CLASSIFICATION OF COMMON PARTIAL DISCHARGE TYPES IN OIL-PAPER INSULATION USING ACOUSTIC SIGNALS

by

Mustafa Harbaji

A Thesis Presented to the Faculty of the
American University of Sharjah
College of Engineering
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Electrical Engineering

Sharjah, United Arab Emirates

February 2014

Approval Signatures

We, the undersigned, approve the Master's Thesis of Mustafa Harbaji.
Thesis Title: Classification of Common Partial Discharge Types in Oil-Paper Insulation Using Acoustic Signals

Signature	**Date of Signature** (dd/mm/yyyy)
_____	_____
Dr. Ayman Hassan El-Hag Associate Professor, Department of Electrical Engineering Thesis Advisor	
_____	_____
Dr. Khaled Bashir Shaban Assistant Professor, Computer Science and Engineering Department Qatar University, State of Qatar Thesis Co-Advisor	
_____	_____
Dr. Ahmed Osman-Ahmed Associate Professor, Department of Electrical Engineering Thesis Committee Member	
_____	_____
Dr. Tamer Jamal Shanableh Associate Professor, Department of Computer Science and Engineering Thesis Committee Member	
_____	_____
Dr. Mohammed El-Tarhuni Head Department of Electrical Engineering	
_____	_____
Dr. Hany El-Kadi Associate Dean College of Engineering	
_____	_____
Dr. Leland Blank Dean College of Engineering	
_____	_____
Dr. Khaled Assaleh Director of Graduate Studies	

Acknowledgments

First of all, I would like to thank Allah, the Most Gracious and the Most Merciful, for making this study possible. In addition, I would like to thank the cooperative people whose help also made this study possible. My acknowledgment goes to Dr. Ayman El-Hag for his understanding, patience, and motivation. Likewise, I would like to express my gratitude to Dr. Khaled Bashir Shaban for his support and for co-advising this thesis. In addition, I would like to thank the Electrical Engineering Department at the American University of Sharjah for their continuous support and the teaching assistantship opportunity. Furthermore, I would like to thank the Qatar National Research Fund for funding this thesis. Last but not least, I would like express my deep appreciation to my family and friends who gave me the moral support to keep going in my study.

To those who lost their

Right to choose…

Abstract

Power transformers' age is highly related to the strength status of its insulation system. Oil-paper insulation system is the main type of insulation for most of the power and distribution transformers in the network. However, such insulation system is subjected to electrical, mechanical and chemical stresses that deteriorate its strength. One of the main sources of such stresses is partial discharge (PD) phenomena. PD within the insulation system can happen at any point from different defect sources like sharp points or voids in the insulation. Knowing the source or type of PD activities provides vital information for maintenance scheduling because PD types have different levels of severity. Since PD activities consume only few mA, they cannot be detected by the typical protective relays. However, PD activities emit energy in different forms that make it possible to detect them using different methods. In practice for power transformers, dissolved gas in oil analysis (DGA), ultra high frequency (UHF) pulses, and acoustic emission (AE) methods are the most common methods used to detect PD. Among these methods, AE has several advantages such as being the most cost effective and easiest to install while the transformer is energized. However, measuring AE in the field can be affected by several measurement challenges. The main contribution of this research is to identify the source of PD under different AE measurement conditions. Four common types of PDs are considered for the classification problem; surface discharge, PD from a sharp point to ground plane, PD from semi parallel plates, and PD from an air void in the insulation. The collected AE signals are processed using pattern recognition techniques to identify their corresponding PD types. For feature extraction and reduction, principle component analysis (PCA) is utilized, whereas k-nearest-neighborhood (KNN) is applied for classification. The measurement conditions include having aged insulation material (oil/paper), a tank size of 1×1×0.5 m dimensions, and high surrounding noise level. In addition, the influence of other practical conditions on the recognition rate is studied including PD location, sensor location, oil temperature, and having a barrier in the line-of-sight between the PD source and the AE sensor. A recognition rate of 94% is achieved while classifying the different PD types measured at the same conditions. In addition, it has been found that PD source location, oil temperature, and barrier

insertion have a significant impact on the recognition rate. However, by including AE samples at different conditions in the training process, a recognition rate of around 90% for all cases is achieved.

Abbreviations

PD	-	Partial Discharge
AE	-	Acoustic Emission
UHF	-	Ultra High Frequency
DGA	-	Dissolved Gas in Oil Analysis
PCA	-	Principle Component Analysis
KNN	-	K-Nearest Neighborhood
DFT	-	Discrete Fourier Transform
QDA	-	Quadratic Discriminant Analysis
SVM	-	Support Vector Machine

Table of Contents

Chapter 1: Introduction 15
1.1 Problem Statement 15
1.2 Thesis Contribution 16
1.3 Thesis Arrangement 18
Chapter 2: Partial Discharge Background 19
2.1 Partial Discharge (PD) 19
2.1.1 PD Definition 19
2.1.2 PD Types and Modeling 20
2.2 PD Sensing Methods 23
2.2.1 Dissolved Gas in Oil Analysis (DGA) 23
2.2.2 Ultra High Frequency (UHF) 25
2.2.3 Acoustic Emission (AE) 26
2.3 Challenges with AE Measurement in the Field 28
2.3.1 Acoustic Sensor Bandwidth 28
2.3.2 Environmental Noises 29
2.3.3 Oil Temperature 30
2.3.4 Barrier inside the Transformer 30
Chapter 3: Experimental Setup 32
3.1 Main Equipment List 32
3.1.1 Oil Tank 33
3.1.2 High Voltage Supply 34
3.1.3 PD Types Electrodes 35
3.1.4 AE Sensor Bandwidth Selection 38
3.1.5 AE Sensor Interfacing 40
3.2 Data Collection 41
3.2.1 PD Location 41
3.2.2 AE Sensor Location 42
3.2.3 Oil Temperature 42

3.2.4 Barrier Insertion *43*
3.3 Experimental Observations 44
Chapter 4: Pattern Recognition **47**
4.1 Pattern Recognition Review 47
4.2 Principle Component Analysis (PCA) 49
4.3 K-Nearest Neighbor Classifier (KNN) 51
4.3.1 Distance Correlation *52*
4.3.2 Other Distance Measures *53*
4.4 Other Feature Extraction Methods 54
4.4.1 Discrete Fourier Transform (DFT) *54*
4.4.2 Wavelet Decomposition *55*
4.5 Other Classifiers 57
4.5.1 Polynomial Classifier *57*
4.5.2 Quadratic Discriminant Analysis (QDA) *57*
4.5.3 Support Vector Machine (SVM) *58*
Chapter 5: Results **59**
5.1 PD Type Classification 59
5.2 Practical Measurement Conditions 62
5.2.1 PD Type Recognition at Different PD Locations *63*
5.2.2 PD Type Recognition at Different sensor locations *66*
5.2.3 PD Type Recognition at Different Oil Temperatures *69*
5.2.4 PD Type Recognition with Barrier Inclusion *73*
Chapter 6: Conclusions and Recommendations **78**
Bibliography **80**
Vita **84**

List of Figures

Figure 1: Surface discharge tracking within a cast resin circuit breaker spout [17]....20
Figure 2: Electrical tree propagation due to air-filled void in insulation [18].....20
Figure 3: Schematic diagram for PD 3 and PD 4 in Table 2 considering an internal gas void presence in the pressboard for PD 4 [7].....22
Figure 4: (a) oil-corona discharge, (b) oil-floating electrode discharge, (c) air cavity discharges, and (d) surface discharge [18].....22
Figure 5: PD models used in H. Ma *et al.* experimental setup [20].....22
Figure 6: Different sensing methods to detect PD in GIS [21].....23
Figure 7: UHF probe for standard oil valve [24].....25
Figure 8: Different types and sizes of AE sensors [25].....26
Figure 9: AE sensor installation.....26
Figure 10: Narrow bandwidth, 20-80 kHz [25].....28
Figure 11: Wide bandwidth, 100-450 kHz [25].....29
Figure 12: Frequency spectra of acoustic measurement using piezoelectric sensor [26].....29
Figure 13: Overall experimental setup.....32
Figure 14: Tank bird's-eye view.....33
Figure 15: Different locations of AE sensor.....33
Figure 16: High voltage source and the control panel.....34
Figure 17: Ground electrode top and side view.....35
Figure 18: Sharp point to ground electrode.....35
Figure 19: Semi parallel plate electrode.....36
Figure 20: Surface discharge electrode.....37
Figure 21: Void discharge electrode.....37
Figure 22: Narrow bandwidth sensor response in quiet lab conditions.....38
Figure 23: Narrow bandwidth sensor response in a noisy condition.....39
Figure 24: High bandwidth sensor response in noisy conditions.....39

Figure 25: Sample AE signal from PD pulse.....40
Figure 26: PD location considered.....41
Figure 27: Oil temperature experiment.....43
Figure 28: Barrier insertion experiment.....43
Figure 29: Typical AE signal at sensor location **b**.....44
Figure 30: AE signal with interferences at sensor location **c**.....45
Figure 31: AE signal with interferences from PD location 9.....45
Figure 32: Pattern recognition steps [31].....48
Figure 33: Influence of PCA component number [33].....49
Figure 34: DFT of an AE sample signal.....55
Figure 35: 3-levels wavelet decomposition [12].....55
Figure 36: Decomposed coefficients of an AE sample.....56
Figure 37: SVM example [41].....58
Figure 38: Tuning number of PCA components for classifying all the four PD types60
Figure 39: Tuning K value of KNN for classifying all the four PD types using PCA features.....61
Figure 40: Void discharge AE sample from PD location 9.....64
Figure 41: Void discharge AE sample from PD location 4.....64
Figure 42: AE sample from semi-discharge at sensor location **b**.....68
Figure 43: AE sample from semi-discharge at sensor location **c**.....68
Figure 44: First three principle components scatter of data measured at 23 ^{0}C.....70
Figure 45: First three principle components scatter of data measured at 70 ^{0}C.....71
Figure 46: First three principle components scatter of data measured without a barrier 75
Figure 47: First three principle components scatter of data measured with a barrier..75

List of Tables

Table 1: Oil-paper insulated transformer failure in the U.S. from 1997 to 2001 [4]...15
Table 2: Ziomek *et al.* PD models [7].....21
Table 3: ANSI/IEEE C57.104 oil test interpretation [22].....24
Table 4: Comparison between practical PD sensing methods within transformer's insulation system [7] [21] 27
Table 5: Speed of AE waves in oil at different oil temperatures [7] 30
Table 6: Speed of AE wave in different mediums [7] 31
Table 7: PD types and locations data.....42
Table 8: Applied voltage for surface discharge at different oil temperatures.....46
Table 9: Measurement conditions of the basic classification problem 59
Table 10: Recognition rates of classifying the four PD types measured at the same conditions.....60
Table 11: Confusion matrix (testing and training with the four PD types measured at the same conditions using PCA followed by KNN with distance correlation).....62
Table 12: Recognition rates when training with data from PD location 4 and testing from PD location 9.....63
Table 13: Recognition rates when training and testing with data from PD location 9 65
Table 14: Recognition rates when training and testing with data equally from PD location 4 and 9.....65
Table 15: Confusion matrix (training and testing with data equally from PD location 4 and 9 using PCA followed by KNN with distance correlation).....66
Table 16: Recognition rates when AE sensor is placed at location **b**.....67
Table 17: Recognition rates when AE sensor is placed at location **c** 67
Table 18: Measurement conditions for the oil temperature experiments 69
Table 19: Recognition rates when testing and training at 23 ^{0}C.....69
Table 20: Recognition rates when testing and training at 70 ^{0}C.....70
Table 21: Recognition rates when training at 23 ^{0}C and at testing at 70 ^{0}C 71

Table 22: Recognition rates when testing and training equally at both 23 ^{0}C and 70 ^{0}C 72
Table 23: Confusion matrix (training and testing with equally from data at 23 ^{0}C and 70 ^{0}C using PCA followed by KNN with distance correlation) 72
Table 24: Recognition rates when training equally at 23 ^{0}C and 70 ^{0}C and testing at 50 ^{0}C.....73
Table 25: Measurement conditions for the barrier experiments 73
Table 26: Recognition rates with barrier inserted in the line-of-site of the PD source74
Table 27: Recognition rates without a barrier inserted in the line-of-site of the PD source 74
Table 28: Recognition rates when training without a barrier and testing with barrier inserted in the line-of-site of the PD source.....76
Table 29: Recognition rates when training and testing equally from data with and without a barrier inserted in the line-of-site of the PD source.....76
Table 30: Confusion matrix (training and testing with equally from data with and without a barrier inserted in the line-of-site of the PD source using PCA followed by KNN with distance correlation).....77

Chapter 1: Introduction

1.1 Problem Statement

Power transformers are the most important assets in any transmission and distribution systems. In most countries, these systems have existed for more than a century [1]. According to a report prepared for the U.S. Department of Energy (DOE) in June 2012, "the average age of installed large power transformers (LPTs) in the United States is approximately 40 years, with 70 percent of LPTs being 25 years or older" [2]. Due to their high cost and critical role in delivering uninterrupted power, large power transformers cannot be easily replaced and usually serve for more than half a century. In case of failure, beside businesses interruption and environmental damage, a typical 100 MVA transformer can cost millions of dollars and an interval of 1.5-2 years to manufacture or substitute [3]. It has been reported that in the U.S., the main cause of power transformer failure and the most costly from 1997 to 2001 was a failure of its insulation system as depicted in Table 1 [4].

Table 1: Oil-paper insulated transformer failure in the U.S. from 1997 to 2001 [4]

Cause of Failure	Number	Total Paid
Insulation Failure	24	$ 149,967,277
Design /Material/Workmanship	22	$ 64,696,051
Unknown	15	$ 29,776,245
Oil Contamination	4	$ 11,836,367
Overloading	5	$ 8,568,768
Fire /Explosion	3	$ 8,045,771
Line Surge	4	$ 4,959,691
Improper Maint /Operation	5	$ 3,518,783
Flood	2	$ 2,240,198
Loose Connection	6	$ 2,186,725
Lightning	3	$ 657,935
Moisture	1	$ 175,000
	94	$ 286,628,811

The aging of a transformer insulation system during its operational life is a natural phenomenon [5]. However, insulation aging can be accelerated when transformers are subjected to abnormal electrical, mechanical, and thermal stresses [6]. An important cause of such stresses on transformer insulation is partial discharge (PD) activities [6], [7]. As a result, it is crucial to continuously assess PD activities as an indicator to potential catastrophic transformer failure by a monitoring system. PD inside an oil-paper insulated transformer can be initiated by different sources that are different in terms of severity. An important aspect of PD monitoring is to identify the source of the PD, which provides vital assets management information to schedule maintenance.

