Determination of Probability of Failure of Power Transformers using Statistical Analysis

Keri Pickster

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, May 2015

Declaration

I declare that this dissertation is my own, unaided work, except where otherwise acknowledged. It is being submitted for the degree of Master of Science in Engineering in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Signed on May 31, 2015.

Keri Pickster

Abstract

In electrical power utilities, there is an ever-growing need for improved asset management. Power transformers are identified as one of the most critical and high impact items of plant within an electric network. For this reason, effective management of transformers is required to reduce the risk to power transfer due to unplanned outages, as well as the high consequential costs associated with catastrophic failure.

The objectives of this work include the evaluation of effectiveness of the current method implemented within Eskom, of evaluating transformers based on their condition/Health Index (HI) to develop replacement strategies, as well as identifying possible improvements to these methods and development of a model that can be utilized for determining the probability of failure of a power transformer based on its HI.

There are two components of the existing model for determining failure probability: the effects of age and HI. Historical failure data was collected for the period 1996 - 2014, including both severe and intermediate failures in the Eskom Transmission network. This included failure mode, demographic information, Dissolved Gas Analysis (DGA) results, oil quality test results and predicted Degree of Polymerization (DP). A data sample of healthy transformers was also collected. The failure data was fitted to a Weibull distribution, and the probability of failure based on age determined. This was compared to the existing distribution parameters and its effectiveness evaluated. Statistical analysis was carried out on the complete data set. Since there are multiple, continuous predictor variables and one dichotomous output variable, a multiple logistic regression model was fitted to the data. This was done for the existing HI, as well as for new HI parameters that were identified as the most significant in predicting the output.

The existing Weibull distribution was found to be ineffective in describing the existing failure data for ages <10 and >50 years. The average age predicted by this model is also unrealistically high and no practical evidence of this is found. An alternative Weibull distribution was found that better described the data. The logistic regression model fitted to the failure data using the existing HI parameters was found to be a poor predictor of probability of failure. An alternative model was found enabling a more accurate prediction, using fewer variables. Due to the large errors in measurements of the predictor variables and in some cases, exponential tolerances, as with DP, inaccuracies are expected within the model. The existing model is found to be ineffective in determining the probability of failure of a power transformer. New HI parameters, an age distribution and logistic regression model were determined, enabling a higher accuracy in predicting failure events and can therefore be utilized in various asset management initiatives and risk mitigation.

Acknowledgements

I wish to thank my supervisor Prof. Ian Jandrell and the research group for their advice and encouragement in this research work. Thank you to Thavenesen Govender for both his academic and motivational assistance. Without his constant motivation and encouragement, I may not have completed this work.

I would like to acknowledge the input from my mentors: Piet Goosen and Roger Cormack, who gave me the support I required both technically and emotionally, as well as my colleagues who are always ready to provide assistance: Sidwell Mtetwa, Nkosinathi Buthelezi, Adesh Singh, Khayakazi Dioka and Carl Wolmarans.

A special thanks goes to Jared for his encouragement and belief in me.

The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data.

~ John Tukey

Contents

List of Figures

List of Tables

Abbreviations

2-FAL	Furfural
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
AUC	Area Under the Curve
BIL	Basic Insulation Level
CCRA	Condition, Criticality and Risk Assessment
CDF	Cumulative Distribution Function
DBDS	DiBenzyl DiSulfide
DC	Direct Current
DGA	Dissolved Gas Analysis
DP	Degree of Polymerisation
DV	Dependant Variable
EAC	Equivalent Annualised Cost
EPP	Emergency Preparedness Plan
FN	False Negative
FP	False Positive
GLM	Generalised Linear Model
HI	Health Index
HMM	Hidden Markov Models
HV	High Voltage
IFT	Inter-Facial Tension
IV	Independent Variable
MAR	Missing At Random
MCAR	Missing Completely At Random
OEM	Original Equipment Manufacturer
OLTC	On Load Tap Changer
PDF	Probability Density Function
ppm	parts per million

pu	per unit
ROC	Receiver Operating Characteristic
SFRA	Sweep Frequency Response Analysis
TDCG	Total Dissolved Combustible Gas
TN	True Negative
TP	True Positive

Chapter 1

Introduction

1.1 Power transformers in the electric utility

Transformers are considered the most crucial and expensive piece of plant within a transmission system. Most transmission systems currently have large populations of aging transformers. With the growing demand for electricity, the loading of transformers is increasing. Current economic strategies call for reduced maintenance as well as capital expenditure. These challenges, which face utilities world-wide, necessitate improved management of transformers. The impact of a power transformer failure can be catastrophic. It would therefore be beneficial to know the risk status of the transformer population in order to facilitate better asset management, optimisation of maintenance, refurbishment and replacement strategies which will ensure maximum asset utilisation and minimise system risk. Statistical analysis on historical transformer failure data is therefore required, in order to obtain a model to determine the probability of failure of the transformers currently in service.

1.2 Transformer population in Eskom

Currently, Eskom Transmission has over 550 power transformers in service. It is of interest to see the age profile of the existing population, in terms of years since manufacture, since reliability is often related to age. From an analysis performed on insurance claims for transformers, it was found that the average age of failure was approximately 15 years [9]. The expected design life of a power transformer is 40 years. This figure is reduced, sometimes substantially, depending on the utilisation of the transformer, i.e. its loading or the environment to which it is exposed.

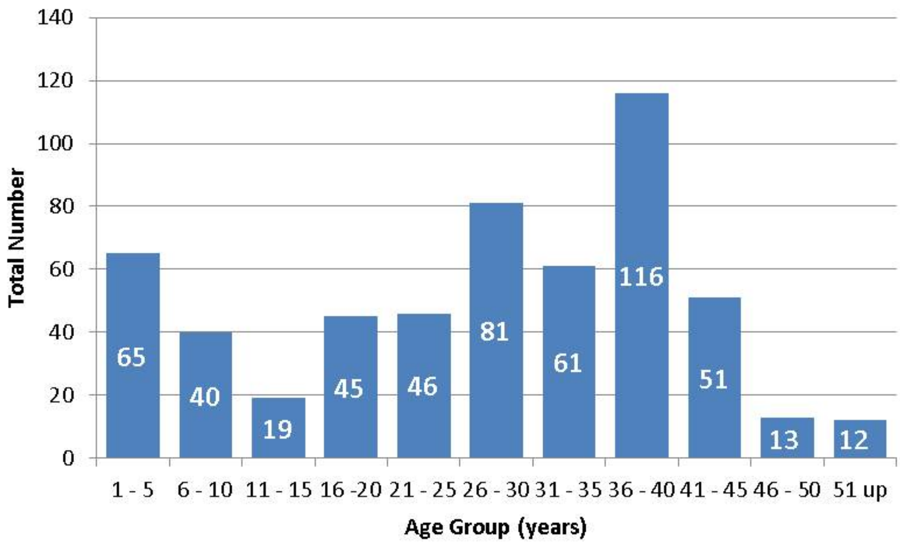

Figure 1.1: Breakdown of transformer age

The breakdown of age within the population of Eskom Transmission transformers is shown in Figure 1.1. From this graph it is evident that the majority (61%) of the population is above the age where probability of failure would be expected to be high and 14% is beyond the expected design life. The average age of transformers in Transmission is currently 29 years, showing an aging, high risk fleet, according to traditional thinking.

With the increased age of the transformers and the following additional risk factors [9]:

- Increased utilisation of equipment
- Deferred capital expenditure
- Reduced maintenance expenses
- Increased power consumption/load demand

it seems inevitable that the transformer failure rate within Transmission can be expected to increase rapidly in the near future. Therefore, it seems prudent to investigate the development of a model that will enable determination of probability of failure of these transformers, and hence improve asset management.

1.3 Asset Management model

Eskom has adopted an asset management tool for optimising asset replacement strategies. The Condition, Criticality and Risk Assessment (CCRA) model enables the phasing and structuring of project plans based on suitable cost-to-benefit ratios. A complete CCRA is required to avoid replacement strategies based on age and condition alone, in an attempt to reduce failure rate. The purpose of this model is to incorporate the consequence of failure as well as the risk into the investment decision making process.

1.4 Previous work: Failure analysis models

There are currently three main methods for determining the probability of failure of power transformers:

- Life estimation based on physical age
- Remaining life estimation based on Degree of Polymerisation (DP), hotspot temperatures and loading
- Analysis of condition data

Various statistical and analytical methods are employed on these base data sets. These methods also vary between considering performance of individual transformers and performance of the transformer population as a whole [9] [10] [11].

The objectives of these models are to determine the number of failures expected per annum in order to determine the financial implications for insurance claims [9] or for determining the necessity of spares [11]. They have also been developed for the purpose of optimising maintenance strategies as well as refurbishment/replacement strategies [12] [13] [14].

1.5 Research objectives

The abovementioned analyses are limited since each model focuses primarily on one facet of the status of the transformer population in question. Each method provides compelling results which can be used for maintenance, asset management and strategic spares holding strategies. The methods that involve detailed knowledge of the transformer's loading, hotspot temperatures, moisture content, etc. since the date of manufacture are unsuitable for the analysis of the existing historical data currently available for the Eskom Transmission population of power transformers.

Methods involving the analysis of age alone are also not considered effective for Transmission's population, since the environmental conditions and loading of these transformers varies throughout the population and therefore, the aging of the transformer, electrically and mechanically, also varies. Age alone cannot be considered, since that would imply that a well-maintained transformer that has been kept in stores since its manufacture will have the same probability of failure as a fully loaded transformer that has been in service for the same number of years. This is clearly not correct.

Methods combining a Bayesian approach and classical statistical analysis of historical failure data are to be considered, since these methods correspond to the methods already being employed in the CCRA model and require the use of data that is readily available. The following analyses comprise the main objectives of this research:

- Verification of the probability of failure based on age currently being utilised in the CCRA model
- Evaluate the significance of the condition data currently being used to

determine the HI in the existing CCRA model. This is to investigate the impact of removing superfluous variables and simplify the HI

- Develop the probability of failure based on condition data to be used as the HI modifier in the CCRA model, based on statistical analysis of Eskom Transmission historical data

The probability of failure based on HI can then be utilised for both the CCRA model, as well as the strategic spares management model.

1.6 Outline of dissertation

This dissertation is structured as follows:

- **Chapter 2: Transformer life assessment**
  The various parameters used in assessing remaining life and condition of a power transformer are outlined in this chapter. Existing methods of ascertaining remaining life and health indexing are discussed.

- **Chapter 3: Condition, Criticality and Risk Assessment Model**
  This chapter is concerned with outlining the existing asset management model. Failure modes and mechanisms, as well as how the probability of failure is related to risk is shown.

- **Chapter 4: Data Management**
  The most important aspect of any statistical analysis is data integrity. In this chapter the various methods used for processing and analysis of the data set used in this study is outlined.

- **Chapter 5: Statistical analysis**
  This chapter provides background into decision theory and the use of statistical analysis, in particular multiple logistic regression, in decision making. The various statistical assumptions and test methods are introduced.

- **Chapter 6: Prediction model**
  The results of the analysis performed on the empirical failure data are discussed in this chapter.

- **Chapter 7: Conclusion and recommendations**
  This chapter summarises the main conclusions drawn from this research work and makes recommendations for future work.

Chapter 2

Transformer life assessment

2.1 Life assessment methods

In the previous chapter the need for an accurate model to determine the probability of failure of a power transformer, relative to its condition was identified. In order to do this, it will be necessary to investigate the various known methods for assessing transformer life and evaluating/monitoring its condition.

There are various ways in which the life of a power transformer is estimated. These are linked to the aging mechanisms of the transformer. The main component determining the remaining life of the transformer is its insulation system. In a power transformer the insulation system is a combination of solid (paper and pressboard) and liquid (insulating oil) insulation. The properties of this insulation combination differs from the properties of the paper and oil separately. The chemical composition of the paper insulation is as shown in Figure 2.1.

Due to the construction of a power transformer, it is expected that once the insulation has reached its end-of-life, so has the transformer. For this reason, life assessment of the transformer becomes life assessment of its insulation system. While there are regeneration methods available that can be used to improve the quality of the oil, once the paper has deteriorated, it can neither be regenerated nor replaced.

The insulating oil serves as a diagnostic tool, as well as an insulating medium.

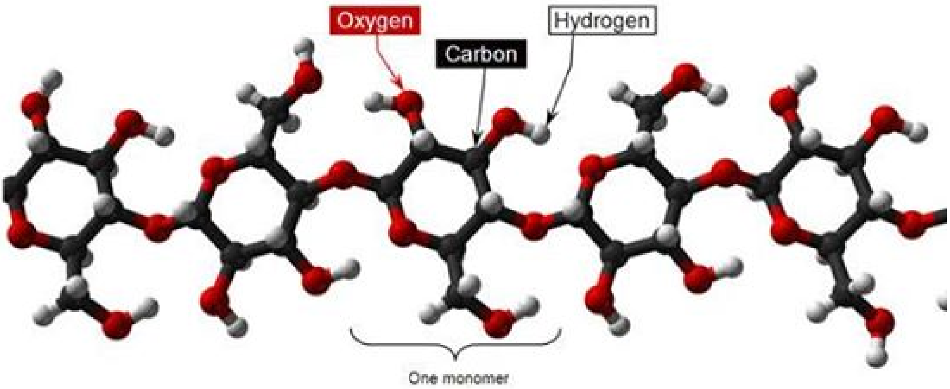

Figure 2.1: Chemical composition of paper insulation [1]

The analogy is often made between transformer oil and blood. Like the blood that is continually flowing and in contact with all parts within a human body, the insulating oil is also continually flowing and in contact with all components within the transformer. A sample of blood can be taken and analysed for pathogens and such a test can assist in the diagnosis of disease. Similarly, any incipient fault within the transformer will result in a chemical breakdown of the oil and/or paper, depending on the nature and severity of the fault, and byproducts of these processes will become dissolved within the oil. The oil can therefore be sampled and analysed to assist in the diagnosis of transformer faults and aging of paper and oil quality. This is done through DGA, DP prediction from Furanic analysis and oil quality measures. Oil sampling, if done correctly, is a non-intrusive procedure and can be performed with the transformer in service.

2.1.1 Aging mechanisms

There are three main factors influencing the aging of cellulose [15]:

- **Thermolysis:** the effect of temperature, usually temperatures $> 120°C$
- **Hydrolysis:** the effect of moisture
- **Oxidation:** the effect of Oxygen

These three factors form a cycle of deterioration as they are both initiators and byproducts of the aging process. it is necessary to evaluate all currently utilised life estimation methods to determine which parameters can be used in predicting failure.

2.1.2 Life estimation

There are currently three main methods for estimating the remaining life or depleted life of power transformers. These methods are then used in determining the expected probability of failure. These methods include:

- Life estimation based on physical age
- Remaining life estimation based on Degree of Polymerisation (DP), hotspot temperatures and loading
- Analysis of condition data

Each of these methods has its benefits and shortcomings.

2.2 Life estimation based on physical age

The simplest, crudest way of assessing the remaining life of a transformer is to look at its age relative to its design life. This can be useful if the transformers operating conditions are well known, however caution must be exercised when making any conclusions based on age alone.

Various analyses have been carried out where a transformer's reliability is determined from the Probability Density Function (PDF), survival function and hazard rates of transformers. The PDF is used to depict probability of failure as a function of age. The survival function, also known as the reliability function, is the probability of a percentage of transformers surviving at least until a certain age. The hazard rate gives the failure rate distribution with age and it is used in the determination of the bathtub curve as shown in Figure 3.2.

One such method was carried out on the Eskom population of transformers in order to determine the probability of failure based on age and historical failure rate [16]. This analysis was then used to evaluate the effect that monitoring the condition of these transformers had on the overall failure rate.

Another statistical analysis based on age is carried out in [11]. Here, three steps in asset management are identified: risk analysis, condition assessment and life cycle decisions: whether to repair, replace or retire the transformer. In this method, the PDF is fitted to a Weibull distribution and analysed to determine the B-lives of the various transformers in order to compare reliabilities. An age at B-life B10 means that at this age 10% of the population will fail and 90% will survive, an age at life B50 is correspondingly the age at which 50% of the population will fail, or the average expected life of the population.

The probability of failure based on age is a conditional probability, e.g. the probability of a transformer failing at the age of 30, given that it has already reached the age of 30.

Weibull distributions, fitted to empirical failure data are commonly used to determine life expectancy. The difficulty with applying such a distribution to power transformers is that often the root cause of failure is not related to the normal aging of the insulation, but rather external factors. Very few failures recorded per annum, resulting in limitations in available data.

2.3 Degree of polymerisation

DP is an estimation of the remaining life of the paper insulation. With a breakdown of the cellulose molecules which make up the paper insulation, there is a reduction in the paper's tensile strength. This equates to a reduction in the paper's mechanical withstand and has very limited impact on its compressive strength and dielectric strength [17]. This means that transformers with a low DP are much more susceptible to failure during short circuit incidents than those with higher DP numbers.

The remaining life of a transformer can be calculated from the temperature and loading of a transformer [18], or by determining the aging rate from the moisture content in the paper [19]. A general rule-of-thumb is that for every 6°C above 98°C that the transformer is operated, the insulation life is halved [18]. This is expected due to the aging/temperature relationship given by the Arrehnius-Dakin equation [20].

A by-product of the aging process of the paper insulation is Furfural (2-FAL). This compound becomes dissolved within the oil, since the paper is continuously in contact with the oil. A sample of insulation oil can be taken and analysed and the concentration of 2-FAL determined in order to approximate the DP of the insulating paper [21]. This approximation of DP is commonly used since direct measurement of DP requires a sample of the paper, which is an intrusive test and in most situations impossible to obtain. Oil sampling is non-intrusive and can be done while the transformer is in service. It is widely accepted that a DP of 900 indicates new insulation, while a DP of 200 indicates paper that is at the end of its life [21].

