in this dataset originate from the assets that survived the time window.

5.4 Data preparation: DNV GL distribution cable Health Index

This section provides a detailed overview of the available utility data on HV instrument transformers (ITs), as presented in table 5.3 and shows how the data is prepared for the HI assessment. This includes the assumptions on data that was not available from the utility and originated from other sources.

5.4.1 Cable circuit information

The data used in this case study covers a region from a distribution cable network of 4150 km (2013, [79]). In the data of this distribution cable network, four types of distribution cables appear, as already shown in table 5.2. The outer layers of the "EDPLK" type cable circuits consists of polyethylene (PE) combined with a wire screen, while in the PILC type circuits, these layers consist of tarred jute yarn with a steel armour outer sheath.

The failure database spans the period 2007-2013. Therefore, from the dataset of PILC cable circuits, only those commissioned before 2007 were selected for this analysis.

5.4.2 Utility data files summary

This section provides an overview of the data files available for the case study. Table 5.5 shows the most important available asset data types for each data file. Appendix B contains a complete version of this table. The case study using this data does not take into account individual joints. Therefore, data file #7 was not used.

Table 5.5: Summary of the most important data files of utility 'H'

#	Data file	Format	Description	Unit
1	Loadflow results	MS Excel	From - To	[-]
			Insulating material	[-]
			Maximum load	[%]
2	Distribution cable circuit diagrams	PDF	From - To	[-]
			Circuit number	[-]
3	Data distribution cables	MS Excel	Conductor material	[-]
			Insulating material	[-]
			Commissioning year	[yyyy]
			Circuit number	[-]
4	Offline PD	PDF	From - To	[-]
			Circuit number	[-]
			Test date	[dd-mm-yyyy]
			Test result	[-]
5	Nestordata failures	MS Excel	From - To	[-]
			Region	[-]
			Component type	[-]
			Failure cause	[-]
			Date of failure	[dd-mm-yyyy]
6	Cable type specs	MS Excel	# conductors	[-]
			Conductor material	[-]
			Insulating material	[-]
			Short circuit current rating (1 s)	[kA]
7	Data distribution joints	MS Excel	Joint object ID	[-]
			Joint asset number	[-]
			Region	[-]
			Type	[-]
			Commissioning year	[yyyy]

5.4.3 Data organization

This section explains the steps taken to put the available asset data in a single file for all assets. Figure 5.1 shows the combination of the data files used during this case study.

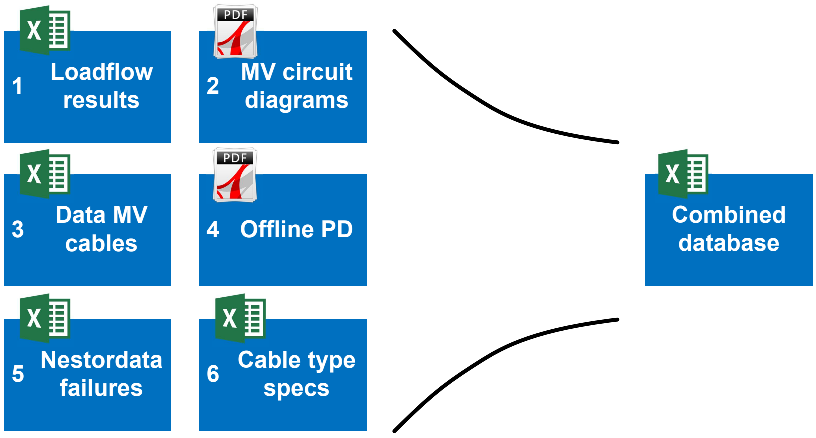


Figure 5.1: Overview of combined data files

As illustrated by figure 5.1, the data of six files in total was combined. This data combination consists of five steps:

1. Combine data file 1 and 2 by manually finding the cable circuit number from the distribution network circuit diagrams.
2. Combine the result of the previous step with file 3 by extracting the commissioning year.
3. Combine this result of the previous step with file 4 by extracting the results of partial discharge measurements
4. Combine this result of the previous step with file 5 by extracting the failures for each cable.
5. Combine this result of the previous step with file 6 by extracting the short circuit current.

For step 3, interpretation of the measurements in the PDF-files was necessary. Table 5.6 gives the three possible PD classification types of the DNV GL distribution cable HI.

Table 5.6: Summary of the PD classification types (based on [80])

Abbreviation	Explanation
Good	All good, no issues
Fair	Minor PDs found, need more intense monitoring
Poor	PDs found: action required, but can be scheduled next maintenance

For this dataset, two different companies have carried out PD measurements and separate reports were available. No direct relation of this information to the DNV GL HI tool was available. Therefore, a relation was constructed. To achieve this, interpretation according to the explanation given in table 5.6 was used. Table 5.7 shows the relation constructed for the types of indication, where the column indicated by a '#' is the condition index number, taken from the measurement reports.

Table 5.7: Interpretation of different PD measurement results.

Type of indication (Company)	#	Explanation	Classification for DNV GL HI: Good	Fair	Poor
Required action (SEBA KMT)	1	Directly replace component			x
	2	Replace component within 1 year			x
	3	New PD measurement within 1 year		x	
	4	No action required within 5 years	x		
Condition code (SEBA KMT)	3	-			x
	4	-		x	
	5	-		x	
	6	New PD measurement within 1 year		x	
	7	New PD measurement within 1 year		x	
	9	New PD measurement within 2-5 years	x		
Priority (KEMA)	1	Dry paper or moisture in insulation			x
	4	No preventive actions necessary	x		

Furthermore, the PD could be classified in two ways, namely global and local. Global PDs are PDs that are found throughout the entire cable circuit. Local PDs are PDs that occur only at single or multiple locations throughout the cable circuit.

For each PD measurement, an expert opinion on the PD occurrence was available in the reports. Only when the report text mentioned that the PDs where found globally, they were interpreted as global PD.

As mentioned in section 7.1.1, it was required to construct load profiles for the week and weekend day. The organisation Energy Data Service the Netherlands (EDSN) provides this type of data [81]. From this data, the average hourly values of the 'E1A' and 'E1B' profiles were taken. These load profiles were normalised to the average estimated worst case load. Figure 5.2 shows the resulting load profiles that were used in the HI.

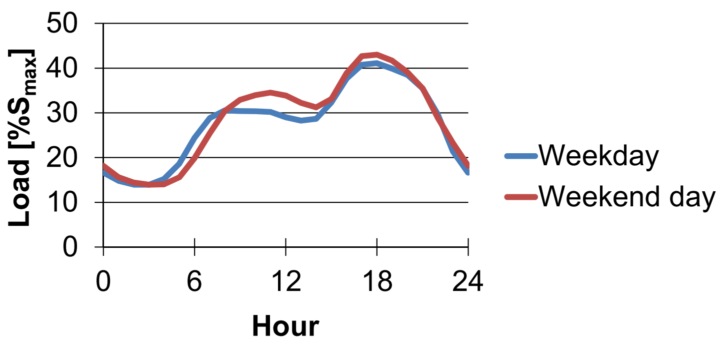


Figure 5.2: Assumed load profiles for the utilities' network

In conclusion, this resulted in an Excel-sheet containing a row for each cable section of a cable circuit. This sheet is presented in appendix C.

5.4.4 HI asset data

For this utility, the data was not yet put in the HI format. Therefore, it was put in the HI format and results for RL were found in 2007. From this data, only the assets with a known commissioning year before 2007 are included. For the distribution cables, the manufacturing year was assumed to be equal to the commissioning year.

Based on the available asset data, subsets of assets were chosen for the analysis:

1. Cable circuits, for which the values for maximum load are available, are selected.
2. Only distribution cables with a known commissioning year before 2007 are selected.
3. For the *XLPE* type, 11 cable circuits are present in the dataset. For the *Mixed* type, 49 cable circuits are present in the dataset. To ensure a homogeneous dataset, only *PILC* type cable circuits are included.

This resulted in a predictions dataset and observations datasets including 886 cable circuits.

Predictions datasets

To create a predictions dataset representative for a standard HI assessment, predictions should originate from the HI based on all available asset data. Therefore, all available asset data was taken into account.

For the assets in the dataset, two main types of asset data can be distinguished: *maximum load* values and *PD* values. For all assets covered in the data, maximum load values were available. For only 18 assets in the dataset, PD values were available. First, the maximum load values in the data are investigated for their relation to failure behaviour. Next, a dummy dataset is used to find which *maximum load* values and *PD* values affect RL predictions of the HI. To conclude, based on the results using the dummy dataset, an additional dataset was constructed by adjusting the load values.

The study presented in [82] suggests that a relation between cable load and failures exists for cable joints. In order to find out to which extent the dataset can show the HI prediction quality based on load values, the relation between maximum load and outages was investigated. Figure 5.3 shows the cable circuits in the dataset. The horizontal axis shows the maximum load in intervals of 10%. The vertical axis shows the number of circuits.

The green coloured bars on top represent the number of circuits in which no failure occurred during the time window (see chapter 3). The red coloured bars on the bottom show the number of circuits for which a failure occurred during the time window. This figure shows that the majority of the circuits experiences a maximum load below 75%.

Figure 5.4 shows, for each load interval, which fraction of the cable circuits have failed during the time window. This figure suggests that for this dataset, a relation between cable circuit load and failure probability is present.

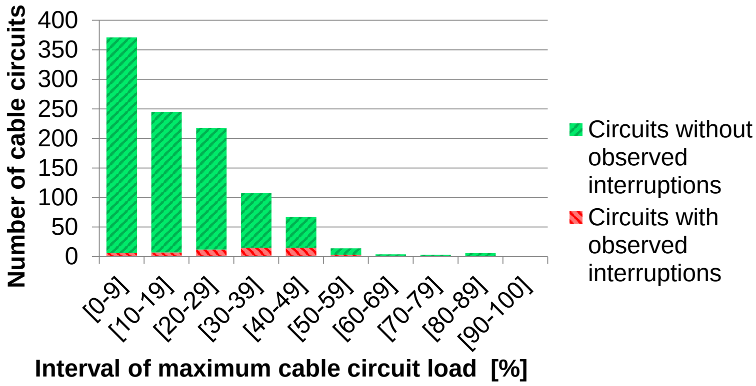


Figure 5.3: Number of cable circuits for each load interval: without observed failures (green); with observed failures (red)

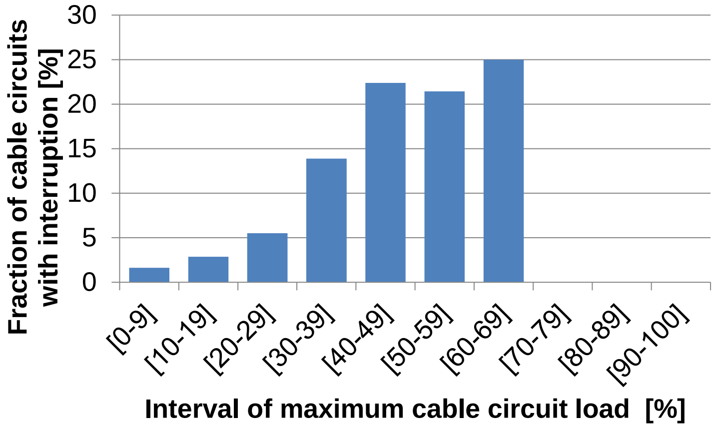


Figure 5.4: Percentage of circuits with failures for each load interval
This figure suggests that a relation between cable load and failures is present.

For the first analysis of the predictions, the load settings as described in section 5.4.3 were applied. The case with these settings is referred to as the *'actual' load* case.

Using the values from the 'actual' load case, the HI was used to find the RLs. Figure 5.5 shows the result for the RL and statistical RL from the 'actual' load case.

In this figure, the horizontal axis shows the asset age at the moment of the HI prediction. The vertical axis of this figure shows the predicted RL. In this graph, two types of predictions are presented:

1. RL predictions for the 'actual' load case

2. RL predictions for the case using only the statistical remaining lifetime function (see section 2.1.1)

As it appears for figure 5.5, the RL predictions of the 'actual' load (blue coloured triangles) overlap exactly with the statistical RL predictions (orange coloured dots), except for one prediction, which has a predicted remaining lifetime of one year. Inspection of this prediction shows that one PD measurement with value 'Poor' has resulted into this predicted RL of one year.

The similarity between the RL predictions of the 'actual' load and the statistical RL predictions suggests that RL predictions using the 'actual' load values are insufficiently taking into account the effect of cable circuit loading on the predicted RL.

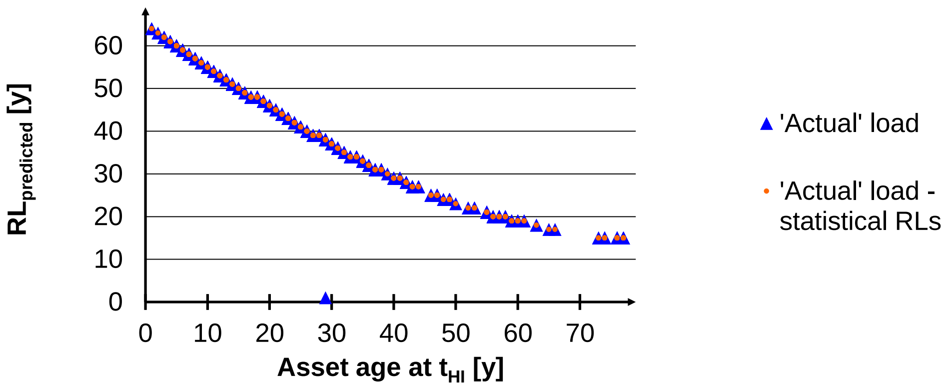


Figure 5.5: RLs of the predictions dataset under 'actual' load compared to the *statistical* RLs of this dataset

Prior to preparing the predictions dataset, a dummy dataset was prepared to determine in what way the condition parameters influence the RL. Figures 5.6a and 5.6b show the results from application of this dummy dataset in the DNV GL distribution cable HI. In both figures, the horizontal axis on the left-hand side shows the *age* at t_{HI} and the vertical axis shows the *predicted RL* (RL_{pred}).

In figure 5.6a, the horizontal axis on the right-hand side shows the *PD value*. In figure 5.6b, the horizontal axis on the right-hand side shows the *maximum load*.

The results presented in figure 5.6a show that for the PD values *fair* and *good*, the same RL_{pred} results for each *age* at t_{HI}. For the PD value *poor*, the RL_{pred} equals one year. This shows that the influence of the PD value for cables with an age of *4* to *64* years is only visible for the PD value *poor*.

The results presented in figure 5.6b show that for the maximum load values *0* through *0.75*, the same RL_{pred} results for each *age* at t_{HI}. For the maximum load value *1*, the RL_{pred} is lower for each *age* at t_{HI}. This shows influence of cable load for cable circuits in this case study with an age of *4* to *64* years only *above* a maximum load value of *75%*.