1.2 Thesis Contribution

Identifying sources of PD activities has been an important research area due to the PD's direct effect on insulation failure. There are many attempts that have addressed achieving high recognition rates for different PD types [8]-[10]. However, most of these attempts were not made to simulate PD in a practical transformer tank environment, as most of them did not take into account the effect of the transformer tank dimensions, the transformer's complex inner structure, or conditions of the oil such as age and temperature. Other studies presented in [6] and [11] reported a high recognition rate for PD inside a transformer, but the mechanisms used were only applicable when the transformer was off-line. A recent work [12] used a low bandwidth acoustic emission (AE) sensor, 20kHz to 80kHz, to study the effect of increasing the tank size, the presence of barriers between the PD source and the AE sensor, and oil age on PD detection capability. The results of this work showed a recognition rate in the range of 96-100% for a small tank using both spectral and statistical features. When barriers were placed between the PD source and the AE sensor, the recognition rate stayed high; however, it dropped significantly when a large tank, 1×1×0.5 m, and old oil were used, with recognition rates in the range of 50-78% and 60-88% respectively. Moreover, the classification problem in this work was for only two classes presented as PD (from sharp electrode) or no PD. In [13], three types of PD were investigated inside a test oil tank of 50×50×80 cm dimensions, namely, PD due to surface discharge in the pressboard, floating metal in the

pressboard, and bubbles in the pressboard. Three AE sensors were used to measure the signals. For feature extraction, time-frequency transformation was proposed from the short-time Fourier transform using seven descriptors [13]. The study shows different descriptor values based on averaging 11 training samples for each PD model. There is no testing provided for classification and the training is performed with a low number of samples. Moreover, there is no consideration of other factors that would affect feature extraction, such as the oil condition.

Recently, a novel attempt to correlate gas formation in oil to common PD types has been investigated [14]. The work is based on dissolved gas in oil analysis (DGA) and PD statistical parameters. The study utilizes phase-resolved partial discharge (PRPD) and canonical correlation analysis (CCA) for feature extraction and classification. The study includes some of the important practical conditions in the simulation like having heated and circulated oil inside the tank. Three common standard defects or PD types in oil-paper insulation are investigated in the study, i.e. corona (sharp discharge), surface discharge, and cavity or void discharge. The results show classification results only for surface and cavity discharges since corona discharge has smaller gas formation. Based on 50 DGA samples collected at different stages of PD and gas formation, the laboratory's results shows a recognition rate of 65% at the early stage, 75% at the mid stage, and 80% at the late stage for surface discharge. Likewise, a recognition rate of 70% at the early stage, 65% at the mid stage, and 100% at the late stage for cavity discharge were achieved based on another 50 DGA samples. However, when the system was tested with data from an actual transformer at an unknown PD stage, the recognition rate dropped to 56% and 52% for surface and cavity discharges, respectively, based on 50 samples for each PD type. Another PD recognition study for oil-paper insulation is presented in [15] which uses ultra-high frequency (UHF) signals. The study uses wavelet analysis for feature extraction and an improved bagging algorithm (IBA) with a back propagation neural network (BPNN) and support vector machine (SVM) for classification. Four common PD types are used in the study; i.e. cavity discharge, surface discharge, corona or sharp discharge, and oil floating electrode discharge. The PD models are simulated in laboratory conditions in a rectangular tank of 9×70×90 cm dimensions. The UHF antenna used is a third-order Peano fractal antenna installed in the inner wall of the

tank. 50 samples are collected per PD type at different applied voltages. The results of the study show recognition rates of 94.56% and 96% using IBA with BPNN and SVM respectively, which has a small improvement of 2-3% when using the traditional bagging algorithm (BA) and non-bagging approach (NBA). However, this study provides no practical conditions for UHF measurements like having a barrier in the line-of-sight of the antenna and PD source or changing the PD location.

In this thesis, the main contribution, besides achieving a high recognition rate for different simulated PD types using AE signals, is to consider practical measurement conditions. The measurement conditions include having aged insulation material (oil/paper), a tank size of 1×1×0.5 m dimensions, and a high surrounding noise level. In addition, other practical conditions' influence on the recognition rate are studied such as different PD locations, sensor locations, oil temperatures, and having a barrier in the line-of-sight between the PD source and the AE sensor. Four common types of PDs are considered for the classification problem; surface discharge, PD from a sharp point to ground plane, PD from semi-parallel plates, and PD from an air void in the insulation.

1.3 Thesis Arrangement

The next chapters of this thesis are arranged as follows. Chapter 2 provides background on PD and the methods used to identify it. Chapter 3 introduces the experiment setup used and the data collection phase. Chapter 4 is an overview of the pattern recognition techniques used to identify the different PD types. Chapter 5 presents the results of PD type recognition on the conducted experiments. Finally, Chapter 6 contains conclusions and recommendations.

Chapter 2: Partial Discharge Background

In this chapter, background information about partial discharge (PD) is presented. The background includes the definition of PD, PD types, and models for oil-paper insulating systems. In addition, background information is presented for PD sensing methods especially for transformers' oil-paper insulating system. As acoustic emission (AE) is the method used to sense PD activities in this thesis, different related issues like challenges and factors effecting AE measurement in on-site conditions are addressed.

2.1 Partial Discharge (PD)

2.1.1 PD Definition

Partial Discharge, or PD, as defined by the International Electro-Technical Commission, IEC 60270, is a "localized electrical discharge that only partially bridges the insulation between conductors." PD activities can be initiated at any point in the insulation system where the breakdown down strength of such points is less than the applied electric field strength. PD is usually observed when the insulation system is aging with time. In addition, improper installation, poor design and/or workmanship, especially in cable joints and terminations that are assembled on-site, can be a main cause of PD to initiate. PD in general can occur externally or internally with respect to the insulation medium. After initiation, PD usually propagates and expands causing further insulation degradation and aging. In power transformers as an example, external PD can occur on the surface of oil impregnated paper because of different insulation medium interfacing that can intensify the electric field and lead eventually to failure in the paper insulation [16]. For internal PD in power transformers, one source can be due to the presence of air voids inside the insulation paper that can cause electric field enhancement inside the insulation, which may lead to a complete breakdown.

2.1.2 PD Types and Modeling

PD inside an oil-paper insulated system can be initiated by different sources that intensify the electric field beyond the strength of the insulation system. It can happen along the interface of different materials with different dielectric constants causing the electric field to be intensified. This type of PD is called surface discharge that usually has high energy, which may accelerate the aging of the insulation system [7]. Figure 1 shows surface discharge tracking within a cast resin circuit breaker spout [17].

Figure 1: Surface discharge tracking within a cast resin circuit breaker spout [17]

Another type of PD that can happen is due to voids in the paper insulation, which intensify the electric field. Here, PD propagates through the insulation in a manner similar to trees' roots. Figure 2 shows the propagation of electric treeing with respect to time [18].

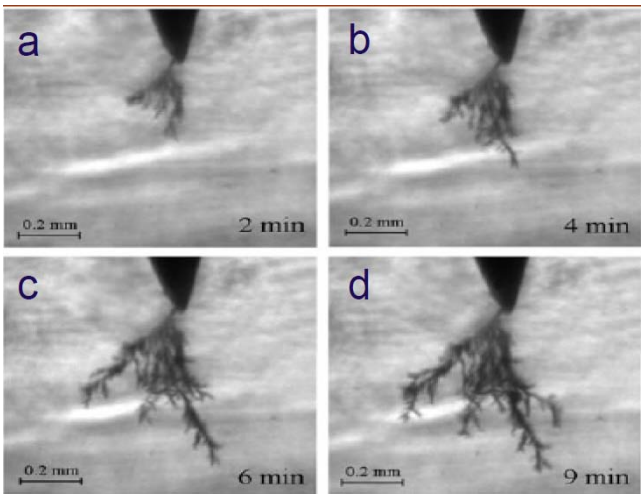


Figure 2: Electrical tree propagation due to air-filled void in insulation [18]

In the literature, several models have been proposed to simulate PD activities inside oil-paper insulation systems [16], [19], [20]. Ziomek *et al.* summarized the important rules to simulate PD activities as follows [7]:

- The construction of models should have the same material structure, shape and proportional geometrical dimensions as the part of the insulation that is modeled; it should reproduce the same mechanism of initiation and development of partial discharges.
- The construction of models should reproduce real distribution of electric field.
- The oil in the model should have similar composition as the real one – content of water, dissolved gases and solid impurities.

In this study, the PD models inside oil-transformers were divided into ten different models as described in Table 2.

Table 2: Ziomek *et al.* PD models [7]

PD1	Surface discharges in uniform electric field, where the normal component of field strength vector is insignificant
PD2	Surface discharges in moderate non-uniform electric field with small normal component of field strength vector
PD3	Surface discharges in non-uniform electric field with large normal component of field strength vector
PD4	Partial discharges in internal gas void
PD5	Partial discharges in gas bubbles in oil
PD6	Pressboard penetrating discharges (puncture)
PD7	Partial discharges from a needle in oil (point-to-plane electrode system)
PD8	Partial discharges from a needle on free potential in oil
PD9	Partial discharges in oil wedge at the oil-pressboard-electrode triple junction
PD10	Turn-to-turn insulation discharges (interturn)

Figure 3 shows the schematic diagram of PD3 and PD4 mentioned in Table 2. Feliciano *et al.*, however, divided PD models in oil-transformers into only four main categories [19]. The categories are corona in oil from sharp points, floating electrode in oil, air-filled cavity, and surface discharge as demonstrated in Figure 4.

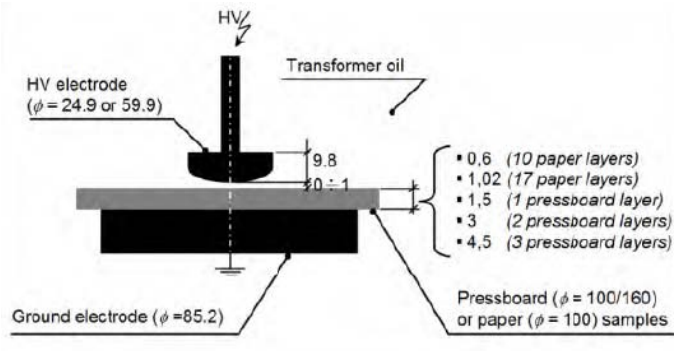


Figure 3: Schematic diagram for PD 3 and PD 4 in Table 2 considering an internal gas void presence in the pressboard for PD 4 [7]

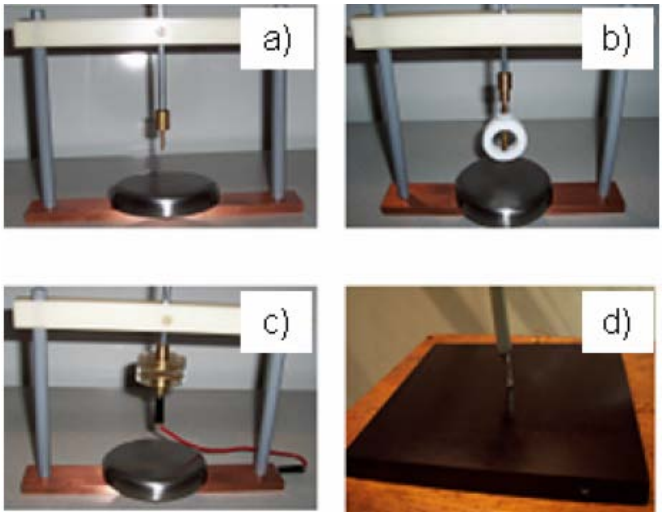


Figure 4: (a) oil-corona discharge, (b) oil-floating electrode discharge, (c) air cavity discharges, and (d) surface discharge [19]

H. Ma *et al.* recently presented five models for different PD types that can happen in oil-paper insulated systems as shown in Figure 5 [21].

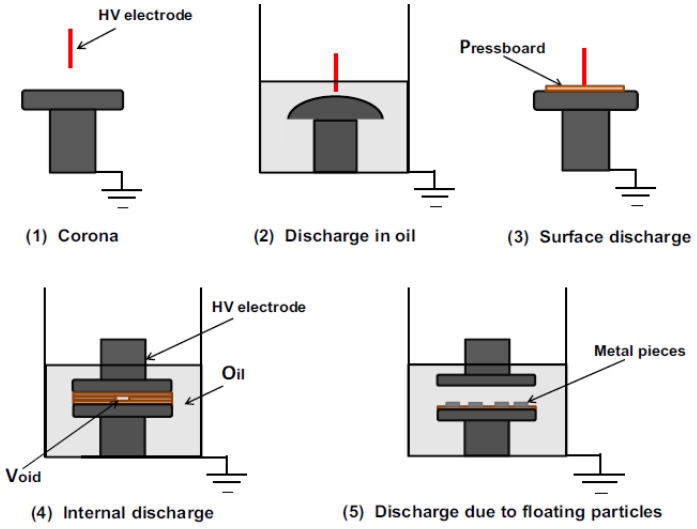


Figure 5: PD models used in H. Ma *et al.* experimental setup [21].

2.2 PD Sensing Methods

When PD activities are initiated, the resulting energy is transformed into different energy forms as mechanical energy represented by acoustic emissions, electrical, thermal, and chemical energy [7]. As a result, a wide range of sensors and techniques can be used to detect PD activities as presented in [22]. Figure 6 shows the use of different sensors to detect PD in a gas-insulated switchgear (GIS).

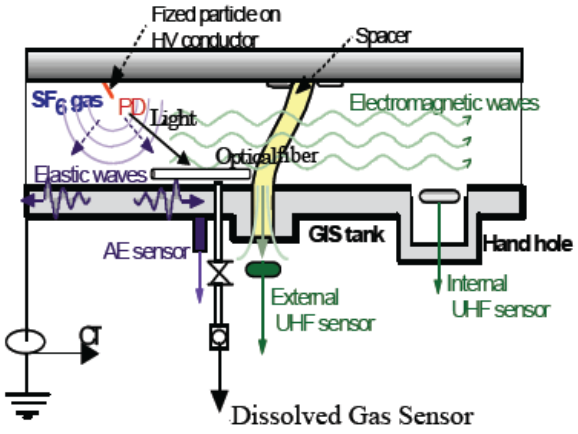


Figure 6: Different sensing methods to detect PD in GIS [22]

However, for oil-paper insulated systems in transformers, not all measurement sensors are applicable due to transformers' complicated internal structure and large outside dimensions [22]. In practice, three types of sensors are commonly used with oil-paper insulated systems for transformers [7], [22]. The methods are dissolved gas in oil analysis (DGA), ultra high frequency (UHF), and acoustic emission (AE). These methods will be explained in the following subsection.

2.2.1 Dissolved Gas in Oil Analysis (DGA)

Different gases can be formed and dissolved in power transformer's oil during the aging of its oil-paper insulation system. These gases tend to form as a result of electrical and thermal abnormalities that can be a result of PD activities or even overloading the transformer. The percentages of these gases in oil are correlated to the source of the abnormalities inside the oil-paper transformer and serve as good indicators to the health status of the insulation system. For example, pyrolysis of hydrocarbons is formed when the insulation system is overheated. In addition, separation of acetylene and hydrogen or ethylene and methane are results of PD

activities in the transformer. Therefore, DGA is utilized by taking oil samples from the transformer periodically and sending them to a specialized laboratory for analysis. There are several standards for sampling, analyzing, and testing the oil samples to interpret and correlate different gas concentrations. These include the ASTM D3613, ASTM D3612, and ANSI/IEEE C57.104. Table 3 shows an interpretation of an oil sample test and the recommended actions where ppm stands for part per million. [23]

Table 3: ANSI/IEEE C57.104 oil test interpretation [23]

Gas Description		Key Gas Concentration (in ppm)		
		Normal Limits* (<)	Action Limits** (>)	Potential Fault Type
Hydrogen	H_2	150	1,000	Corona, Arcing
Methane	CH_4	25	80	Sparking
Acetylene	C_2H_2	15	70	Arcing
Ethylene	C_2H_4	20	150	Severe overheating
Ethane	C_2H_6	10	35	Local Overheating
Carbon monoxide	CO	500	1,000	Severe overheating
Carbon dioxide	CO_2	10,000	15,000	Severe overheating
Total Combustibles	TDCG	720	4,630	
* As the value exceeds this limit, sample frequency should be increased with consideration given to planned outage in near term for further evaluation. ** As value exceeds this limit, removal of transformer from service should be considered.				
This table is derived from information provided within ANSI/IEEE C57.104				

However, DGA has some limitations and drawbacks. According to SIEMENS, a leading manufacturer of oil-paper power transforms, DGA cannot help in finding the fault location, detecting acute faults that develop within seconds or minutes, or detecting temperatures that are below 150°C degrees for a long time as the one caused by faulty cooling operation and leads to the degradation of the paper and oil [24]. In addition, DGA is a lengthy and time-consuming process. Furthermore, the fact that it is done periodically might give misleading information about the health status of the transformer where abnormalities might happen right after the sampling process. Moreover, the DGA sampling process is done usually when the transformer is offline, which is extremely inconvenient for both the utility and the customer. Although some manufactures have developed on-site DGA units, which are suitable for integration with the new smart grid concept, there is a high price tag associated with them.