The expected remaining life can be calculated by taking the rate of change of DP as well as the relative saturation and temperature into consideration [15].

Vashishtha et al [22] investigate methods of determining the remaining life of a transformer based on the moisture content of the paper insulation and the effect of moisture on the rate of decrease of DP, i.e. the mechanical strength, of the paper insulation. This is a well-defined method and is also related to the loading of the transformer as outlined in [18]. Both methods do however require extensive knowledge and reliable history of the transformer to be kept. It is necessary to capture measurements of these parameters from the day of energisation of the transformer. This is a difficult task for a brand new population and an impossible one for an aged one.

Another method being utilised entails determining the change in DP over time by analysing transformer hotspot temperatures [8], as per [18]. In this method, change in DP versus time is modelled, including the errors associated with the uncertainties in measurements of both hotspot temperature and DP.

The actual DP value, as well as the cut-off threshold are known with limited accuracy, since exact determination of DP is an intrusive test. The likelihood of obtaining a representative sample of paper in order to perform the test within the laboratory is also low. A drawback of DP measurements is the uncertain results they produce due to different structures, manufacturers, loadings, and maintenance histories, as well as interference of measurements [23]. This method is however focused on population reliability rather than individual transformer reliability and the decrease in probability of failure that can be expected should some of the transformers within the population be replaced with new ones.

In [23] a method is described where measurements of DP, and condition parameters are taken during scheduled maintenance activities. A conditional probability of failure is then determined by using an Artificial Neural Network (ANN) to correlate the insulation degradation with measurements taken during maintenance activities.

A method used to optimise maintenance schedules is developed in [12], based on the probability of failure of the components of a transformer, such as paper insulation, bushings and tapchangers, etc. and the impact such a failure would have on the system as a whole. Occasionally, the probability of failure of a transformer is extremely low just before the failure occurs. This occurs when a large deviation from the expected behaviour occurs towards the end of the lifetime of the transformer. For this reason, the transformer health data is based on extreme value distribution theory and Monte Carlo simulations are used for this method. In this way, the probability of failure is the ratio of the number of transformer failures and the number of simulations.

2.4 Oil quality

Various tests can be performed on the insulating oil to determine its condition. The most commonly used methods include: moisture content, IFT, dielectric strength, acidity, tanδ, colour/appearance and sludge.

2.4.1 Moisture content

A relationship exists between moisture in oil and moisture in paper. Moisture migrates between the paper insulation and the insulating oil. The transformer's paper insulation is highly hygroscopic and as a result most of the moisture within the transformer will be found within the paper. At higher temperatures, the moisture tends to move out of the paper and into the oil. For this reason, it is crucial to note the transformer oil temperature at the time of sampling the oil in order to have any meaningful estimatation of the moisture content in the paper. The relationship between moisture in oil, moisture in paper and temperature is shown in Figure 2.2.

Moisture in the transformer is an important parameter to control. It is both an initiator and a byproduct of the aging process. Care must be taken to ensure the moisture levels within the transformer are kept at a minimum, with the use of silicone breathers and online drying systems.

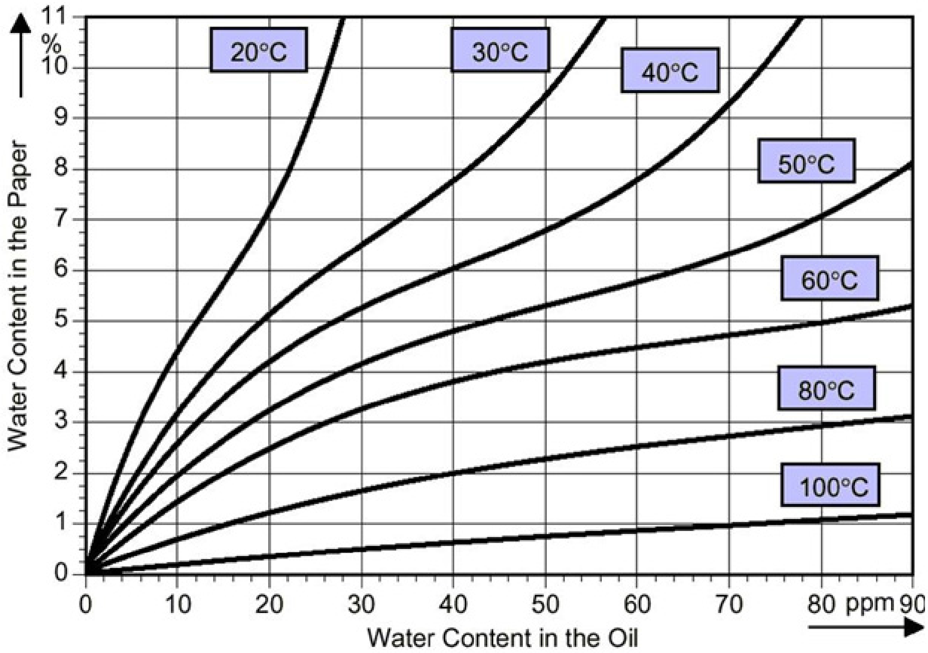

Figure 2.2: Equilibrium curve showing relationship between moisture content in oil and paper at different temperatures [2]

Moisture in oil originates from both internal, aging processes and also from the external environment. Moisture content in oil can be reduced by purification methods, such as: hot oil circulation, oil dehydration, degasification and filtration. These methods are however, not effective for the reduction of moisture content of the cellulose insulation.

2.4.2 Interfacial tension

Dissipation factor is a measure of the physical properties of the oil and is measured by measuring the surface tension of the oil against that of water. A high value of IFT is expected for new oil, while a low IFT is indicative of deterioration or oxidation of the oil.

An IFT of >35 mN/m is expected for new oil and should be monitored regularly if this value drops below 33 mN/m. Once IFT drops to <25 mN/m, action should be taken.

2.4.3 Dielectric strength

The dielectric strength of the oil is the oil's ability to withstand electrical stress. Moisture and fine particles within the oil result in a decreased dielectric strength. Dielectric strength is measured by applying a power frequency voltage between two submerged electrodes, as shown in Figure 2.3.

Dielectric strength values > 60 kV$_{rms}$ are considered acceptable for in service transformers. Dielectric strength can be improved by purification or filtration processess.

2.4.4 Acidity

The presence of acidic compounds is an indication of oxidation of the oil. High values of acidity will result in the oil becoming corrosive. New oil should not contain acidic compounds and should have a neutralisation number <0.01 mg KOH/g. Once the acidity increases to >0.2 mg KOH/g, then the acid should be removed by regeneration processes, or replaced, depending on other quality parameters.

Figure 2.3: Laboratory measurement of dielectric strength

2.4.5 Dissipation factor

The dissipation factor is a measure of the loss angle, or the percentage leakage current flowing through the oil under high voltage stress. New oil should have a low tanδ, which increases with deterioration of the oil, due to impurities and oxidation of the oil. A tanδ value of less than 0.005 is expected for new oil.

Oil with high tanδ values can only be improved by oil filtration or regeneration processes.

2.4.6 Colour/appearance

This is not a critical test, but it is useful as a quick, comparative evaluation of oil deterioration. An oil that is green colour indicates the presence of arcing, as expected in a tapchanger sample. As the oil colour varies from clear to yellow, to orange, to brown and black, the condition varies from new oil, to extremely bad/deteriorated or contaminated oil. The presence of free water, insuluble sludge, carbon, cellulose or fibres, dirt, etc. will usually give the oil a cloudy appearance.

Figure 2.4: Variations in insulating oil colour, for different oil conditions

2.4.7 Sludge

Sludge is a thick sediment or deposit which is primarily formed during the oxidation of the transformer oil. It comprises an insoluble, resinous, polymeric substance that is conductive, hygroscopic and is a heat insulator. The presence of sludge reduces the oil's dielectric strength and can aggressively increase the aging rate of both the oil and the cellulose insulation. Sludge accumulates on the cooling surfaces and between the windings which can restrict efficient cooling of the transformer [1].

The presence of sludge is undesirable and can be removed with hot oil, filtration and chemical processing of the oil.

2.5 Analysis of condition data

There are primarily three ways in which to assess the condition of a power transformer. These include:

- Dissolved Gas Analysis (DGA)
- Electrical testing
- Visual inspections

With these methods, incipient faults and defects can be identified.

2.6 Dissolved Gas Analysis

DGA is a well established diagnostic method, where oil samples are taken routinely and the composition of the gases dissolved within the oil analysed. There are mainly eight gases of interest for diagnostics, outlined in 2.1

Table 2.1: Diagnostic gases

Gas	**Chemical formula**
Hydrogen	H_2
Methane	CH_4
Ethane	C_2H_6
Ethylene	C_2H_4
Acetylene	C_2H_2
Carbon Monoxide	CO
Carbon Dioxide	CO_2
Oxygen	O_2
Nitrogen	N_2

The gas concentrations are measured and a trend of gas production is recorded. A developing fault within the transformer will lead to larger quantities of gases being generated and dissolved within the oil. The speed at which the fault gases are produced is an indication of the magnitude of the fault. For this reason, the rate of production is often considered to be of greater importance than the actual gas concentrations. The fault gases are also generated in small quantities during normal operation, due to natural aging. For this reason, relatively high concentrations of gas can also not be alarming if those concentrations are not increasing rapidly.

There are various methods of interpreting the gas results. These include well known methods such as: Total Dissolved Combustible Gases (TDCG), Duval's triangle, Roger's ratios, etc. [24] [25]. An additional method developed by analysing the DGA signatures of failed transformers and relating the identified patterns to gassing transformers in service to determine the fault cause and potential failure mode has been developed in [26].

The fault gases are produced in different quantities over a range of temperatures, due to the breakdown of the solid and oil insulation. An indication of how the gases are produced relative to temperature is shown in Figure 2.5. The temperature of the fault is an indication of the type of fault present. A partial discharge fault will generate relatively low temperatures and therefore, the main gases indicating this type of fault are Hydrogen and Methane. Arcing will generate extremely high temperatures and therefore, Acetylene is the prominent gas for this type of fault. Bare metal faults generate heat in the range 300-700°CṠo in this case, the gases that are expected to present would be Ethane and Ethylene. Since the paper is made of cellulose molecules, it is expected that any fault involving paper covered components would yield higher concentrations of Carbon Monoxide and Carbon Dioxide. CO/CO_2 is a ratio that can be used to determine whether the fault is thermal or dielectric [25]. Since these gases are naturally occurring within the atmosphere, some caution must be exercised when interpreting these results.

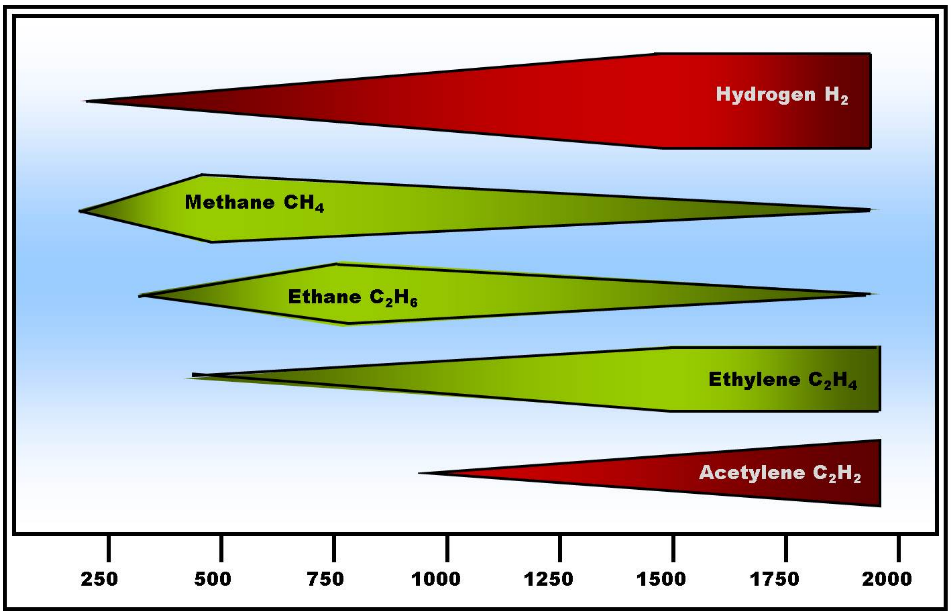

Figure 2.5: Gas concentrations at different temperatures [3]

Pathak et al [10] investigate failure probability and expected time to failure using Hidden Markov Models (HMM) and TDCG of oil samples. The disadvantage of this method is that only TDCG condition data is considered. This is of limited value in DGA since it accounts for the total concentration of all combustible gases and the relative concentrations and rates of production are not clearly visible and hence substantial detail required for a thorough analysis is missing. In general TDCG is not considered a good indicator of condition.

In work done by Cigré working group A2-111 [27], it was found that reliable statistical data is often difficult to get since relatively few failures of power transformers are experienced over long periods of time. A method of life determination based on measurements of condition data and mechanisms of breakdown and failure, established from physics and chemistry tests, has been developed. The condition data is obtained from forensic analysis of failed units on site and a "Degradation Model" based on DGA and oil condition. A "Life Model" developed using the influence of temperature on the degradation of paper insulation and the model developed in [20] is then used. Failure rates based on condition and service time are then determined.

Failure analysis using three tests: oil analysis, Furan derivatives analysis and HMM analysis is described in [28]. Oil condition and diagnostic gases (used in DGA, laboratory and online monitoring), as well as Furanic analysis (DP) are used to determine performance while HMM are used to determine probability of failure. In this method, similar to [10], the state within the HMM is hidden and the outcome, which is dependent on state is visible. Hidden variables are co-related through a Markov process to determine the outcome, rather than independent of each other. In this way, the model is trained in a similar way to an ANN.

2.7 Electrical testing

The main tests of interest are listed in Table 3.2. These are tests performed to evaluate the components of the active part and identify defects.

2.7.1 Sweep Frequency Response Analysis

Sweep Frequency Response Analysis (SFRA) is a method used to evaluate the mechanical integrity of core, windings and clamping structures of power transformers.

This test is performed by injecting a variable frequency, low voltage signal over a wide frequency range, into each winding terminal of a transformer and measuring its transfer function. The frequency response is unique for each power transformer and is often referred to as a fingerprint due to the complex resistances, self-inductances, ground capacitances, coupling inductances and series capacitances that comprise the core/winding assembly. This fingerprint is then used as reference data to be compared with future measurements to identify possible faults [4]. An example showing a healthy transformer and a transformer where the measurements do not correspond to the fingerprint, indicating the presence of a fault is shown in Figure 2.6

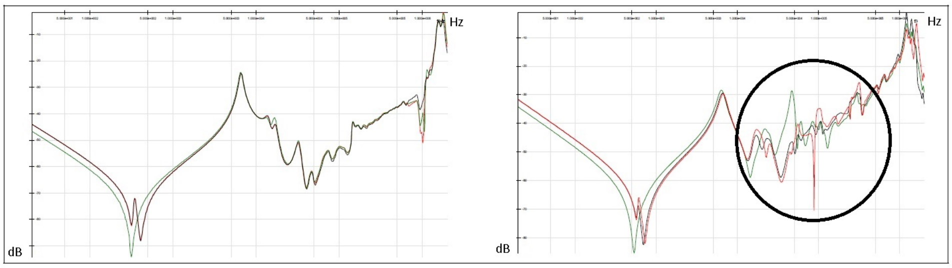

Figure 2.6: SFRA results: Healthy transformer(left), faulty transformer(right) [4]

Interpretation of results requires expertise and experience and results should be reviewed with caution since this is a comparitive test and somewhat subjective. It is also necessary for the tester to be experienced when performing this test since the test procedure is very specific and various test conditions: test leads, grounding, noise and interference, etc. can affect the results, and hence repeatability can be jeopardised and lead to confusion when interpreting results.

2.7.2 DC resistance

DC resistance tests are performed in the factory, to determine the I^2R losses and end temperature in a temperature rise test, and in the field for assessing possible damage. The purpose of this test is to identify faults that occur due to poor design, assembly, handling, poor environments, overloading or poor maintenance [1].

The resistances of different windings are calculated by injecting a DC current, in the range of 10-20 A depending on transformer size, into the test winding and measuring the volt drop across the winding. The results of the different windings are then compared with each other to identify differences in the windings and any loose connections. If the transformer is fitted with a tapchanger, tapchanger faults will also be identified.

Resistance measurements are made phase to phase and if the readings are within 1% of each other, then the test results are considered acceptable.

2.7.3 Dissipation factor and winding capacitance

This test is performed to provide information about movement and leakage losses within the power transformer. This test is used to check the integrity of the insulation between windings and earth and check for the presence of contaminants [1]. Ideal insulation will have only capacitive current but due to aging and impurities in the insulation, leakage current will begin to flow. This current has both capacitive and resistive components.

A test voltage is applied across the winding under test and the tanδ value is calculated from the ratio of resistive to capacitive components of current as shown in Equation 2.1. The test voltages are then increased to 1.2 and 2 pu. The angle defined by δ is known as the loss angle.

$$\tan\delta = \frac{I_R}{I_C} \tag{2.1}$$

The results are evaluated in two ways. The first method is the comparison of test results with previous test results to determine if deterioration of insulation has occurred. The second is to examine the results over the range of test voltages: deteriorated insulation will have a tanδ value that increases with increasing voltage and good insulation will have approximately the same values over the entire voltage range.

2.7.4 Insulation resistance

The insulation resistance test provides information about the core circulating current and unintentional short circuits across the insulation. This test is performed by short circuiting all untested terminals and applying a voltage in the range of 500 V - 2.5 kV between the test terminal and earth for one minute [1].