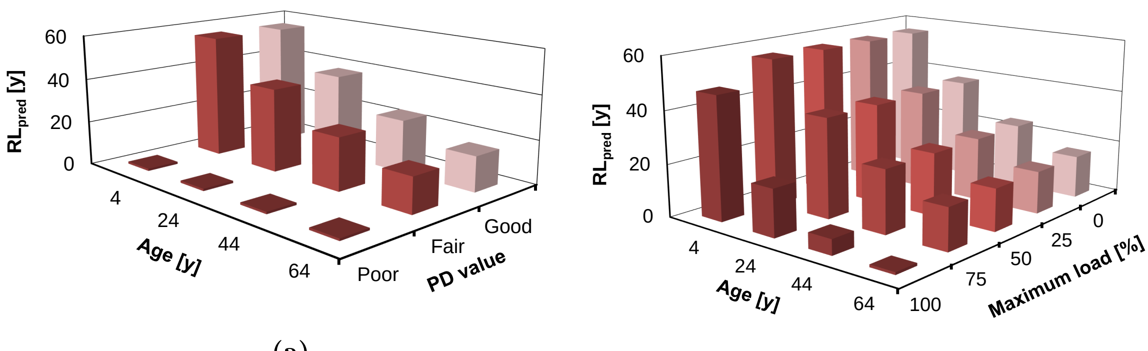


Figure 5.6: Results using a dummy dataset for finding the influence on predicted remaining lifetime of: (a) the PD value; (b) the maximum load value
This figure shows, for cable circuits in this case study with an age of 4 to 64, that: (a) only the PD value 'poor' influences predicted RL; (b) only maximum load values above 75% influence the predicted RL

Only load values from *75%* to *100%* influence the RL_{pred} and figure 5.3 showed that the majority of the cable circuits experiences a maximum load value below *75%*. To investigate the effect of load on the results, two datasets were prepared, based on the settings for maximum load:

1. 'Actual' maximum load dataset
2. 'Tweaked' maximum load dataset

For the 'actual' maximum load dataset, the settings as described in section 5.4.3 were applied.

For the 'tweaked' maximum load dataset, two settings of the 'actual' maximum load dataset were changed:

1. 'Tweaked' maximum load [%] = 3/4 + 1/4 × 'actual' maximum load [%]
2. Load profile (weekday and weekend day) = 100% continuous for each hour

The result after applying these settings is shown in figure 5.7. In this figure, the predictions for the 'tweaked' load (orange coloured dots) show that, when looking at different ages, the predicted remaining lifetime differs from the 'actual' load (blue coloured triangles). This shows that the predictions dataset resulting from the settings for 'tweaked' load, clearly includes the effect of ageing by load.

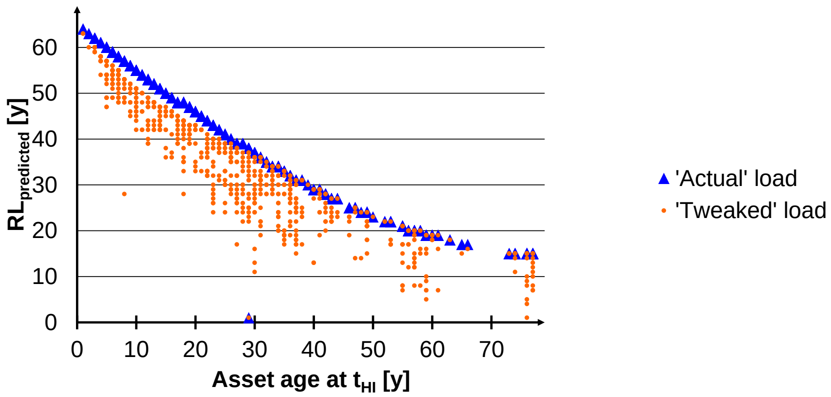


Figure 5.7: RLs of the predictions dataset under 'actual' load compared to the RLs under 'tweaked' load.

Observations dataset

As shown in [76], the topology of cable circuits in distribution cable networks changes over time. Cable joints or cable sections that (have been predicted to) fail are in most cases replaced by a new cable section and two new joints. These continuous changes in topology cause the reliability of cable circuits, apart from its ageing-related changes, to change over time. These changes can influence the observed reliability, which follows from the observations dataset, both positively and negatively. A higher reliability can be observed due to the new cable sections and joints in the normal life period that have survived the period of infant mortality (see section 4.2.1). However, the new cable sections and joints that are in the period of infant mortality can introduce a lower reliability.

Interruptions in the observations dataset originate from the cable circuits present in the failure database that experienced an interruption during the period 2007-2014. For this analysis, the failure moment is taken equal to the first observed interruption of the cable circuit. As an interruption in a cable circuit is not a direct indication for end-of-life for this cable circuit, this assumption is pessimistic and therefore, on the safe side. This has a negative influence on the observed reliability, caused by the observations dataset.

Suspensions in this dataset are the cable circuits, for which no interruption has occurred during the time window.

Chapter 6

Case study: DNV GL instrument transformer Health Index

This chapter aims to show an application of the methods presented in chapter 3 using data of HV instrument transformers.

In this particular case study, it will be the application of method 5, life data analysis, as described in section 3.5.5. The case study focuses on instrument transformers (ITs) of a Dutch TSO using the DNV GL instrument transformer HI.

For the ease of reading, in this chapter, the DNV GL IT HI will be denoted as just HI.

The chapter is divided into five parts. Firstly, in section 6.1, an explanation is given of the items in this HI required for the case study. Secondly, section 6.2 presents the assumptions for this case study and shows the hypothesis regarding the outcome of this case study. Thirdly, section 6.3 shows the results of the case study. To conclude, sections 6.4 and 6.5 present the conclusions and recommendations.

6.1 DNV GL IT HI

This section describes the HI version used for this case study. Firstly, section 6.1.1 presents the available asset data and the settings. Secondly, section 6.1.2 explains the remaining lifetime definition. Thirdly, section 6.1.3 shows the used HI output representation. To conclude, section 6.1.4 presents the software and hardware used by the HI.

6.1.1 Asset data

This section describes the asset data the HI can use to perform a HI assessment. As described in section 2.1.1, even when part of this data is missing, it is still possible to determine asset health. However, section 2.1.1 also showed that the amount of data that is available for such calculation has a clear influence on the uncertainty in the output.

Table 6.1 provides an overview of asset data types used by this HI. Parameters #1 through #7 identify the individual asset.

Table 6.1: Asset data types for the DNV GL IT HI

Category	#	Parameter
Asset identification	1	Substation
	2	Section
	3	Component identification number
	4	Manufacturer
	5	Type
	6	Commissioning year
	7	Rated voltage

6.1.2 HI remaining lifetime definition

As already introduced in section 3.3, different definitions are used among the HIs for the remaining lifetime (RL) of assets. For the HI of this case study, the RL is defined as the period, starting from the predictions moment, during which at most *1%* of the assets with the same values for asset data is expected to fail [83].

6.1.3 HI output representation used for the comparison

As described in section 2.1.1, different output representations exist for the HI. For this HI version, asset health is represented by a *RL*, a *colour code* and an *intensity*. From these three values, the *HI prediction quality quantification methods* only use RL. Therefore, the representations used for this case study are limited to the values for RL.

6.1.4 Software and hardware

This HI is based on a spreadsheet in Microsoft Excel 2003 using Visual Basic for Applications (VBA) code. This HI uses the number of available asset data types to determine uncertainty. The value of t_{HI} was set to 2009 in the Excel sheet.

The calculations where performed on multiple computers. Using a computer equipped with an Intel i7 processor, the computational time was approximately 20 minutes to find the HI for around 13000 assets.

6.2 Assumptions and hyphothesis

The hypothesis for this case study is summarized in two main points:

1. For this case study, the assumptions on the utility data, as shown in chapter 5 are used.
2. Using the life data analysis method, it is possible to show the validity of the DNV GL IT HI.

6.3 Results

This section describes the results of the case study for the DNV GL IT HI using method 5, life data analysis (see section 3.5.5). As described in section 3.2.1, two datasets are necessary for *HI prediction quality quantification* and *validation*: predictions and observations. In the life data analysis, the predicted RLs were translated to predicted failures and predicted suspensions. The observed lifetimes were translated to observed failures and observed suspensions. For this analysis 90% confidence bounds (see section 4.4.1) were applied.

Sections 6.3.1 and 6.3.2 show the results of the failure distribution fitting for the predictions and observations datasets. Section 6.3.3 shows the comparison of these curves for the *HI prediction quality quatification* and *validation* of this HI.

6.3.1 Predictions

Figure 6.1 shows the result of life data analysis for the predictions dataset. In this figure, the horizontal axis shows the asset age on a logarithmic scale. The vertical axis shows the probability of failure, also on a logarithmic scale. For this analysis, both predicted failures and predicted suspensions were taken into account (see section 3.5.5).

The blue dots in this figure show the *data points* of failures, the blue triangles show the *suspension* points. The striped dark blue line is the probability line. The striped light blue lines are the 90% confidence bounds. These three lines almost follow the same line, showing narrow confidence bounds. The majority of the data points are outside these confidence bounds, i.e. the majority of the data points are so-called *outliers* (see section 4.4). This large number of outliers shows that the data-quality is low. Therefore, the data is not representative for the predicted failure behaviour of the assets and cannot be used in drawing a firm conclusion on the validity of this HI.

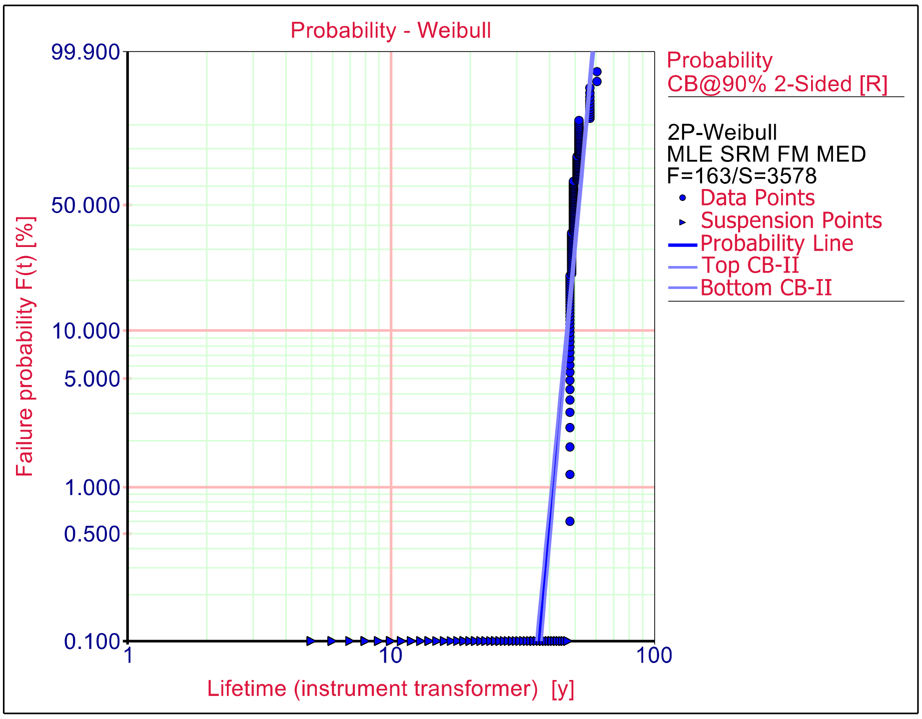


Figure 6.1: 2P Weibull fit on the predictions dataset assuming both predicted failures and predicted suspensions

6.3.2 Observations

Figure 6.2 shows the result of life data analysis for the predictions dataset. In this figure, like figure 6.1, the horizontal axis shows the asset age on a logarithmic scale. The vertical axis shows the probability of failure, also on a logarithmic scale. For this analysis, both observed failures and observed suspensions were taken into account (see section 3.5.5).

Like in figure 6.1, the green dots in this figure show the *data points* of failures, the green triangles show the *suspension* points. The striped dark green line is the probability line. The striped light green lines are the 90% confidence bounds. Only one of these points is outside confidence bounds, i.e. only one *outlier* is present in the dataset (see section 4.4). This would suggest the data quality to be high. However, only eight failures are present in the dataset, which explains the wide confidence bounds. As stated in 5.1.1, if possible, a minimum of ten failures should be obtained to describe the failure behaviour of the assets. Therefore, this utility data cannot be used in drawing a firm conclusion on the validity of this HI.

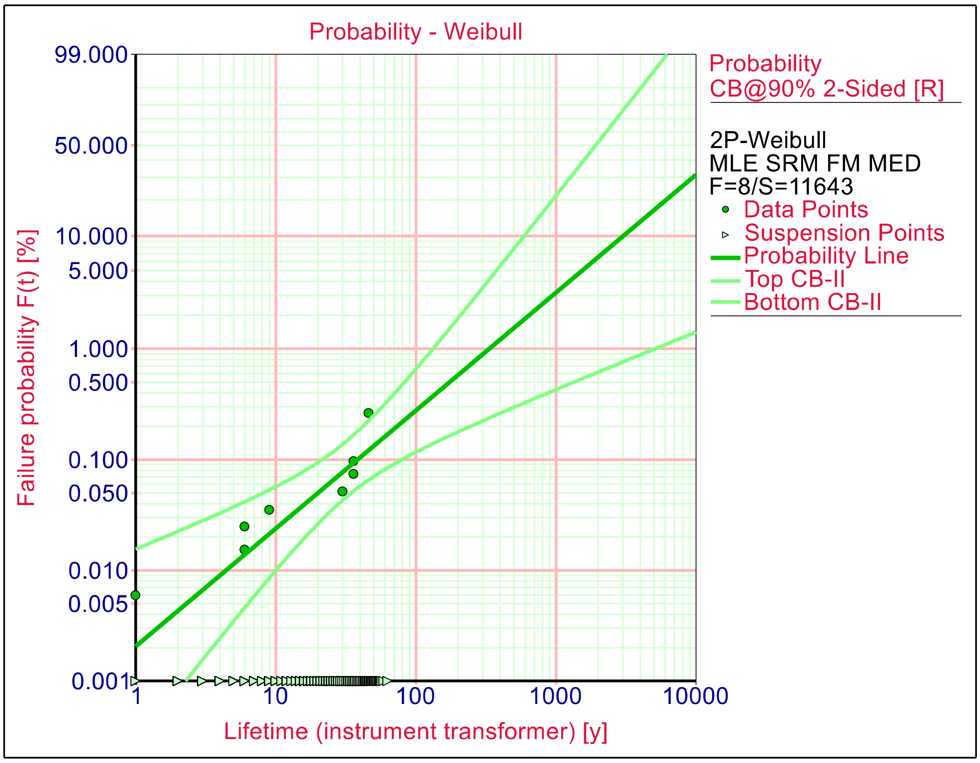


Figure 6.2: 2P Weibull fit on the predictions dataset assuming both predicted failures and predicted suspensions

6.3.3 HI prediction quality quantification and validation by predictions and observations comparison

This section shows the results of the *HI prediction quality quantification* and *validation* by comparing the life-curves shown in the previous sections.

The comparison between predictions and observations is done using method 5 as explained in section 3.5.5. Figure 6.3 shows the life data analysis for the predictions (in blue colour) and observations (in green colour) of figures 6.1 and 6.2 in a single graph.

The horizontal striped line points towards the average predicted lifetime for this group of assets. As shown in section 6.1.2, this lifetime equals the lifetime for which *at most* 1% of the assets have failed. The first row of table 6.2 shows this lifetime, including its confidence bounds.

The vertical arrow points towards the observed failure probability for this lifetime. The second row of table 6.2 shows the observed failure probability for this lifetime, including its confidence bounds.