2.2.2 Ultra High Frequency (UHF)

Ultra high frequency (UHF) components and electromagnetic waves (EM) in form of pulses are emitted as a result of PD activities in the insulation system of the power transformer. UHF probes are utilized to sense such signals; a typical UHF probe is shown in Figure 7. UHF probes usually take advantage of the transformer tank structure that serves as a Faraday cage, which minimizes the interference of the noises outside the tank. Therefore, UHF probes are inserted inside the transformer tank through customized windows or often through the oil valves of the transformer tank, which limit the number of probes that can be used. In addition, installment of such probes is done usually when the transformer is offline, which is not convenient for the utility or the customers.

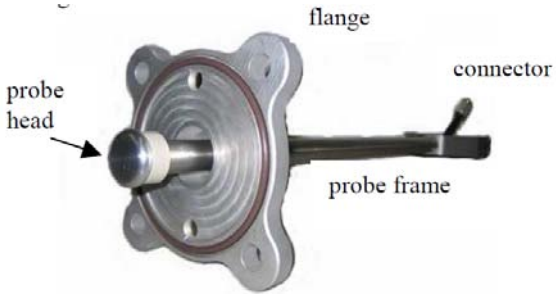


Figure 7: UHF probe for standard oil valve [25]

A UHF probe is a very sensitive sensor that can give sometimes false alarms, which is non-desirable for cost effective maintenance scheduling. However, this high sensitivity can be used to trigger other less-sensitive sensors such as an acoustic emission sensor to start capturing. Furthermore, signals received by UHF probes can be correlated to the type or source of PD activities. In addition, with the help of multiple UHF probes, the location of the PD activities can be determined. The typical frequency range that such probes can capture is 300 MHz to 3 GHz. Such a high frequency range makes it very expensive to process in terms of the equipment specifications and time requirements.

2.2.3 Acoustic Emission (AE)

Acoustic emission (AE) is a mechanical energy form emitted among other energy forms when different PD activities are initiated within the insulation system. AE sensors utilize such emissions or waves through a piezoelectric material that converts them into electrical signals, which are easy to process. AE sensors are usually very small in size as they can reach up to 200 grams and 30x60 mm in dimension [26]. Figure 8 shows different types and sizes of available AE sensors.

Figure 8: Different types and sizes of AE sensors [26]

The installation of AE sensors is quite easy for power transformers, where they can be mounted on a magnetic holder that can stick on the metallic walls of the transformer while it is online, as depicted in Figure 9.

Figure 9: AE sensor installation

Besides being small and easy to install, AE sensors are inexpensive compared to other types of sensors, which makes it very convenient to use multiple sensors around the tank for better coverage. Furthermore, the frequency range of AE waves is from 20 kHz up to 1 MHz, which is much less than the UHF range, which makes AE sensors cost-effective in terms of equipment specifications and time requirements. Table 4 summarizes and compares online PD sensing methods for power transformers.

Table 4: Comparison between practical PD sensing methods within transformer's insulation system [7] [22]

	Method		
	DGA	**UHF**	**AE**
Measured Quantity	Dissolved gases in oil	Electromagnetic wave	Sound
Detection Sensitivity	High (depending on PD activity time)	High (depending on distance and location of the PD source)	Moderate (depending on the location of the PD source)
Intensity Measurement	No	Limited	Limited
PD Identification	Limited	Yes	Yes
Frequency Band	NA	300 MHz to 3 GHz	20 kHz to 1 MHz
PD Location	No	Yes	Yes
Installation Difficulty	Moderate (transformer must be turned off)	Moderate (through oil valve, transformer under load /through dielectric window, transformer must be turned off and opened)	Low (transformer under load)
Number of Sensors	1 (due to high cost)	Limited by number of oil valves or dielectric windows	Open structure, typically 1-16

Due to the clear advantages of AE over the other possible methods, AE has been chosen for this thesis work. Next, further concentration is presented on AE practical measurement in the field.

2.3 Challenges with AE Measurement in the Field

There are several challenges that can affect measuring AE signals in the field environment. The following subsections explore the main challenges such as selecting AE sensor bandwidth, field mechanical noises, barriers inside the transformer tank, and variation in oil temperatures.

2.3.1 Acoustic Sensor Bandwidth

AE sensors, as stated earlier, convert mechanical acoustic waves into electrical signals through a piezoelectric material. The typical frequency range of the AE sensors used to register PD activities is divided into two bandwidths; narrowband sensors (20-100 kHz) and wideband sensors (100 kHz up to 1MHz). Selection of the optimal acoustic emission sensor type for registering PD activities is a tradeoff between sensitivity and PD types the sensor can cover. Narrowband sensors are recommended by some professionals because of their higher sensitivity in the narrowband bandwidth [7]. On the other hand, wideband sensors are recommended by others as they are capable of detecting all types of PD activities in oil-paper insulated transformers [7]. However, this type of sensor is less sensitive to surface discharges [7]. Figure 10 and Figure 11 show typical narrowband and wideband sensors' bandwidth versus their sensitivity, respectively.

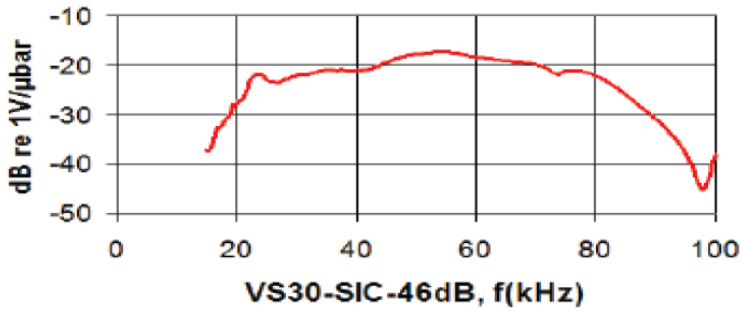


Figure 10: Narrow bandwidth, 20-80 kHz [26]

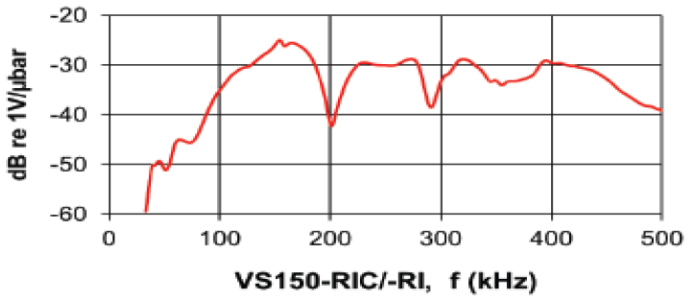


Figure 11: Wide bandwidth, 100-450 kHz [26]

2.3.2 Environmental Noises

Another concern in using AE sensors, especially on-site, is the presence of other mechanical disturbances or noises. The following list shows the main disturbance sources that might be presence on-site [7]:

- Switching of on-load tap changer
- Thermal faults of transformer's active part
- High-voltage switchgear operations near the investigated transformer
- Environmental noises (thunderstorms, rain, wind)
- Core magnetostriction noise (Barkhausen effect)
- Loose shielding connection in transformer tank

Fortunately, most of the AE sensors' bandwidth for PD detection is of good nature as PD has higher spectral components than mechanical environmental noises as wind or rain as seen in Figure 12 [27]. This also gives a good hint for choosing the proper bandwidth for the AE sensor.

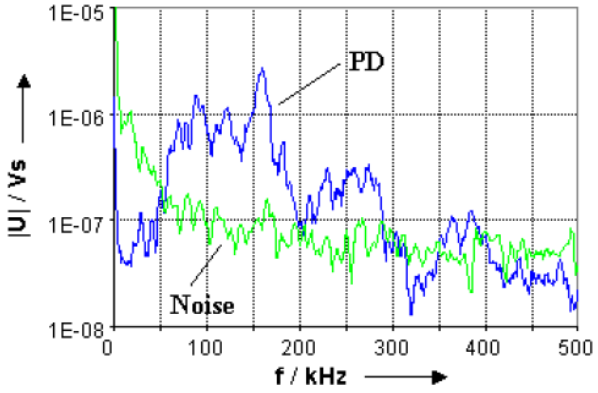


Figure 12: Frequency spectra of acoustic measurement using piezoelectric sensor [27]

2.3.3 Oil Temperature

Besides insulation, transformer oil serves an important role in cooling the transformer into safe operation limits. Power transformer's oil has different temperatures based on its loading. The oil temperature during heavy loading can reach 80 ^{0}C using high efficiency copper conductors [28]. AE waves in oil have different propagation speeds at different oil temperatures. Table 5 shows the effect of increasing the oil temperature on the acoustic wave speed.

Table 5: Speed of AE waves in oil at different oil temperatures [7]

Oil Temperature (ºC)	AE wave speed (m/s)
20	1413
50	1300
80	1200
110	1100

The speed of AE waves has a direct relationship with the acoustic impedance (Z) of the medium. The acoustic impedance, which is analogous to the electrical impedance, determines the ease of AE propagation through the medium [29]. The acoustic impedance can be expressed as the product of the equilibrium density of the medium (ρ), and the AE wave speed (V) [29]:

$$Z = \rho V \quad (1)$$

Therefore, the acoustic impedance will vary depending on the AE speed at the different oil temperatures.

2.3.4 Barrier inside the Transformer

Power transformers contain different materials like oil, pressboard, and metallic core and conductors. The speed of the AE wave varies depending on the material it propagates through. Table 6 shows the acoustic wave speed in different transformer materials. Consequently, the acoustic field inside the transformer tank is very complex because of wave reflection and diffraction caused by having an acoustic impedance mismatch in the different mediums (steel, copper, pressboard, oil) [7], [29].

Table 6: Speed of AE wave in different mediums [7]

Medium at 20 ºC	**AE wave speed (m/s)**
Transformer oil	1413
Impregnated pressboard	1500
Copper	3570
Steel	5100

Chapter 3: Experimental Setup

In this chapter, the experimental setup is described. The description includes the explanation of the equipment used, the PD models simulated, and the choice of the AE sensor type. In addition, the experiment details are described for each of the simulated practical conditions. The practical conditions simulated (the data collection phase) is based on varying the PD location, changing the sensor location, altering the oil temperature, and finally inserting a metallic barrier in the line-of-site between the PD source and the AE sensor.

3.1 Main Equipment List

Figure 13 shows the overall experimental setup used throughout the study.

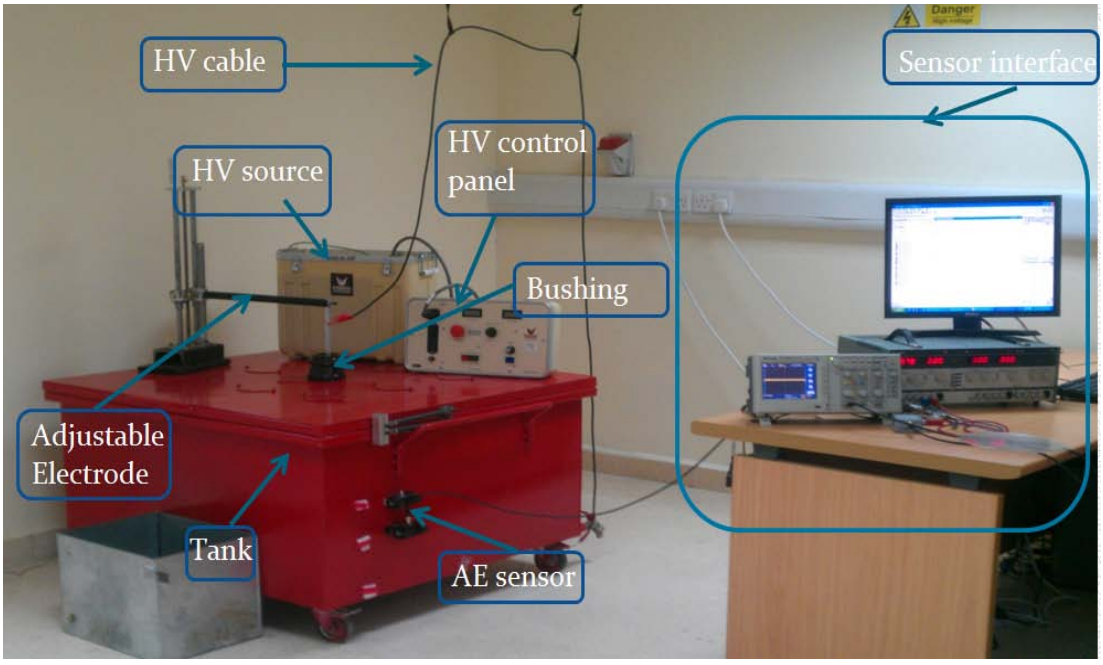


Figure 13: Overall experimental setup

The following subsections give a description for each item used in the setup.

3.1.1 Oil Tank

The oil tank used has a dimension of 1x1x0.5 m and is filled with aged oil received from a local utility company. The tank top cover has 9 symmetrical holes that can be used as possible PD locations as shown in Figure 14.

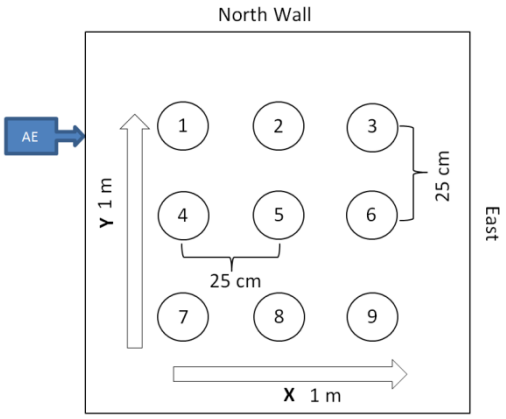


Figure 14: Tank bird's-eye view

The oil level inside the tank is about 20 cm where the AE sensor can be placed at three different depths as depicted in Figure 15. However, the sensor is always fixed on the tank's west wall close to hole or PD location 1 as shown in Figure 14. The AE sensor is fixed on the tank's wall with a magnetic holder, Figure 9. In addition, silicone grease was applied between the tank surface and AE sensor to increase the quality of the transferred AE signal and to reduce the reflections on the contact surface.