The exact insulation resistance value may vary depending on various factors and two similar transformers can have completely different insulation resistance values. For this reason, there is no standard value for acceptance of this test, although it is commonly accepted that a value > 1 MΩ per kV is acceptable for units in service. For new insulation, values in the GΩ range are expected.

2.7.5 Infrared scanning

Infrared scanning provides information about external connections, internal connections, bushing oil levels, cooling system blockages and hot spots. This test should only be performed by experienced operators since background temperatures, emissivity values of different materials, etc. all play a crucial role in interpretation of results.

It is not the absolute temperatures of different equipment that is concerning, but often the differences between different components that is a strong indicator of a fault. IR scanning should be carried out regularly but particularly when equipment is heavily loaded.

2.7.6 Magnetising current

The magnetisating current test provides information about the presence of core faults, inter-turn faults or unintentional loops in the earthing structure. This test is performed by applying the test winding with balanced three-phase voltage while all other winding terminals are left open. The current in each phase is measured when the voltage is applied to each phase. The measurements are taken with three tapchanger positions: nominal, extreme positive and extreme negative positions.

This is a comparative test that can be done in three ways:

- Measured values are compared to previous test results
- Measured values are compared with those of a sister unit, one with the same design
- The measured values of each phase are compared with each other

The test is considered successful if the measured results are within 30% of the reference results.

2.8 Visual inspections

Visual inspections are performed routinely by the operator on site and as part of maintenance activities. These include checks and tests on the auxiliary components of the transformer.

2.8.1 Bushings

All bushings routinely have tanδ and capacitance measurements taken. The values of these measurements should be within acceptable limits as stipulated by the OEM.

The bushings are inspected for damage to the bushing bodies and insulator sheds. They are checked to be free of chips/tears, radial cracks, flashover burns, copper splash and copper wash. The cementing and fasteners of the bushings are checked to be secure. The bushings are checked for evidence of oil leaks and correct oil levels. Should any defect be identified, corrective action is taken.

2.8.2 Tapchanger

All tapchangers have the following checks performed:

- Oil tests: moisture and dielectric strength are checked to be within limits stipulated by the OEM
- Speed test: results are checked to be within the limits stipulated by the OEM
- Contact thickness test: results are checked to be within the limits stipulated by the OEM
- Transition resistance: results are checked to be within the limits stipulated by the OEM
- Number of operations: number of operations determine the maintenance intervals for the tapchanger

2.8.3 General inspections

The bushing-metal interfaces, gaskets, weld seals, flanges, valve fittings, gauges and monitors are checked for oil leaks and moisture ingress. Should any of these be identified, corrective action is taken.

The transformer tank, marshalling kiosk and tapchanger mechanism box are checked for rust or corrosion. Cabinets are checked for evidence of condensation, moisture or insect/rodent ingress. Weld seals, flanges, valve fittings, gauges and monitors are checked for rust or corrosion. Seals, condensation heaters and locking mechanisms are checked for damage.

The conservator/Oil Preservation System is inspected for rust, corrosion and paint damage on the tank body. The weld seals, flanges, valve fittings, gauges and monitors are checked for rust, corrosion and evidence of moisture ingress.

The cooling system is inspected for rust, corrosion or oil leaks on the body of the radiators or pipework. Fan and pump enclosures are checked to ensure they are free of rust, corrosion and oil leaks and securely mounted in position with no signs of vibration. Fan and pump bearings are inspected to ensure they are in good condition and fan controls are operating as per design.

The overall physical condition of the transformer is inspected to ensure that it is externally clean and corrosion free. The condition of all primary and secondary connections is checked. The condition of all monitoring, protection and control, pressure relief, gas accumulation and silica gel devices and auxiliary systems (including online DGA monitoring and drying systems) that are mounted on the power transformer, is checked. External evidence of overheating or internal overpressure are inspected. Maintenance and service records are checked.

2.9 Conclusion

In this chapter, the various methods currently in use for assessing transformer life and evaluating its condition were introduced. These parameters will be used in the development of a model for the determination of probability of failure based on HI.

Chapter 3

Condition, Criticality and Risk Assessment Model

3.1 Asset management model

An asset management model (CCRA) has been adopted by Eskom for optimising investment decisions. The outcome of this model is dependent on the determination of probability of failure based on a HI that is developed using the life assessment parameters discussed in the previous chapter. It is therefore necessary to evaluate the basis of this model, as well as the use of the HI parameters.

3.2 Overview of methodology

Eskom has adopted an asset management tool for optimising asset refurbishment/replacement/retirement strategies. The Condition, Criticality and Risk Assessment (CCRA) model enables the phasing and structuring of project plans based on suitable cost-to-benefit ratios.

The inputs to this model include plant age, condition/HI, probability of failure and consequences of failure. The output of the model is the overall risk allowing analysis of cost/benefits of replacement/refurbishment. The timing of replacement is optimised according to the strategy shown in Figure 3.1.

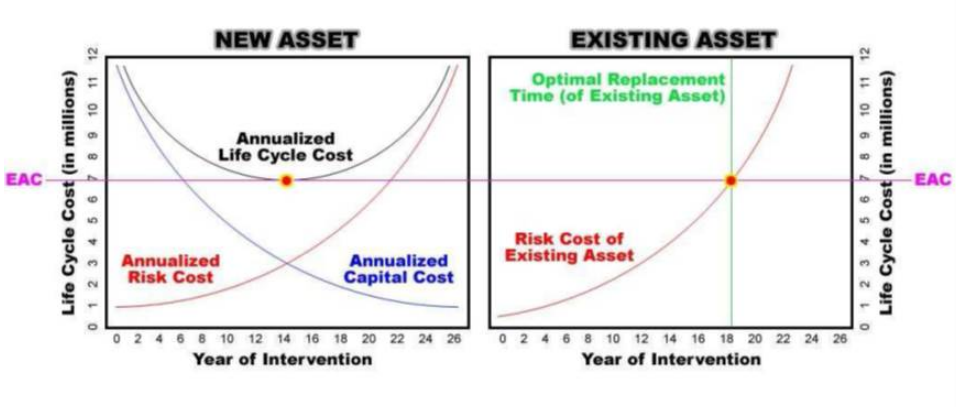

Figure 3.1: Optimisation of asset replacements [5]

Bathtub curves, as shown in Figure 3.2, developed from Weibull distributions depict the probability of failure of the plant as a function of age (P(age)). There are three distinct regions in a bathtub curve:

1. **Infant mortality**: indicated by a high failure rate in the first few years after manufacture, that decreases over time. These failures are generally attributed to inherent design defects or manufacturing errors.
2. **Random failures**: indicated by a flat region (constant failure rate) over the age range in the middle of the expected design life, where random failures are expected throughout the population.
3. **Wear out**: indicated by an increasing failure rate with an increase in age. This region has the highest failure rate. This area is related to wear out and is the area of most interest in this study.

A HI for each plant type is determined, based on the condition of the plant, relative to end-of-life. The HI is a quantification of condition measurements that are taken on the transformer and an overall score is obtained. This is then used to determine a probability of failure as a function of HI (P(HI)). Bathtub curves, developed from Weibull distributions depict the probability of failure of the plant as a function of age (P(age)). If P(HI) > P(age), then P(age) is modified according to P(HI), i.e. the bathtub curve becomes steeper, indicating an increased rate of aging, and a new probability of failure is determined as a function of both age and HI (P(age, HI)). This final probability of failure is

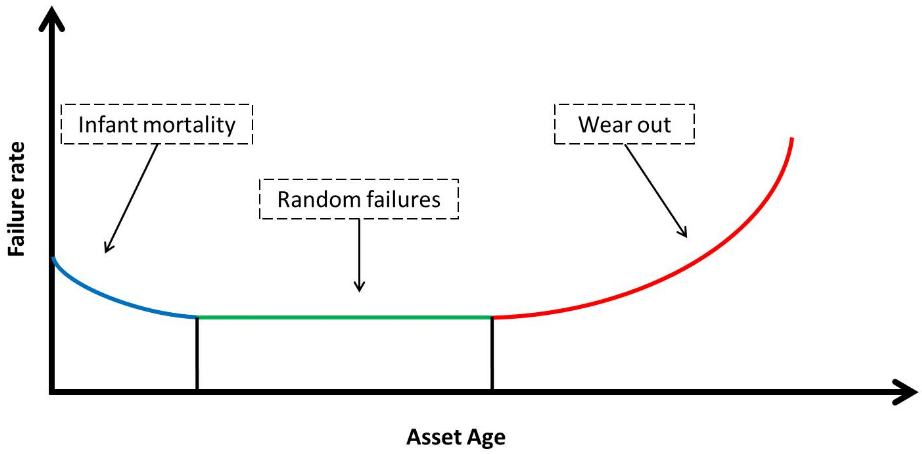

Figure 3.2: Example of equipment bathtub failure curve [5]

then used for further analysis. This then comprises the "Condition" portion of the CCRA model.

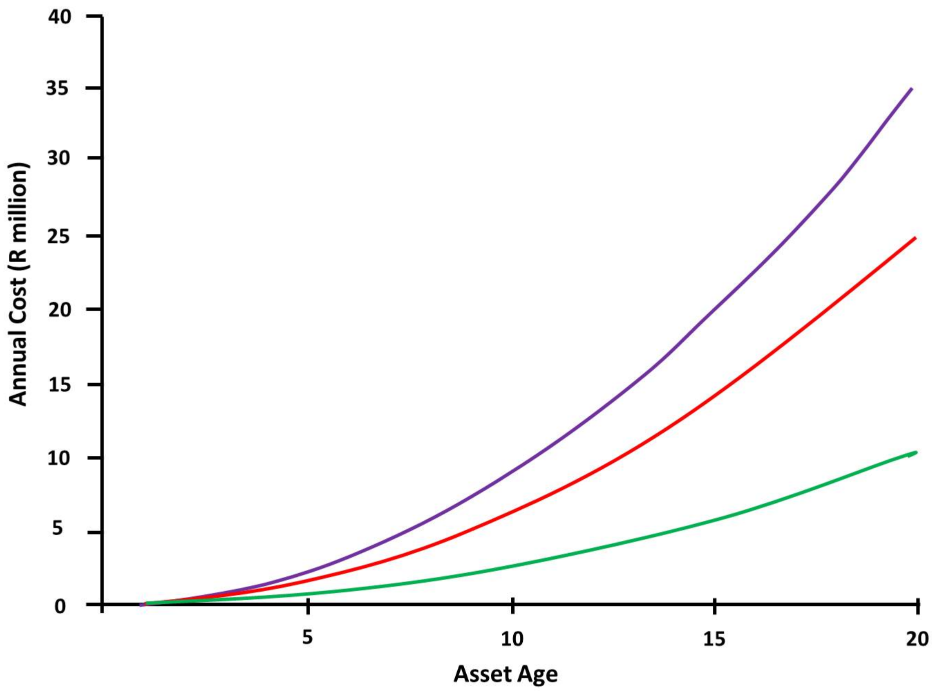

Figure 3.3: Effect of Health Index modifier on probability of failure curve [6]

The consequence (or impact) of failure of the plant is then determined, to ascertain the impact such a failure would have on the system, should it occur. This is done by analysing:

- the position of the plant in the network

- whether it has (N - 1) redundancy
- whether or not contingencies are available
- what consequential damage could occur or is expected should the failure of the plant be catastrophic
- the impact of adjacent plant failing at the same time, or during the outage of the initial failure

Different scenarios are drawn up for each failure event. These scenarios are developed from the Emergency Preparedness Plans (EPP) of each substation in which the transformers are installed. EPPs outline the processes to follow in the case of an emergency or loss of plant, to recover load. The consequences include the following:

- Replacement cost of failed plant
- Cost of emergency repairs, if applicable
- Cost of damage to adjacent plant
- Cost of unserved energy or customer interruption, should it occur
- Safety or environmental costs

Each scenario is then analysed, including its probability of occurrence. This then comprises the "Criticality" portion of the CCRA model and is a monetary value.

Risk is defined as the product of the probability of an event occurring and the consequences associated with that incident or, the frequency and severity of the losses [29].

A method similar to that employed in the CCRA model is that of a time-dependent failure probability, based on available condition data [13]. The primary objective of this method is for the optimisation of required maintenance interventions. A probability of failure based on the commonly used Weibull distribution is used to determine equipment time-to-failure. A Bayesian approach is used with this model due to limitations in empirical data available.

By using a Bayesian approach, uncertainties due to lack of information are expressed via probability distributions. Unknown parameters, such as shape β and scale η in the Weibull distribution are considered as random variables. In this way, Baye's theorem is used to determine the posterior probabilities based on condition data. This acts as a modification of the expected prior probability of failure that is assumed, based on age and failure rate, as done in [16].

$$Risk = Probability \times Consequence \tag{3.1}$$

Risk is defined as the product of probability and consequence or, the frequency and severity of the losses [29]. The risk of each item of plant is calculated from the probability of failure that was calculated as a function of age and HI and the consequence cost that was calculated from the impact of the failure. This comprises the "Risk" portion of the CCRA model and is also a monetary value.

Due to limitations in financial resources, it is necessary to perform some prioritisation in justifiable projects. This is done in order to justify spending alternatives. The benefit versus the risk of implementing a project this year rather than delaying it by a year or more is determined. In order to do this, the benefit to cost ratio is evaluated. The risk value is compared to the cost of refurbishment/replacement of the asset. If this ratio is very high, the refurbishment/replacement strategy is prioritised. If it is very low, the refurbishment/replacement strategy will more than likely be deferred to a later time. This is useful in determining the impact on the business of a decision to delay projects.

An accurately defined probability of failure, relevant to the population of transformers being analysed is critical, since it forms the basis of this calculation. With a probability of failure that is not representative of the population being analysed, the risk analysis is inaccurate and decision-making will be flawed and unoptimised. This leads to an increased risk of wasteful expenditure and an attempt to decrease the asset failure rate, rather than the real risk associated with the impact of the failure, which is the ultimate goal.

3.3 Failure definition

Asset management tools are refined based on probability and subsequent consequences of failure. Failure can range from anything between a minor defect which can be repaired on site, to a catastrophic event, necessitating the replacement of the failed transformer, as well as adjacent plant and possibly interruption of supply. Failure of a power transformer is therefore defined as per the definition of failure found in [30]:

> *"Failure: the termination of the ability of a circuit, bay or item to perform a required function"*

Failures are then separated into three levels which are defined as follows:

- **Severe:** The transformer requires replacement or removal from site to facilitate repair within a factory. In both instances, a new transformer will be installed to return the circuit to service.
- **Intermediate:** The transformer requires repair, but this can be implemented on site. This is usually intrusive work to restore the transformer to working condition and return it to service.
- **Minor:** these are trip events that remove the transformer from service temporarily. No work is required in order to return to the plant to service, since the transformer's major components have not been affected.

Failures of components that are critical to the operation of the transformer are also considered failures of the transformer. For example, the failure of an HV bushing, or an OLTC will be considered failures of the transformer since the transformer cannot operate in the absence of those components.

For the purpose of this study, only failures classified as severe or intermediate are considered since the minor failures have no direct impact on the end-of-life of the transformer, or the decision to refurbish/replace/retire it.

3.4 Modes of failure

As outlined in Cigré working group 12.05 report [7], international surveys are performed on failures of large power transformers and failure statistics are reported with 10 year intervals. These statistics are based on standard modes of failure with the following root causes:

- Core
- Windings
- Bushing
- Tapchangers
- Main tank and oil system
- Auxiliaries
- Other

The breakdown of failures, by cause, experienced within Eskom Transmission during the period 1996 - 2014 is as shown in Figure 3.4. From this graph, it can be seen that the vast majority (56%) of all failures experienced during this period is related to external components such as bushings and tapchangers.

3.5 Model parameters

The existing HI model for power transformers consists of a number of measurable condition parameters. These are separated into four categories, each with different weightings. The high level composition of the HI is shown in Tables 3.1, 3.2, 3.3 and 3.4.

Table 3.1 provides the parameters related to the most influential component of the power transformer impacting the end-of-life and expected remaining life. These parameters provide information about both the solid and liquid insulation.

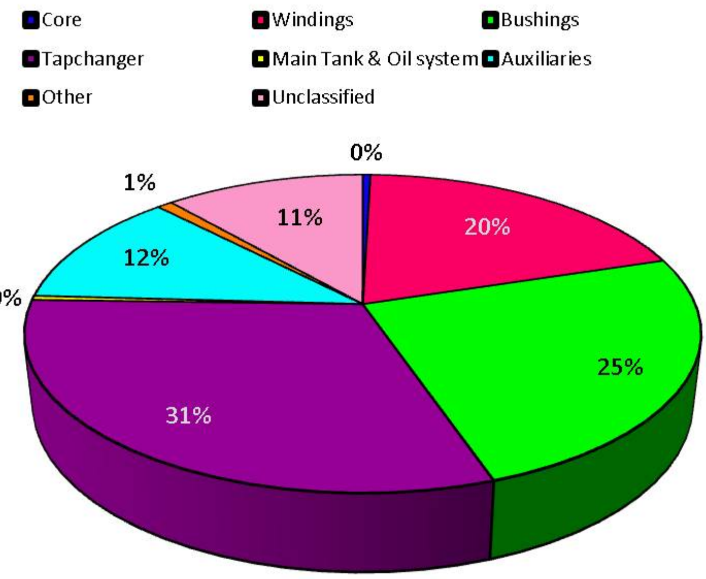

Figure 3.4: Breakdown of failures in Eskom Transmission as per Cigré reporting structure [7]

Table 3.1: Overview of power transformer HI insulation parameters

Insulation	**Oil test**
Furanic analysis	DP
Oil quality	Moisture
	IFT
	Dielectric strength
	Acidity
	Colour/appearance
	tan δ
	Sludge

Table 3.2 provides the electrical tests used to identify potential faults in various components of the active part of the transformer. Faults in any of these components can lead to catastrophic failure of the transformer.