Table 6.2: Derivation of the observed failure probability for the average predicted lifetime

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq$ 1%)	[y]	51.2	51.6	51.9
2	Observed failure probability for predicted lifetime	[%]	0.26	0.14	0.07

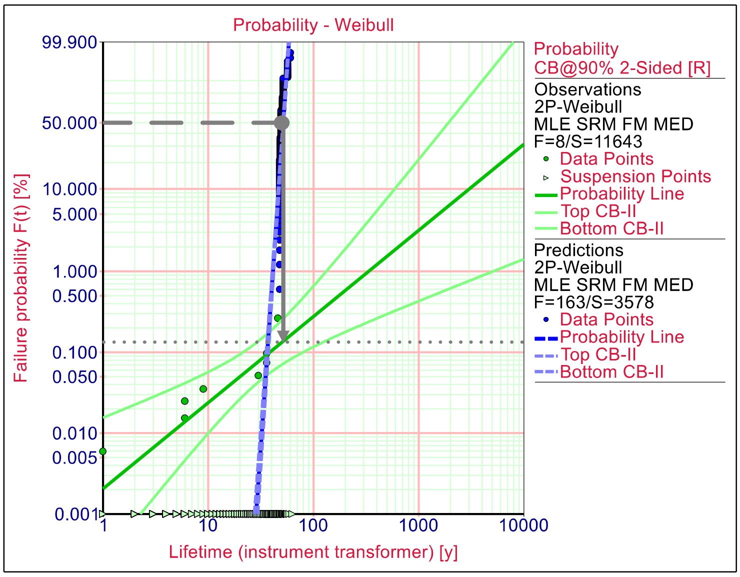


Figure 6.3: Comparing the results of the dataset of predictions to the results of the observations dataset (90% confidence bounds)

Figure 6.4 is a copy of figure 6.3, adding elements to show *validation* and *HI prediction quality*.

The *red box* in top of this figure shows the area, in which an observed failure probability for the predicted lifetime shows that the HI prediction is *non-valid*.

Conversely, the *green box* on the bottom of this figure shows the area, in which an observed failure probability for the predicted lifetime shows that the HI prediction is *valid*. For this HI, the observed failure probability for the predicted lifetime is in the green box. Therefore, in this case the HI prediction is considered *valid*, as shown in the third row of table 6.3.

The three arrows in the right of this figure show the distance of the observed failure probability for the predicted failure probability including the 90% confidence bounds. This *distance* (in percent) is inversely proportional to the *HI prediction quality* and is defined by:

$$HI\ prediction\ quality = 1 - \frac{distance\ [\%]}{5\%} \tag{6.1}$$

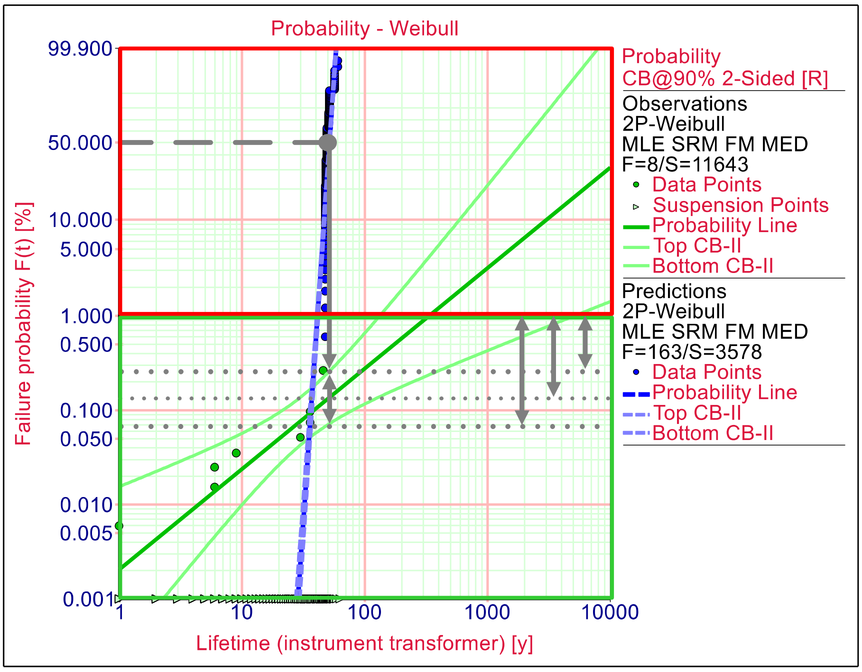


Figure 6.4: Comparing the results of the dataset of predictions to the results of the observations dataset (90% confidence bounds)

Table 6.3: Derivation of the observed failure probability for the average predicted lifetime

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq$ 1%)	[y]	51.2	51.6	51.9
2	Observed failure probability for predicted lifetime	[%]	0.26	0.14	0.07
3	Validity predicted lifetime	{Valid, Non-valid}	Valid	Valid	Valid
4	HI prediction quality	(0-1)	0.26	0.14	0.07

6.4 Conclusions

From this case study, the following conclusions can be drawn:

1. This case study has shown the application of a method for quantifying prediction quality of a HI and validation of a HI. The hypothesis, that it is possible to quantify HI prediction quality and to validate this HI (see section 6.2) is therefore not rejected by this case study.

2. Both for distribution fitting of the predictions and observations datasets in this case study, the data quality was too low to draw firm conclusions on the validity of this HI.

3. No firm conclusion can be drawn regarding the finding of this case study on the DNV GL instrument transformer HI that the predictions were found to be *too optimistic* and therefore *non-valid*. This finding followed from the value of the mean observed age, which is higher than the value of the mean predicted age.

6.5 Recommendations

For this case study, no condition data was available at t_{HI}. Therefore, only the *HI prediction quality* of the statistical assessment functions was determined for this HI for *validation* of this HI. To enable better *HI prediction quality quantification* for this HI at a moment in the future, the following is recommended:

1. For each failed asset, the utility should link the asset's serial number to the asset in the HI.

2. The *time window* should cover a time period during which sufficient failures have occurred. When at least ten failures are considered the minimum and when extrapolating the number of failures in the time window of this dataset, the minimum time window would be 6 years wide, allowing for an analysis in 2015. For this specific case study, it was found that majority of the condition indicators were available starting from 2013. Therefore, a failure database covering failures until at least the year 2015 are necessary for this dataset.

Chapter 7

Case study: DNV GL distribution cable Health Index

This chapter describes the case study for a population of 10 kV distribution cables.

The chapter is divided into five parts. Firstly, in section 7.1, an explanation is given of the items in this HI required for the case study. Secondly, section 7.2 presents the assumptions for this case study and shows the hypothesis regarding the outcome of this case study. Thirdly, section 7.3 shows the results of the case study. To conclude, sections 7.4 and 7.5 present the conclusions and recommendations.

7.1 DNV GL distribution cable HI

This section describes the HI version used for this case study. Firstly, section 7.1.1 presents the available asset data and the settings. Secondly, section 7.1.2 explains the remaining lifetime definition. Thirdly, section 7.1.3 shows the used HI output representation. To conclude, section 7.1.4 presents the software and hardware used to fin the HI.

7.1.1 Asset data and settings

This section describes the asset data and settings the HI uses to perform a HI assessment. As described in section 2.1.1, even when part of this data is missing, it is still possible to determine asset health. Notwithstanding, the amount of data that is available for such calculation has a clear influence on the reliability of the assessment.

Table 7.1 provides an overview of asset data types used by the DNV GL distribution cable HI that were available for this case study. Appendix C shows an extended version of this table, including the parameters of this HI that were not available for the case study. Parameters #1 through #3 identify the individual cable circuit. Based on the cable type(s), each cable circuit belongs to a group of cables with particular failure behaviour. From this information, the statistical remaining lifetime is calculated. The batch number selects a distribution which represents the failure behaviour of the cable circuit.

Parameters #4 through #18 contain more detailed information on the cable circuit. Appendix C explains these parameters in more detail.

Table 7.1: Summary of the main available HI asset data

Category	#	Parameter	Unit
Asset identification	1	Source	[-]
	2	Destination	[-]
	3	Batch number	[#]
General information	4	Maintenance section	[-]
	5	Cable insulation main cable	{XLPE, PILC, Mixed}
	6	Commissioning year	[yyyy]
Design details	7	Rated current	[A]
	8	Rated short-circuit current (1 s)	[A]
	9	Number of cores per cable	[#]
	10	Earthing	{Single, Double, CrossB}
	11	Waterblocking screen/conductor	{Yes,No}
Utilization data	12	Max loading	[MVA]
	13	Historical maximum loading	[%]
	14	Number of short circuit currents since last maintenance	[#]
Partial discharge	15	Date of test	[dd-mm-yy]
	16	Test result	{Normal, Fair, Poor}
	17	PD occurrence	{Local, Global}
Failure data	18	Number of spontaneous failures in cable	[#]

In addition to the asset data, this HI allows for application of more general settings that apply to (groups of) the dataset. Table 7.2 provides an overview of these settings. The cells that describe asset data that was not available or applicable for the case study are marked grey.

Items #1 through #3 are the ageing related settings. Items #4 through #6 are defined for each batch. Together, they define which distribution type and parameters are used to derive the statistical RL. Item #7 contains minimum values for dissolved gas analysis concentrations. Item #8 contains two hourly load profiles: one for modelling the weekdays and one for modelling the weekend days.

Table 7.2: Summary of the main HI parameters
Items coloured grey were not available or applicable.

Category	#	Parameter	Unit
Ageing	1	Load growth factor	[%/y]
	2	Maximum leakage (oil-filled cables)	[L/month]
	3	Minimum degree of polymerization (paper-based cables)	[-]
Batch	4	Distribution type	[-]
	5	Distribution parameter 1	[-]
	6	Distribution parameter 2	[-]
DGA parameters	7	Minimum follow-up concentrations	[ppm]
Load profile	8	Weekday,weekend day (24 values)	[%]

7.1.2 HI remaining lifetime definition

As already introduced in section 3.3, different definitions are used among the HIs for the remaining lifetime (RL) of assets. For the HI of this case study, the RL is defined as the period, starting from the predictions moment, during which at most *5%* of the assets with the same values for asset data is expected to fail [80].

7.1.3 HI output representation used for the comparison

As described in section 2.1.1, different output representations exist for the HI. For this HI version, asset health is represented by a *RL*, a *time to additional maintenance*, a *colour code* and an *intensity*. From these three values, the *HI prediction quality quantification methods* only use RL. Therefore, the representations used for this case study are limited to the values for RL.

7.1.4 Software and hardware

This HI is based on a spreadsheet in Microsoft Excel 2003 using Visual Basic for Applications (VBA) code. This HI uses the number of available asset data types to determine uncertainty. The value of t_{HI} was set to 2009 in the Excel sheet.

The calculations where performed on multiple computers. Using a computer equipped with an Intel i7 processor, the computational time was approximately 20 minutes to find the HI for around 13000 assets.

7.2 Assumptions and hyphothesis

For this case study, the assumptions on the utility data, as shown in section 5.4, are used. The following hypothesis is formulated:

> Using the life data analysis method, it is possible to test the validity of the DNV GL distribution cable HI.

7.3 Results

This section describes the results of the case study for the DNV GL distribution cable HI using method 5, life data analysis (see section 3.5.5). As described in section 3.2.1, two datasets are necessary for *HI prediction quality quantification* and *validation*: predictions and observations. In the life data analysis, the predicted RLs were translated to predicted failures and predicted suspensions. The observed lifetimes were translated to observed failures and observed suspensions. For this analysis 90% confidence bounds (see section 4.4.1) were applied.

Sections 7.3.1 and 7.3.2 show the results of the failure distribution fitting for the predictions and observations datasets. Section 7.3.3 shows the comparison of these curves for the *HI prediction quality quantification* and *validation* of this HI.

7.3.1 Predictions

Predictions 'actual' load

Figure 7.1 shows the result of life data analysis for the predictions dataset of the 'actual' load case, as shown in section 5.4.4. In this figure, the horizontal axis shows the asset age on a logarithmic scale. The vertical axis shows the probability of failure, also on a logarithmic scale. No failures were predicted in the time window (see section 3.5.5). Including suspensions would lead to only predicted suspensions. As for failure distribution fitting, failures are necessary, only predicted failures were taken into account.

The blue dots in this figure show the *data points* of failures. The striped dark blue line is the probability line. The striped light blue lines are the 90% confidence bounds. These three lines almost follow the same line, showing narrow confidence bounds. The majority of the data points are outside these confidence bounds, i.e. the majority of the data points are so-called *outliers* (see section 4.4). This large number of outliers suggests that the data-quality is low.

Besides, the distribution does not properly fit the data. Using a distribution different from the 2P Weibull distribution would possibly result in a slightly better fit, but still a large part of the data points would consist of outliers. This large number of outliers would still suggest that the data-quality is low. As shown in section 3.5.5, only the lifetime for 50% failure probability is used from this probability line. By visual inspection, a slightly better distribution fit would lead to a slightly different lifetime for 50% failure probability with still a low data quality.

Therefore, the data is not representative for the predicted failure behaviour of the assets and cannot be used in drawing a firm conclusion on the validity of this HI.

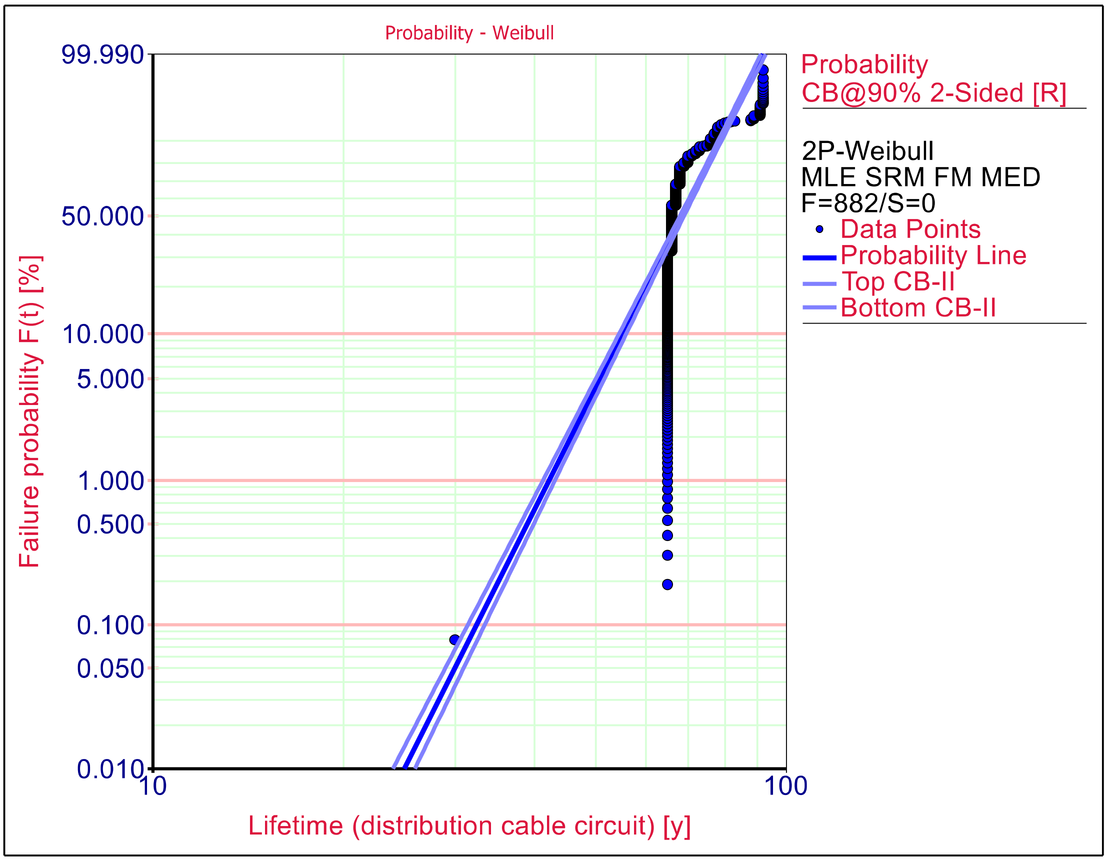


Figure 7.1: 2P Weibull fit on the predictions dataset of 'actual' load, assuming only predicted failures

Predictions 'tweaked' load

Figure 7.2 shows the result of life data analysis for the predictions dataset of the 'tweaked' load case, which was prepared to include the effect of cable circuit load on failure behaviour (see section 5.4.4). In this figure, the horizontal axis shows the asset age on a logarithmic scale. The vertical axis shows the probability of failure, also on a logarithmic scale. For this analysis, both predicted failures and predicted suspensions were taken into account (see section 3.5.5).