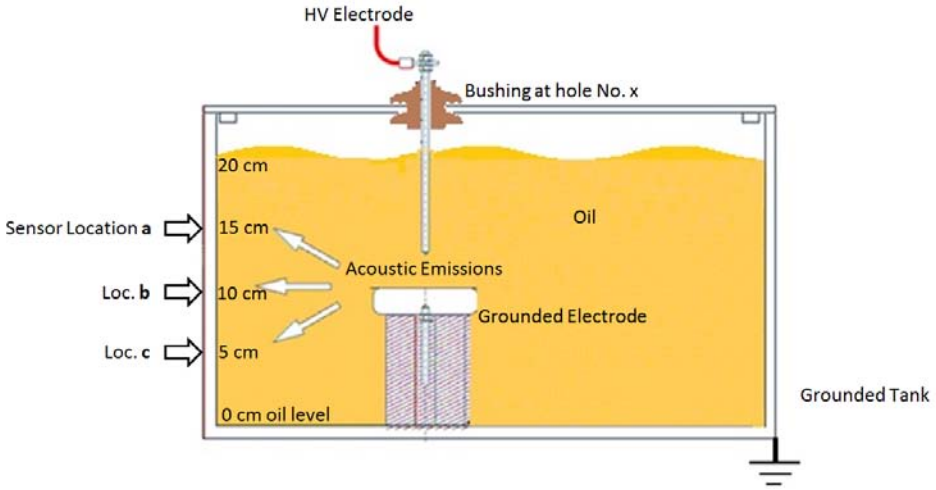


Figure 15: Different locations of AE sensor

3.1.2 High Voltage Supply

The high voltage source used in the setup is a 40 kV, 10mA 50/60 Hz AC tester. The high voltage source is accompanied with a control panel to adjust the desired voltage level and trip the source at a preset current limit. The control panel also has digital meters to accurately give the readings of both the voltage in kV and current in mA. The control panel and the high voltage source shown in Figure 16 can be connected together through a 3 meter long cable for a safe distance to the operator.

Figure 16: High voltage source and the control panel

The high voltage source is connected to the high voltage electrode through a cable. The cable is suspended in the air to prevent it from being close to the earthed tank, which would introduce undesired PD between the cable and the tank. In addition, a ceramic bushing is used to provide insulation between the earthed tank and the immersed high voltage source going inside the tank through one of the possible holes on its top cover. However, the available bushing can efficiently insulate only up to 15 kV, which limits the possibility to use the full range of the high voltage device. Nevertheless, the 15 kV is sufficient to produce the PD types examined in this thesis. Finally, the grounding reference for the high voltage source is connected to the system common ground with the earthed tank.

3.1.3 PD Types Electrodes

The selection of PD types in this work is based on the most common types found in the literature as previously described in subsection 2.1.2. The ground electrode has a height of 8 cm and a disk diameter of 10 cm as shown in Figure 17. Both the high voltage and ground electrodes are fully immersed in the tank's oil as depicted in Figure 15. In addition, the location of the ground electrode can be adjusted to be aligned with the high voltage electrodes coming through any hole in the tank's top cover.

Figure 17: Ground electrode top and side view

The following is a description of the four PD electrode types used in this study:

1- Sharp electrode to ground plane: This type of PD occurs mostly when there are defects in the manufacturing or assembling of the transformer windings or tank structure. The sharp point intensifies the electric field around it to a level beyond the insulation system strength that initiates dangerous PD activities. Figure 18 shows the electrode arrangement used in this study.

Figure 18: Sharp point to ground electrode

2- Semi parallel plate electrode: electrode arrangement simulates PD due to semi-uniform electric field. Such an arrangement is less severe compared to the point-plane arrangement. Figure 19 shows the electrode setup used.

Figure 19: Semi parallel plate electrode

3- Surface discharge: This type of PD happens mostly along the interface of different materials with different dielectric constants causing the electric filed to be intensified. This type of PD is very dangerous as it has high energy associated with it, which accelerates the aging of the paper insulation and affects the oil strength. Figure 20 shows the electrode setup used with oil-impregnated aged insulation paper. The reason to choose a sharp electrode with the surface setup is to reduce the voltage required to initiate PD activities and keep it within the 15kV limit. In addition, such a configuration represents a worst-case scenario.

Figure 20: Surface discharge electrode

4- Void discharge. This type of PD happens mostly due to defects in the manufacturing phase of the paper insulation. This type of PD has a significant effect on the life of the paper insulation system as it propagates with time and expands, which ultimately leads to catastrophic failure. To simulate a void in the insulation, three PVC discs are glued with special material together, where an air-filled hole or void is inserted in the center of the middle disc. The reason to choose PVC is because it has a very close dielectric constant to an impregnated insulation paper in oil [30]. Figure 21 shows the electrode setup used. PVC discs have bigger diameter than the ground electrode to insure PD is only initiated within the inserted void and not due to surface discharge.

Figure 21: Void discharge electrode

3.1.4 AE Sensor Bandwidth Selection

As previously discussed, there are two types of AE sensors. The first type is the narrow band sensors with a bandwidth ranging from 20 kHz-100 kHz. These sensors are sensitive to surface discharge but not immune to acoustic noise. On the other hand, the second type is the wide band sensor with a bandwidth of 100 kHz to 1MHz. The major advantages of the wide band sensors are their ability to detect all types of PD signals. Moreover, they are immune to most of the acoustic noise as depicted in Figure 12 where PD signals above 100 kHz have considerable higher spectral components than mechanical environmental noises as wind or rain. The use of two different sensors in this research has been investigated. The first sensor has a bandwidth of 20-80 kHz and the second has a bandwidth of 100-450 kHz. The former sensor performs extremely undesirably with a great response when a small level of noise is introduced in the level of 50dB, which is the level of noise in a quiet suburb or at around 30.5 meters from a large electrical transformer [31]. Figure 22 shows the narrow bandwidth sensor response in quiet lab conditions around 35dB to 40dB, whereas Figure 23 shows its response at 50dB noisy conditions. Using an AE sensor with a narrow bandwidth would be extremely unreliable in field conditions.

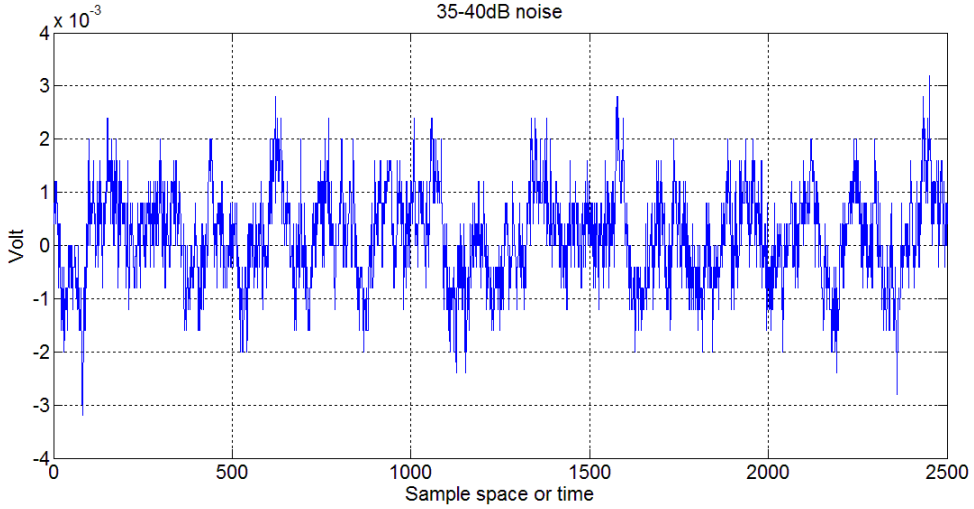


Figure 22: Narrow bandwidth sensor response in quiet lab conditions

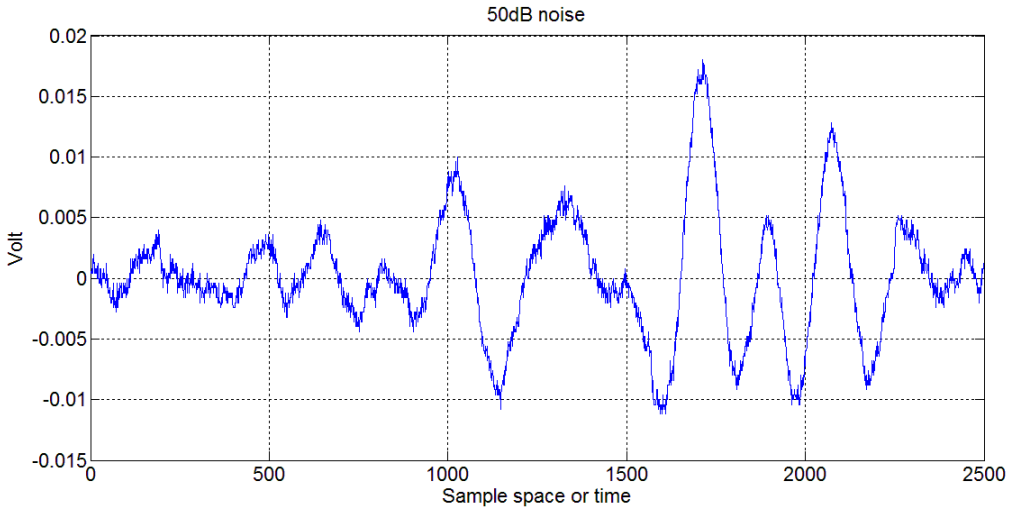


Figure 23: Narrow bandwidth sensor response in a noisy condition

The later sensor, with a bandwidth of 100 kHz to 450 kHz, was tested against a noise level of up to 80 dB, which is the level of noise in an average factory [31]. Figure 24 shows the high bandwidth sensor performance in around 80dB noisy conditions. The sensor shows an immune response to the noise introduced. Therefore, the high bandwidth AE sensor, 100 kHz to 450 kHz, is chosen for data collection in this thesis.

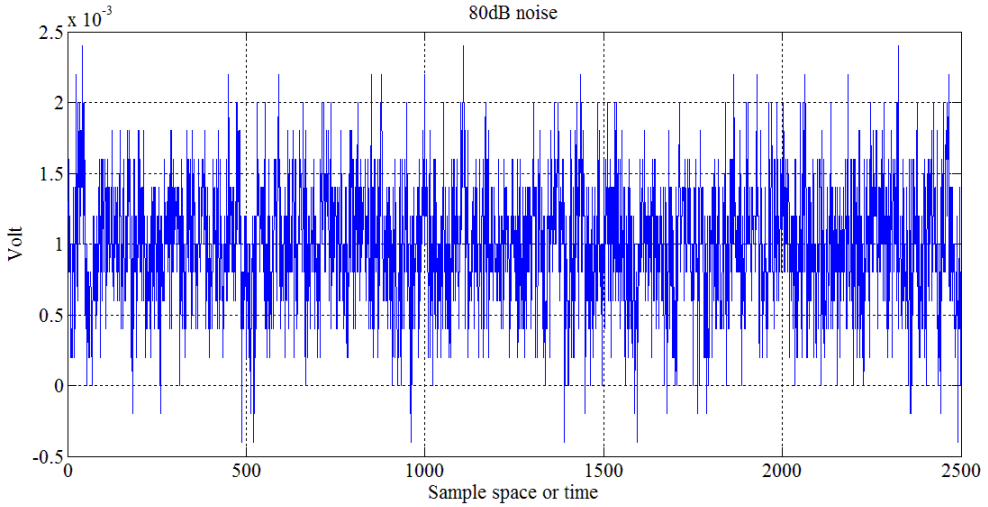


Figure 24: High bandwidth sensor response in noisy conditions

3.1.5 AE Sensor Interfacing

The AE sensor chosen has a bandwidth that spans from 100 kHz to 450 kHz. Therefore, the least sampling frequency should be equal to or greater than double the maximum frequency limit of the sensor, according to the Nyquist theorem. However, in practice and by viewing a sample PD pulse signal on the scope, a sampling frequency of 10 MHz sample per second is sufficient to view a complete pulse signal, as shown in Figure 25.

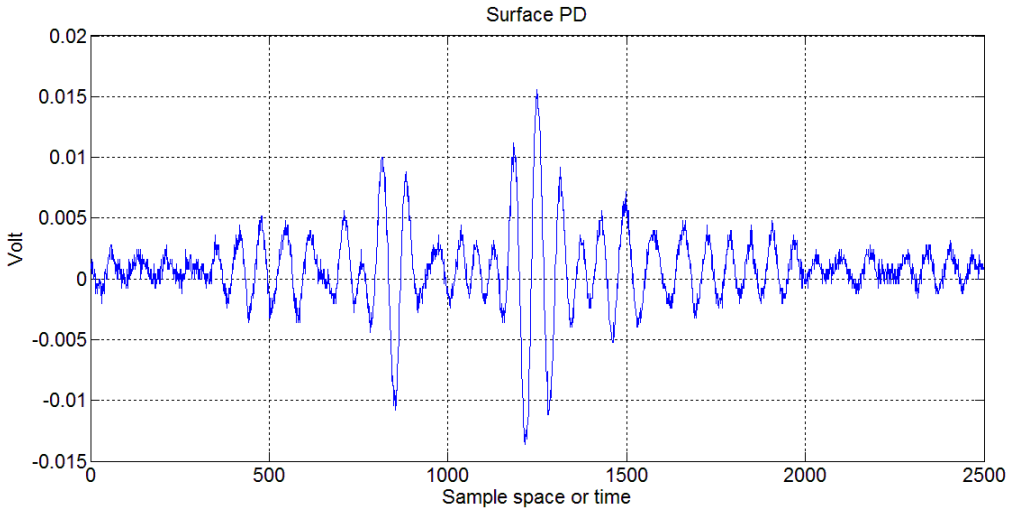


Figure 25: Sample AE signal from PD pulse

The data acquisition equipment used is a 60 MHz bandwidth oscilloscope that can only save a 2500 point per viewed signal. Therefore, with a 10 MHz sampling frequency, the viewed signal has a length of 250 micro seconds. The scope has been interfaced with Matlab to register the signals directly in specified directories on a computer.

Before connecting the AE sensor to the oscilloscope, an amplifier stage of 46dB into 50 ohms is installed. The signal is then fed into the oscilloscope that has a setting to filter or reject any high frequency noises that arise, especially the ones that are induced from the high voltage source. The next subsection discusses the collected data of the conducted experiments.

3.2 Data Collection

In this work, data from the four different PD types have been captured at different simulated field conditions. These conditions can affect the measured AE signals emitted from PD activities. The simulated field conditions are:

a) Changing PD location.
b) Changing sensor location.
c) Changing oil temperature.
d) Having a barrier in the line-of-sight between the PD source and the AE sensor.

The following subsections give descriptions about the experiments conducted to simulate these conditions.

3.2.1 PD Location

PD activities can be initiated at any point in the insulation system for power transformers. Therefore, different PD types have been simulated at three different locations in the tank model as demonstrated in Figure 26. Table 7 summarizes the experiments done with the different PD types at the different locations.