Table 3.3 outlines the various visual inspections, checks and tests that are performed during routine inspections and maintenance activities. The purpose of these checks is to identify deterioration that has the potential of developing into a transformer fault. Any defects identified during these inspections can be corrected with maintenance activities.

Table 3.2: Overview of power transformer HI electrical test parameters

Electrical test	**Sub-test**
SFRA	N/A
DC resistance	N/A
tan δ and winding capacitance	N/A
Core insulation resistance	N/A
Infrared scanning	N/A
Magnetising current	N/A

Table 3.3: Overview of power transformer HI visual inspection parameters

Component	**Visual inspection**
Bushings	tan δ and capacitance
	Visual inspection
Tapchanger	Oil test
	Speed test
	Contact thickness
	Transition resistance
	Number of operations
Visual inspections	Oil leaks
	Conservator condition
	Cooling system condition
	Tank & overall physical condition

Table 3.4 shows the DGA results that are used in condition assessment evaluations. The combustible gas concentrations as well as the rates of production are of interest and are included.

Of the four HI categories, only two are used in this study, namely: Insulation and DGA. The reason is that this study is concerned with the determination of probability of failure relative to end-of-life of the transformer and not all of these components are related to end-of-life determination.

The primary purpose of the electrical tests is to diagnose faults once the transformer has been removed from service. A deviation in the test results is indicative of an immediate threat to the transformer and would be repaired as necessary and retested, prior to re-energisation. After such repair, the test results would again be satisfactory. For this reason, the test results are not

Table 3.4: Overview of power transformer HI DGA parameters

Parameter	**Gas**
Concentrations	H_2 (Hydrogen)
	CH_4 (Methane)
	C_2H_6 (Ethane)
	C_2H_4 (Ethylene)
	C_2H_2 (Acetylene)
	CO (Carbon Monoxide)
	CO_2 (Carbon Dioxide)
Daily rate of production	ΔH_2
	ΔCH_4
	ΔC_2H_6
	ΔC_2H_4
	ΔC_2H_2
	ΔCO
	ΔCO_2

considered useful in terms of condition monitoring or indicative of long term plant health deterioration. The requirement/reason for repair is more useful for identification of potential faults and investigation of root cause of failure.

The visual inspections are related to maintenance activities and identified deficiencies can be addressed relatively easily. Again, should any fault be identified within the tapchanger or bushings, these will be repaired/replaced. Although these components have a large impact on the failure of power transformers, the routine condition monitoring/checks are not useful from a modeling perspective since they have binary condition values rather than the continuous values which are indicative of slow deterioration.

3.6 Conclusion

In this chapter, the basis of the CCRA model was discussed. The condition parameters to be included, as outlined in more detail in the previous chapter were grouped into four categories for further evaluation and use in the determination of the probability of failure of power transformers. The probability of failure was identified as one of the critical inputs into the CCRA model, required for determining overall risk and optimisation of refurbishment strategies. In order to determine the probability of failure, a suitable statistical method to achieve this needs to be identified.

Chapter 4

Statistical analysis

4.1 Decision theory

All improvements and developments are made as a result of intelligent decision making. Decision making is the process of arriving at a conclusion or resolution based on the consideration of various alternatives. Decision making can only be improved with better understanding of the problem.

Decision theory is defined as the study of principles and algorithms for making correct decisions. While simple decisions can be made without a theory, often complex decisions, ones involving high risk, levels of uncertainty and time dependency, will require mathematical or statistical models to produce optimised outcomes. This is a probabilistic approach which moves away from heuristic decision making. The deficiency with this method is that uncertainty related to known unknowns is taken into account and the extreme influences of the unknown unknowns are ignored.

Although many robust statistical methods are available for mathematical approximation of real situations, the results should always be analysed critically. Limitations within these models are ever-present and should not be relied upon blindly or unquestioningly. This is unfortunately, common practice and is referred to as the Ludic Fallacy [31] and is advised against. Decisions should therefore be made with the best possible information at hand, while minimising risk (risk cannot be completely eliminated).

4.1.1 Decision making

In Bayesian decision theory or Bayesianism [32],

> *"the aim is to reduce a Decision Maker's incoherence, and to make the Decision Maker approximate the behaviour of the hypothetical Rational Agent, so that after aiding he should satisfy Maximizing Expected Utility"*

Bayesian decision making is summarised by the following four principles [33]:

- The Bayesian subject has a coherent set of probabilistic beliefs and these beliefs are in compliance with mathematical laws of probability
- The Bayesian subject has a complete set of probabilistic beliefs, all alternatives have a degree of belief
- The Bayesian subject changes their beliefs in accordance with their conditional probabilities, on presentation of new evidence
- The Rational Agent chooses the option with the highest expected utility

A statistical model is developed by gathering evidence and testing the model's effectiveness against a defined hypothesis. As new evidence is presented, belief in the hypothesis changes accordingly, as per point 3 above and the decision is potentially changed.

4.1.2 Decision classification

The purpose of the statistical model in this study is to assign various transformers to two different classes: healthy and failed, and to define their degree of belonging to each class. Each assignment is viewed as a decision and needs to be evaluated accordingly. This is done by means of a confusion matrix, as shown in Table 4.1. The risk associated with errors and misclassifications can be more severe in some instances.

In this case:

- a True Positive (TP) is a failed transformer correctly classified
- a False Positive (FP) is a healthy transformer classified as a failed one
- a False Negative (FN) is a failed transformer classified as a healthy one
- a True Negative (TN) is a healthy transformer classified correctly

Table 4.1: Confusion matrix for decision making

	Predicted	
Actual	**True**	**False**
True	TP	FP Type I error
False	FN Type II error	TN

The accuracy of the model is evaluated with Equations 4.1, 4.2 and 4.3 respectively [34]. Equation 4.1 gives an indication of the accuracy of the model in general.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4.1}$$

The Fall-out or False Positive Rate is given by Equation 4.2. The implication of a high Fall-out is a large number of healthy transformers being replaced unnecessarily. With the high capital cost of a power transformer, this can have a large negative financial impact on the power utility.

$$False\ Positive\ Rate = \frac{FP}{FP + TN} \tag{4.2}$$

The Miss Rate or False Negative Rate is given by Equation 4.3. The implication of a high Miss Rate, is not being able to pre-emptively replace the failed transformers and avoid the potential damage to adjacent plant, safety and environmental risks should the transformer fail catastrophically. Loss of income and reputation should power outages result from the failure are also a consequence.

$$False\ Negative\ Rate = \frac{FN}{TP + FN} \tag{4.3}$$

Both Fall-out and Miss Rate have negative implications that cannot be ignored. Analysis of the individual accuracies is however not conclusive in the final decision making process. These error rates need to be evaluated in terms of risk and not solely probability.

4.2 Methods of statistical analysis

Statistical analysis will be performed on a portion of the condition data set only, in order to determine the weightings, and statistical significance of each condition parameter. Once the significance of each parameter has been determined from historical failure data, analysis will be carried out on the data set containing relevant data about both failed transformers, as well as healthy ones. A sensitivity analysis will be performed to determine whether or not the model is better at predicting failure than simply predicting by chance.

Care must be exercised in determining the correct statistical test to use for a specific analysis. Depending on the type and number of both the dependent and independent variables in question, different tests and methods are used.

In this case, there is one dependent variable: transformer failure, which is a categorical, dichotomous variable since the transformer can only be in one state at a time. It can either be failed or healthy, not both simultaneously. There are multiple dependent variables and these variables, as outlined in Table 3.1 and Table 3.4 of the condition data are all continuous variables.

According to the summary of statistical tests that can be performed for any given analysis given in [35], there are two types of statistical models that can be developed that fit the data to be analysed in this study. These are: multiple logistic regression and discriminant analysis.

4.3 Multiple logistic regression

Multiple logistic regression is much the same as regular logistic regression, but with more than one independent variable used to determine the dependent variable.

This method assumes that there are only two groups to which an individual sample can belong. Each sample can only belong to one group at a time [36]. In this study, there are only two groups, namely: failed transformers and healthy transformers. A specific transformer can be classified as failed or healthy, not both simultaneously.

The purpose of logistic regression is to determine the linear relationship between y and x. Where y is the natural logarithm of the odds ratio of the probability of failure and x is the input vector $x = (x_1, x_2, x_3, ..., x_n)$ which contains different features of the individual [36]. The output y is not itself a useful quantity, but is useful in determining p, which is the probability of failure.

The logistic regression model equation is shown in Equation 4.4:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \tag{4.4}$$

and denotes the relationship between y and x. Where $(\beta_1, ..., \beta_n)$ are weighted coefficients that determine the magnitude of the influence each variable in x has on the populations of each group [37] and β_0 is the intercept.

Two probabilities are calculated:

- **Prior probability:** This is the probability that an individual is more likely to belong to one group than another [37]. This is useful in reducing the probability of misclassifying an individual. For example, at any point in time there is a much higher percentage of the transformer population that is healthy than has failed. Therefore, the probability of picking a healthy transformer at random from the current population is much greater than picking one that is about to fail.

- **Post probability:** This is the probability that an individual belongs to a specific group [37]. This is useful in determining the probability of a transformer failing, since for the transformer to belong to the group of failed transformers, the transformer would have had to have already failed, in which case prediction is not necessary as the outcome is already known.

The cost of misclassification can also be determined using this model. This is important since the model is not assumed to be ideal and misclassifications can be expected. The cost of misclassification will be different based on the individual that is misclassified. For example, the cost of classifying a high risk transformer as healthy will be far greater than the cost of classifying a healthy transformer as high risk. Care must also be taken to not overstate risk, since, while the cost of misclassifying a healthy transformer as high risk may be higher than the opposite, if too many healthy transformers are missclassified, the cost of any action taken to mitigate this risk could outweigh any potential benfit.

The model should also be tested in order to determine whether or not the classification based on p is more accurate/reliable than a classification that was made based on chance alone.

4.4 Statistical assumptions

The following assumptions that are required for normal Linear Regression and General Linear models, based on ordinary least squares algorithms are not required for Logistic Regression models:

- **Linearity:** no linearity between IVs and DV is required
- **Normality:** there is no requirement for IVs to be multivariate normal
- **Homoscedasticity:** random variables need not have homogeneity of variance
- **Variable type:** the logistic regression can handle nominal and ordinal data as independent variables, not only interval and ratio

For binary Logistic Regression, based on maximum likelihood estimates, the following assumptions are required [38]:

- The dependent variable must be binary
- Independence of error terms is required. There should be independence between data points
- No multicollinearity must be present, i.e. independent variables must be independent of each other
- Linearity between the independent variables and log odds is required, although it is not required between the independent variables and the dependent variable

4.5 Model evaluation/Goodness of fit

Various methods exist for evaluating the goodness of fit of a statistical model. The tests used for ordinary linear regression are different to those used for logistic regression. Interpretation of the various test statistics is different for both models. Goodness of fit of the models developed in this study is evaluated by examining both likelihood ratio tests and pseudo-R^2.

4.5.1 Data separation

The optimal evaluation of goodness of fit of a statistical model requires three data sets, namely: training data, validation data and test data [39]. The training and validation data sets comprise the data set used in developing the final statistical model. With these data sets, both the model inputs and output are known. The test data set is comprised of new data where only the inputs are known and the outputs are determined using the statistical model. These results are then analysed.

The training data is used with statistical methods to develop the prediction model. The model is then tested using the validation test set, where the

IV data are input into the prediction model and the output of the model is compared with the known outputs in order to determine the degree of accuracy of the model. Once the model has been found sufficiently accurate, the test data is input into the model and the results are then analysed. This process is outlined in Figure 4.1 below.

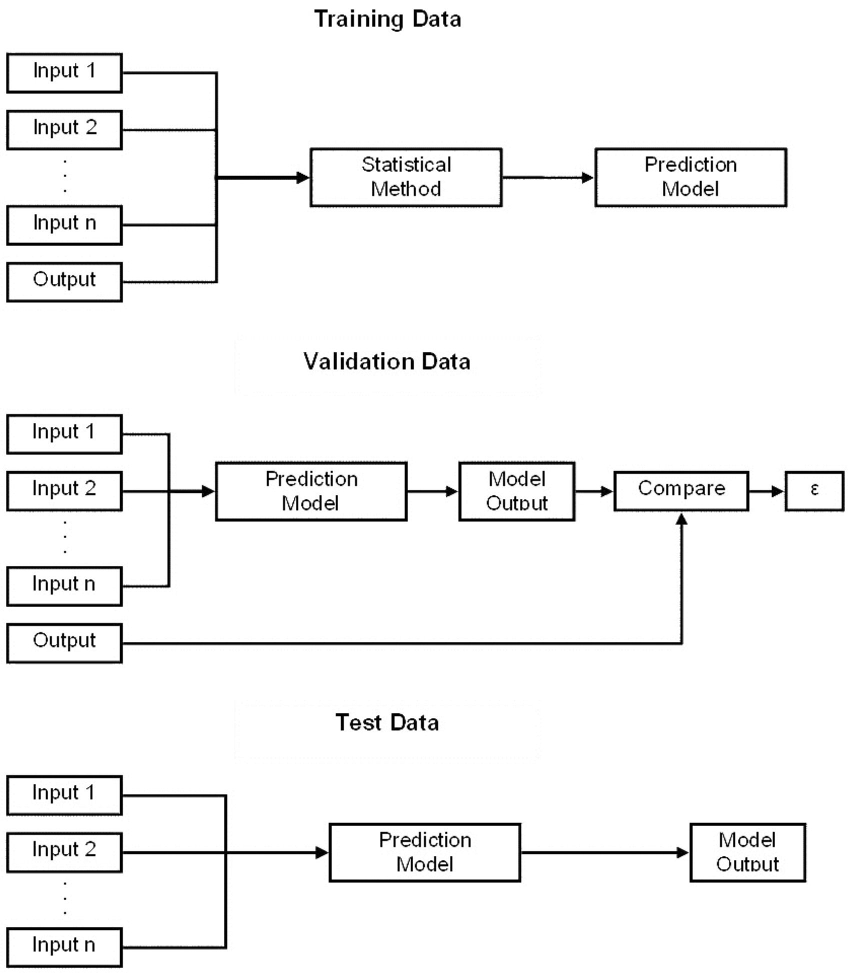

Figure 4.1: Relationship between different datasets and prediction model

This method is however used in the ideal case where a large data set is available. In cases such as this study, the limited size of the data set this methodology and other model goodness-of-fit methods will be utilised.

4.5.2 Deviance and likelihood ratio tests

The likelihood ratio test is shown in Equation 4.5 [38]. The deviance of the model is the difference between the fitted values and the expected values. For this reason, the smaller the deviance of the fitted model compared to the deviance of the null model, the better the model is at predicting the outcome.

$$D_{fitted} - D_{null} = -2 \ln \frac{likelihood\ of\ fitted\ model}{likelihood\ of\ null\ model} \tag{4.5}$$

where:

D_{fitted} is the deviance of the fitted model

D_{null} is the deviance of the null model

4.5.3 Pseudo-R^2

In a linear model, R^2 is used to evaluate goodness of fit of a model by the proportion of variance in the dependent variable explained by the independent variables. Since Logistic Regression is heteroscedastic, the proportionate reduction in error is not constant across the range of predicted outcomes. The R^2 statistic is therefore not interpreted in the same was as for linear regression where very low values are expected. [40].

The pseudo-R^2 statistic is the proportion of variance of the latent variable, inferred from other variables but not directly observed which is explained by the covariate, the variable affecting the relationship between the IVs and DV, and is shown in Equation 4.6.

$$R_L^2 = \frac{D_{null} - D_{fitted}}{D_{null}} \tag{4.6}$$

where D_{fitted} and D_{null} are as defined for Equation 4.5.

4.5.4 Hosmer-Lemeshow test

The Hosmer-Lemeshow test [39] is a χ^2 statistic that is calculated on data that are grouped into groups with approximately the same number of observations per group (usually 10 groups). This test statistic is given by Equation 4.7.

$$H = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{N_g \pi_g (1 - \pi_g)} \tag{4.7}$$

where:

O_g are the observed events

E_g are the expected outcomes

N_g are the observations

π_g is the expected risk

G is the number of groups

The disadvantages of this test are that is has a large dependence on the number of observations grouped, as well as the number of groups. This test also has reduced accuracy in predicting certain types of lack of model fit. For this reason, this test is not implemented in this study.

4.6 Conclusion

In this chapter, the need for statistical analysis in the decision making process required to implement the CCRA model, that was introduced in the previous chapter, was discussed.

An outline of the statistical method of Multiple Logistic Regression which was used for this study was outlined, along with the statistical assumptions made. All models require testing and various goodness-of-fit tests are available for evaluating logistic regressions. These methods are outlined in this chapter. The use of the deviance and likelihood ratio tests, as well as pseudo-R^2 tests are used in this study.

Statistical analysis is often an iterative process that is highly dependent on the integrity of the data used in the analysis. Methods of data processing are therefore introduced and applied to the data used in this study.

Chapter 5

Data Management

5.1 Source data

The source data to be used in this study was obtained from the Eskom databases used for storing reliability and condition data. Since this study involves the statistical analysis of power transformer failures, it was necessary to obtain historical data for both failed transformers, as well as "healthy" ones.

Power transformers do not have a high failure rate, with a high average failure rate being approximately only ten per annum. For this reason, failure data is scarce and in order to obtain a reasonable data sample, it was necessary to use data over an 18 year period (1996 - 2014).

Failure data was obtained from failure records and reports. Maintenance reports and factory records were also interrogated to obtain maximum information. Only failures satisfying the definition of failure as per Section 3.3 were considered.