The cyan blue dots in this figure show the *data points* of failures, the cyan blue triangles show the *suspension* points. The striped cyan blue line is the probability line. The cyan light blue lines are the 90% confidence bounds. These three lines almost follow the same line, showing narrow confidence bounds. Some of the data points are outside these confidence bounds, i.e. the majority of the data points are so-called *outliers* (see section 4.4). These three data points could be removed. However, this would leave only eight failures. By visual inspection, a distribution fit excluding these outliers would lead to a slightly different lifetime for 50% failure probability.

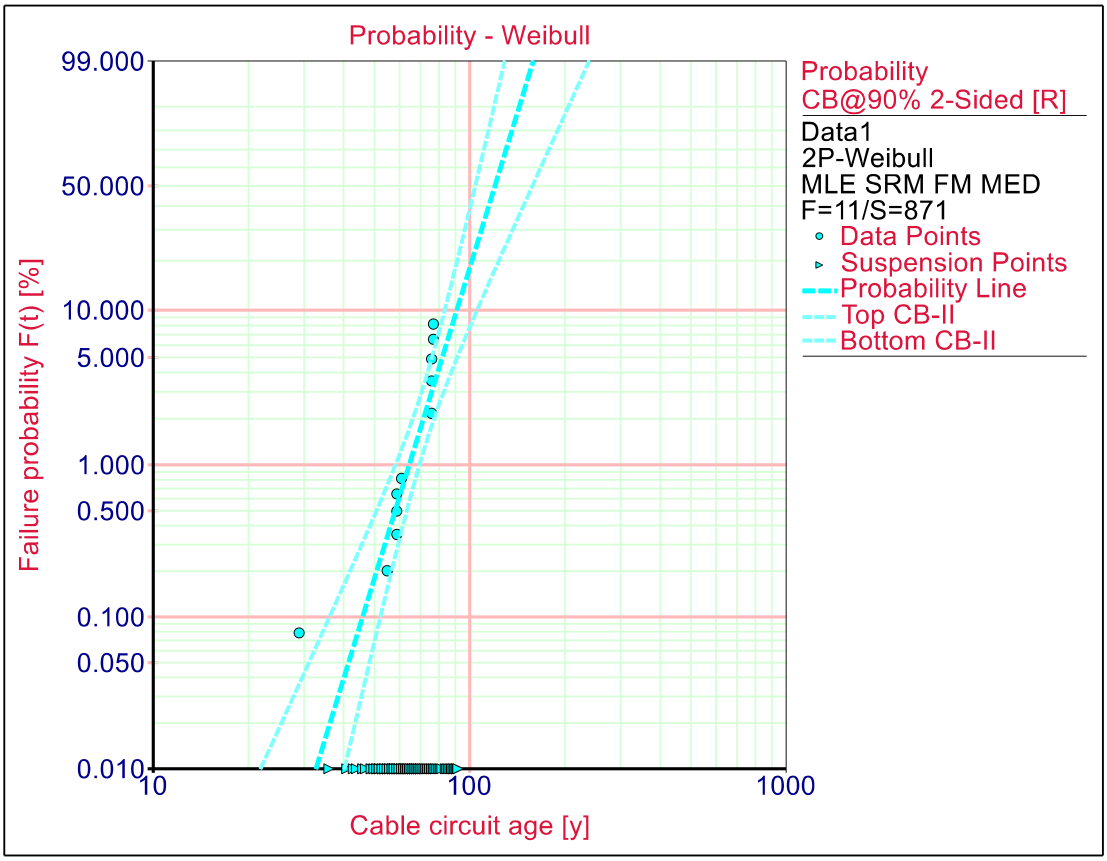


Figure 7.2: 2P Weibull fit on the predictions dataset of 'tweaked' load, assuming both predicted failures and predicted suspensions

7.3.2 Observations

Figure 7.3 shows the result of life data analysis for the predictions dataset. In this figure, like in figures 7.1 and 7.2, the horizontal axis shows the asset age on a logarithmic scale. The vertical axis shows the probability of failure, also on a logarithmic scale. For this analysis, both observed failures and observed suspensions were taken into account (see section 3.5.5).

Like in figures 7.1 and 7.2, the green dots in this figure show the *data points* of failures, the green triangles show the *suspension* points. The striped dark green line is the probability line. The striped light green lines are the 90% confidence bounds. These three lines almost follow the same line, showing narrow confidence bounds. A large number of the data points are outside these confidence bounds, i.e. a large number of the data points are so-called *outliers* (see section 4.4). This large number of outliers shows that the data-quality is low.

Furthermore, changes in the topology of cable circuits (see section 5.4.4) can influence the observed reliability, which follows from the observations dataset, both positively and negatively. Next, the definition of a failure for cable circuits has a negative influence on the observed reliability. Both the changes in topology as the failure definition can influence the results of this case study significantly.

Therefore, the data is assumed to be not representative for the observed failure behaviour of the assets and cannot be used in drawing a firm conclusion on the validity of this HI.

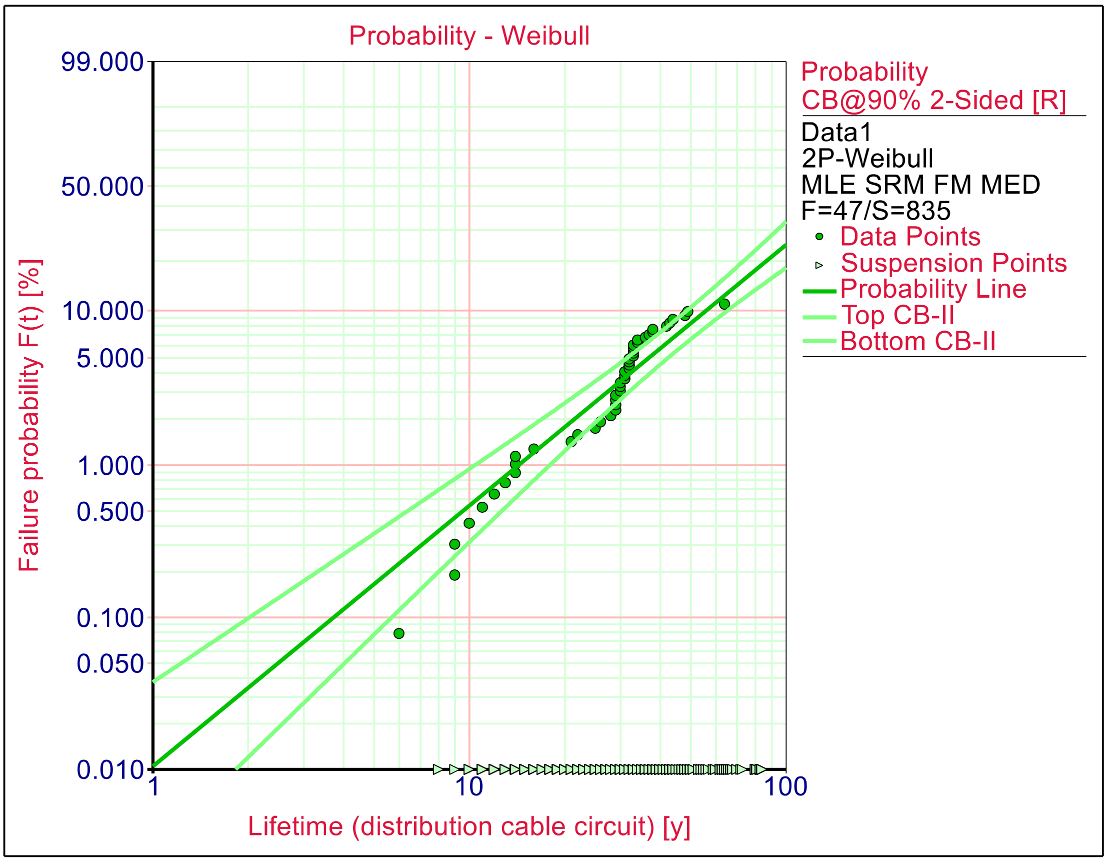


Figure 7.3: 2P Weibull fit on the observations dataset assuming both observed failures and observed suspensions

7.3.3 HI prediction quality quantification and validation by predictions and observations comparison

This section shows the results of the *HI prediction quality quantification* and *validation* by comparing the life-curves shown in the previous sections.

The comparison between predictions and observations is done using *life data analysis*, as explained in section 3.5.5. Firstly, the dataset of 'actual' load predictions is compared to the observations. Secondly, the dataset of 'tweaked' load predictions is compared to the observations.

Comparison predictions 'actual' load to observations

Figure 7.4 shows the life data analysis for the predictions (in blue colour) and observations (in green colour) of figures 7.1 and 7.3 in a single graph.

The horizontal striped line points towards the average predicted lifetime for this group of assets. As shown in section 7.1.2, this lifetime equals the lifetime for which *at most* 5% of the assets have failed. The first row of table 7.3 shows this lifetime, including its confidence bounds.

The vertical arrow points towards the observed failure probability for this lifetime. The second row of table 7.3 shows the observed failure probability for this lifetime, including its confidence bounds.

Table 7.3: Derivation of the observed failure probability for the average predicted lifetime

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq 5\%$)	[y]	67.9	68.4	68.4
2	Observed failure probability for predicted lifetime	[%]	10.6	13.8	17.7

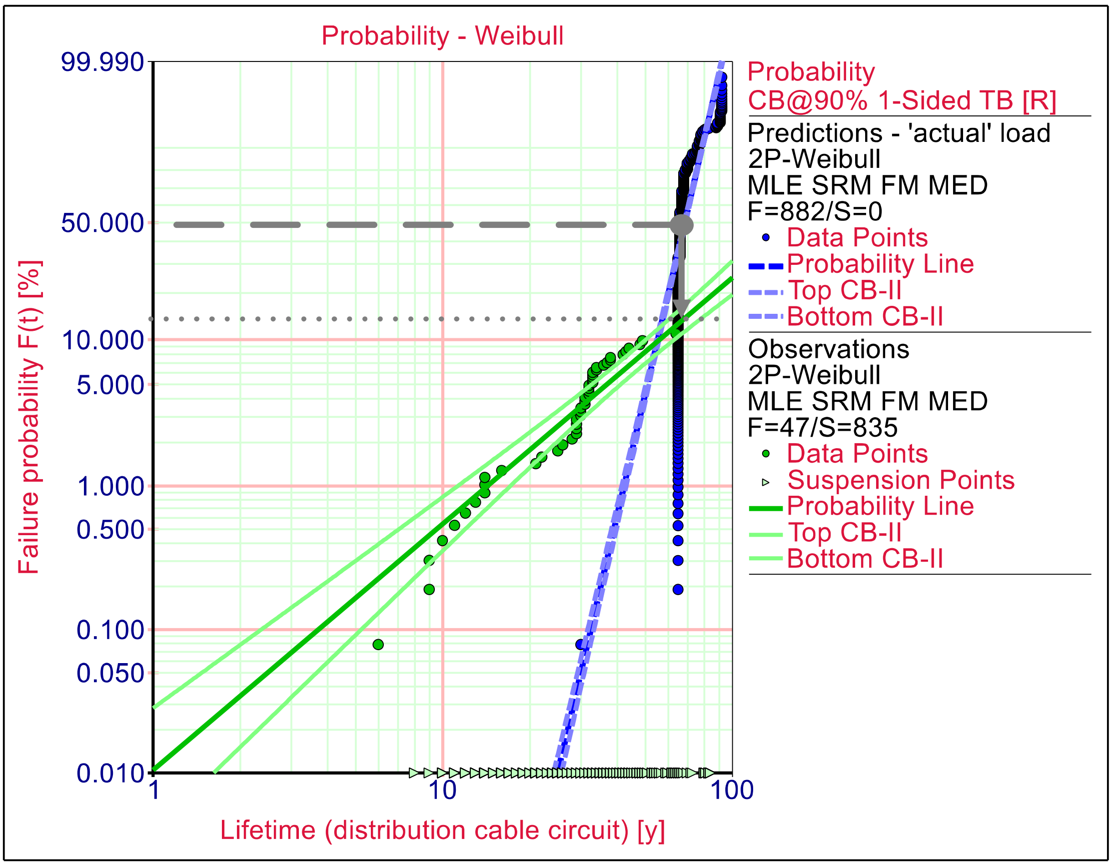


Figure 7.4: Comparing the results of the dataset of predictions of 'actual' load to the results of the observations dataset (90% confidence bounds)

Figure 7.5 is a copy of figure 7.4, adding elements to show *validation* and *HI prediction quality.*

The *red box* in top of this figure shows the area, for which an observed failure probability for the predicted lifetime shows that the HI prediction is *non-valid.*

Conversely, the *green box* on the bottom of this figure shows the area, in which an observed failure probability for the predicted lifetime shows that the HI prediction is *valid.* For this HI, the observed failure probability for the predicted lifetime is in the *red box.* Therefore, in this case the HI prediction is considered *non-valid*, as shown in the third row of table 7.4. This non-valid HI prediction, cause the *HI prediction quality* to be equal to zero, which is shown in the fourth row of table 7.4.