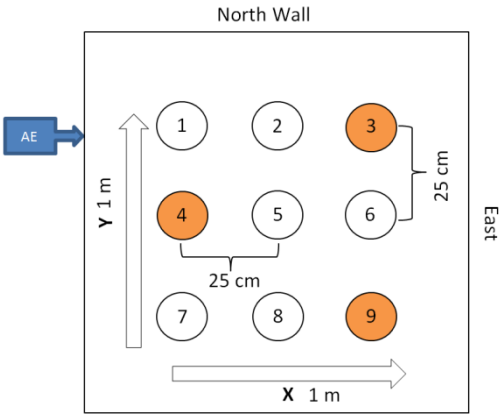


Figure 26: PD location considered

Table 7: PD types and locations data

PD Type	PD Location	Applied Voltage (kV)	Measured current (mA)
Sharp Point	4,9	8	0.20
Surface Discharge	3,4,9	11.8	0.25
Semi Uniform	4	9.5	0.23
Void	3,4,9	6	0.18

The line-of-sight distance between the AE sensor and PD at locations 3, 4, and 9 are 75, 36, and 95 cm, respectively.

3.2.2 AE Sensor Location

The sensor, as discussed earlier, can be placed at three different depths as depicted in Figure 15. However, the sensor is always fixed on the tank's west wall close to hole or PD location 1, as shown in Figure 26. PD location 4 has data for all PD types at the three different sensor depth locations.

3.2.3 Oil Temperature

Power transformer's oil has different temperatures based on the loading condition. In addition as discussed, AE waves propagate at different oil temperatures at different speeds. To study the temperature effect, different data sets were taken at around 23 ^{0}C, 50 ^{0}C, and 70 ^{0}C. All data were taken at PD location 3 and at sensor location **b**. Only surface discharge and void discharge experiments were conducted at the different temperature levels. Figure 27 shows the equipment used in the heating experiments as the thermometer and one of the heaters.

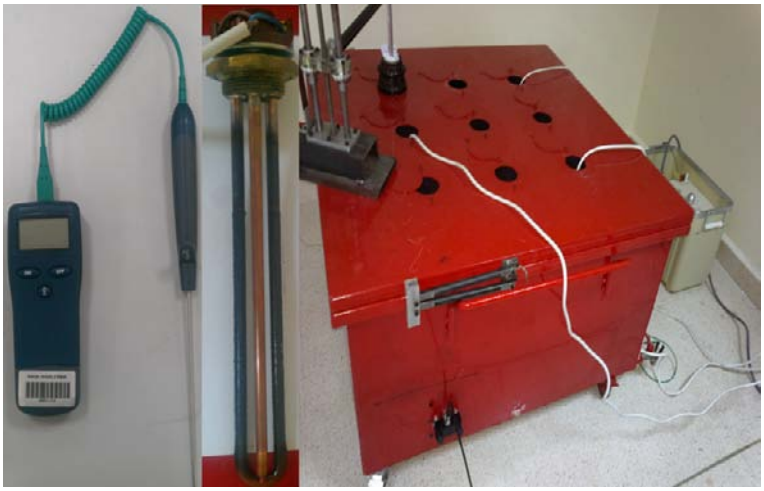


Figure 27: Oil temperature experiment

3.2.4 Barrier Insertion

AE waves have different propagation speeds through the transformer's different materials. The core of the transformer can act as a barrier between the AE waves emitted from a PD source and the AE sensor. To simulate the barrier effect, a metallic obstacle made of transformer core silicon steel is inserted in the line-of-sight between the AE sensor and the PD source. The experiment configuration is shown in Figure 28.

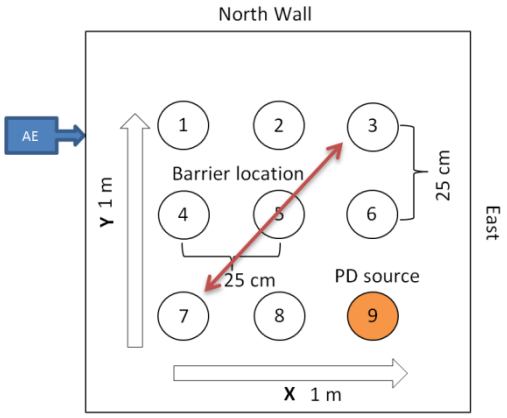


Figure 28: Barrier insertion experiment

The PD measurements were conducted at PD location 9 with the AE sensor fixed at location **a**. Only surface and void discharge data had been collected. However, because of attenuation caused by the barrier, the AE samples could only be detected by increasing the applied voltage to around 4 kV compared to Table 7. This attenuation might be one of the main disadvantages of the AE method.

3.3 Experimental Observations

In this subsection, experimental observations are listed.

1- The PD signal tends to be more repetitive when the high voltage source is applied for a longer time. This agrees with the fact that the insulation loses its strength faster with time when PD activities are present.

2- Some PD signals registered at sensor location **c** contain reflection interferences. This can be explained by the fact that the sensor at location **c** is very close to an edge with the tank's base that may result in multiple reflections. Figure 29 and Figure 30 show an AE signal at PD location 4 with the sensor fixed at location **b** and **c** respectively.

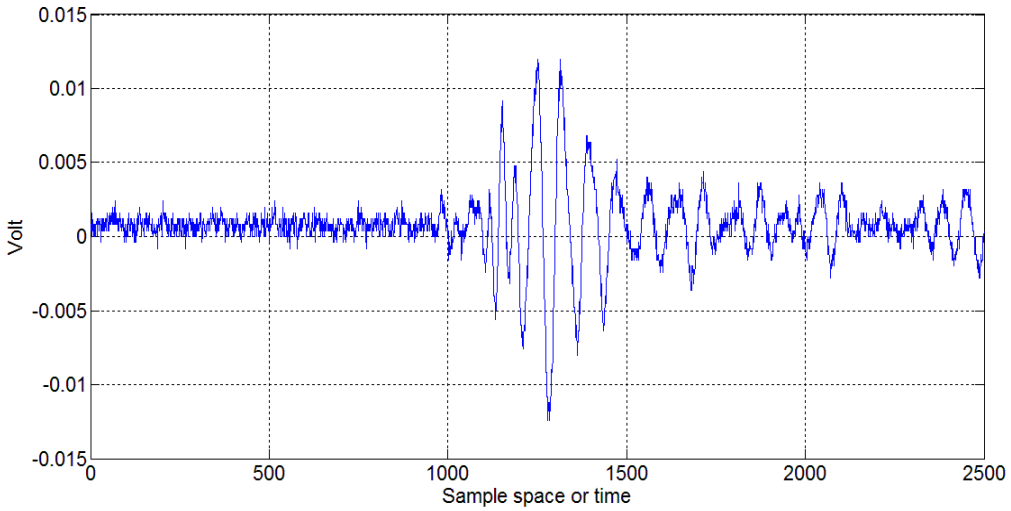


Figure 29: Typical AE signal at sensor location **b**

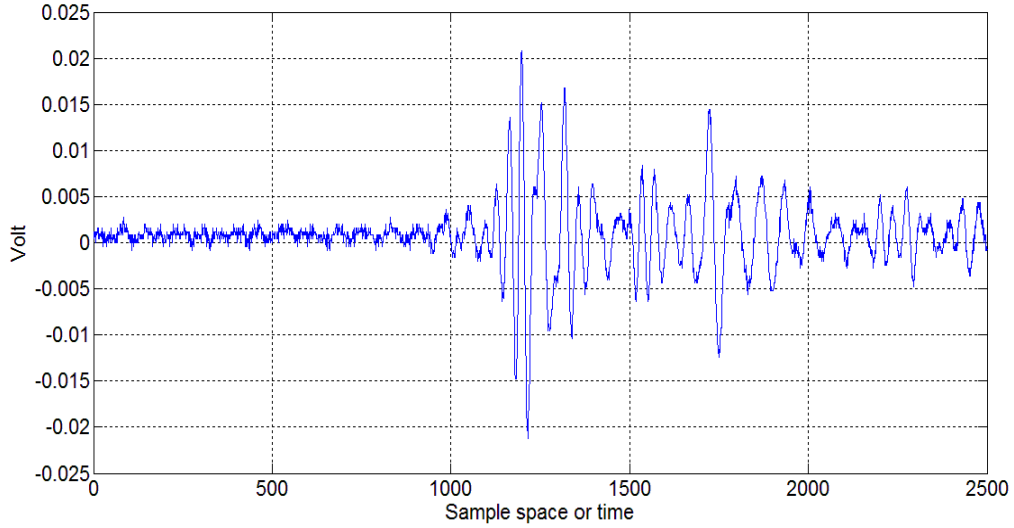


Figure 30: AE signal with interferences at sensor location **c**

3- Most PD signals received from PD location 9, which has a 95 cm line-of-sight distance from the AE sensor, were registered with some reflection interferences. This can be explained by the relatively long distance the signals are travelling through oil and hence the high possibility of the AE signal to interfere with other AE signals traveling at different speeds in the tanks' metallic walls. Figure 31 shows a signal registered from PD location 9 with the AE sensor fixed at location **b**.

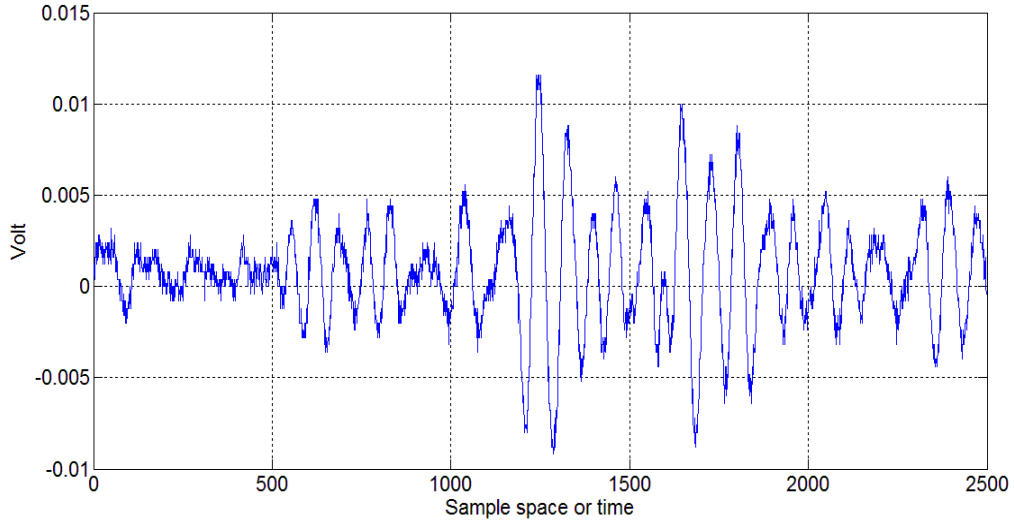


Figure 31: AE signal with interferences from PD location 9

4- In the oil heating experiments, PD activities were initiated at a lower applied voltage for heated oil than at room temperature. Table 8 shows the oil temperature and the corresponding applied voltage for the surface discharge experiment at PD location 9. This indicates that the insulation system strength gets weaker as the temperature rises. This can be explained by the fact that free electrons at higher temperatures have higher energy to initiate PD activities.

Table 8: Applied voltage for surface discharge at different oil temperatures

Oil Temperature (ºC)	**Applied Voltage (kV)**
23	11.8
70	6.9

Chapter 4: Pattern Recognition

This chapter provides a background on pattern recognition in general, and gives emphasis to the selected feature extraction methods and classifiers used in this study.

4.1 Pattern Recognition Review

Pattern recognition can be defined as the process that assigns labels to unlabeled inputs based on discriminating properties they possess, usually referred to as “features”. The assigned labels correspond to distinct groups usually referred to as “classes”. The system that does the pattern recognition is referred to as the “classifier”. There are a wide range of features to choose from and classifiers to select. This selection is usually dependent on the type of data that needs to be recognized. For example, in voice recognition, if the data that needs to be classified are the different language phonemes, frequency features might be ideal for such kind of data. In addition, the classifier selection is usually based on the distribution shape of the data. In general, there are two types of classifiers, i.e. parametric and non-parametric. In parametric classifiers, the distribution shape of the different classes is usually known. Therefore, with sufficient training data, parameters of distribution can be estimated such as the mean and variance, and classifiers as maximum likelihood (ML) can be used. On the other hand, non- parametric classifiers follow no theoretical assumptions nor have information about the data distribution shape, which is the case with most of the “real world” or practical data.

Pattern recognition processes usually follow certain procedures and steps. Figure 32 shows the complete structure of a typical pattern recognition system [32].

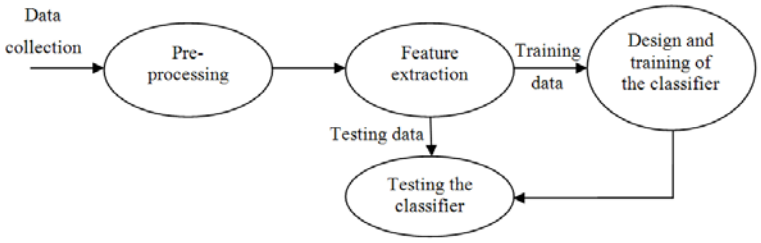


Figure 32: Pattern recognition steps [32]

The data collection phase is responsible for gathering representative raw data for each class under real conditions. If the environmental conditions presented at the data collection phase are not realistic, the pattern recognition system usually achieves poor results when tested in practice. For example, in voice recognition, this effect is very clear when training with data without noise, as in a quiet room and test in a noisy environment as a busy street. The second phase, pre-processing, is responsible for making the data more representative, such as by removing outliers from the data set. The feature extraction phase, as stated earlier, depends on the data type and aims to obtain the most distinguished features for the different classes. Furthermore, the classifier design depends on the distribution shape of the data. Usually large or sufficient pre-known class data is used as part of the classifier design, and is referred to as the "training data". The last step is to validate and test the design with data unseen by the system before, which are referred to as the "testing data".

In this thesis, different feature extraction methods were tested including principle component analysis (PCA), discrete Fourier transform (DFT), and wavelet decomposition. In addition, different classifiers were used with the extracted features such as K-nearest neighbor (KNN), polynomial classifier, quadratic discriminant analysis (QDA), and support vector machine (SVM). Compared to the other combinations, PCA followed by KNN proved to have high and stable recognition rates when applied at different data conditions. The next section describes the pattern recognition system components used in this thesis. PCA and KNN are considered as the main feature extraction method and classifier, respectively; hence, they are described first.

4.2 Principle Component Analysis (PCA)

Principle Component Analysis (PCA) or the Karhunen-Lo`eve expansion [33] is a reduction technique that preserves data information in the reduced space with minimum loss [34], [35]. It does so by projecting the data on the direction of the largest variance in a lower dimension, which maximizes the scatter of all projected samples [36]. Therefore, PCA analysis can be a good feature extractor for classification only if the direction of the maximum variance carries distinguished class information. However, not all dimensions can carry such discriminant information. Figure 33 demonstrates how projecting two classes' data, green and red, into 2 principle components dimension can be separable and ideal for classification, while applying PCA further causes the classes to not be separable.

Figure 33: Influence of PCA component number [34]

Therefore, the choice of the new dimension order is critical for achieving a high recognition rate.