The definition of a healthy transformer follows the assumption: if the transformer has been in uninterrupted service since 1999, then in 1999 it was healthy. All data for "healthy" transformers was therefore collected for transformers within the existing population from 1999. This presents some limitations, since complete historical condition data is not available for all transformers.

A total of 512 transformers were obtained through this process comprising 193 failed and 319 healthy transformers.

A number of challenges arise when handling data for analysis. These include: missing data, detection of outliers, transformation of data as well as sample bias. Methods of overcoming these challenges are outlined below.

5.2 Data types

Data can be classified into four main types as proposed by Stevens [37], namely: nominal, ordinal, interval and ratio. These are explained in Table 5.1 below.

Table 5.1: Steven's measurement system

Type	**Explanation**	**Example**
Nominal	There are a number of distinct categories that the variable can be classified into.	Names Religion
Ordinal	There are a number of distinct categories that the variable can be classified into, and the categories have a known order.	Service ranking Factory ratings
Interval (discrete /continuous)	This is an ordinal variable with an equal distance between successive values.	Temperature Calender dates
Ratio	Interval variables with fixed zero measurement points, hence preserving ratios independent of the unit of measurement.	Height Difference in time

The data used in this study are comprised of three of these data types. These include: nominal, interval and ratio. Each data type is handled differently since different information is available from each variable.

5.3 Missing data

When performing statistical analyses, the issue of missing data is always a concern. Due to the fact that this study is based on historical data (18 years old), it is apparent that some of the data required in this study will be missing. This may be due to a number of factors including: operating practices changing, carelessness of data storage and inability to recover certain portions of transformer records, etc.

There are three categories of missing data [41]: Missing Completely At Random (MCAR) where the data are missing independently of the DV or IV, Missing At Random (MAR) where the data are missing dependent on one or more IVs but independently of the DV, and non-ignorable where the missing data is dependent on both the IVs and DV. There are a number of ways of handling the missing data. These are outlined below.

5.3.1 Listwise/Casewise deletion

This is the simplest method of dealing with missing data. Any record that contains a missing variable is deleted from the sample. This can cause substantial reduction in sample size and lead to large biases [42].

5.3.2 Pairwise data deletion

With this method, data records with missing data variables are used in the analysis only when the analysis does not involve the missing variable. Again this method can produce large biases and unequal sample sizes [42].

5.3.3 Mean substitution

Mean substitution involves substituting each missing variable with the mean of all corresponding variables within the entire data set. This is problematic since it reduces the variance of the variables substituted in this way, which can lead to underestimating the spread of the data [43].

5.3.4 Hot deck imputation

In this instance, the record that is most similar to that with a missing variable is found and the value of the variable in this record that corresponds to the missing variable is substituted for the missing value [44]. The difficulty that arises is in defining similarity, since this is contextual and is therefore not a simple task [45]. This method also does not account for uncertainty in the approximation.

5.3.5 Regression methods

A missing variable is predicted using a regression model which is determined from the other complete variables, i.e. the missing variable becomes the response variable and the other variables become the predictor variables [38]. This leads to a complete data set with a reduced standard error.

5.3.6 Expectation maximisation and Raw maximum likelihood

These methods can be used to handle data that is MAR. In these methods estimates are found of the most likely value that the missing variable might have. A vector of means and a covariance matrix are developed that are superior to those that are developed from the previous methods of approximating missing data that are mentioned above [41], [46]. The disadvantage of these methods is that large sample sets are required.

5.3.7 Multiple imputation

Multiple imputation has a number of advantages over other methods of missing data approximation. This method involves having more than one estimate for a given missing variable, computed using other values within the data set. In this way, the variance between the estimates gives information about the uncertainty of the imputation. In this way, biases in the data are also reduced. Since there are a number of estimates for each missing variable, after multiple imputation, there are a number of complete data sets instead of just one. Each data set is analysed individually and the results are then compressed to form only one final solution [41]. This method is also disadvantageous since large data sets are required.

Since the data set available for this study exhibits non-ignorable missing data and is of limited size, the prefered methods of handling the missing data are not practical. For this reason, the listwise deletion method was utilised. This leads to significant reduction in sample size and the data needs to be evaluated for biasing.

5.4 Data visualisation

Visualisation of the raw data set is a simple method for determining the integrity of the data. By plotting the raw data, it is easy to visualise any discrepancies within the data, for example, if you have a dichotomous variable, but have values lying at points between 0 and 1, then it is obvious that the data set contains bad data. It is prudent to perform this quick check, prior to any data processing or analysis to confirm that the data is in fact as expected.

5.4.1 Gas concentrations and production rates

The concentrations of the different gases was plotted against the failures for the data set to be used in this study. Since the data was collected for only failed and healthy transformers, the failed variable can only have a value of either "0" (indicating a healthy transformer) or "1" (indicating a failed transformer)

The data is examined to determine conformity to an expected pattern. In this case, the data is expected to form two distinct groups, one at low concentrations of gas in healthy transformers and another at high concentrations in

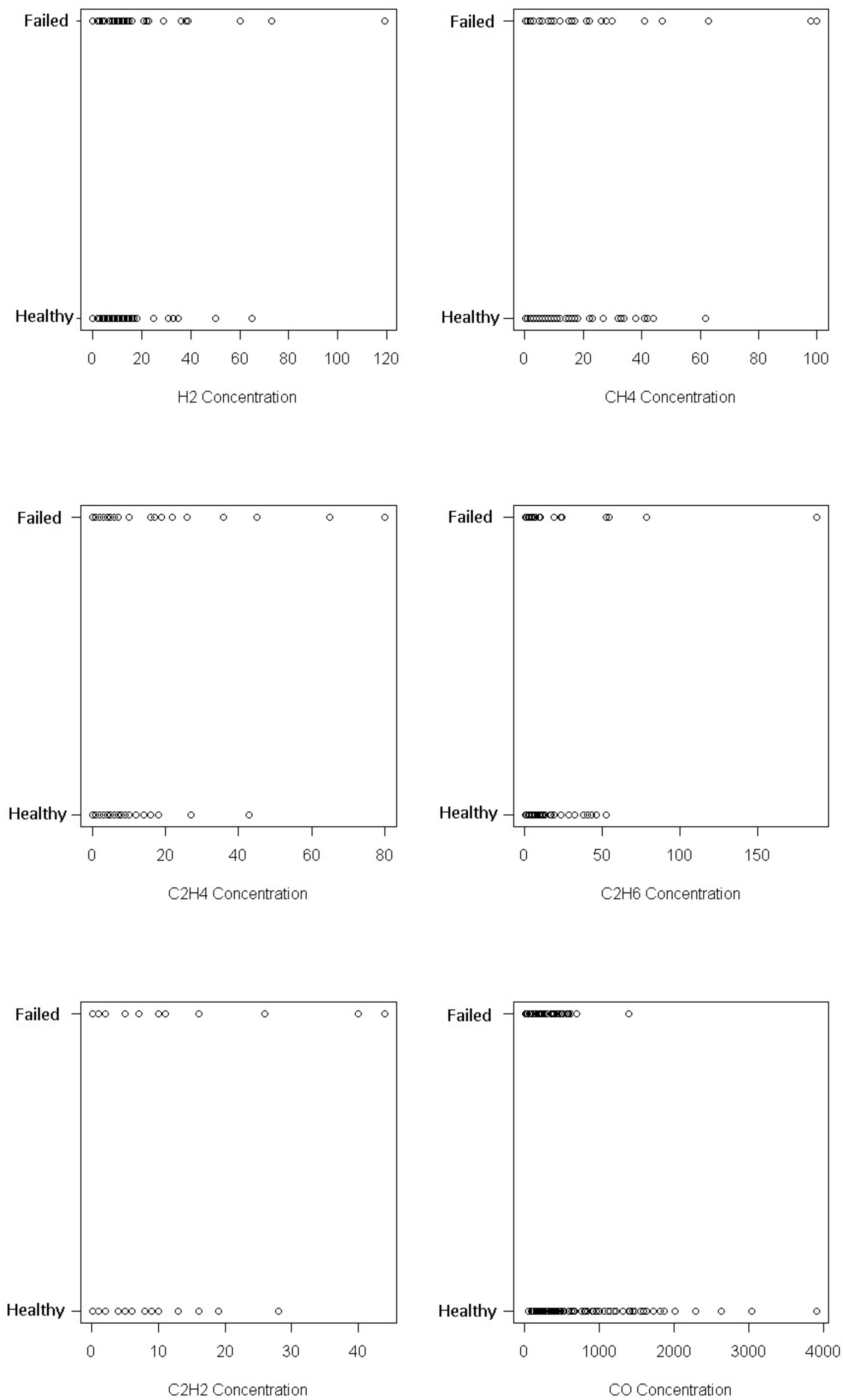

Figure 5.1: Plot of gas concentrations vs failures

failed units.

From the results shown in 5.1 it can be seen that, with the exception of Carbon Monoxide, the expected trend in gas concentrations is evident, with higher gas concentrations present in the failed transformers and lower concentrations in the healthy transformers. The opposite trend is present in the CO data, which is not expected.

The daily rates of production of the different gases was plotted against the failures for the data set to be used in this study. From the results shown in Figure 5.2 it can be seen that, with the exceptions of Carbon Monoxide and Ethylene, the expected production rates are evident, with higher production rates in the failed transformers and lower production rates in the healthy ones.

Similarly to the gas concentration data, the daily rate of production data is examined to determine conformity to an expected pattern. In this case, the data is expected to form two distinct groups, one at low rates of gas production in healthy transformers and another at high rates of gas production in failed units.

The trends of CO and C_2H_4 show higher rates of decrease in concentrations for failed transformers. This decrease in production rate of CO is congruent with the concentrations of the gas in failed transformers.

5.4.2 DP and oil quality

The values of the DP and oil quality tests were plotted against the failures for the data set to be used in this study.

The results shown in Figure 5.3 indicate that there are higher moisture and lower dielectric strength measurements in the sample of failed transformers as expected.

The measurements of acidity and DP indicate values that are fairly evenly distributed across their ranges for failed and healthy transformers alike. Lower DP values and higher acidity values in failed units would have been expected.

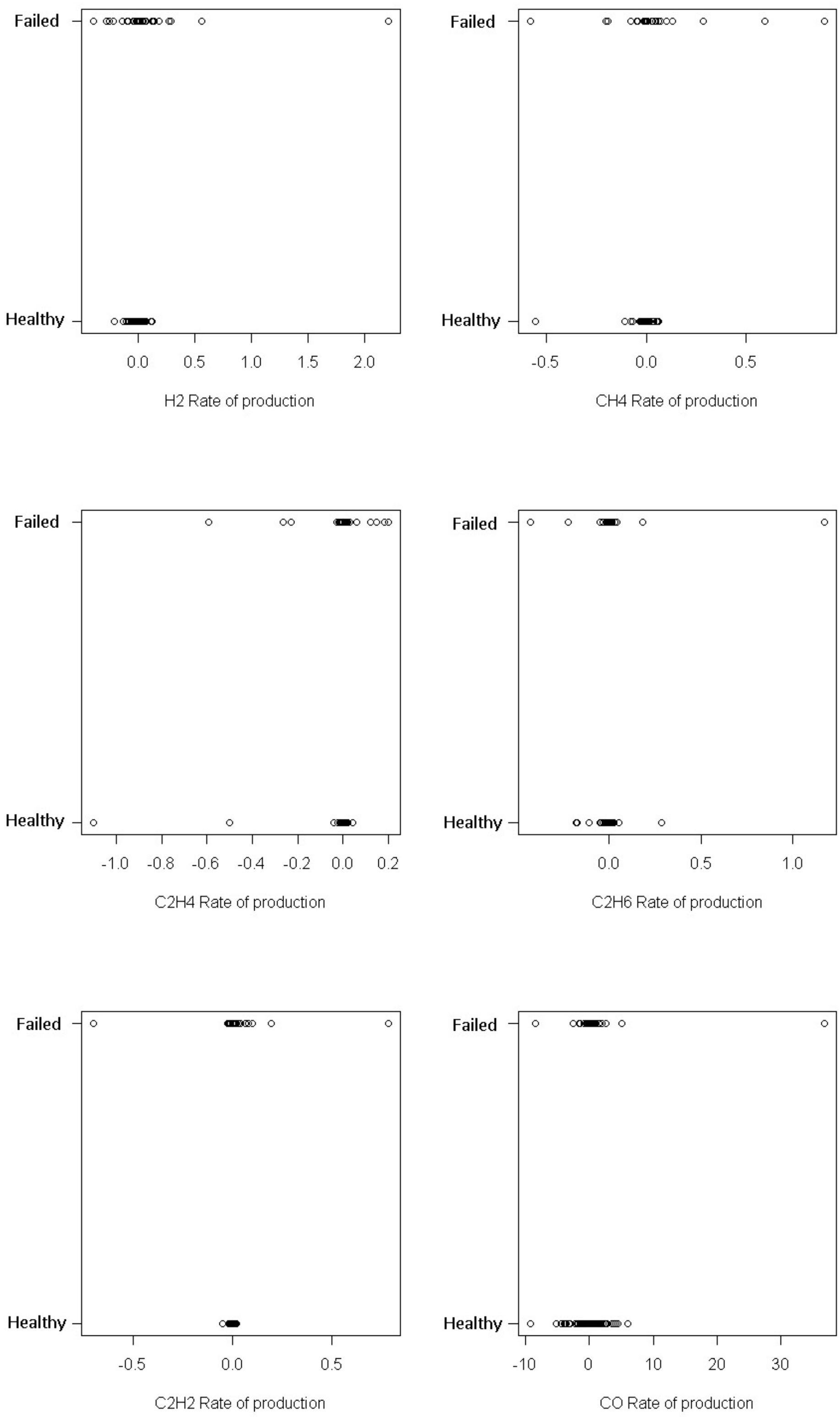

Figure 5.2: Plot of gas rates of production vs failures

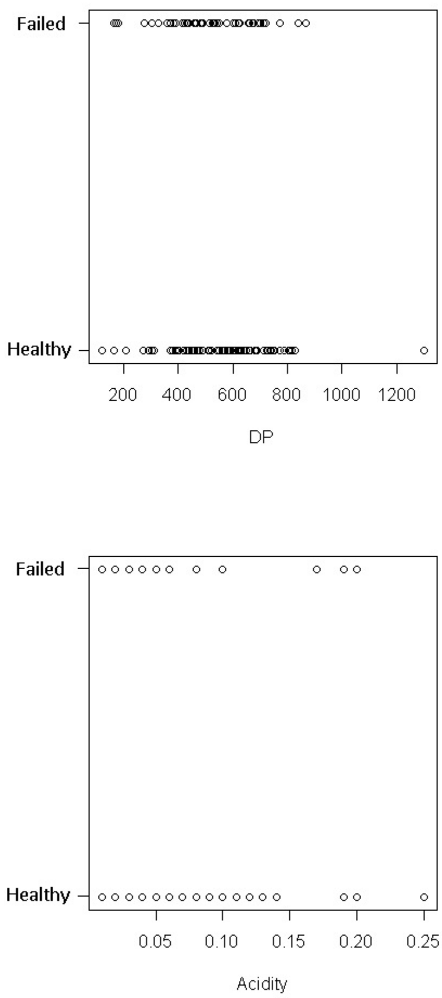

Figure 5.3: Plot of insulation parameters vs failures

5.4.3 Health Index

The values of the DGA and Insulation HI scores were plotted against the failures for the data set to be used in this study.

The results shown in Figure 5.4 indicate the expected trend in scores for failed transformers is lower than for healthy ones in both insulation and oil quality.

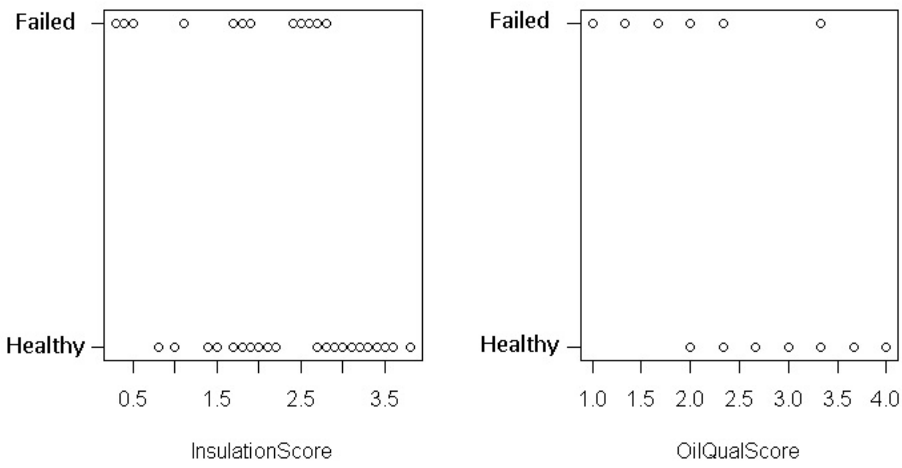

Figure 5.4: Plot of HI scores vs failures

5.5 Transformation of data

Data are usually normalised according to the min-max minimisation algorithm into a range [0 1] according to equation Equation 5.1 below.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{5.1}$$

where:

x_{norm} is the normalised value in range [0 1]
x is the value to be normalised
x_{max} is the maximum possible value for a particular observation x
x_{min} is the minimum possible value for a particular observation x

Various other transformation techniques can be employed to normalise/linearise the data as required by the specific statistical analysis being performed. Normalisation is however not a requirement for logistic regression and such data processing is therefore not required for this data set. The logit function used in the Logistic Regression linearises the DV.

5.6 Outliers

Box plots are then used to determine the skewness of the data, and identify possible statistical outliers [47]. All identified outliers are included in the sample analyses, unless they have been confirmed to be erroneous values.