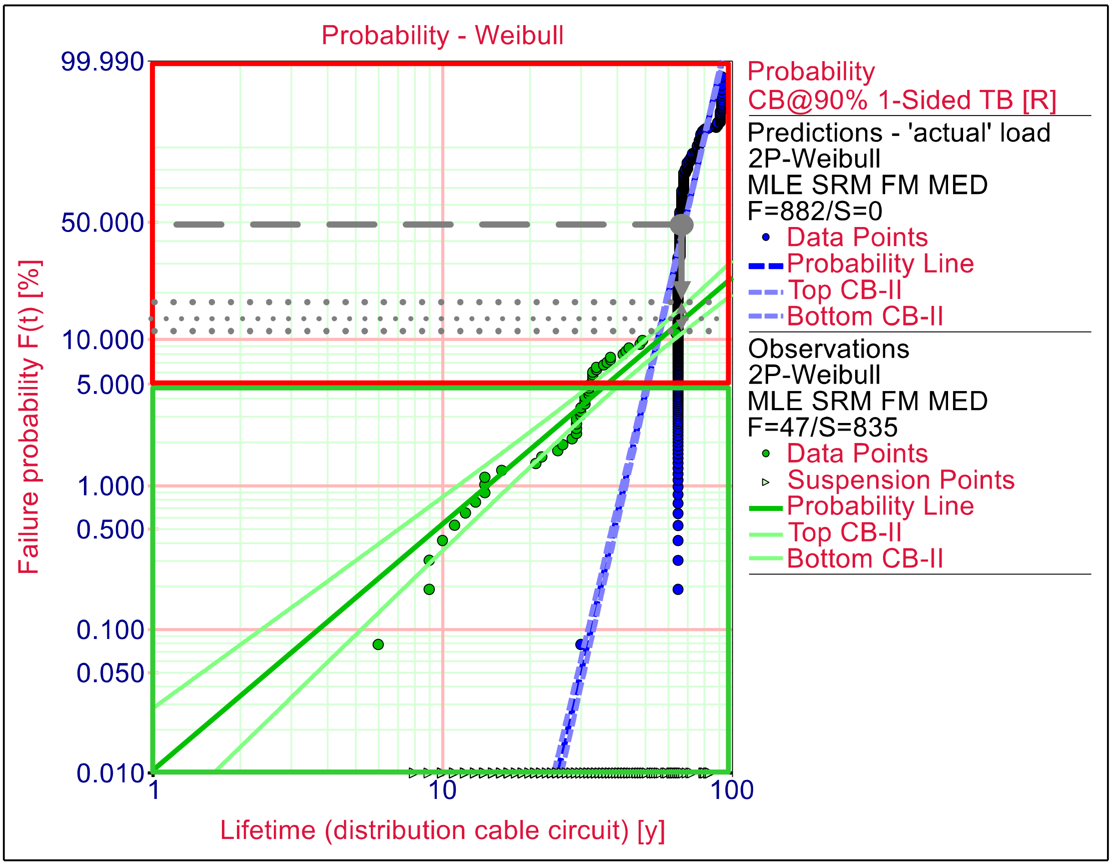


Figure 7.5: Derivation of validity and HI prediction quality

Table 7.4: Derivation of the observed failure probability for the average predicted lifetime

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq$ 5%)	[y]	67.9	68.4	68.4
2	Observed failure probability for predicted lifetime	[%]	10.6	13.8	17.7
3	Validity predicted lifetime	{Valid, Non-valid}	Non-valid	Non-valid	Non-valid
4	HI prediction quality	(0-1)	0	0	0

Comparison predictions 'tweaked' load to observations

Figure 6.3 shows the life data analysis for the predictions (in blue colour) and observations (in green colour) of figures 7.1 and 7.3 in a single graph.

The horizontal striped line points towards the average predicted lifetime for this group of assets. As shown in section 7.1.2, this lifetime equals the lifetime for which *at most* 5% of the assets have failed. The first row of table 7.5 shows this lifetime, including its confidence bounds.

The vertical arrow points towards the observed failure probability for this lifetime. The second row of table 7.5 shows the observed failure probability for this lifetime, including its confidence bounds.

Table 7.5: Derivation of the observed failure probability for the average predicted lifetime

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq$ 5%)	[y]	99.7	120.2	144.8
2	Observed failure probability for predicted lifetime	[%]	18.2	32.4	56.9

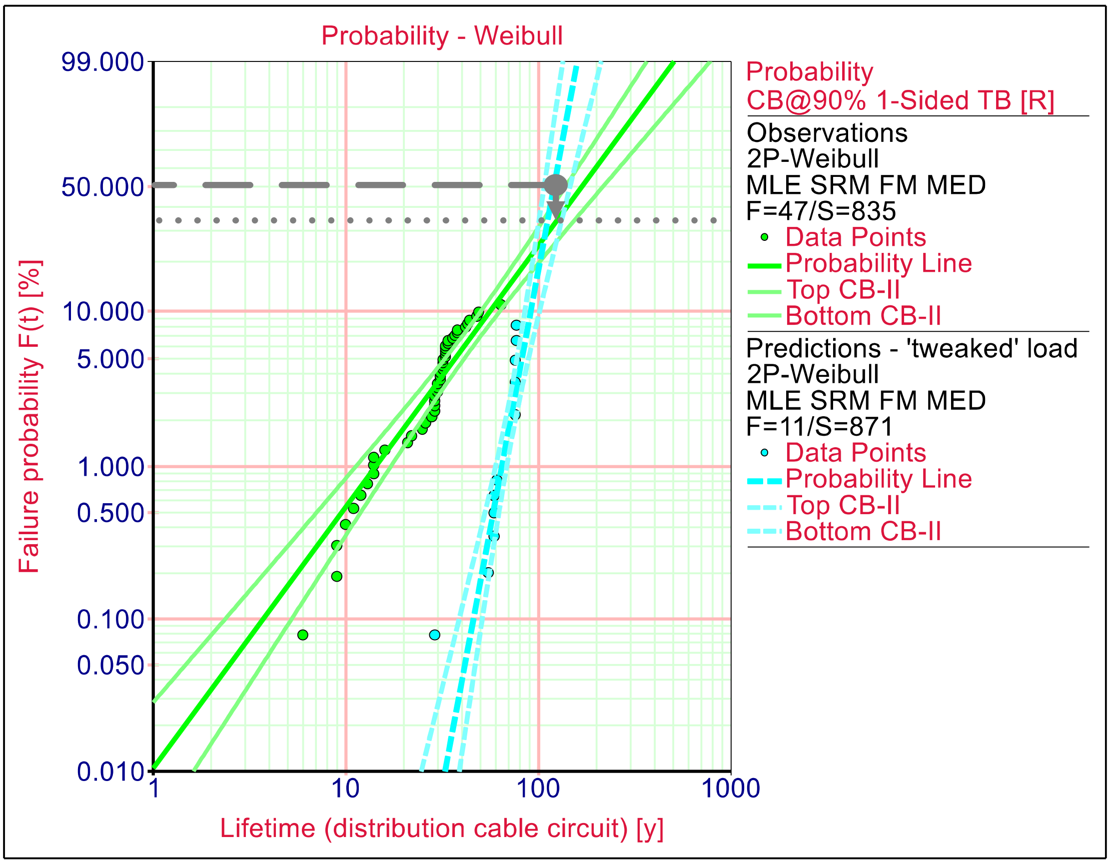


Figure 7.6: Comparing the results of the dataset of predictions of 'actual' load to the results of the observations dataset (90% confidence bounds)

Figure 7.7 is a copy of figure 7.6, adding elements to show *validation* and *HI prediction quality*.

The *red box* in top of this figure shows the area, for which an observed failure probability for the predicted lifetime shows that the HI prediction is *non-valid*.

Conversely, the *green box* on the bottom of this figure shows the area, for which an observed failure probability for the predicted lifetime shows that the HI prediction is *valid*. For this HI, the observed failure probability for the predicted lifetime is in the *red box*. Therefore, figure 7.7

suggests that the HI prediction is *non-valid*, as shown in the third row of table 7.6. Therefore, the *HI prediction quality* is equal to zero, which is shown in the fourth row of table 7.6.

Furthermore, compared to the results using the 'actual' load dataset, the results using the 'tweaked' load dataset suggest that the observed failure probability for the predicted lifetime has a larger distance from the defined 5 % HI observed failure probability. This would suggest that these predictions are even less accurate.

However, as shown in section 5.4.4, the values of the RL predictions for 'tweaked' load are in general lower than the RL predictions for the 'actual' load. Therefore, it could be expected that the 'tweaked' load would show that the HI predictions more accurate than for the 'actual' load case. Only the 'tweaked' load case included predicted suspensions, which can be the cause for this finding.

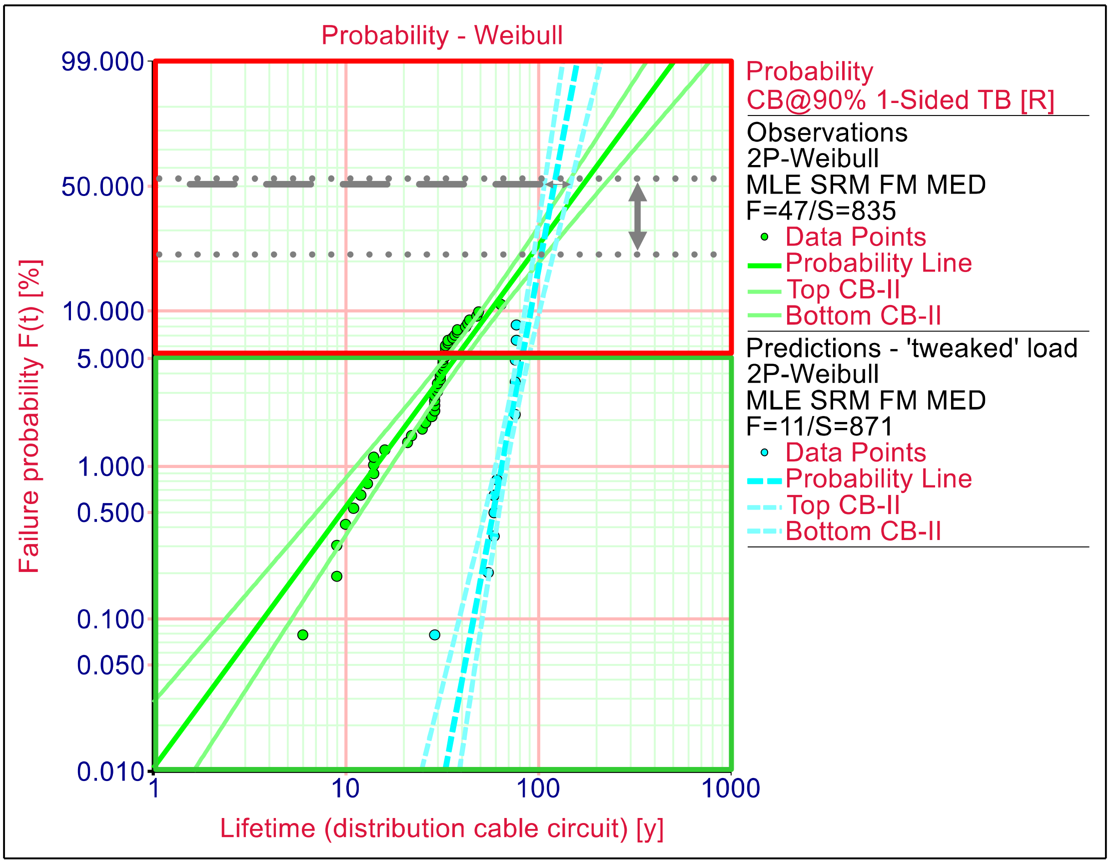


Figure 7.7: Comparing the results of the dataset of predictions to the results of the observations dataset (90% confidence bounds)

Table 7.6: Derivation of validity and HI prediction quality

			90% bound	Lifetime	90% bound
1	Predicted lifetime (failure probability $\leq$ 5%)	[y]	99.7	120.2	144.8
2	Observed failure probability for predicted lifetime	[%]	18.2	32.4	56.9
3	Validity predicted lifetime	{Valid, Non-valid}	Non-valid	Non-valid	Non-valid
4	HI prediction quality	(0-1)	0	0	0

7.4 Conclusions

From this case study the following conclusions can be drawn:

1. This case study has shown the application of a method for quantifying prediction quality of a HI and validation of a HI. The hypothesis, that it is possible to quantify HI prediction quality and to validate this HI (see section 7.2) is therefore not rejected by this case study.
2. Both for distribution fitting of the predictions and observations datasets in this case study, the data quality was low. Furthermore, the observations dataset introduces large uncertainty in observed reliability, caused by the definition of a failure and the changes in topology of cable circuits. Therefore, no firm conclusions can be drawn on the validity of this HI.
3. No firm conclusion can be drawn regarding the finding of this case study on the DNV GL distribution cable HI that the predictions were found to be *too optimistic* a case using 'actual' load values and a case using 'tweaked' load values and therefore *non-valid.*

7.5 Recommendations

For this case study, only maximum load data was sufficiently available at t_{HI}. Therefore, only the *HI prediction quality* of the statistical assessment functions and utilization assessment functions was determined for this HI for *validation* of this HI. To enable better *HI prediction quality quantification* for this HI using the assets of this utility, at a moment in the future, the following is recommended:

- The *time window* should cover a time period during which a sufficient number of failures have occurred. When at least ten failures are considered the minimum and when extrapolating the number of failures in the time window of this dataset, the minimum time window would be 6 years wide, allowing for an analysis in 2015. For this case study, it was found that majority of the condition indicators were available starting from 2013. Therefore, a failure database covering failures until at least the year 2015 are necessary for this dataset.

Chapter 8

Conclusions and recommendations

To ensure a reliable electricity network today and in the future, asset managers of utilities rely on predictions of asset management software tools in their asset management decisions. The DNV GL Health Index (HI) is one of these tools. This thesis has shown methods to quantify the *HI prediction quality* for *validation* of the DNV GL HI. Practical case studies were performed to use additional utility data in the application of one of these methods.

In this chapter, the conclusions and recommendations are presented.

8.1 Conclusions

The conclusions are presented according to the three main questions addressed in this thesis:

1. In which way can the *HI prediction quality* of the DNV GL HI be quantified?

2. In which way can *HI prediction quality quantification* of the DNV GL HI be used to *validate* the DNV GL HI?

3. Using additional data, which statements can be made on *prediction quality* and *validation* of the DNV GL HI?

Firstly, five *HI prediction quality quantification methods* have been presented in this thesis.

- Two *HI prediction quality quantification methods* presented in this thesis enable *HI prediction quality quantification* of the DNV GL HI. These methods, as described in chapter 3, consist of a method that quantifies *prediction error* for every *single* asset and a method that quantifies the *prediction error* for a *group* of assets. These methods use, compared to the three methods that are described, most information from the predicted asset health. Furthermore, they complement each other with respect to weaknesses introduced by their working principle: on *group level* and on *single asset level.* This makes the combination of these methods suitable for *HI prediction quality quantification.*

Secondly, it was found that four out of five *HI prediction quality quantification methods* can be used for *validation* of HIs.

- Due to the remaining lifetime definition in the HI, only one *HI prediction quality quantification method* presented in this thesis is suitable for *HI prediction quality quantification* and *validation* of the DNV GL HI.

Thirdly, a limited number of failures, data on a only a limited number of condition data parameters and limited utilization data was gathered during the collection of the utility data. In other words, only *limited* data was available. Furthermore, the quality of the data in both case studies was too low to use life data analysis in drawing firm conclusions on the validity of the HIs.

- The practical case study for the DNV GL instrument transformer HI presented in this analysis using one *HI prediction quality quantification method* does not reject the hypothesis that it is possible to quantify *HI prediction quality* using utility data.
- No firm conclusion can be drawn regarding the finding of this case study on the DNV GL instrument transformer HI that the predictions were found to be *too pessimistic* and therefore *valid*.

The practical case study for the DNV GL distribution cable HI included two additional assumptions on the observations dataset (see the observations data in section 5.4.4): the first interruption in a cable circuit were taken equal to a failure and the changing cable circuit reliability by changes in network topology due to maintenance were not taken into account.

- The practical case study for the DNV GL distribution cable HI presented in this analysis using one *HI prediction quality quantification method* does not reject the hypothesis that it is possible to quantify *HI prediction quality* using utility data.
- The practical case study for the DNV GL distribution cable HI presented in this analysis has revealed that for *HI prediction quality quantification* of cable circuits, besides a required data quality, the influence of the definition of a failure and changing reliability by changes in network topology should be incorporated.
- No firm conclusion can be drawn regarding the finding of this case study on the DNV GL distribution cable HI that the predictions were found to be *too optimistic* and therefore *non-valid*.