The PCA method works by processing all training samples regardless of class. Assume we have **m** training feature vectors or samples from different classes as $\{x_1, x_2, \ldots., x_m\}^T$ with each sample x_i of size **d**. The first step of the method is to subtract the sample mean from the data as follows:

$$z_i = (x_i - \mu) \tag{2}$$

where the sample mean is defined as follows:

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{3}$$

Moreover, let's assume a linear transformation matrix **E** of **d×k** size to project the samples from **d** dimension to **k** dimension, where **k** < **d**. The new reduced feature vectors $\boldsymbol{y_i}$ of size **k** are defined as follows:

$$y_i = z_i\, E \quad where\ i = 1,2, \dots, m \tag{4}$$

The transformation matrix **E** is derived from the data scatter matrix that is defined as:

$$S = \sum_{i=1}^{m} z^T{}_i z_i \tag{5}$$

The PCA, as stated earlier, maximizes the scatter of all projected samples. Therefore, the transformation matrix **E** is:

$$E = [e1\ e2 \dots ek] \tag{6}$$

where $\{ei | i = 1,2, \dots, k\}$ is the eigenvectors corresponding to the **k** largest eigenvalues of the scatter matrix. The testing samples should be shifted by the same sample mean of the training data before using the obtained transformation matrix **E** to project them into the new reduced diminution **k**.

One obstacle with PCA is the size of the scatter matrix **S**, where it is clear from Equation (5) that it would be a **d×d** matrix. However, this can be tackled using the snapshot method proved in [37]. For an **m×d** matrix containing all the training samples, the maximum number of non-zero eigenvalues that the matrix can have is min (**m-1, d-1**). The number of training samples (**m**) is usually much less than the number of the row features **d**. Therefore, the most non-zero eigenvalues that can be found are equal to **m-1**. So we can calculate eigenvalues of $\boldsymbol{z_i\, z^T{}_i}$ (an **m×m** matrix) instead of $\boldsymbol{z^T{}_i z_i}$ (a **d×d** matrix). It is obvious that the dimension of $\boldsymbol{z^T{}_i z_i}$ is much lower than $\boldsymbol{z_i\, z^T{}_i}$. The eigenvectors of the snapshot matrix are related to the first **m** eigenvectors corresponding to the largest eigenvalues of the scatter matrix as follows:

$$norm\,(E\ scatter_{first\ m\ vectros}) = norm(\, z^T\ E\ snapshot) \tag{7}$$

4.3 K-Nearest Neighbor Classifier (KNN)

Although simple and easy to implement, K-Nearest Neighbor (KNN) is a classification algorithm that has been reported as one of the top 10 algorithms in data mining [38]. In addition, it has been claimed by some researchers that KNN can achieve better results for multi-model classes than SVM, which is by far a much more complicated algorithm [38]. The KNN method is a non-parametric technique and that makes it very convenient for practical real-world data that do not follow theoretical assumptions as being linearly separable or mixtures of Gaussian distributions. The implementation of KNN is based on comparing each test sample with the entire training data set, and then making a classification decision based on certain algorithmic parameters. The first of these parameters is the choice of the value ***K*** that decides the number of the neighbor training samples that should be used to decide the class of a test sample. The selection of the best ***K*** value can be tricky; however, there are several conventions that give a good hint. For example, the classification rate might be very sensitive to noise if the value of ***K*** is chosen to be too small when the neighborhood has many data from different classes. However, in this work, ***K***= 1 and 2 show similar high results with PCA, while the recognition rate drops by increasing ***K*** beyond ***K***= 2.

Another important parameter for the KNN algorithm is the choice of the comparison method, which is usually referred to as the distance measure. There are several types of distances that can be used to calculate the similarity or distance between two points; nevertheless, it is always preferable to choose the distance type that can achieve the smallest distance that implies a larger likelihood of being the correct class. The choice of the best distance measure usually depends on the nature of the data. For example, for high dimensional data, it is advisable not to choose the Euclidean distance, as it tends to be less discriminating for large feature vectors. The following is a step-by-step algorithm for building a typical KNN classifier.

Given a training data set ***D*** and a test sample $z = (\tilde{x}_i, \tilde{y}_i)$, the KNN algorithm calculates the distances or similarities between the test point and all training points $(x, y) \in D$ to determine its nearest neighbor list, D_z, where:

- $\tilde{x}_i$ is the test data while $\tilde{y}_i$ is its class
- x is the training data set and *y* is the corresponding class set.

The test data is classified based on the majority class of its nearest neighbors as follows. First, calculate $d(\tilde{x}_i, x)$or the distance between the testing data z and each training sample $(x, y) \in D$. After that, select $D_z \subseteq D$, the set of the closest $\boldsymbol{K}$ training objects to z**.** Finally**,** the algorithm should output the class $\tilde{y}$ of each testing data as:

$$\tilde{y} = (\arg max)_v \sum_{(x_i, y_i) \in D_z} I(v = y_i) \qquad (8)$$

where v is a class label, y_i is the class label of the *i*th nearest neighbor, and the function $I(\cdot)$ is an indicator function that returns the value 1 if the argument is true or 0 if it is false.

4.3.1 Distance Correlation

As stated in the previous section, one of the most important parameter settings for KNN classifiers is the choice of the distance measure. A proper choice for this measure is the one that yields minimum value, which implies higher probability of being near to a sample from a similar class. This subsection describes distance correlation while the next subsection briefly describes the other distance measures used.

The distance correlation measures the statistical dependence between two random variables or vectors [39]. The main difference between the classical definition of correlation, Pearson's correlation, and distance correlation is the type of relationship between the random vectors when the correlation is zero. A couple of random vectors is said to be uncorrelated if correlation=0, but this does not necessarily mean they are independent. However, if the distance correlation=0 between two random vectors, the two vectors are said to be independent. Distance correlation is derived mainly from other statistical measures such as variance, standard deviation, and distance covariance. A typical formulation for distance correlation is as follows:

$$dCor(x_g, y_p) = \frac{dCov(x_g, y_p)}{\sqrt{dVar(x_g) dVar(y_p)'}} \qquad (9)$$

Or

$$dCor(x_g, y_p) = \frac{(x_g - \bar{x}_g)(y_p - \bar{y}_p)}{\sqrt{(x_g - \bar{x}_g)(x_g - \bar{x}_g)'}\sqrt{(y_p - \bar{y}_p)(y_p - \bar{y}_p)'}} \quad (10)$$

$$\bar{x}_g = \frac{1}{n}\sum_j x_{g_j} \quad (11)$$

$$\bar{y}_p = \frac{1}{n}\sum_j y_{p_j} \quad (12)$$

Next are some of the major properties of distance correlation:

i. $0 \leq dCor(x_g, y_p) \leq 1$.

ii. $dCor(x_g, y_p) = 0$ if and only if x_g and y_p are independent.

If the correlation distance or $dCor(x_g, y_p) = 1$, then x_g and y_p are completely dependent which implies that they are from the same class if KNN is used for classification. However, since it is preferable to have the nearest distance in KNN be correct, and since $0 \leq dCor(x_g, y_p) \leq 1$, the distance correlation can be modified as such:

$$dCor(x_g, y_p)_s = 1 - \frac{(x_g - \bar{x}_g)(y_p - \bar{y}_p)}{\sqrt{(x_g - \bar{x}_g)(x_g - \bar{x}_g)'}\sqrt{(y_p - \bar{y}_p)(y_p - \bar{y}_p)'}} \quad (13)$$

For example, if $dCor(x_g, y_p) = 0$, which implies that x_g and y_p are completely independent, then the distance, according to Equation (13), $dCor(x_g, y_p)_s$ is maximum and equal to 1, which implies a minimum chance for x_g and y_p to be from the same class. The next subsection briefly describes the formulas of the other tested distance measures.

4.3.2 Other Distance Measures

The following list describes the formulas of Euclidean, cityblock, and cosine distance measures:

1- The Euclidean distance is the basic distance measure described by the Pythagorean formula as follows:

$$d(x_g, y_p) = \sqrt{\sum_{k=1}^{d} \left(x_g^{(\kappa)} - y_p^{(\kappa)}\right)^2} \quad (14)$$

2- The cityblock or Manhattan distance is an approximation of the Euclidean distance. A typical formula for this distance measure is:

$$d(x_g, y_p) = \sum_{k=1}^{d} \left|x_g^{(\kappa)} - y_p^{(\kappa)}\right| \quad (15)$$

3- The cosine distance is based on computing the inner product between two vectors, which measures the similarity between them. The following is the cosine distance formula:

$$d(x_g, y_p) = \frac{x_g^T . y_p}{\|x_g\| \|y_p\|} \quad (16)$$

The next section briefly describes the tested feature extraction methods and classifiers, namely DFT and wavelet decomposition for feature extraction, and polynomial, QDA, and SVM for classification.

4.4 Other Feature Extraction Methods

4.4.1 Discrete Fourier Transform (DFT)

Discrete Fourier Transform (DFT) is a well-established technique for analyzing the frequency content of discrete time domain signals. A typical formulation to obtain the DFT coefficients X_k of a discrete signal $x_n = [x_1\ x_2\ \dots\ x_{N-1}]$ is:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \quad (17)$$

In this thesis, a small subset of the frequencies corresponding to the largest amplitudes of the DFT coefficients $|X_k|$ is used as feature vectors for the different PD classes. Figure 34 demonstrates applying DFT on an AE sample signal.

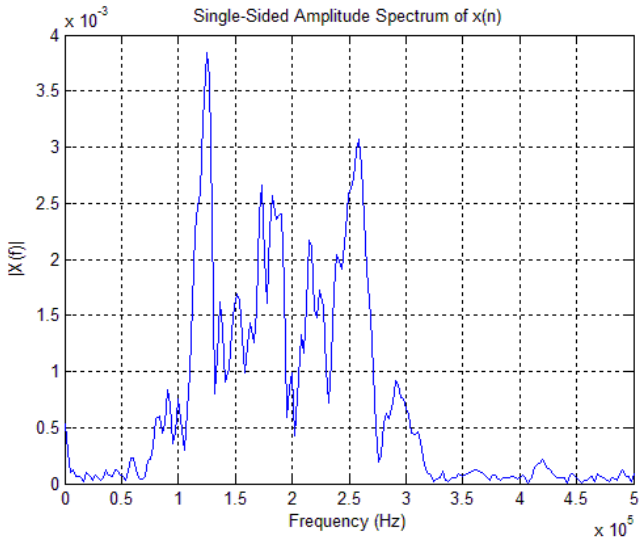


Figure 34: DFT of an AE sample signal

4.4.2 Wavelet Decomposition

Wavelet decomposition is a very useful technique for analyzing non-stationary or transient signals whose frequency response varies with time. The method is usually used for denoising, signal compression, and feature extraction. It is based on signal windowing within a variable-sized region. In other words, it processes a time domain signal with consecutive high-pass and low-pass filters. The high-pass filters reveal the high frequency contents referred to as the detail coefficients, whereas low-pass filters reveal the low frequency contents referred as approximation coefficients. Figure 35 illustrates three-level wavelet decomposition algorithm where S, H, L, ai, and di stand for signal, high-pass filter, low-pass filter, signal approximation at level *i*, and signal detail at level *i*, respectively.

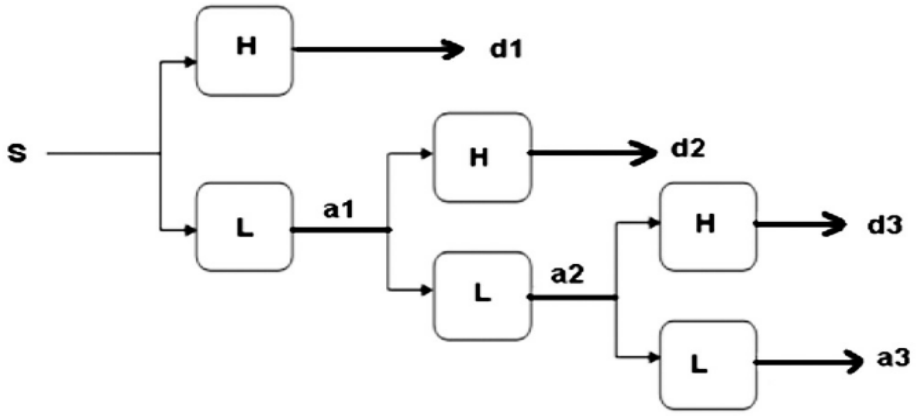


Figure 35: 3-levels wavelet decomposition [12]

The wavelet decomposition is usually represented by scales and time axes. The scales are related to frequencies by a relation that is governed mainly by the mother wavelet used. As in [12], the mother wavelet used is ‘db 15’ with five-level decomposition. Figure 36 demonstrates some of the decomposed coefficients of an AE sample x (n).

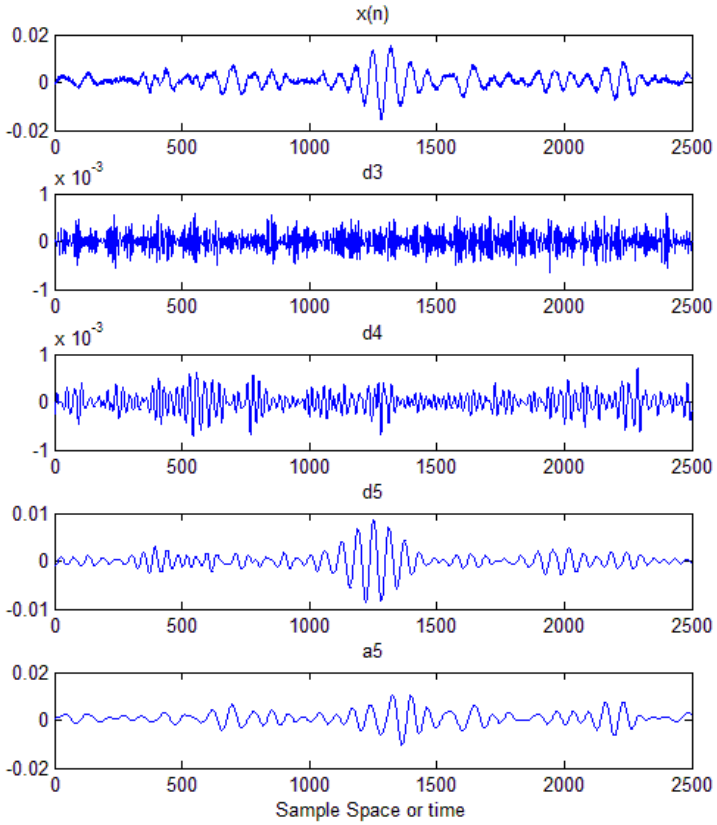


Figure 36: Decomposed coefficients of an AE sample

In this thesis, the extracted features from the decomposed coefficients are based on sub-band entropy, which is “a statistical measure of the energy dispersion among different spectral bands” [12]. The sub-band entropy used is log-energy described by the following formula:

$$E_i = \sum_i \log\left(s_i^2\right) \qquad (18)$$

where s_i is a decomposed coefficient.