This method uses examination of the statistical percentiles of the data set. The statistics of interest in this analysis are: maximum, minimum, median, mean, 1^{st} quartile (25^{th} percentile) and 3^{rd} quartile (75^{th} percentile). These statistics are then plotted as follows:

- 1^{st} and 3^{rd} quadrant form the top and bottom of each box for each variable. 50% of the data is represented within this box, with the length of the box being the interquartile range.
- The median is represented by the horizontal line within the box. A line that is not perfectly centered is an indication of skewness of the data which is a measure of assymmetry about the sample mean. Skewness in the data is an indication that the data are not normally distributed. This is not of concern in this study since normality is not a requirement of logistic regression.
- The maximum and minimum of the sample are represented by lines (sometimes referred to as "whiskers") extending from the top and the bottom of the box.

A general assumption is that an outlier is a value that falls more than 1.5 times the interquartile range away from either the top or the bottom of the box [47]. Therefore, if no outliers are present, the maximum value within the sample falls on the top of the upper "whisker" and the minimum falls on the bottom of the lower "whisker". Potential outliers are indicated by the presence of data points either above or below the upper and lower "whiskers" respectively.

While outliers may be identified statistically, the data samples need to be evaluated critically since the "outliers" may in fact just be extreme values that have a critical impact on the analysis and should not be removed without just cause.

5.6.1 Gas concentrations and production rates

Both Carbon Monoxide and Carbon Dioxide have higher concentrations and daily rates of production than the hydrocarbon gases. For this reason, the values of all gases were scaled to pu values for the purpose of graphical comparison. The boxplots in Figure 5.5 show the gas concentrations of the DGA data both with and without identified outliers.

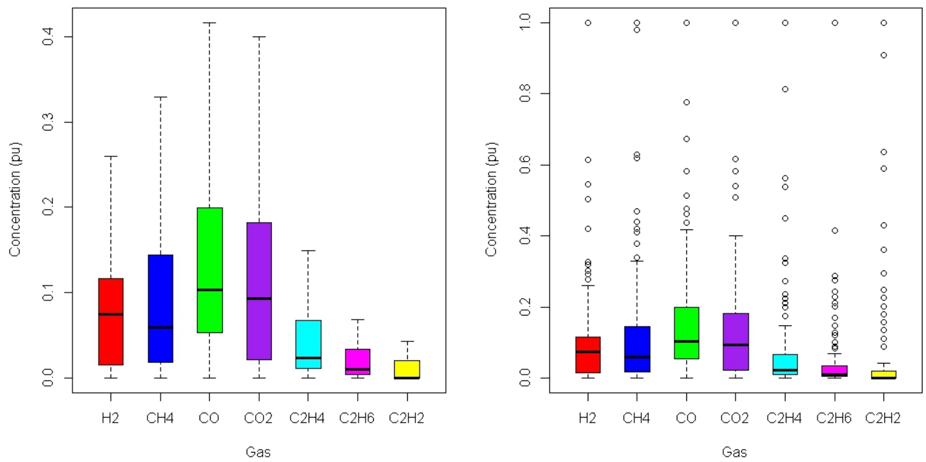

Figure 5.5: DGA concentrations(left), including potential outliers(right)

These plots indicate a number of potential outliers in the data by the points above the top whisker. Table 5.2 shows the actual values of the percentile statistics for each gas. From these values, it can be seen that none of the gases have excessively high maximum values and in fact the maximums of each gas are still considered low. Generally these values would be discarded as outliers, however, the outliers that have been identified result from failed transformers' data and can therefore not be excluded. In this case, valuable information would be lost if these outliers were excluded from analysis. No obvious, real outliers are evident.

The gas production rates were also scaled to per unit values for graphical comparison. The boxplots in Figure 5.6 show the gas concentrations of the DGA data both with and without identified outliers.

These plots indicate the presence of outliers in both the upper and lower regions. On examination of the actual values of the percentile statistics shown

Table 5.2: Percentile summary of DGA concentrations

	H_2	CH_4	CO	CO_2	C_2H_4	C_2H_6	C_2H_2
Min	0.1	0.1	6	0.1	0.1	0.1	0.1
1^{st} **Quad**	2	2	215.5	244	1	1	0.1
Median	9	6	409	1041	2	2	0.1
Mean	11.78	11.79	611	1384.6	6.2	8.3	2.7
3^{rd} **Quad**	14	14.5	783.5	2040.5	5.5	6.5	1
Max	119	100	3912	11200	80	188	44

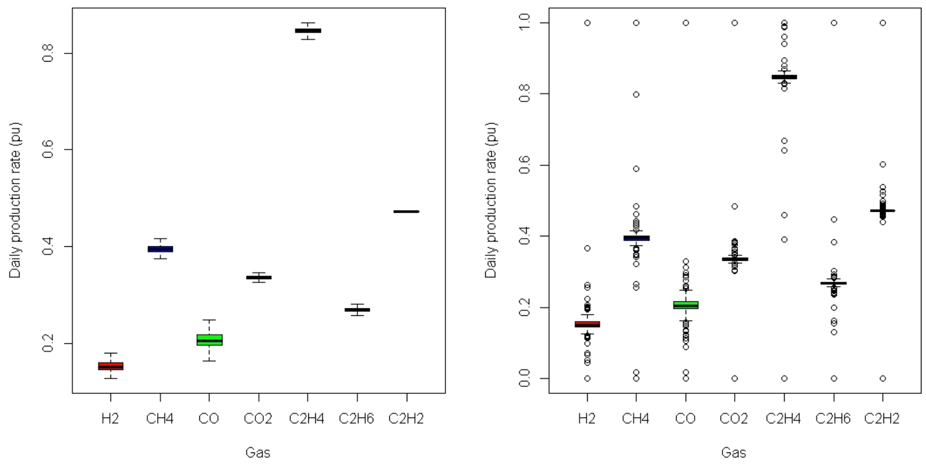

Figure 5.6: DGA rate of production(left), including potential outliers(right)

in Table 5.3, none of the values are unrealistic and consequently cannot be eliminated as outliers.

Some uncertainty in interpretation of the production rate values is present due to the method of calculation employed. Manual oil samples do not lend themselves to reliable production rate calculations.

5.6.2 DP and oil quality

The measurements recorded for DP and oil quality were reduced to per unit values for the purpose of graphical comparison. The boxplots indicating distribution of the DP and oil quality data both with and without possible outliers are shown in Figure 5.7.

Table 5.3: Percentile summary of DGA production rates

	H_2	CH_4	CO	CO_2	C_2H_4	C_2H_6	C_2H_2
Min	-0.394	-0.583	-9.298	-137.4	-1.101	-0.432	-0.703
1^{st} **Quad**	-0.018	-0.009	-0.271	-0.520	-0.006	-0.006	0
Median	0	0	0.135	0.125	0	0	0
Mean	0.019	0.002	0.266	2.255	-0.014	0.001	0.004
3^{rd} **Quad**	0.021	0.007	0.742	1.708	0.006	0.004	0
Max	2.212	0.892	36.838	273.9	0.199	1.174	0.788

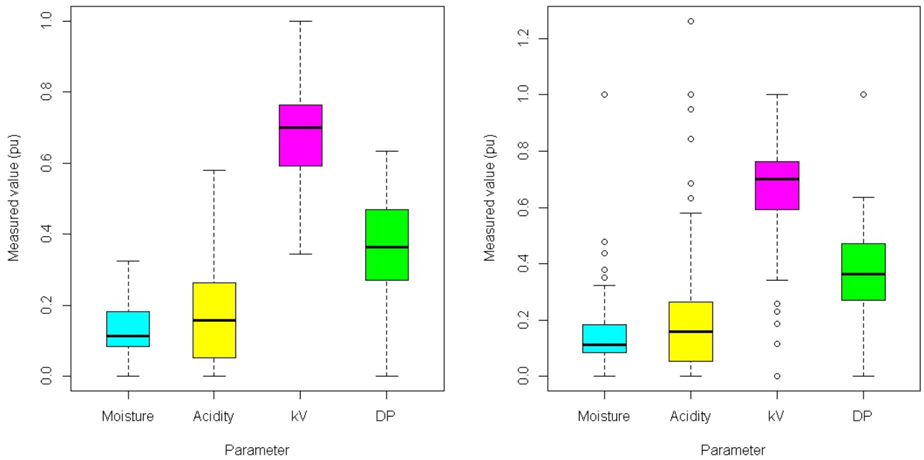

Figure 5.7: Oil quality and DP(left), including potential outliers(right)

These plots indicate possible outliers in the positive extreme for moisture, acidity and DP and in the negative extreme for dielectric strength.

Table 5.4: Percentile summary of Oil Quality and DP measurements

	DP	**Acid**	**kV**	H_2O
Min	121	0.01	25	1
1^{st} **Quad**	440	0.02	67	7
Median	550	0.04	74	9
Mean	578	0.05	72	11
3^{rd} **Quad**	676	0.06	78	14
Max	1300	0.25	95	72

On examination of the actual values of the statistical percentile analysis shown in Table 5.4, it can be seen that the range of data values is reasonable, with the exception of the maximum value for DP.

This value appears to be unrealistically high since the expected value of DP for brand new paper is 1200 and once it has been processed during the manufacture of a transformer it has reduced to approximately 900. This value is therefore considered an outlier and removed from the dataset.

5.6.3 Health Index

The values of HI score are uniform for all components, in the range (0-5). For this reason, it was not necessary to reduce the values to per unit equivalents. The boxplots of the DGA and Insulation HI scores, both with and without potential outliers are shown in Figure 5.8. These plots indicate a number of potential outliers.

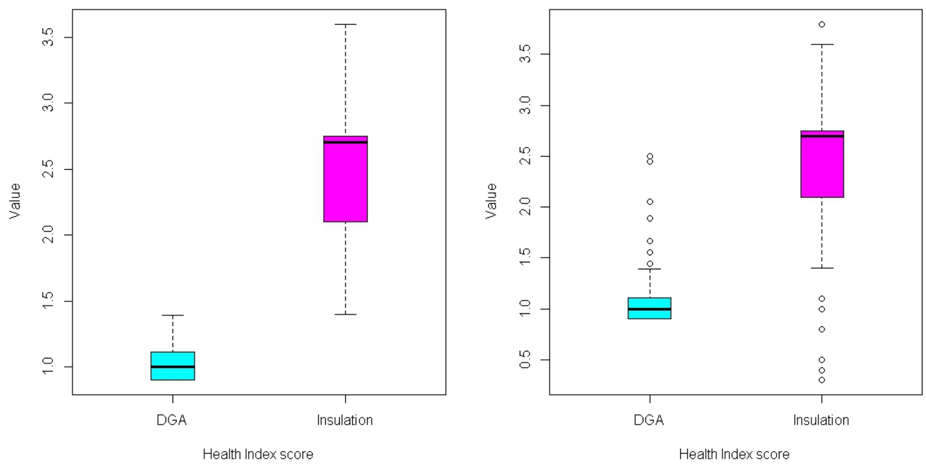

Figure 5.8: HI parameters(left), including potential outliers(right)

Table 5.5: Percentile summary of HI parameters

	DGA	**Insulation**
Min	0.9	0
1^{st} **Quad**	0.9	2
Median	1	3
Mean	1.1	2.6
3^{rd} **Quad**	1.1	3
Max	2.5	4

Examination of the statistical percentiles in Table 5.5, shows values that are reasonable and cannot be eliminated as outliers. This is congruent with the decision not to remove outliers in the base data used to calculate the HI scores.

5.7 Conclusion

In this chapter, the data set to be used in this study was defined. The data was evaluated visually and found to display the expected trend. The data set was found to contain records with missing data and these records were then deleted, since the data set is too small to utilise methods of missing data approximation. The data was examined graphically, using boxplots to identify potential outliers. In this case, only the outliers identified in the DP measurements were removed since there was no evidence supporting the removal of the outliers identified in the other variables. The data set processed, using the methods outlined in this chapter, will be used in the development of the statistical model for probability of failure determination.

Chapter 6

Prediction model

6.1 Research objectives

As outlined in Chapters 1 and 3, there are two main components in determining the probability of failure of a power transformer:

- Probability of failure based on age
- Probability of failure based on HI

Verification of the existing model parameters, as well as validation of the input variables is performed.

6.2 Probability of failure based on age

As outlined in Chapter 3, bathtub curves are used to estimate the probability of failure of a piece of equipment based on its physical age. This is commonly done using a Weibull distribution shown in Equation 6.1.

$$f\left(t\right) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{6.1}$$

where:

β is the shape parameter (slope)
η is the scale parameter (characteristic life)

6.2.1 Distribution parameters

The Weibull distribution currently utilised in the CCRA model in Eskom is using a shape parameter $\beta = 3$ and a scale parameter of $\eta = 40$, representative of the desired design life of 40 years. This distribution was determined based on other models in use internationally for other equipment.

Statistical analysis was carried out to fit a Weibull distribution to Eskom's empirical failure data. This resulted in different distribution parameters. A best fit distribution was found to have shape parameter $\beta = 1.8$ and a scale parameter of $\eta = 24.7$.

A plot of the Cumulative Density Function (CDF) for the failure data, the existing CCRA distribution and the new fitted distribution is shown in Figure 6.1.

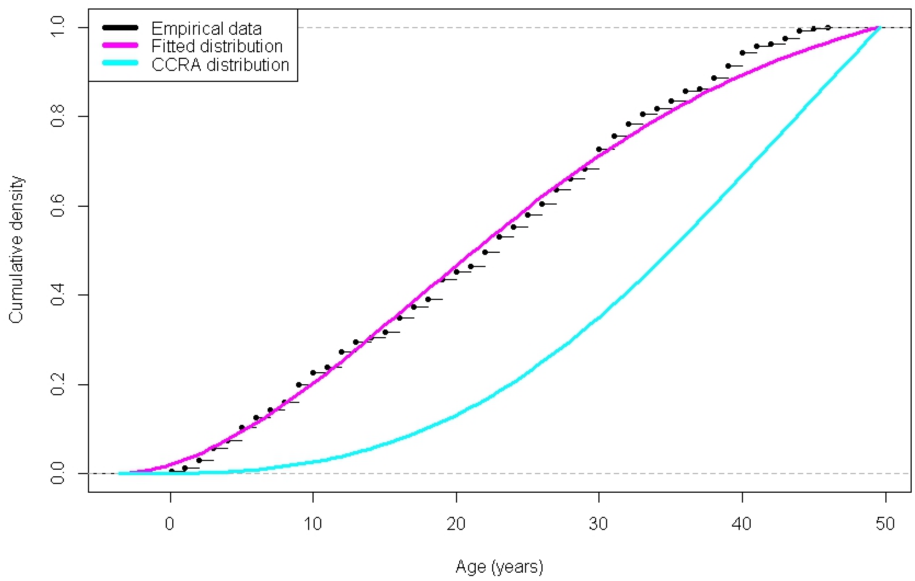

Figure 6.1: Cumulative Density Function of failure data, CCRA and fitted Weibull distributions

6.2.2 Goodness of fit

A simple graphical method of determining goodness of fit of a theoretical distribution is to examine a quantile-quantile (q-q) plot of two data sets. A q-q plot is a plot of the quantiles of the first data set against the quantiles of the second data set. This is useful in determining whether or not the two data sets have common distributions.

Two q-q plots are generated:

1. Comparing the quantiles of the empirical failure data with the quantiles of the CCRA Weibull distribution
2. Comparing the quantiles of the empirical failure data with the quantiles of the fitted Weibull distribution

The quantiles are plotted with a 45° reference line. Two datasets from the same distribution will fall uniformly around this line. A large deviation from this line is evidence that the two data sets come from populations with different distributions.

A 95% confidence envelope is also plotted. The q-q plots for the CCRA and fitted Weibull distributions are shown in Figures 6.2 and 6.3 respectively.

From Figure 6.2, it can be seen that the Weibull distribution of the CCRA model fits the empirical data in the age group (8-52) and is skewed, with the 45° reference line not passing through the origin. The average age of failure for this distribution is 40 years, which is not evident in the empirical data. The average age of failure is 24 years. This is indicative of a poor fit.

Figure 6.3 shows the empirical data fits the fitted Weibull distribution well for the age group (0-42). This is expected since no data is available for failures of transformers older than 45. The reference line passes through the origin and the average age of failure for this distribution is 24.7 years, which is congruent with the empirical data. This distribution is found to be a better representation of the data.