8.2 Recommendations

8.2.1 Recommendations for DNV GL

A limited number of failures, data on a only a limited number of condition data parameters and a limited amount of utilization data was gathered during the collection of the utility data. In order to collect utility data which encompasses these data types, the following is recommended:

- An investigation should be performed looking for a platform that enables quick data input for utilities and keeps the effort for development of assessment functions in the present tool to a minimum. A start could be to look at existing maintenance software tools (e.g. DNV GL's own maintenance management software CASCADE).

The case studies have demonstrated a step-wise approach in using utility data for *HI prediction quality quantification* and *validation* for two asset types covered by the DNV GL Health Index. When this step-wise approach is automatized and embedded in the Health Index, this will ensure continuous enhancement of the tool's quality.

- DNV GL should consider to embed this step-wise approach in its health index.

Remaining lifetime in the DNV GL HI is currently defined as the period after which the value of failure probability does not exceed a given reference value. In order to compare predicted health and observed health on a *single asset* level, the following is recommended:

- A feasibility study should be performed in including a remaining lifetime definition for the moment when the failure probability equals 50%. This feasibility study includes the proof that when validity of the health index using this definition is shown, it also holds for the health index using the remaining lifetime definition for the moment when the failure probability equals X%. This enables to apply the methods based on single asset level, which complement the methods based on group level. Therefore, it is expected that, when methods are applied both on group and single asset level combined with asset data of sufficient quality and quantity, stronger statements on the validity of the Health Index become possible.

8.2.2 Recommendations for further research

The data of the case studies in this analysis was insufficient for drawing firm conclusions. The cable circuits appeared to introduce two additional assumptions: the first interruption in a cable circuit was taken equal to a failure and the changing cable circuit reliability by changes in network topology due to maintenance was not taken into account. For a more extensive case study in the future, it is recommended to take this into account during the choice of the asset type. Therefore, it is recommended to:

- investigate assets with a more firmly defined end-of-life than distribution cables:
  - Instrument transformers
  - Power transformers
  - Circuit breakers

Limited advice from statisticians was used to test the methods presented in this thesis on their statistical correctness. Together with the aforementioned recommendation, the following is recommended:

- The methods presented in this thesis need further investigation with respect to the statistical correctness of their statements on *HI prediction quality* and *validation*. This investigation should focus on the extending the definition of *validation* for the *prediction error* and *life data analysis* methods, as presented in sections 3.5.4 and 3.5.5. An extended definition for *validation* using these methods combined with data without the limitations of the data used in this thesis is expected to result in validation of the DNV GL HI.

To conclude, this research has been a contribution on the road towards a health index that can provide asset managers with more reliable predictions to support their asset management decisions. This will ensure a more reliable, sustainable and cost-effective electricity network for today and the future, in the Netherlands and across its borders.

Bibliography

[1] A. G. Jongepier. Ageing networks: Consume, prolong or replace. In *Electricity briefing paper Leonardo Energy*, 2007.

[2] BSI/IAM. Pas 55-1:2008 asset management. specification for the optimized management of physical assets. In *ed: British Standards Institution*, pages v–x, 2008.

[3] A. Soemeer. Optimized maintenance planning for transmission power systems. improvement of the condition indexing process. 2013.

[4] D. Hughes, P. Pschierer-Barnfather, and T. Pears. Condition based risk management (cbrm): Bringing excellence in asset management into the boardroom. In *Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*, pages 1–4, June 2009.

[5] D. Hughes. Condition based risk management (cbrm) 2014; enabling asset condition information to be central to corporate decision making. In *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*, pages 1–5, June 2005.

[6] D. Hughes. The use of "health indices" to determine end of life and estimate remnant life for distribution assets. In *Electricity Distribution, 2003. CIRED 2003. 17th International Conference and Exhibition on*, May 2003.

[7] D. Hughes and T. Pears. Cbrm, a process to link engineering knowledge to investment planning. In *Electricity Distribution (CIRED 2007)*, May 2007.

[8] J. J. Smit and C. M. Ackerman. Diagnostic maintenance management system for mv and hv substations. In *Electricity Distribution. Part 1: Contributions. CIRED. 14th International Conference and Exhibition on (IEE Conf. Publ. No. 438)*, volume 1, pages 19/1–19/5 vol.1, 1997.

[9] G. J. Cliteur and J. M. Wetzer. Condition assessment of power transmission and distribution components. In *Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482)*, volume Summaries, pages 47–47, 2001.

[10] A. van Schijndel, J. M. Wetzer, and P. A. A. F. Wouters. Remaining lifetime modelling for replacement of power transformer populations. In *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, pages 597–600, April 2008.

[11] M. E. Vermeer, A. Sharma, J. M. Wetzer, P. Zonneveld, and D. Boender. Asset management decision support modeling, using a health index, for maintenance and replacement planning. In *Cigré GCC POWER 2013*, November 2013.

[12] G. A. Bloemhof and J. Knijp. Squeezing margins: reducing costs while balancing risks. In *Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482)*, volume Summaries, pages 304–304, 2001.

[13] A. van Schijndel, J. M. Wetzer, and P. A. A. F. Wouters. Forecasting transformer reliability. In *Electrical Insulation and Dielectric Phenomena, 2006 IEEE Conference on*, pages 577–582, Oct 2006.

[14] A. van Schijndel, J. M. Wetzer, and P. A. A. F. Wouters. Modelling technique for power transformer replacement waves. In *Probabilistic Methods Applied to Power Systems, 2008. PMAPS '08. Proceedings of the 10th International Conference on*, pages 1–5, May 2008.

[15] G. Bajracharya. *Multi-Agent Model-Based Optimization for Future Electrical Grids*. PhD thesis, Delft university of Technology, 2014.

[16] B. J. Grotenhuis, J. E. Jaspers, A. Kerstens, A. H. van der Wey, and F. H. De Wild. Increasing the capacity of cable systems using cable asset management based on thermal and mechanical properties. In *Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482)*, volume Summaries, pages 202–202, 2001.

[17] T. Koltunowicz. *Accelerated insulation aging due to thermal and electrical stresses in future power grids*. PhD thesis, Delft university of Technology, 2014.

[18] A. van Schijndel, J. M. Wetzer, and P. A. A. F. Wouters. Validation of reliability forecasting for power transformers. In *Probabilistic Methods Applied to Power Systems, 2008. PMAPS '08. Proceedings of the 10th International Conference on*, pages 1–4, May 2008.

[19] J. M. Wetzer, G. J. Cliteur, W. R. Rutgers, and H. F. A. Verhaart. Diagnostic- and condition assessment-techniques for condition based maintenance. In *Electrical Insulation and Dielectric Phenomena, 2000 Annual Report Conference on*, volume 1, pages 47–51 vol.1, 2000.

[20] B. Gorgan, P. V. Notingher, J. M. Wetzer, H. F. A. Verhaart, P. A. A. F. Wouters, A. van Schijndel, and G. Tanasescu. Calculation of the remaining lifetime of power transformers paper insulation. In *Optimization of Electrical and Electronic Equipment (OPTIM), 2012 13th International Conference on*, pages 293–300, May 2012.

[21] P. A. A. F. Wouters, A. van Schijndel, and J. M. Wetzer. Remaining lifetime modeling of power transformers: individual assets and fleets. *Electrical Insulation Magazine, IEEE*, 27(3):45–51, May 2011.

[22] J. M. Wetzer. Maintaining future (electrical) power systems. In *Future Power Systems, 2005 International Conference on*, pages 6 pp.–6, Nov 2005.

[23] M. Pomp and V. Shestalova. Switching costs in netherlands energy markets: Can liberalisation bring benefits to small customers? *The Economist*, 155(3):305–321, 2007.

[24] E. Hamilton and D. Hughes. Cbrm, the practical application of intelligent asset management to united utilities distribution network, 2004.

[25] D. Hughes, T. Pears, and Y. Tian. Linking engineering knowledge and practical experience to investment planning by means of condition based risk management. In *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, pages 539–542, April 2008.

[26] D. Hughes, P. Barnfather, and T. Pears. Quantifying asset risk to optimise investment, a realistic prospect? In *Electricity Distribution (CIRED 2010)*, June 2010.

[27] T. Hjartarson, B. Jesus, D. T. Hughes, and R. M. Godfrey. Development of health indices for asset condition assessment. In *Transmission and Distribution Conference and Exposition, 2003 IEEE PES*, volume 2, pages 541–544 vol.2, Sept 2003.

[28] D. Hughes, P. Barnfather, and A. Harrison. Combining load and condition based risk models for integrated investment planning. In *Electricity Distribution (CIRED 2013), 22nd International Conference and Exhibition on*, pages 1–5, June 2013.

[29] D. T. Hughes and T. Richards. The use of condition information in strategic risk management, presented at EuroTechCon, Birmingham (UK), October 2002.

[30] I. Hjartarson, T. Khan, and M. Godfrey. Health indices for substation asset condition assessment. In *EDIST Conference, Markham, Ontario, Canada, January 2004*, January 2004.

[31] E. Goddard. Network output measures workshop. May 2008.

[32] J. Endrenyi, G. J. Anders, and A. M. Leite da Silva. Probabilistic evaluation of the effect of maintenance on reliability. an application [to power systems]. *Power Systems, IEEE Transactions on*, 13(2):576–583, May 1998.

[33] G. J. Anders, J. Endrenyi, and C. Yung. Risk-based planner for asset management [of electric utilities]. *Computer Applications in Power, IEEE*, 14(4):20–26, Oct 2001.

[34] G. J. Anders, J. M. Braun, S. Rizzetto, and D. Casselman. Real time monitoring system for 230 kv cables in tunnels. In *Transmission and Distribution Conference and Exposition, 2003 IEEE PES*, volume 3, pages 886–891 vol.3, Sept 2003.

[35] G. Anders, S. Otal, and T. Hjartarson. Deriving asset probabilities of failure: effect of condition and maintenance levels. In *Power Engineering Society General Meeting, 2006. IEEE*, 2006.

[36] A. Naderian, S. Cress, R. Piercy, F. Wang, and J. Service. An approach to determine the health index of power transformers. In *Electrical Insulation, 2008. ISEI 2008. Conference Record of the 2008 IEEE International Symposium on*, pages 192–196, June 2008.

[37] T. Hjartarson. Next generation asset management. In *Smart Grid Newsletter*, December 2010.

[38] T. Hjartarson and S. Otal. Predicting future asset condition based on current health index and maintenance level. In *Transmission Distribution Construction, Operation and Live-Line Maintenance, 2006. ESMO 2006. IEEE 11th International Conference on*, pages –, Oct 2006.

[39] A. Jahromi, R. Piercy, S. Cress, J. Service, and W. Fan. An approach to power transformer asset management using health index. *Electrical Insulation Magazine, IEEE*, 25(2):20–34, March 2009.

[40] J. Endrenyi and G. J. Anders. Aging, maintenance, and reliability - approaches to preserving equipment health and extending equipment life. *Power and Energy Magazine, IEEE*, 4(3):59–67, May 2006.

[41] Y. Tsimberg. Use of asset management risk assessment methodology for public safety analysis. February 2010.

[42] Y. Tsimberg. Assessing condition of overhead lines using best asset management practices. July 2009.

[43] B. Quak, E. Gulski, J. J. Smit, F. J. Wester, and E. R. S. Groot. Database support for condition assessment of distribution power cables. In *Electrical Insulation, 2002. Conference Record of the 2002 IEEE International Symposium on*, pages 468–471, Apr 2002.

[44] E. Gulski, E. R. S. Groot, J. J. Smit, and B. Quak. Knowledge sharing of asset condition data and performances. In *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*, pages 1–4, June 2005.

[45] R. Jongen, E. Gulski, P. H. F. Morshuis, J. J. Smit, and A. L. J. Janssen. Statistical analysis of power transformer component life time data. In *Power Engineering Conference, 2007. IPEC 2007. International*, pages 1273–1277, Dec 2007.

[46] B. Quack. *Information Strategy for Decision Support in Maintaining High Voltage Infrastructures*. PhD thesis, Delft University of Technology, 2007.

[47] J. J. Smit, O. M. Piepers, and E. Gulski. Diagnostic decision support system with condition indexing for asset managers of electrical infrastructures. In *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, pages 603–608, Oct 2007.

[48] R. A. Jongen, P. H. F. Morshuis, J. J. Smit, A. L. J. Janssen, and E. Gulski. Failure analysis of in service failed resin cable joints by means of a statistical approach. In *Electrical Insulation and Dielectric Phenomena, 2006 IEEE Conference on*, pages 517–520, Oct 2006.

[49] R. A. Jongen, J. J. Smit, and A. L. J. Janssen. Application of statistical analysis in the asset management decision process. In *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, pages 1284–1287, April 2008.

[50] J. J. Smit, O. M. Piepers, and E. Gulski. Diagnostic decision support system with condition indexing for asset managers of electrical infrastructures. In *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, pages 603–608, Oct 2007.

[51] E. Gulski, J. J. Smit, F. J. Wester, and J. W. van Doeland. Condition assessment of high voltage power cables. In *Power System Technology, 2004. PowerCon 2004. 2004 International Conference on*, volume 2, pages 1661–1666 Vol.2, Nov 2004.

[52] R. A. Jongen, P. H. F. Morshuis, E. Gulski, J. J. Smit, and A. L. J. Janssen. Service availability assessment of mass-insulated power cables by means of statistical analysis. In *Electrical Insulation, 2008. ISEI 2008. Conference Record of the 2008 IEEE International Symposium on*, pages 707–711, June 2008.

[53] J. J. Smit, B. Quak, and E. Gulski. Integral decision support for asset management of electrical infrastructures. In *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on*, volume 3, pages 2622–2628, Oct 2006.

[54] R. A. Jongen, E. Gulski, P. H. F. Morshuis, J. J. Smit, and A. L. J. Janssen. Statistical analysis of power transformer component life time data. In *Power Engineering Conference, 2007. IPEC 2007. International*, pages 1273–1277, Dec 2007.

[55] R. A. Jongen, E. Gulski, J. J. Smit, and A. L. J. Janssen. Statistical analysis of life time data of service aged components for asset management decision process support. *Proceedings of the 16th International Symposium on High Voltage Engineering*, 2009.

[56] B. Quak, J. J. Smit, and E. Gulski. Decision support for maintenance management of high voltage components. In *Properties and applications of Dielectric Materials, 2006. 8th International Conference on*, pages 262–266, June 2006.

[57] E. Gulski, F. J. Wester, Ph. Wester, E.R.S. Groot, and J. W. van Doeland. Condition assessment of high voltage power cables. *Cigré*, 2004.

[58] R.A. Jongen. *Statistical lifetime management for energy network components*. PhD thesis, Delft University of Technology, The Netherlands, 2012.

[59] R. A. Jongen, E. Gulski, and J. J. Smit. Statistical analysis of diagnostic- and life time data of high voltage components. In *Power Engineering Conference, 2007. IPEC 2007. International*, pages 355–359, Dec 2007.

[60] E. Gulski, S. Meijer, J. J. Smit, F. de Vries, R. Leich, H. Geene, R. Koning, L. Lamballais, T. J. W. H. Hermans, E. R. S. Groot, J. Slangen, M. J. M. Boone, J. Kanters, A. Bun, and A. L. J. Janssen. Condition assessment and am decision support for transmission network components. *Cigré*, 2006.

[61] R. A. Jongen, P. H. F. Morshuis, J. J. Smit, and A. L. J. Janssen. Influence of ambient temperature on the failure behavior of cable joints. In *Electrical Insulation and Dielectric Phenomena, 2007. CEIDP 2007. Annual Report - Conference on*, pages 643–646, Oct 2007.