4.5 Other Classifiers

4.5.1 Polynomial Classifier

The polynomial classifier can be considered as an approximation to the optimal Bayes classifier [32]. The classifier expands an incoming feature vector by adding all pairwise products of the individual elements. For example, a quadratic expansion for **X** containing 2-dimensional training feature vectors of different classes can be expressed as:

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \quad (19)$$

Hence the augmented features are defined as

$$\mathbf{X_{aug}} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{11}x_{12} & x_{11}^2 & x_{12}^2 \\ 1 & x_{21} & x_{22} & x_{21}x_{22} & x_{21}^2 & x_{22}^2 \\ 1 & x_{33} & x_{32} & x_{31}x_{32} & x_{31}^2 & x_{32}^2 \end{bmatrix} \quad (20)$$

Assuming each row of **X** corresponds to a different class, the target matrix can be defined as:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (21)$$

The weight matrix **W** is achieved by multiplying the Pseudo-inverse of $\mathbf{X}_{\mathbf{aug}}$ by the target matrix. Each incoming test feature vector **Y** has to undergo the same expansion of the training data. The class label of the test vector is then determined using the obtained weight matrix as follows:

$$C = {Y_{aug}}^T W \quad (22)$$

4.5.2 Quadratic Discriminant Analysis (QDA)

Quadratic discriminant analysis (QDA) is a typical classifier used in supervised learning problems [40]. The method is based on modeling the different classes as Gaussian distributions. The Gaussian distribution parameters can be estimated using ML on the training data of the different classes. After that, the posterior distributions are used to estimate the class of any testing sample. QDA is used instead of linear discriminant analysis (LDA) when the there is no assumption that the different classes have the same covariance. A typical formula for the quadrature discriminant function can be described as:

$$g_i = x^t W x + w^t x + w_{io} \quad (23)$$

where W, w^t, and w_{i0} are derived from the Gaussian distributions parameters. As can be seen from the first term of Equation (23), the function is quadratic in terms of **x**. Therefore, the resulting decision boundaries are quadratic, such as ellipses and parabolloids.

4.5.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) classifier, developed in 1995, aims to construct the best hyperplane that separates the data corresponding to two classes [41]. The best hyperplane constructed by SVM is the one with the biggest margin among the different classes. The margin represents the maximal width of the slab parallel to the hyperplane that does not contain any internal data points. The data points located on the boundary of the slab are referred to as support vectors. Figure 37 demonstrates the previous descriptions where "-" and "+" represents data points from two different classes [42].

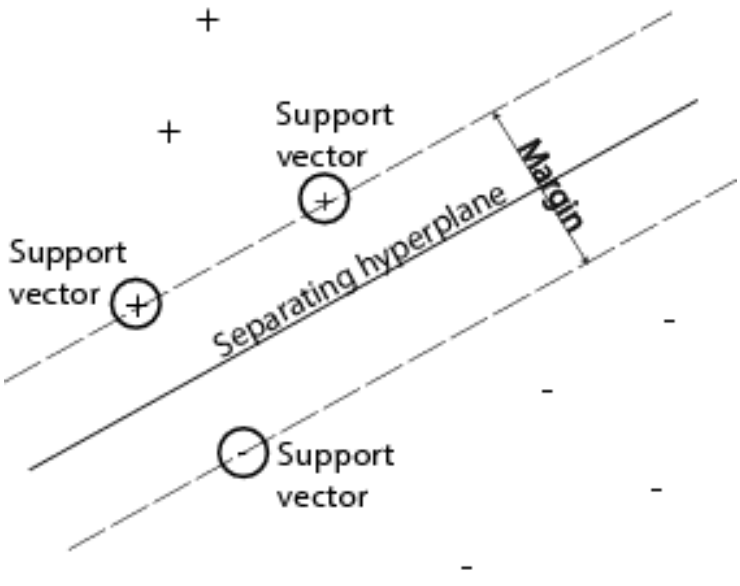


Figure 37: SVM example [42]

The concept of SVM is mainly intended for two-class problems; however, it can be expanded to multi-class problems by reducing it to several binary problems. A complete formulation for the algorithm and code can be retrieved from [42].

Chapter 5: Results

In this chapter, the pattern recognition techniques discussed in Chapter 4 are applied on the data collected from the experiments described in Chapter 3. As described earlier, each AE signal sample consists of 2500 raw points, which can be considered as a raw feature vector. In this work, different feature extraction techniques are applied directly on the AE signals like DFT, wavelet decomposition, and PCA. In addition, the different feature extraction techniques are combined with polynomial, QDA, SVM, and KNN classifiers. PCA extraction followed by KNN showed high and stable recognition rates in all studied cases. A similar effective combination of PCA extraction followed by a KNN classifier has been reported in [21] and [43]. The following results are based on the described methods.

5.1 PD Type Classification

The main contribution of this thesis work is to study the effects of some main practical AE measurement conditions on recognizing common PD types. In addition, the proposed pattern recognition system should be able to achieve high recognition rate, above 90%, as in the previously reviewed literature in [15] and [21].

In this work, the basic classification problem includes all the four PD classes measured at the same conditions as depicted in Table 9.

Table 9: Measurement conditions of the basic classification problem

Condition	Comment
PD location	4
Sensor Location	**a**
Oil temperature	23 ^{0}C
Training samples per class	70
Testing samples per class	40

The data are processed using different combinations of feature extraction techniques and classifiers, which were discussed in Chapter 4. The recognition rates of the different combinations are summarized in percentages in Table 10.

Table 10: Recognition rates of classifying the four PD types measured at the same conditions

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	57	87	81
QDA	62	88	91
SVM	61	79	87
KNN with Euclidean distance	86	85	89
KNN with cityblock distance	86	84	90
KNN with cosine distance	62	69	91
KNN with correlation distance	25	63	94

As can be seen from Table 10, PCA resulted in higher recognition rates using the different classifiers. This implies that applying PCA resulted in a separable dimension suitable for classification. In this work, the dimension or the number of principle components used was tuned to achieve the highest recognition rate. Figure 38 depicts the tuning process where the best number of principle components to classify all the four PD types is 9.

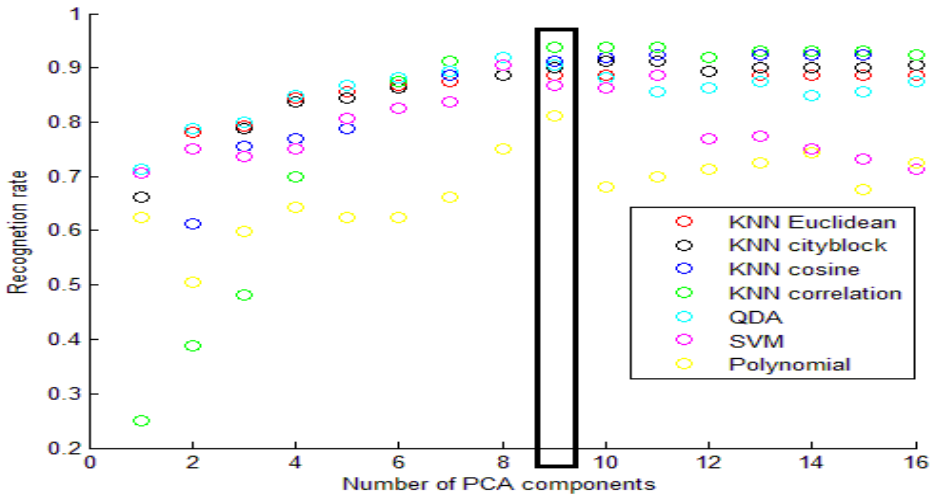


Figure 38: Tuning number of PCA components for classifying all the four PD types

PCA followed by KNN with distance correlation achieved the highest recognition rate in Table 10 when using ***K***=1. The value of ***K***=1 with the different distance measures showed the highest results when PCA was followed by KNN as shown in Figure 39.

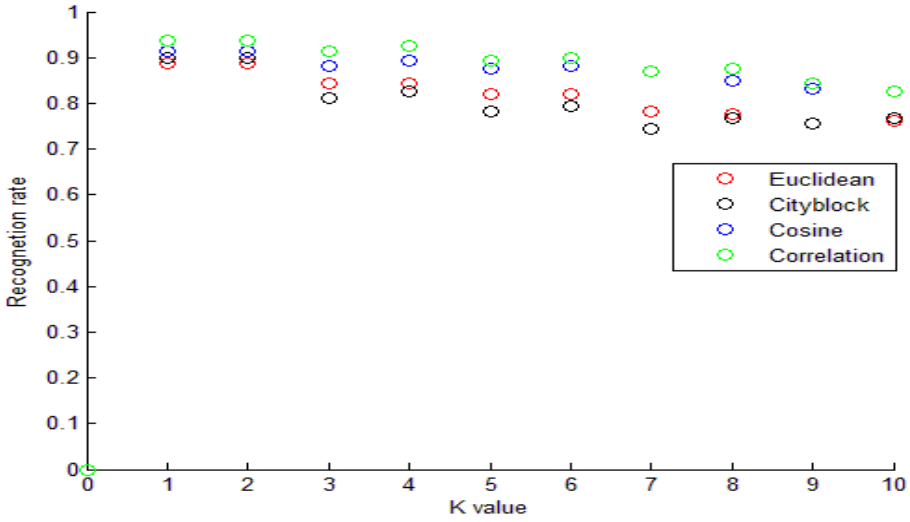
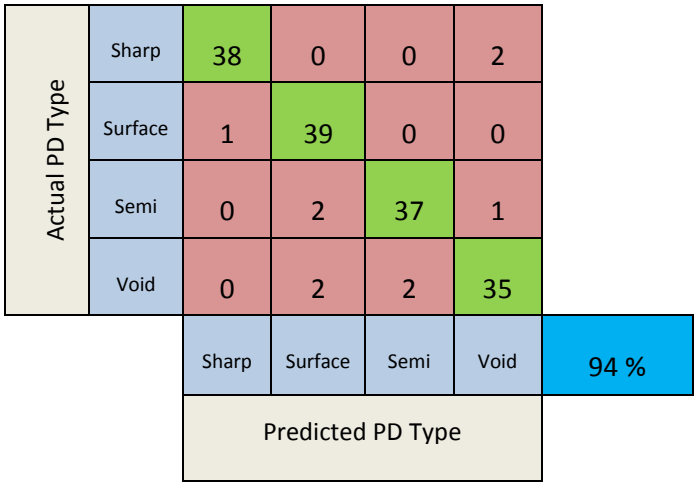


Figure 39: Tuning K value of KNN for classifying all the four PD types using PCA features

In addition, DFT followed by Euclidean KNN and wavelet decomposition followed by a QDA classifier showed good results, which indicates that the different PD types have different AE frequency patterns. The above recognition rates are obtained using confusion matrices. A confusion matrix or contingency table is a method for visualizing the performance of classification techniques to predict the correct classes. The diagonal elements represent correct predictions while off-diagonal elements represent misclassification or confusion. The total recognition rate of the confusion matrix is the percentage of the correct predictions with respect to the overall number of testing samples. Table 11 demonstrates the confusion matrix when testing and training with the four PD types measured at the same conditions using PCA followed by KNN with distance correlation.

Table 11: Confusion matrix (testing and training with the four PD types measured at the same conditions using PCA followed by KNN with distance correlation)



5.2 Practical Measurement Conditions

The next subsections examine the impact of some of the practical AE measurement conditions on the recognition rate. The simulated practical conditions as mentioned earlier are:

a) Changing PD location
b) Changing sensor location
c) Changing oil temperature
d) Having a barrier in the line-of-sight between the PD source and the AE sensor.

To study the influence of PD location, the recognition system is trained with data at one location and then tested with data at different location. Likewise, the effect of oil temperature on the recognition rate is examined by training the recognition system at a certain oil temperature and then testing it with data at a higher temperature. Similarly, the impact of the barrier is inspected by training with data taken when no barrier is inserted and testing with data taken when the barrier is inserted. In addition, the effect of the sensor location on the recognition rate is investigated.

5.2.1 PD Type Recognition at Different PD Locations

In subsection 5.1, the classification problem used training and testing data from the same PD location. In this subsection, the effect of training the system from one location and testing it with another is addressed. Therefore, the recognition systems were trained with data at PD location 4 and tested at PD location 9 keeping all other conditions the same; only three types of PD were investigated in this study: sharp, surface and void PD's. The recognition rates of the different classification method combinations for this condition are summarized in Table 12.

Table 12: Recognition rates when training with data from PD location 4 and testing from PD location 9

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	18	52	32
QDA	21	50	53
SVM	19	38	32
KNN with Euclidean distance	13	65	39
KNN with cityblock distance	11	64	35
KNN with cosine distance	26	64	35
KNN with correlation distance	38	60	43

As can be seen from Table 12, the recognition rate dropped significantly when the system was tested with data from another location. AE signals form PD location 9 had to undergo a long propagation path that caused them to suffer either additive or subtractive interferences. For example, such interferences are obvious when comparing a void discharge at PD location 9 as depicted in Figure 40 with another sample taken from PD location 4 as shown in Figure 41.

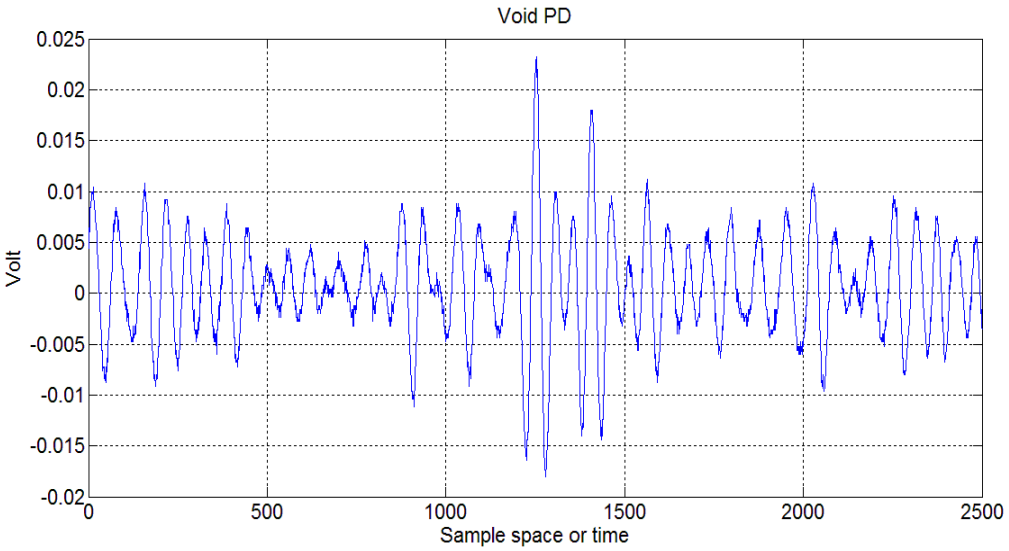


Figure 40: Void discharge AE sample from PD location 9

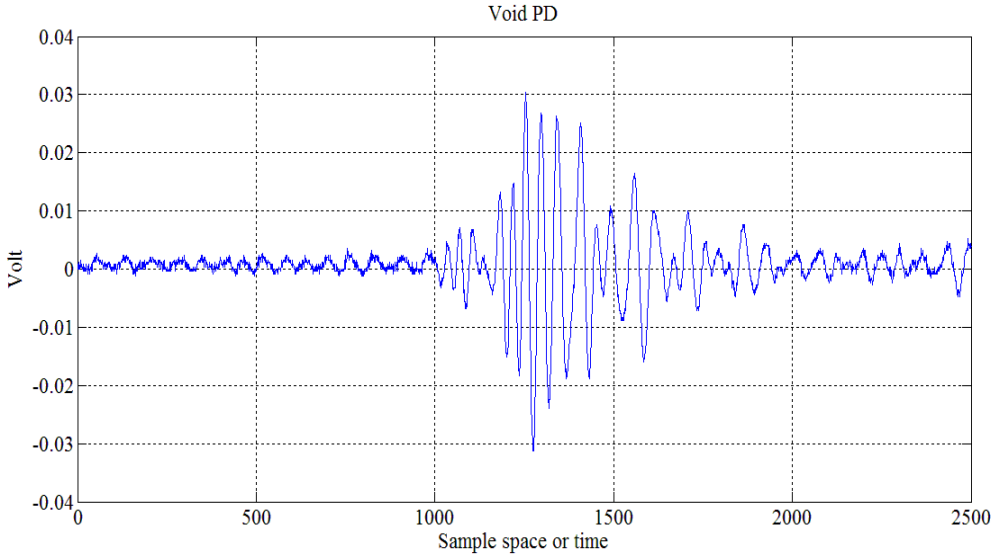


Figure 41: Void discharge AE sample from PD location 4

To make sure that data of PD location 9 was not the main reason for misclassification, the classifiers are trained and tested only from that location. Most combinations classify the three PD types effectively as shown in Table 13.