Both q-q plots show evidence of distinct patterns where the data shifts above

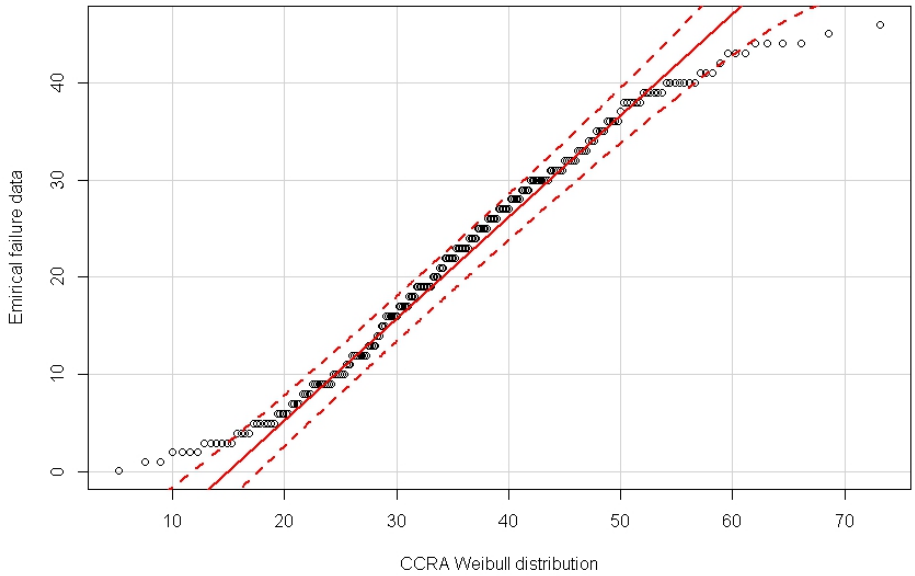

Figure 6.2: Existing CCRA model bathtub curve based on a theoretical Weibull distribution

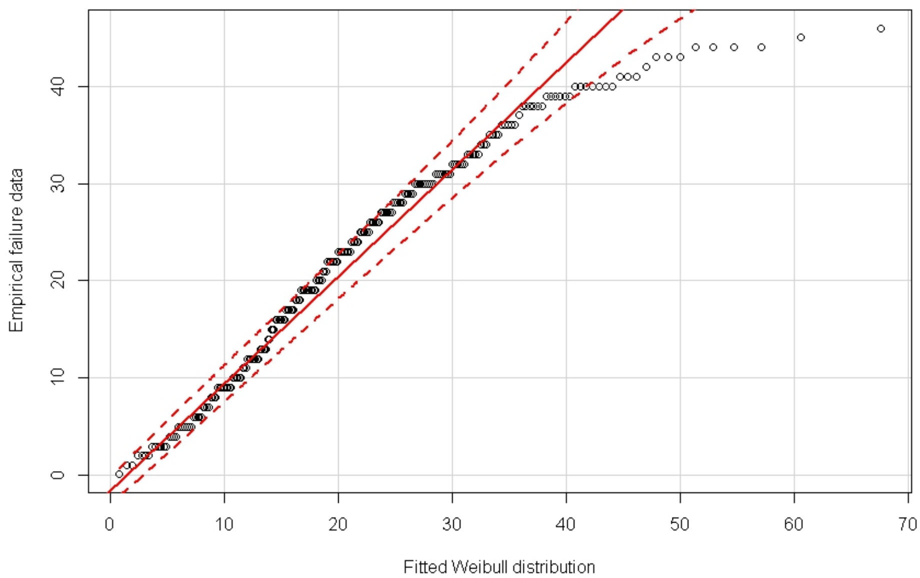

Figure 6.3: Bathtub curve based on a Weibull distribution developed from empirical failure data

and below the reference line. This is indicative of a poor fit. An ideal fit would show the data randomly distributed around the reference line across the entire range.

6.3 Probability of failure based on HI

The probability of failure based on HI is derived with a logistic regression model. This regression was performed using the Generalised Linear Model (GLM) function, with the binomial distribution linked by the logit function in R. This analysis was performed with the existing HI scores, as well as the raw data as IVs.

6.3.1 Methodology

The data set was processed to remove missing data and outliers, as outlined in Chapter 5. A logistic regression model was then found to best fit the data using the HI scores for DGA and Insulation as IVs.

An iterative process is then carried out with all the raw data to evaluate the influence each input variable had on the overall output prediction.

Once the most significant variables have been identified, the logistic regression is run again to find the best fit model for the data set.

Each model is then evaluated for goodness of fit and compared with each other and the null model, which is the model with no predictors, i.e. an intercept only.

6.3.2 Health Index model

A logistic regression was run with the HI DGA and Insulation parameters. From the results shown in Figure 6.4, it can be seen that both the Insulation score and DGA score are both statistically significant, with p-values lower than 0.05.

The mathematical function equation that results is shown in Equation 6.2. The coefficients of this equation can be interpreted similarly to an ordinary regression, as follows:

```
Deviance Residuals:
Min       1Q        Median    3Q        Max
-1.7791   -0.7815   -0.6819   1.1374    1.8357

Coefficients:
                    Estim     Std. Err   z value   Pr(>|z|)
(Intercept)          0.5570    1.2740    0.437     0.661923
DGAHI                2.0746    0.9207    2.253     0.024251 *
InsulationScore     -1.3943    0.3664   -3.805     0.000142 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 181.60  on 138  degrees of freedom
Residual deviance: 156.75  on 136  degrees of freedom
AIC: 162.75

Number of Fisher Scoring iterations: 4
```

Figure 6.4: R output of logistic regression with HI parameters as IVs

- For every one unit increase in DGA score, the log odds ratio of failure increases by 2.07
- For every one unit increase in Insulation score, the log odds ratio of failure decreases by 1.39

This in itself is not very useful and in order to obtain information regarding the probability of failure itself, Equation 6.2 must be exponentiated.

$$\ln\left(\frac{p_{failure}}{1-p_{failure}}\right) = 0.557 + 2.07\,(DGA) - 1.39\,(Insulation) \tag{6.2}$$

The next statistics that can be read from the output in Figure 6.4, are the model and null deviances. The null model is the model including only the intercept, i.e. no predictor variables included. Therefore, if the model deviance is lower than the null deviance, the model is better at predicting the outcome than an empty model.

The deviances and degrees of freedom are analysed with likelihood ratio tests to yield a result of 4.003×10^{-6}, which is lower than 0.001 indicating a significantly better prediction than the null model.

The regressions for the HI parameters: DGA and Insulation, are visualised in Figure 6.5. The predicted probabilities for each HI parameter is plotted for the developed logistic regression model. A failure probability is calculated across the expected data range for each parameter, with the other parameter constant at its sample mean. Each parameter is plotted with a 95% confidence interval.

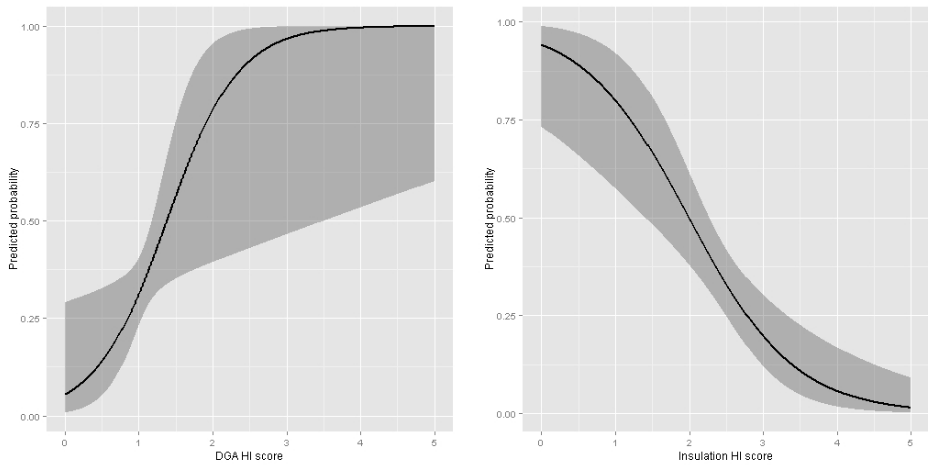

Figure 6.5: Regression plots for HI parameters

The plots show the expected trend for both parameters in predicting failure. There is large uncertainty present in both parameters, particularly in DGA, with $> 40\%$ uncertainty for scores above 2. This is to be expected since the model was trained with low DGA scores.

6.3.3 Raw data

Logistic regressions were run for various predictor variables, using the raw data instead of weighted averages, as was done for the HI parameters. This was done in an iterative process in order to identify the statistically significant variables and exclude all extraneous variables whose only contribution is in making the model unnecessarily complicated and noisy.

Models were run for the following variable, in various combinations:

- Gas concentrations
- Gas production rates
- Gas ratios (Duval, Roger's, IEC, %TDCG)
- Oil quality
- Insulation

The models were run with only the raw data, except for the gas ratios. The gas ratios were analysed since various relationships between the gases have been identified and found to be useful in analysis of dissolved gases.

The results of the models using production rates and gas ratios, indicated that none of these variables were statistically significant, with all variables having p-values > 0.05. For this reason, these variables were excluded from further study.

Regressions run with only gas concentrations indicated that of all the gases, only CO and CH_4 were statistically significant. Similarly with the oil quality and DP variables, the regressions indicated the only variables of significance were acidity and DP. Both regressions were then run again without the extraneous variables and the statistical relationship was found to increase.

A logistic regression was then run with the four variables that were identified as significant: CO, CH_4, DP and acidity. From the results shown in Figure 6.6, it can be seen that both the CO concentration, CH_4 concentration, acidity and DP values are statistically significant, with p-values lower than 0.05.

The mathematical function equation that results is shown in Equation 6.3. The coefficients of this equation can be interpreted similarly to an ordinary regression, as follows:

- For every one unit increase in CO concentration, the log odds ratio of failure decreases by 0.0037
- For every one unit increase in CH_4 concentration, the log odds ratio of failure increases by 0.046

```
Deviance Residuals:
Min        1Q         Median    3Q        Max
-1.7788    -0.8355    -0.2297   0.7512    2.6071

Coefficients:
               Estimate     Std. Error   z value   Pr(>|z|)
(Intercept)    3.307e+00    8.750e-01    3.779     0.000157 ***
CH4Conc        4.572e-02    1.555e-02    2.940     0.003283 **
COConc        -3.688e-03    9.189e-04   -4.014     5.98e-05 ***
Acid          -1.800e+01    6.074e+00   -2.963     0.003049 **
DP            -3.335e-03    1.123e-03   -2.970     0.002977 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 181.60  on 138  degrees of freedom
Residual deviance: 127.63  on 134  degrees of freedom
AIC: 137.63

Number of Fisher Scoring iterations: 6
```

Figure 6.6: R output of logistic regression with selected IVs

- For every one unit increase in DP value, the log odds ratio of failure decreases by 0.0033
- For every one unit increase in acidity, the log odds ratio of failure decreases by 18

This in itself is not very useful and in order to obtain information regarding the probability of failure itself, Equation 6.3 must be exponentiated.

$$\ln\left(\frac{p_{failure}}{1-p_{failure}}\right) = 3.31 + 0.046\,(CH_4) - 0.0037\,(CO) - 18\,(Acid) - 0.0033\,(DP) \tag{6.3}$$

The next statistics that can be read from the output in Figure 6.6, are the model and null deviances. The null model is the model including only the intercept, i.e. no predictor variables included. Therefore, if the model deviance is lower than the null deviance, the model is better at predicting the outcome than an empty model.

The deviances and degrees of freedom are analysed with likelihood ratio tests to yield a result of 5.320×10^{-11}, which is lower than 0.001 indicating a significantly better prediction than the null model.

The regressions for the raw data variables: CO, CH_4, DP and acidity, are visualised in Figure 6.7. The predicted probabilities for each variable is plotted for the developed logistic regression model. A failure probability is calculated across the expected data range for each variable, keeping all other variables constant at their sample means. Each variable is plotted with a 95% confidence interval.

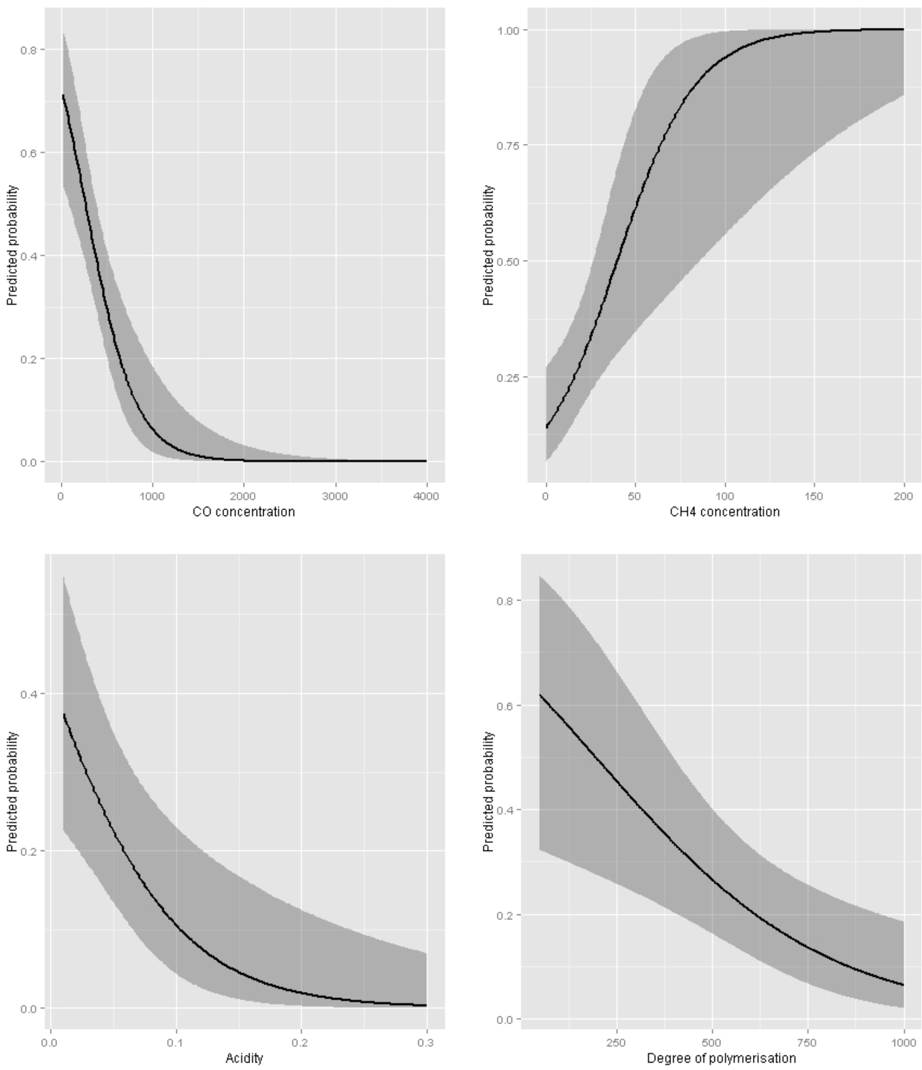

Figure 6.7: Regression plots for raw data parameters

The plots show the expected trend for the four variables in predicting failure. There are large uncertainties present in all variables, which are expected due to the nature of the variables being analysed and the limited data samples.

6.3.4 Model comparison

The two models can be compared to each other by performing the likelihood ratio test, as was done with each model and the null model previously.

The results of a χ^2 anova analysis, using the HI regression model as the null hypothesis yields a p-value of 4.748×10^{-07}. This would suggest that the new model using the four raw data variables is substantially better at predicting failure than the model using the existing HI parameters.

Another method of comparing the fit of two models is to compare their respective Akaike Information Criteria (AIC) since it is used for reporting the trade-off between fitting (likelihood) and parsimony (number of parameters) of the model. The AIC values for both models are shown in Figures 6.4 and 6.6 respectively.

In this case the HI model AIC is 162.75 while the raw data model AIC is 137.63, indicating a substantial improvement in model fit using the raw data, despite the increased parsimony.

The ROC curves for both the HI and raw data models are shown in Figure 6.8. Analysis of the Area Under the Curve (AUC) is a common way of showing the discrimination ability of a statistical model. The specificity and sensitivity are calculated as per Equations 4.2 and 4.3.

A ROC curve which goes closer to the top left hand corner of the plot indicates a model with ideal discrimination ability, whereas a ROC curve close to the 45° line indicates a model with no discrimination ability. The AUC has a range of 0.5 (no discrimination) to 1 (perfect discrimination).

Analysis of the curves in Figure 6.8, shows the raw data model with an AUC of 0.848 and the HI model with an AUC of 0.774. This again shows that the raw data model is substantially better at predicting failure than the HI model.

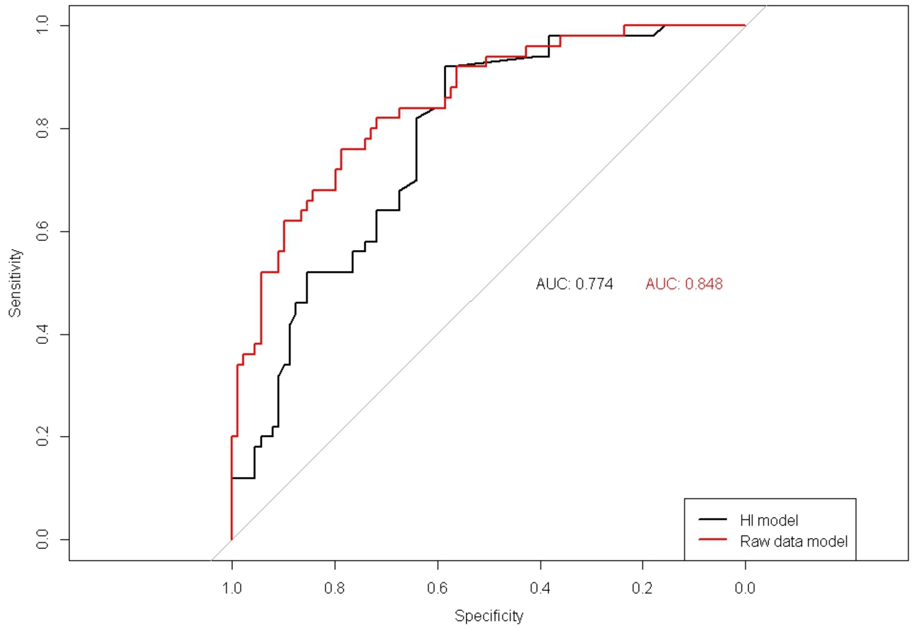

Figure 6.8: ROC curves for HI and raw data models

6.4 Error sources/uncertainty

No model will be perfect, due to the uncertainty that is introduced by the known and unknown unknowns as well as errors in the variables utilised in the study. In this study, there are a number of sources of error in the predictor variables being considered. These errors translate to increased uncertainty and tolerances in the output of the model and, as such, should be limited as far as possible.

6.4.1 Degree of Polymerisation

The DP value has uncertainty introduced in mainly two ways: measurement/ approximation methods and by its exponential nature.