[62] R. A. Jongen, E. Gulski, P. H. F. Morshuis, J. J. Smit, and A. L. J. Janssen. Application of statistical tools for assessing service availability of high voltage power cables. In *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, pages 546–550, April 2008.

[63] L. Chmura, H. Jin, P. Cichecki, J. J. Smit, E. Gulski, and F. D. Vries. Use of dissipation factor for life consumption assessment and future life modeling of oil-filled high-voltage power cables. *Electrical Insulation Magazine, IEEE*, 28(1):27–37, January 2012.

[64] R. A. Jongen, P. H. F. Morshuis, J. J. Smit, and A. L. J. Janssen. Failure prediction in 50kv mass-insulated power cable systems. *Proceedings of 7th International Conference on Insulated Power Cables (Jicable '07)*, pages 170–173, 2007.

[65] B. Quak, E. Gulski, Ph. Wester, and J. J. Smit. Fundamental aspects of information processing and the decision process to support asset management. In *Properties and Applications of Dielectric Materials, 2003. Proceedings of the 7th International Conference on*, volume 3, pages 982–985 vol.3, June 2003.

[66] R. A. Jongen, P. H. F. Morshuis, E. Gulski, J. J. Smit, J. Maksymiuk, and A. L. J. Janssen. Application of statistical methods for making maintenance decisions within power utilities. *Electrical Insulation Magazine, IEEE*, 22(6):24–35, Nov 2006.

[67] P. Hafkamp and G. Schufters. Designing and implementing a maintenance management system with an expert support system: methodology, implementation, expert it-system, achieved cost reductions. In *Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482)*, volume Summaries, pages 46–46, 2001.

[68] Ph. Wester, J. J. Smit, J. Oestergaard, and J. Corbett. Development of asset management services. In *Cigré session Paris 2002*, pages 23–308, 2002.

[69] L. Chmura, P. H. F. Morshuis, E. Gulski, J. J. Smit, and A. L. J. Janssen. Statistical analysis of subcomponent failures in power transformers. In *Electrical Insulation Conference (EIC), 2011*, pages 216–220, June 2011.

[70] E. Gulski, F Wester, W. Boone, N. van Schaik, E. F. Steennis, E. R. S. Groot, J. Pellis, and B. J. Grotenhuis. Knowledge rules support for cbm of power cable circuits. *Cigré*, 2002.

[71] E. Gulski, B. Quak, and E. R. S. Groot. Asset management integral decision support model for distribution power cables. In *Power Engineering Conference, 2005. IPEC 2005. The 7th International*, pages 788–793 Vol. 2, Nov 2005.

[72] R. A. Jongen, E. Gulski, and J. J. Smit. Statistical analysis of diagnostic- and life time data of high voltage components. In *Power Engineering Conference, 2007. IPEC 2007. International*, pages 355–359, Dec 2007.

[73] G. Bajracharya. Optimization of maintenance for power system equipment using a predictive health model. Technical report, Delft University of Technology, 2009.

[74] G. Bajracharya, T. Koltunowicz, R. R. Negenborn, Z. Papp, D. Djairam, B. de Schutter, and J. J. Smit. Optimization of maintenance for power system equipment using a predictive health model. In *PowerTech, 2009 IEEE Bucharest*, pages 1–6, June 2009.

[75] ReliaSoft Corporation. *Life Data Analysis Reference*. Tucson, Arizona, 2014.

[76] R. P. Y. Mehairjan. Application of statistical life data analysis for cable joints in mv distribution networks. Master's thesis, Delft University of Technology, The Netherlands, 2010.

[77] E. J. de Haan. High voltage asset performance modeling. Master's thesis, Delft University of Technology, The Netherlands, 2011.

[78] Iec/ieee guide for the statistical analysis of electrical insulation breakdown data (adoption of ieee std 930-2004). *IEC 62539 First Edition 2007-07 IEEE 930*, 2007.

[79] Utility 'H'. Kwaliteits- en capaciteitsdocument, 2014.

[80] DNV GL. Interpretation - ahrm - information request (2014-08-04), 2014. Information on the health index parameters, internal document of DNV GL.

[81] EDSN. profielen-elektriciteit-2012-versie-1.00.zip, 2012.

[82] E. F. Steennis, P. Soepboer, J. Mosterd, P. Buys, P. Oosterlee, L. Bokma, and R. Meier. The effect of high current loads on joints in mv cable systems. June 2011.

[83] DNV GL. Concept rapport conditiebepaling utility ‘g’ 211209, 2009. Information on the health index parameters, internal document of DNV GL.

[84] Specialist asset management at utility ‘H’, 2014. Information forthcoming from phone and e-mail conversations.

Appendices

Appendix A

Additional health index requirements, asset data and comparison

This chapter lists, for each HI SoT, the additional HI requirements to those mentioned in characteristic 1 of section 2.1. Furthermore, it shows, for each HI SoT, the entire tables for the asset data types and comparison in characteristic 2 of section 2.1.

A.1 DNV GL HI SoT

A.1.1 Additional requirements DNV GL HI SoT

Compatibility

- A HI is required to achieve optimal solutions. Therefore, it should be part of a uniform system with a systematic and consistent approach [12].
- The HI needs to be applicable to multiple types of assets and configurations [16].
- The decision support software that contains the HI should:
  - support condition based maintenance (CBM) [67].
  - be able to interface with other software [67, 8].

Implementation

- During the implementation phase of the decision support software, a step-wise approach needs to be followed consisting of the following parts [8]:
  - Selecting components by investment analysis.
  - Perform a system functional analysis using statistics and FMECAs.
  - Build ageing models.
  - Include the system in the maintenance activities.
  - Use feedback to improve the ageing model.

Content/information handling

- A HI should contain databases to store the information from measurements.
- Translating partial discharge (PD) data into asset health makes use of pattern recognition and fingerprint techniques. This requires reference databases to be included in the HI [19].

Capabilities

- HIs should be created with the following objectives in mind [8]:
  - Control and reduce maintenance costs.
  - Maintain acceptable reliability and safety levels.
  - Prevent ageing-related failures.
  - Apply lifetime extension to reduce costs if feasible.
  - Preserve the expert knowledge of the ageing workforce.
  - Enhance the knowledge of maintenance activities.
- It should be possible to update the HI according to utilities needs [8].

A.1.2 Asset data types DNV GL HI SoT

Table A.1: (Number of) asset data types as found in literature for the DNV GL Energy HI SoT

Category	Asset type	#	(Number of) asset data types	Reference
Transformer	Power transformers	1	Bushings loss angle	[9]
		2	Core loss angle	[9]
		3	Core no load losses	[19]
		4	Degree of Polymerisation	[18]
		5	Dissolved Gas Analysis	[20, 19, 9]
		6	Furfural analysis	[9]
		7	Historical ambient temperature	[18]
		8	Historical hourly loading pattern	[18]
		9	Installation date	[18]
		10	Loading	[18]
		11	Oil analysis	[9]
		12	OLTC drive power	[19]
		13	OLTC dynamic resistance	[19]

Table A.1: (Number of) asset data types as found in literature for the DNV GL Energy HI SoT (continued)

Category	Asset type	#	(Number of) asset data types	Reference
		14	Paper insulation furfural analysis	[19]
		15	Partial Discharge analysis	[20]
		16	Resistance measurements	[9]
		17	Thermal imaging	[9]
		18	Type of paper	[18]
Switchgear	General	19	Corrosion	[9]
		20	Deformation	[9]
		21	Drive contact position	[19]
		22	Drive contact velocity	[19, 9, 8]
		23	Drive trip coil current	[19, 9, 67]
		24	Drive vibrational analysis	[19, 9]
		25	Gas analysis (SF_6)	[9]
		26	Leakage	[9]
		27	Leakage current	[8]
		28	Number of switching operations	[9]
		29	Oil analysis	[9]
		30	Oil leakage	[67]
		31	Oil level	[67]
		32	On-line partial discharge measurement	[19]
		33	Partial discharge measurement	[9, 8]
		34	Resistance measurement	[9]
		35	Secondary system trip coil current	[19]
		36	Sound	[9]
		37	Thermal imaging	[9]
		38	Vacuum leakage test	[9]
		39	Vegetation	[9]
		40	Vibration of bushings	[9]
Power cables	General	41	Dielectric spectroscopy: capacitance	[19]

Table A.1: (Number of) asset data types as found in literature for the DNV GL Energy HI SoT (continued)

Category	Asset type	#	(Number of) asset data types	Reference
		42	Dielectric spectroscopy: loss angle	[19]
		43	Distributed temperature sensing	[16]
		44	Partial discharge localization	[19]
		45	Partial discharge on-line VHF detection	[19]
		46	Partial discharge 0,1 Hz off-line detection and localization	[19]
		47	Partial discharge single and double sided measurements	[19]
Generators	Stator and windings	48	Capacitance	[19]
		49	Conductor resistance	[19]
		50	High voltage tests	[19]
		51	Insulator resistance	[19]
		52	KEMA Inspection System for Retaining Rings	[19]
		53	Loss angle	[19]
		54	Partial discharge measurement	[19]
		55	Polarization index	[19]
		56	Video endoscopy	[19]
		57	Condition of wedges	[19]

A.1.3 Detection method classification DNV GL HI SoT

Table A.2: Classification of the detection methods of the DNV GL Energy HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table contains a classification for the individual detection methods.

Subject	Subject value	Detection methods for subject value
Costs	Low (often already available)	7,8,9,10,18
	High	1,2,3,4,5,6,11,12,13,14,15,16,17,19, 20,21,22,23,24,25,26,27,28,29,30,31, 32,33,34,35,36,37,38,39,40,41,42,43, 44,45,46,47,48,49,50,51,52,53,54,55, 56,57
Out of service required	No	5,6,7,8,9,10,11,17,19,20,25,26,28,29, 30,31,32,36,37,39,40,43,44,45,46,47, 48,49,50,51,52,53,54,55,56,57
	Yes	1,2,3,4,12,13,14,15,16,18,21,22,23, 24,27,33,34,35,38,41
Maturity	Mature	1,2,3,4,5,6,7,8,9,10,11,12,15,16, 17,18,19,20,21,26,27,28,29,30,31, 36,37,39,40,46,47,50,52,53,57
	Development phase	13,14,22,23,24,25,32,33,34,35,38, 41,42,43,44,45,48,49,51,54,55,56
Ageing detection	Long term ageing	5,6,7,8,9,10,11,28,29,30,34,38, 39,56,57
	Short term ageing	1,2,3,4,12,13,14,15,16,17,18,19, 20,21,22,23,24,25,26,27,31,32, 33,35,36,37,40,41,42,43,44,45, 46,47,48,49,50,51,52,53,54,55
Accuracy	Good	1,2,3,4,5,6,7,8,9,10,11,12,13,14, 16,17,18,19,20,21,22,23,24,25,26, 27,28,29,30,31,34,35,36,37,38,39, 40,41,42,43,48,49,50,51,52,53,55, 56,57
	Bad	15,32,33,44,45,46,47,54

A.2 EA Technology HI SoT

A.2.1 Additional requirements EA Technology HI SoT

Capabilities

- A HI should represent the overall asset health indicating the suitability for continued service from multiple condition indicators [27].

EA Technology uses the term "Condition Based Risk Management" (CBRM) for the framework their HI is a part of. As a consequence of cooperation with utilities, the initial versions of this framework were developed bearing in mind the following (from [5]):

- Utilities should be able to use them for achieving their short term objectives. This required the first HI predictions to aim for the short term.
- Developing a representation of asset health within a short time period allowing use for potential investment plans. This time pressure caused to focus on the use of existing data.

A.2.2 Asset data types EA Technology HI SoT

Table A.3: (Number of) asset data types as found in literature for the EA technology HI SoT

Category	Asset type	(Number of) asset data types	Reference
Transformer	Distribution transformers	6	[30]
	Power transformers	7	[30]
	Reactors	7	[30]
	Tap changers	4	[30]
Transformer	Distribution transformers	6	[30]
	Power transformers	7	[30]
	Reactors	7	[30]
	Tap changers	4	[30]
Switchgear	Air blast circuit breakers	17	[30]
	Air magnetic circuit breakers	16	[30]
	Circuit switchers	10	[30]
	Disconnecting switches	10	[30]
	Fuses	2	[30]
	GIS Systems	28	[30]
	Metal clad switchgear	35	[30]
	Oil circuit breakers	11	[30]
	Reclosers	7	[30]
	SF_6 circuit breakers	17	[30]
	Vacuum circuit breakers	16	[30]
-	Capacitors	9	[30]
Instrument transformer	Capacitive voltage transformers	8	[30]
	Oil filled current transformers	12	[30]
	Oil filled potential transformers	11	[30]
	SF_6 filled current transformers	11	[30]
Auxiliaries	Batteries	10	[30]
	Buildings	3	[30]

Table A.3: (Number of) asset data types as found in literature for the EA technology HI SoT (continued)

Category	Asset type	(Number of) asset data types	Reference
	Cables and potheads	6	[30]
	Chargers	7	[30]
	Drainage and geo-technical systems	6	[30]
	Fences	6	[30]
	Fire protection systems	6	[30]
	Grounding systems	10	[30]
	High pressure air systems	44	[30]
	Insulators	6	[30]
	Mobile unit substations	17	[30]
	Power line carrier	6	[30]
Overhead lines	Wood pole overhead lines	Cross arm	[27, 6]
		Fittings	[27, 6]
		Guys	[27, 6]
		Insulators	[27, 6]
		Wood pole	[27, 6]
	Steel tower overhead lines	Not given	[6]

A.2.3 Detection method classification EA Technology HI SoT

As no asset data types were found in literature of the EA Technology HI SoT, no asset detection method classification was possible for this HI SoT.

A.3 Kinectrics HI SoT

A.3.1 Additional requirements Kinectrics HI SoT

Capabilities

- As stated in [36], well-established methods are not yet available to quantify asset health based on all available data. Therefore, HIs should be clear on their capabilities and limitations.

A.3.2 Asset data types Kinectrics HI SoT

Table A.4: (Number of) asset data types as found in literature for the Kinectrics HI SoT

Category	Asset type	(Number of) asset data types	Reference
Transformer	Distribution transformers	6	[30]
	Power transformers	7	[30]
	Reactors	7	[30]
	Tap changers	4	[30]
Switchgear	Air blast circuit breakers	17	[30]
	Air magnetic circuit breakers	16	[30]
	Circuit switchers	10	[30]
	Disconnecting switches	10	[30]
	Fuses	2	[30]
	GIS Systems	28	[30]
	Metal clad switchgear	35	[30]
	Oil circuit breakers	11	[30]
	Reclosers	7	[30]
	SF_6 circuit breakers	17	[30]
	Vacuum circuit breakers	16	[30]
-	Capacitors	9	[30]
Instrument transformer	Capacitive voltage transformers	8	[30]
	Oil filled current transformers	12	[30]
	Oil filled potential transformers	11	[30]
	SF_6 filled current transformers	11	[30]

Table A.4: (Number of) asset data types as found in literature for the Kinectrics HI SoT (continued)

Category	Asset type	(Number of) asset data types	Reference
Auxiliaries	Batteries	10	[30]
	Buildings	3	[30]
	Cables and potheads	6	[30]
	Chargers	7	[30]
	Drainage and geo-technical systems	6	[30]
	Fences	6	[30]
	Fire protection systems	6	[30]
	Grounding systems	10	[30]
	High pressure air systems	44	[30]
	Insulators	6	[30]
	Mobile unit substations	17	[30]
	Power line carrier	6	[30]
Overhead lines	Wood pole overhead lines	Cross arm	[27, 6]
		Fittings	[27, 6]
		Guys	[27, 6]
		Insulators	[27, 6]
		Wood pole	[27, 6]
	Steel tower overhead lines	Not given	[6]

A.3.3 Detection method classification Kinectrics HI SoT

Table A.5: Classification of the detection methods of the Kinectrics HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table contains a classification for the individual detection methods.