The rate at which DP changes with time, given a constant temperature, is given by Equation 6.4 [18]. The aging rate of insulation is a function of time, temperature, moisture, Oxygen and acid content. According to figures published in [18], normal insulation life of a well-dried, oxygen-free, thermally upgraded insulation system at the reference temperature of 110 °C, the expected operating life of a transformer retaining DP of 200 is 150 000 hours or 17.12 years.

$$\frac{1}{DP_t} - \frac{1}{DP_{initial}} = A \times \exp\left(\frac{-E_a}{RT}\right) t \tag{6.4}$$

Figure 6.9 shows the relationship of DP with age in years, including the thresholds for a 95% confidence interval, at an operating temperature of 98°C, the normal operating temperature for a fully loaded transformer. It can be seen that for higher values of DP, new insulation, the tolerance on expected life consumption is small. At lower DP values, old insulation, the tolerance on expected life consumption is large. At 200 DP, the generally recognised value for insulation end-of-life, the expected age ranges from 10 to 29 years. This is a 19 year uncertainty. Therefore, this measure is of limited value at the lower DP values, since a variance of 19 years on an age of 17 years is more than 100%.

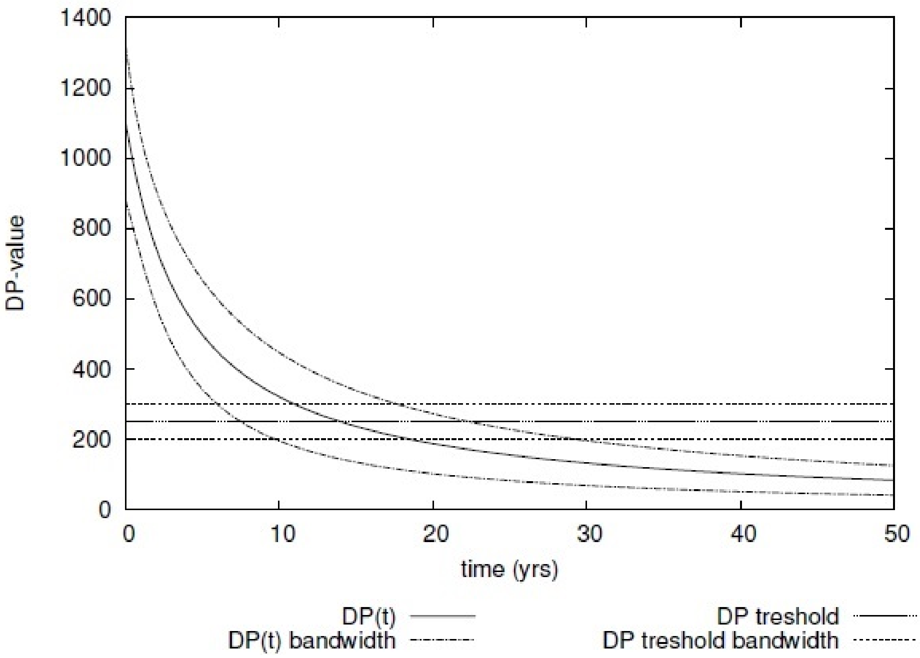

Figure 6.9: DP value and threshold with accuracy bandwidth versus time [8]

This is congruent with the breakdown mechanisms of polymers. This level of uncertainty at lower DP values could also be indicative of more leeway when operating a transformer at these values since breakdown occurs much more slowly than at higher values.

Additional uncertainty is introduced in the measurement methods used for determining DP. The optimum is to take a paper sample and test the tensile strength in the laboratory. This is often not possible since it is an intrusive test. Common practice is to take an oil sample and test for furanics and estimate DP from the results. This measurement is prone to the error associated with oil sampling, as well as the oil test. The test is often not repeatable with large variances in the results of multiple tests on a single sample.

6.4.2 Design

Over the course of 18 years, there would have been substantial differences in design philosophies, oil types used, maintenance philosophies, management strategies, loading, etc. All these aspects affect the life expectancy and failure probability of the power transformer.

- **Design**: Different internal structures, better understanding of design leading to optimised design parameters, differences/improvements in transformer cooling systems.
- **External components**: Changes in types of bushings, design and insulation and tapchangers, arc-quenching media, design and maintenance activities required.
- **Insulation**: Improvements and changes in the type of insulation used and its breakdown characteristics.
- **Oil**: Improvements in oil specifications, changes in additives which change the aging characteristics of the transformer oil. Examples such as inhibitors which inhibit oxidation of the oil, but lead to accelerated degradation on depletion and Dibenzyl Disulfide (DBDS), also known as corrosive sulphur was previously added to the oil resulting in the breakdown of the copper within the transformer.
- **Online filtration**: Various devices can be connected to the transformer for monitoring or moisture filtration and control. These devices also remove additives from the oil, as well as dissolved gases which will have an implication on DGA and diagnosis of incipient faults.

- **Maintenance**: Maintenance activities that are required vary. Any intrusive work to be performed on a transformer is a potential for changing the condition of the transformer, ideally for the better, but occasionally it can be detrimental to the health of the transformer due to human error. Compliance with required maintenance intervals and requirements are also not always managed optimally.
- **Loading**: Transformers undergo various changes in loading patterns over their lives, these can range from only 20% utilisation to overloading at the design limits. In general the Transmission population of transformers are loaded at 50% or less. Increases in loading due to failure of adjacent units is also present. The loading is determined by management philosophies, planning objectives, operational requirements, etc.

All these aspects will have an impact on the uncertainty in the predicted probability of failure of a transformer.

6.4.3 DGA

Uncertainty is present in the DGA results for various reasons. These are primarily due to the oil sampling methods and experience and competence of the person taking the sample. Invalid results can be obtained for samples that can have adverse impacts on the overall trend of the transformer's gas profile.

The distribution of gases within the transformer is not uniform, samples from different sample points on the transformer yield differing results and should not be compared directly. Temperature and loading also have an impact on the amount of gas that is dissolved within the oil. For these reasons, the results of the samples taken need to be carefully analysed for validity. Some errors in measurement, if performed repeatedly can lead to results that appear valid and even alarming. Others can lead to results that appear stable when a fault is present.

The method of computing daily rate of production is another source of error. With manual samples it is simply the average daily rate of production over the time interval between two samples. This is not a robust calculation since any spikes in gas production within the sample interval can be masked. Negative production rates can also be present, indicating a reduction in gas concentration. This could also be attributed to a poor sample being taken or measurement errors in the analysis process. This is the suspected reason for the lack of statistical significance of this variable in the regression models.

A way in which this error can be reduced is by using more frequent samples, such as obtained from an online analyser and calculating the rate of production using mean squares or splines.

6.5 Discussion of results

The results obtained from the study show that, with the existing available data, a statistical model can be developed that can be used to predict probability of failure of a power transformer to some degree that is better than an empty model.

The existing Weibull distribution for determining the probability of failure based on age was found to be ineffective at describing the empirical failure data that was available. Alternative shape and scale parameters were found that better described the data. The model based on the HI parameters was found to be less effective at predicting probability of failure than the model based on the the four raw data variables.

Various factors were identified for reducing the accuracy of the model, including: limitation of available failure data, unavailability of some condition data and measurement uncertainties in some of the data that was used.

While both models were found to be statistically more significant than an empty model, neither was concluded to be optimal. However, a model better suited to the asset management model can be developed with various improvements made with the availability of more data.

Chapter 7

Conclusion and recommendations

7.1 Conclusions

The current asset management model was reviewed and deficiencies in the determination of the probability of failure were found, which could result in incorrect risk calculations and potentially poor decision making. Deficiencies were found in the determination of probability of failure based on both age and HI.

Transformer life estimation and condition monitoring methods were investigated and the optimum variables to be used in a statistical analysis were identified. These were limited by availability of these parameters in the empirical failure data that was analysed.

The Weibull distribution currently being used for determination of probability of failure based on age was found to be deficient in describing the existing failure data and an alternative distribution with different shape and scale parameters was found. The alternative model was found to be effective in describing the data in the age group 0 - 42 with an average age of failure being 24, which is congruent with experience and significantly different to the 40 years suggested by the original distribution.

A model was developed based on the existing HI parameters and was found to be statistically more significant than an empty model. The data was then interrogated to identify which of the measured variables held the most significance and if any extraneous variables were present. Results of this analysis indicated that of all the variables, only four: CO concentration, CH_4 concentration, DP and acidity, held any significance. These variables were then used in the development of an alternative model and found to be more effective at predicting the probability of failure than the HI variables.

Although an alternative model was found that was better at predicting probability of failure than the model currently in use, this model is also not optimal. Using either of these models will lead to undesired results, since a large number of transformers requiring replacement will not be replaced, while others that are replaced could have remained in service for many years. In effect, neither the risk that these transformer failures pose to the business, nor the failure rate of the transformers within the existing population will be reduced significantly using these models.

The limited data and large tolerances on each variable's data, lead to significantly reduced confidence in the conclusions that can be drawn from this data. For this reason, these models are not optimised.

The risk to the business of having no model available for asset management decisions can be lower than the risk of using an unoptimised model. This is illustrated in Table 7.1 with a simple example.

Table 7.1: Example comparison of risk to business using different asset management models

	No model	**Unoptimised model**
Model accuracy (%)	0	50
Failure rate	10 / annum	5 / annum
Number of transformers replaced	10 (failed)	5 (failed) + 10 (predicted)
Equipment replacement cost	10 × R 20 million	15 × R 20 million
Consequence cost	10 × R 5 million	5 × R 5 million
Total cost	R 250 million	R 325 million

Although the actual problem is more complicated than indicated in the example, the implication of overestimating the value of simply reducing failure rate, not taking overall risk into consideration is illustrated. In the example, the asset management model had a 50% accuracy on the prediction of failures. In this research, neither model provided such a high accuracy, resulting in a <50% improvement. In order to substantially reduce the overall cost to the company, a model with a very high accuracy will be required.

It is therefore concluded that, if risk management is the objective, the existing models should not be used as a basis for investment decisions.

The fact that the models discussed in this research are identified as being better predictors of probability of failure than simply basing the decision on chance alone is compelling evidence that an optimised model, capable of allowing reliable predictions, can be developed in the future with the availability of more reliable data. For this reason, further research should be performed to develope a model with a prediction accuracy that can be used for risk reduction.

7.2 Recommendation for future work

Using the statistical methods outlined in this work, the failure data should be analysed in terms of system data. This would include demographics as well as location and operating/system conditions. While these factors are not influencial in the decision for replacement/refurbishment of a transformer, they do contribute to failures and their interactions should be noted. Interactions of variables may be useful in specification and design of transformers in the future.

Recently, within the past 5 years, a large number of online gas analysers have been installed on a large percentage of the transformer population in Eskom. It would be useful to utilise the online data in the analysis process. This data is significantly more accurate in identifying trends in gas production. The errors associated with human interaction and measurement are eliminated since the devices operate online.

The calculation methods for determining the values of the variables to be used in the model should be investigated and improved to yield more reliable data for analysis.

All Models are wrong, but some are useful
˜ George Box

References

[1] Eskom and ABB Powertech. *Theory, Design, Maintenance and Life Management of Power Transformers, Power Series Volume 5.* Y-Land, 2008.

[2] T.V Oommen. Moisture equilibrium charts for transformer drying practice. *IEEE Transactions on Power Apparatus and Systems*, PAS-103(10): 3063–3067, 1984.

[3] P Gervais and J Aubin. On-line monitoring of key fault gasses in power transformers and shunt reactors. In *Proceedings of 62nd International Doble clients conference*, 1995.

[4] A Kraetge, M Kruger, H Viljoen, and A Dierks. Aspects of the practical application of sweep frequency response analysis (sfra) on power transformers. In *6th Southern Africa Regional Conference*, 2009.

[5] D Marlow, P Davis, D Trans, D Beale, S Burn, and A Urquhart. *Remaining Asset Life: a state of the Art Review. Strategic Asset Management and Communication.* WERF, 2009.

[6] *CCRA Economic Life Manual*, 2009.

[7] A Bossi. An international survey on failures of large power transformers in service, final report of cigre working group 12.05. *Cigre Electra*, No 88, 1983.

[8] A van Schijndel, J.M Wetzer, and P.A.A.F Wouters. Forecasting reliability of transformer populations. In *CIRED 19th International Conference on Electricity Distribution*, 2007.

[9] W.H. Bartley. Analysis of transformer failures, imia wgp33(03). In *36th Annual Conference of the International Association of Engineerin Insurers, Stokholm*, 2003.

[10] J Pathak, Y Jiang, V Honavar, and J McCalley. Condition data aggregation with application to failure rate calculation of power transformers. In *IEEE Computer Society, Proceedings of the 39th Annual Hawaii International Conference on System Sciences, (HICSS 06)*, 2006.

[11] R Jongen, P Morshuis, J Smit, A Janssen, and E Gulski. A statistical approach to processing power transformer failures. In *CIRED, 19th International Conference on Electricity Distribution*, number Paper 546, 2007.

[12] C Osorio and N Sawant. *Transformer Lifetime Prediction*. Intelligent Energy Projects, ee292k edition.

[13] Z Zhang, Y Jiang, and J.D McCalley. *Bayesian Analysis of Power Transformer Failure Rate based on Condition Monitoring Information*. Iowa State University.

[14] J Endrenyi, G Anders, and L da Silva. Probabilistic evaluation of the effect of maintenance on reliability an application. *IEEE Transactions of Power Systems*, 13(2):576–583, 1998.

[15] L Lundgaard, W Hansen, D Linhjell, and T Painter. Aging of oil-impregnated paper in power transformers. *IEEE Transactions on Power Delivery*, 19(1):230–239, 2004.

[16] J.N Jagers, J Khosa, P.J de Klerk, and C.T Gaunt. Transformer reliability and condition assessment in a south african utility. In *XVth International Symposium on High Voltage Engineering, Slovenia*, 2007.

[17] ABB Power Technologies Management Ltd. *Transformer Handbook*. ABB, 2004.

[18] *IEC 60076-7: Loading Guide for Oil-Immersed Power Transformers*, 2008.

[19] G Urbani and R Brooks. Using recovery voltage method to evaluate aging in oil-paper insulation. In *IEEE International Conference on Conduction and Breakdown in Solid Dielectrics*, pages 93–97, 1998.

[20] T Dakin. Electrical insulation deterioration treated as a chemical rate phenomenon. *AIEE Transactions*, 67, 1948.

[21] *IEC 60450: Measurement of the Average Viscometric Degree of Polymerisation of New and Aged Cellulosic Electrically Insulating Materials*, 2007.

[22] Y Vashishtha, P Ascione, and Q Su. *The Uncertainty in Power Transformer Fault Diagnostics Using Conventional Testing Methods*. Power and Water Authority, Monash University.

[23] E Susemihl. Transformer reliability estimation. In *Energy Tech Magazine*, 2005.

[24] *IEC 60599: Mineral oil impregnated electrical equipment in service - Guide to interpretation of dissolved and free gasses analysis*, 1999.

[25] *IEEE std. C57.104: IEEE Guide for the interpretation of gasses generated in oil-immersed transformers*, 2008.

[26] S Govender and A Singh. *Toolkit for Transformer Life Management - Report on: Condition Assessment Tools and Methodologies*. Trans-Africa Projects, 2009.

[27] E Gockenbach, X Zhang, Z Liu, H Chen, and L Yang. Life prediction of power transformers with condition monitoring. In *Cigre, A2-111*, 2012.

[28] M.A.R Uzair, M Mohiuddin, and M.K Shujauddin. Failure analysis of power transformers. *International Journal of Emerging Technology and Advanced Engineering (IJETAE)*, 3(9), 2013.

[29] R.B Jones. *Risk-Based Management: A Reliability-Centered Approach, 1 edition*. Routledge, 1995.

[30] *IEC 60050: International Electrotechnical Vocabulary*, 1999.

[31] N Taleb. *The Black Swan: The impact of the highly improbable*. New York: Random House, 2007.

[32] A Freeling. A philosophical basis for decision aiding. *Theory and Decision*, 16(2):179–206, 1984.

[33] LJ Savage. The theory of statistical decision. *Journal of American Statistical Association*, 46(253):55–67, 1951.

[34] T Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

[35] J.D Leeper. Choosing the correct statistical test, 2014. URL `http://bama.ua.edu/~jleeper/627/choosestat.html`.

[36] G.J McLachlan. *Discriminant analysis and statistical pattern recognition*. Wiley series in Probability and Mathematical Statistics, 1992.

[37] A Afifi, V Clark, and S May. *Computer-Aided Multivariate Analysis, 4th Edition*. Chapman and Hall/CRC, 2004.

[38] L Aiken and S West. *Multiple regression: Testing and Interpreting Interactions*. Sage Publications, 1991.

[39] D Hosmer and S Lemeshow. *Applied Logistic Regression, 2nd Edition*. Wiley Series in Probability and Statistics, 2000.

[40] B Hu, J Shao, and M Palta. Pseudo-r^2 in logistic regression model. *Statistica Sinica*, 16:847–860, 2006.

[41] D.B Rubin. Multiple imputations in sample surveys a phenomenological bayesian approach to nonresponse. In *The Proceedings of the Survey Research Methods Section of American Statistical Association*, pages 20–34, 1978.

[42] R.L Carter. Solutions for missing data in structural equation modeling. *Research and Practice in Assessment*, 1(1), 2006.

[43] S Burke. Missing values, outliers, robust statistics and non-parametric methods. *Statistics and Data Analysis*, 1:19–24, 2001.

[44] P.D Allison. *Missing Data*. Sage Publications, 2002.

[45] M Hu, S Savucci, and M Choen. Evaluation of some popular imputation algorithms. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 1:308–313, 1998.

[46] H Shafri, A Suhaili, and S Mansor. The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis. *Journal of Computer Science*, 3(6):419–423, 2007.

[47] N Schwertman, M Owens, and R Adnan. A simple more general boxplot method for identifying outliers. *Computational Statistics and Data Analysis*, 47:165–174, 2004.