Subject	Subject value	Detection methods for subject value
Costs	Low (often already available)	12,28,34
	High	1,2,3,4,5,6,7,8,9,10,11,13,14,15,16, 17,18,19,20,21,22,23,24,25,26,27,29, 30,31,32,33,35,36,37,38,39,40,41,42, 43,44,45
Out of service required	No	1,2,3,5,6,7,8,10,12,13,14,15,16,17,18, 20,21,22,23,24,25,27,28,30,31,32,33, 34,37,38,39,40,41,42,44,45
	Yes	4,9,11,19,26,29,35,36,43
Maturity	Mature	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16, 17,18,19,20,22,23,24,25,26,27,28,29, 30,31,32,33,34,35,36,37,38,39,40,41, 42,44,45
	Development phase	21,43
Ageing detection	Long term ageing	4,5,6,9,11,12,13,14,15,16,17,18,19, 21,35,36,38,41,43,44,45
	Short term ageing	1,2,3,7,8,10,20,22,23,24,25,26,27,28, 29,30,31,32,33,34,37,39,40,42
Accuracy	Good	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20,22,23,24,25,26,27,28, 29,30,31,32,33,34,35,36,37,38,39,40, 41,42,43,44,45
	Bad	21

A.4 TU Delft HI SoT

A.4.1 Additional requirements TU Delft HI SoT

Content/information handling

- Information on stresses in service, maintenance experience, failures and defects, laboratory tests and exit materials could be collected for use in health indexing (Gulski, [60]).
- Within a HI, a selection should be made regarding the parameters indicating component health (Quak, [56]).
- An increasing amount of condition data enhances the quality of the assessment as performed by a HI (Gulski, [44]).
- A HI needs to cover all of its relevant disciplines and aspects (Gulski, [71, 47]).
- To prevent excluding the most profitable scenario, a HI should be able to calculate the asset health in all types of maintenance scenarios (Quak, [56]).
- A HI should take into account all types of information and different levels in the decision process (Jongen, Quak, [56, 72]).
- A HI should support the decision process of the asset manager (Quak, [56]).

Capabilities

- For a HI incorporating reliability centred maintenance or risk based maintenance, the failure behaviour, probability of failure and failure rate of the asset should be known (Jongen, [64]).

A.4.2 Asset data types TU Delft HI SoT

Table A.6: (Number of) asset data types as found in literature for the TU Delft HI SoT

Category	Asset type	#	(Number of) asset data types	Reference
Transformer	General	1	Dielectric losses, tanδ behaviour (up to 1.7×U_0)	[36, 39, 35]
		2	Dielectric response (return voltage amplitude, shape, time behaviour and linearity)	[63]
		3	Failure data	[69, 45]
		4	In-service data	[69, 45]
		5	OLTC dynamic contact resistance	[63]
		6	Partial discharge inception voltage	[63]
		7	Partial discharge location/patterns	[63]

Table A.6: (Number of) asset data types as found in literature for the TU Delft HI SoT (continued)

Category	Asset type	#	(Number of) asset data types	Reference
		8	Partial discharge magnitudes (up to 1.7×U_0)	[63]
	Power transformer	9	Dissolved gas analysis	[60, 72]
		10	Frequency response analysis	[60]
		11	Functional check	[60]
		12	Infrared scan	[60]
		13	Oil analysis	[60]
		14	On-line partial discharge analysis; localisation inside the transformer	[60]
		15	On-line partial discharge analysis; phase resolved pattern analysis	[60]
		16	On-line partial discharge analysis; time domain analysis	[60]
		17	On-line partial discharge analysis; VHF/UHF spectral analysis	[60]
		18	Tap-changer diagnostic	[60]
		19	Visual inspection	[60]
		20	Ambient temperature pattern	[15]
		21	Load pattern	[15]
Gas Insulated Switchgear	General	22	Partial discharge inception voltage	[63]
		23	Partial discharge location/patterns	[63]
		24	Partial discharge magnitudes (up to 1.7xU0)	[63]
		25	VHF/UHF partial discharge detection	[60]
Cable	General	26	Partial discharge diagnosis with Oscillating Wave Test System Methodology	[60, 70, 51]
		27	Failure data	[64]
		28	Life time data	[64]
		29	Return voltage with CDS Methodology	[60, 57]
	Power cables	30	Partial discharge extinction voltage	[46, 51]
		31	Partial discharge inception voltage	[46, 51]

Table A.6: (Number of) asset data types as found in literature for the TU Delft HI SoT (continued)

Category	Asset type	#	(Number of) asset data types	Reference
		32	Partial discharge intensity	[46, 51]
		33	Partial discharge magnitude	[46, 51]
		34	Partial discharge pattern	[46, 51]
	Distribution cables	35	Partial discharge detection (50 Hz off-line, U_0)	[70, 51]
		36	Partial discharge detection (50 Hz on-line, U_0)	[70]
		37	VHF partial discharge detection (50 Hz off-line, U_0)	[70]
		38	VHF partial discharge detection (50 Hz on-line, U_0)	[70]
		39	VLF partial discharge detection (0.1 Hz off-line, U_0)	[70]
	HV oil-filled paper insulated	40	Check oil pressures	[51]
		41	Dielectric losses, tanδ behaviour (up to 1.7×U_0)	[63, 57, 72, 51]
		42	Dissolved gas analysis (DGA)	[57, 72]
		43	Inspection of hydraulic system	[51]
		44	Load current	[63]
		45	Oil analysis	[51]
		46	Oil analysis: AC electric strength	[51]
		47	Oil analysis: Dielectric dissipation factor	[51]
		48	Oil analysis: dissolved gas analysis	[51]
		49	Paper quality	[51]
		50	Partial discharge inception voltage	[63, 46]
		51	Partial discharge location/patterns	[63]
		52	Partial discharge magnitudes (up to 1.7×U_0)	[63]
		53	Relative tanδ	[63]
		54	Visual inspection cable terminals	[51]
		55	Δtanδ	[63]
	HV oil-filled paper insulated/mass impregnated	56	DC sheath test	[51]

Table A.6: (Number of) asset data types as found in literature for the TU Delft HI SoT (continued)

Category	Asset type	#	(Number of) asset data types	Reference
		57	Determination of impregnation coefficient	[51]
		58	g-value measurement (soil thermal resistance)	[51]
		59	HV tests on cable samples (destructive)	[51]
		60	Inspection of earthing system	[51]
		61	Lead sheath analysis (destructive)	[51]
		62	Visual inspection of accessories	[51]
	HV XLPE accessories	63	VHF/UHF partial discharge detection	[60]
	Soil parameters	64	Thermal conductivity (κ)	[60]
		65	Thermal diffusivity (Dh)	[61]
		66	Volume heat capacity (C)	[61]
	External parameters	67	Ambient temperature	[61]

A.4.3 Detection method classification TU Delft HI SoT

Table A.7: Classification of the detection methods of the TU Delft HI SoT for five subjects
The classification of detection methods was done qualitatively using the author's knowledge. This table contains a classification for the individual detection methods.

Subject	Subject value	Detection methods for subject value
Costs	Low (often already available)	3,4,25,26,65
	High	1,2,5,6,7,8,9,10,11,13,14,15,16,17,18, 19,20,21,22,23,24,27,28,29,30,31,32, 33,35,36,37,38,39,40,41,42,43,44,45, 46,47,48,49,50,51,52,53,54,55,56,57, 58,59,60,61,62,63,64
Out of service required	No	3,4,12,13,14,15,16,17,18,23,25, 26,38,41,42,43,44,45,46,52,56, 60,61,62,63,64,65
	Yes	1,2,5,6,7,8,9,10,11,19,20, 21,22,24,27,28,29,30,31, 32,33,34,35,36,37,39,40, 47,48,49,50,51,53,54,55, 57,58,59
Maturity	Mature	1,2,3,4,5,9,10,11,12,13,18,19, 20,21,27,28,29,40,41,42,43,44, 45,46,47,51,52,53,54,55,56,57, 58,59,60,62,63,64,65
	Development phase	6,7,8,14,15,16,17,20,21,22,23,24,28, 29,30,31,32,33,34,35,36,37,48,49,50, 61
Ageing detection	Long term ageing	1,2,3,4,9,10,13,25,26,38,39,40, 42,47,51,54,55,58,65
	Short term ageing	5,6,7,8,11,12,14,15,16,17,18,19,20,21, 22,23,24,27,28,29,30,31,32,33,34,35, 36,37,38,41,43,44,45,46,48,49,50,52, 53,56,57,59,60,61,62,63,64
Accuracy	Good	1,2,3,4,5,9,10,11,12,13,18,19, 25,26,27,38,39,40,41,42,43,44, 45,46,47,51,52,53,54,55,56,57, 58,59,60,62,64,65
	Bad	6,7,8,14,15,16,17,20,21,22,23,24,28, 29,30,31,32,33,34,35,36,37,48,49,50, 61

Appendix B

Extended tables with asset data

Table B.1: Summary of the data files of utility 'H'

#	Data file	Format	Description	Unit
1	Loadflow results	MS Excel	From	[-]
			To	[-]
			Length	[m]
			# conductors	[-]
			Conductor material	[-]
			Conductor surface area	[mm^2]
			Insulating material	[-]
			Nominal voltage (U_{nom})	[kV]
			Nominal current (I_{nom})	[A]
			Maximum power (P)	[kW]
			Maximum reactive power (Q)	[kvar]
			Maximum apparent power (S)	[kVA]
			Maximum current (I)	[A]
			Maximum load	[%]
2	Distribution cable network circuit diagrams	PDF	From	[-]
			To	[-]
			Circuit number	[-]
3	Data distribution cables	MS Excel	Cable object ID	[-]
			Cable asset number	[-]
			Surface area	[mm2]
			Conductor material	[-]
			Insulating material	[-]
			Rated voltage (U_{rated})	[kV]
			Location	[x_1,y_1,x_2,y_2]

Table B.1: Summary of the utility's data files (continued)

#	Data file	Format	Description	Unit
			Commissioning year	[yyyy]
			Region	[-]
			Circuit number	[-]
4	Offline PD	PDF	From	[-]
			To	[-]
			Circuit number	[-]
			Test date	[dd-mm-yyyy]
			Test result	[-]
			PD occurrence	[-]
5	Nestordata failures	MS Excel	From	[-]
			To	[-]
			Region	[-]
			Component type	[-]
			Component age bin	[y_1-y_2]
			Failure cause	[-]
			Date of failure	[dd-mm-yyyy]
6	Cable type specs	MS Excel	# conductors	[-]
			Conductor material	[-]
			Conductor surface area	[mm^2]
			Insulating material	[-]
			Short circuit current rating (1 s)	[kA]
7	Data distribution joints	MS Excel	Joint object ID	[-]
			Joint asset number	[-]
			Start date	[dd-mm-yyyy]
			End date	[dd-mm-yyyy]
			Region	[-]
			Type	[-]
			Commissioning year	[yyyy]
			Location	[x,y]

Appendix C

Extended tables with asset data

Table C.1: Main HI asset data, items marked grey were not available

Category	#	Parameter	Unit
Asset identification	1	No.	[-]
	2	Source	[-]
	3	Destination	[-]
	4	Voltage level	[kV]
	5	Batch number	[#]
General information	6	Maintenance section	[-]
	7	Cable insulation main cable	{XLPE, PILC, Mixed}
	8	Commissioning year	[yyyy]
Environmental asset	9	Mechanical conditions (forces like sinking soils or vibrations)	{Vibrations, Sinking, No}
Design details	10	Rated current	[A]
	11	Rated short-circuit current	[A]
	12	Rated short-circuit period	[s]
	13	Number of cores per cable	[#]
	14	Earthing	{Single, Double, CrossB}
	15	Waterblocking screen	{Yes, No}
	16	Waterblocking conductor	{Yes, No}
Utilization data	17	Max loading	[MVA]
	18	Historical maximum loading	[%]
	19	Historical loading (entire lifetime, in 15 min intervals)	[A]
	20	Actual short circuit current	[A]
	21	Actual short circuit period	[ms]

Table C.1: Main HI asset data, items marked grey were not available (continued)

Category	#	Parameter	Unit
	22	Number of load variations more than 50% Inom	[#/year]
	23	Number of short circuit currents since last maintenance	[#]
	24	Level of oil leakage	{Good, Fair, Poor}
	25	Sheath test results	{Good, Fair, Poor, Critical}
	26	Damage level of terminations	{No, Minor, Major, Severe}
Partial discharge	27	Date of test	[dd-mm-yy]
	28	Test result	{Normal, Fair, Poor}
	29	PD occurrence	{Local, Global}
Dissolved gas analysis	30	Overall lab code	{Green, Amber, Red}
	31	DGA joints (H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO and CO_2)	[ppm]
	32	DGA terminations (H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO and CO_2)	[ppm]
Additional tests	33	Paper DP test	[DP Value]
	34	Tan delta	{good, fair, poor}
	35	Sheath crystallization test	{Passed, Failed}
	36	Metal sheath damage	{Yes, No}
	37	OWTS results	{Good, Fair, Poor, Critical}
Auxiliaries	38	General alarm cable condition	{Good, Poor, Critical}
Maintainability	39	Number of spare cores	[#]
	40	Availability of spare parts	{Yes, No}
	41	Availability of maintenance knowledge	{Yes, No}
Failure data	42	Number of spontaneous failures in cable	[#]

In conclusion, this resulted in an excel-sheet containing a row for each cable section of a cable circuit. For use in the HI, a single row per cable circuit was required.

This was achieved by combining the columns:

1. Circuit name: used for combination
2. From: used for combination

3. To: used for combination
4. Core material: not used in HI
5. Surface area: not used in HI
6. Length: sum of all the lengths of a circuit
7. Load: one of the values (all equal)
8. Cable type: PILC, XLPE, mixed. Only PILC cables selected
9. Surface area: not used in the HI
10. U_{nom}: value (final dataset: all 10 kV).
11. S_1: maximum load, one of the values (all equal). This value was converted into an expected maximum load value of 2007, based on a publication (KCD) of the utility [79].
12. Load: maximum load [%], one of the values (all equal) This value was converted into an expected maximum load value during 2007. This was done taking into account a 1,9% load growth per year, based on a publication of the utility [79].
13. Failed joints: first occurred failure, the total number of failures and the number of failures before 2007.
14. Off-line PD: per circuit date, PD level and PD occurrence.
15. Commissioning year: lowest year is taken from all circuit parts.
16. I_{sc} (1s): lowest values of allowed short circuit current were taken.
17. Water-blocking screen: all cables include a water-blocking screen and water-blocking conductor [84].
18. Number of cores per cable circuit: only for XLPE or mixed, one core per cable was observed (22 circuits). All PILC cables consist of three cores per cable.

Intelligent Electrical Power Grids /
High-Voltage Technology and Management
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Address
Mekelweg 4
2628 CD Delft
The Netherlands

Postal address
P.O. box 5031
2600 GA Delft

Telephone: +31 15 278 2870
Website: http://iepg.ewi.tudelft.nl/