Table 13: Recognition rates when training and testing with data from PD location 9

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	69	90	89
QDA	63	91	89
SVM	74	78	92
KNN with Euclidean distance	84	89	90
KNN with cityblock distance	85	90	91
KNN with cosine distance	74	90	92
KNN with correlation distance	33	91	89

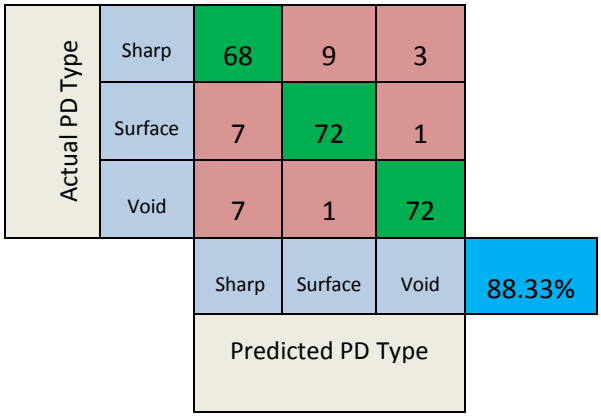
Since different combinations can effectively classify data trained and tested from the same location, a comprehensive system can be made if it is trained equally from both PD locations 4 and 9. Table 14 indicates the improvement on the recognition rates when training and testing equally from both locations.

Table 14: Recognition rates when training and testing with data equally from PD location 4 and 9

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	57	80	77
QDA	48	84	83
SVM	55	76	83
KNN with Euclidean distance	76	75	90
KNN with cityblock distance	76	76	89
KNN with cosine distance	71	69	88
KNN with correlation distance	42	73	88

PCA results followed by KNN show the highest results in Table 14. Table 15 illustrates the performance of PCA followed by KNN with distance correlation.

Table 15: Confusion matrix (training and testing with data equally from PD location 4 and 9 using PCA followed by KNN with distance correlation)



Based on the presented results, it can be concluded that PD location has an impact on recognition rate. However, including training data from different PD locations can reduce the PD location effect.

5.2.2 PD Type Recognition at Different sensor locations

In this subsection, the effect of changing the sensor location is examined. All PD types were measured under the same conditions as in Table 9 except for changing the sensor location. The sensor locations as depicted in Figure 15 are varied at three different locations: **a**, **b**, **c**, which are 15, 10, and 5 cm respectively above the tank base. Similar to the results presented in Table 10, the recognition rates with the sensor placed at location **b** shows high recognition rates as shown in Table 16. In addition, most of the recognition rates with the sensor placed at location **c** presented in Table 17 are high, but a bit lower than the results at location **b**. This could be justified since location **c** is close to the edge near the tank base, which caused some PD signals from different types to suffer small interference caused by reflections. This can be illustrated by comparing a sample taken from a sensor fixed at location **b** and another sample suffering from interference taken from sensor location **c** for semi discharge as shown in Figure 42 and Figure 43 respectively.

Table 16: Recognition rates when AE sensor is placed at location **b**

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	54	93	90
QDA	53	92	84
SVM	73	90	94
KNN with Euclidean distance	79	90	91
KNN with cityblock distance	79	88	92
KNN with cosine distance	71	84	93
KNN with correlation distance	21	71	93

Table 17: Recognition rates when AE sensor is placed at location **c**

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	60	78	84
QDA	59	74	88
SVM	63	61	71
KNN with Euclidean distance	76	61	88
KNN with cityblock distance	76	65	86
KNN with cosine distance	70	63	91
KNN with correlation distance	22	66	91

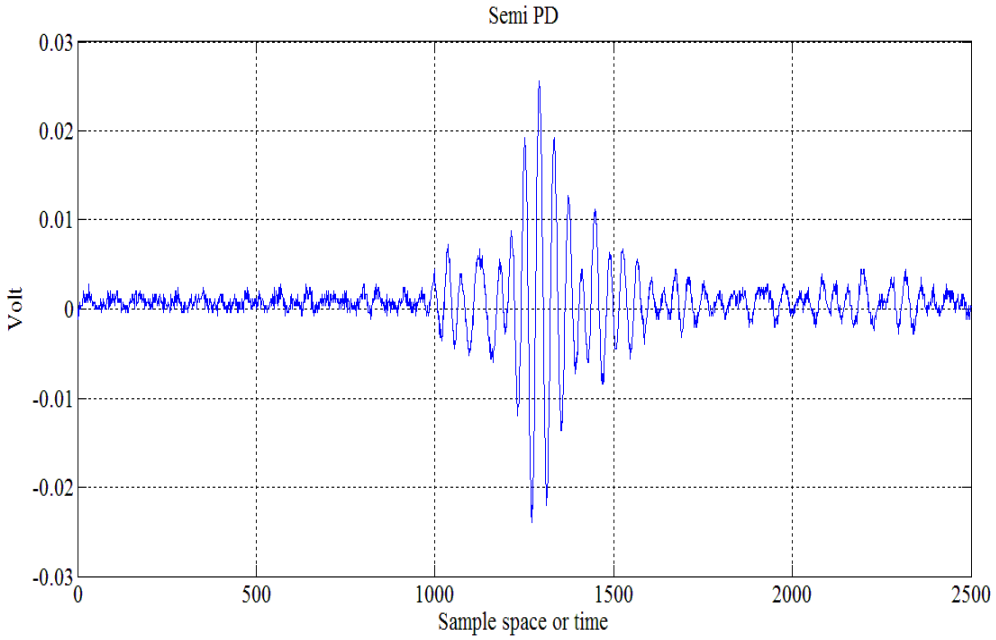


Figure 42: AE sample from semi-discharge at sensor location **b**

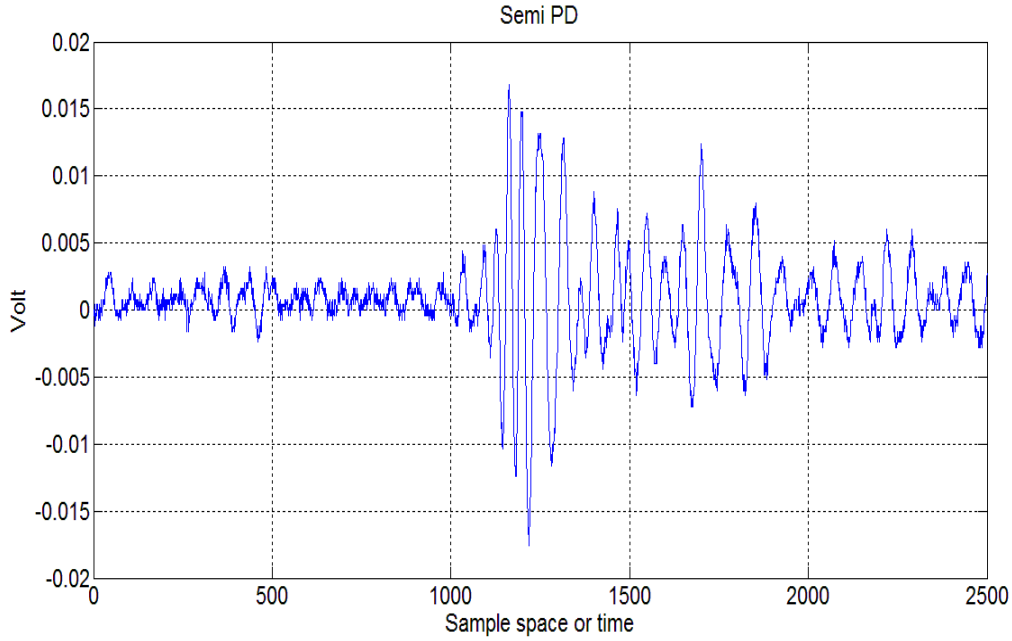


Figure 43: AE sample from semi-discharge at sensor location **c**

By comparing the recognition rate at the different sensor locations, it can be concluded that the recognition system is less affected by the sensor depth location. However, it would be recommended to keep the sensor away from any edges to avoid signal reflections and interferences.

5.2.3 PD Type Recognition at Different Oil Temperatures

In this subsection, the effect of heating the oil insulation on PD's from surface and void discharges is examined. AE waves in oil, as discussed earlier, have different propagation speeds at different oil temperatures, which have a direct relationship with the acoustic impedance (Z) of the oil. The measurement conditions for this investigation are shown in Table 18 . The first step in the analysis of the temperature effect is to test the recognition system when the system is tested and trained at the same oil temperature. The following recognition rates are presented when:

1- Training and testing with data taken at 23 ^{0}C as illustrated in Table 19.
2- Training and testing with data taken at 70 ^{0}C for void and surface discharges as shown in Table 20.

Table 18: Measurement conditions for the oil temperature experiments

Condition	Comment
PD location	3
Sensor Location	**b**
Oil temperature	23, 50, and 70 ^{0}C
Training samples per class	135
Testing samples per class	80

Table 19: Recognition rates when testing and training at 23 ^{0}C

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	88	93	79
QDA	88	94	88
SVM	85	87	80
KNN with Euclidean distance	85	88	91
KNN with cityblock distance	85	88	90
KNN with cosine distance	74	92	90
KNN with correlation distance	50	88	91

Table 20: Recognition rates when testing and training at 70 ^{0}C

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	56	80	96
QDA	63	84	99
SVM	56	71	96
KNN with Euclidean distance	84	70	98
KNN with cityblock distance	85	64	98
KNN with cosine distance	84	67	98
KNN with correlation distance	50	68	98

PCA combinations with the different classifiers effectively recognized the PD types when it was trained and tested at the same temperature. However, different data scatters can be noticed when comparing the first main three principle components of PCA for the data at 23 ^{0}C depicted in Figure 44 and the components for the data at 70 ^{0}C presented in Figure 45.

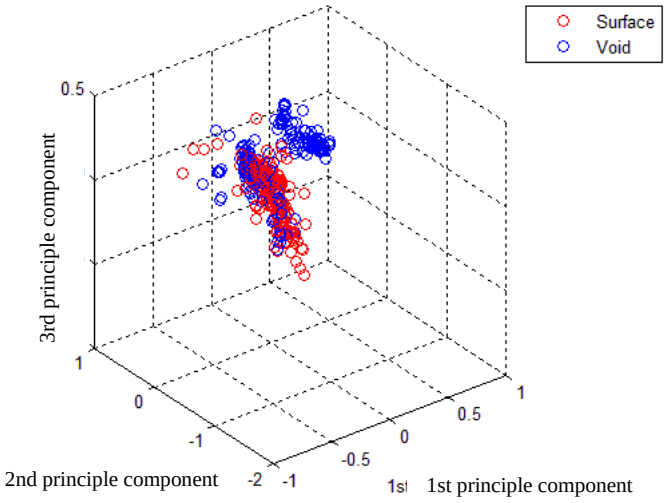


Figure 44: First three principle components scatter of data measured at 23 ^{0}C

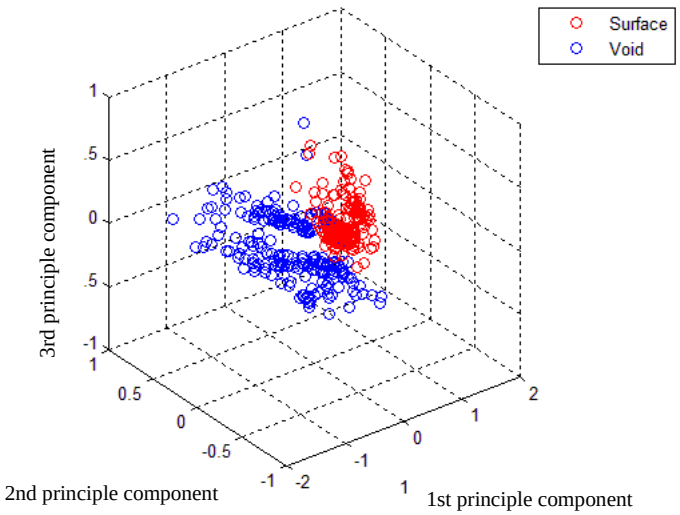


Figure 45: First three principle components scatter of data measured at 70 ^{0}C

Therefore, it would be expected to have a drop in the recognition rate if the system is trained from one temperature and tested at the other. Table 21 confirms this expectation when training at 23 ^{0}C and testing at 70 ^{0}C.

Table 21: Recognition rates when training at 23 ^{0}C and at testing at 70 ^{0}C

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	44	58	73
QDA	44	58	75
SVM	42	41	76
KNN with Euclidean distance	49	50	55
KNN with cityblock distance	49	50	56
KNN with cosine distance	51	46	58
KNN with correlation distance	50	48	58

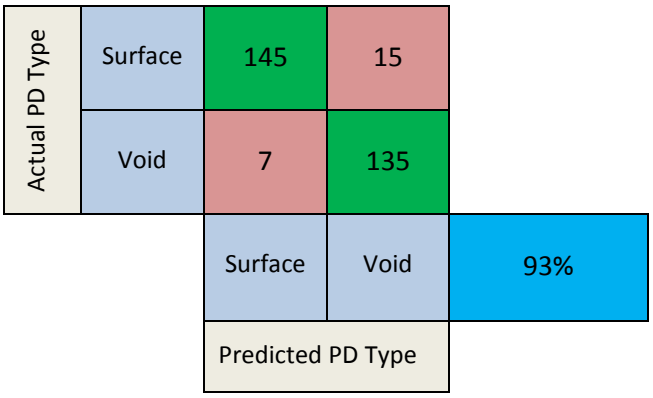
As previously discussed, minimizing the recognition error can be achieved by training with data at the different temperatures. Table 22 presents enhancement on the recognition rates when training and testing equally at the different temperatures.

Table 22: Recognition rates when testing and training equally at both 23 ^{0}C and 70 ^{0}C

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	63	70	94
QDA	76	78	93
SVM	51	74	82
KNN with Euclidean distance	79	70	93
KNN with cityblock distance	79	69	92
KNN with cosine distance	75	68	93
KNN with correlation distance	50	72	93

The confusion matrix of PCA followed by KNN with distance correlation in Table 22 is shown in Table 23.

Table 23: Confusion matrix (training and testing with equally from data at 23 ^{0}C and 70 ^{0}C using PCA followed by KNN with distance correlation)



A further analysis is to investigate whether it is necessary to include data samples measured at more temperatures in the training process. Therefore, a set of 150 testing samples from surface discharge at 50 ^{0}C are classified when training the classifiers at both 23 ^{0}C and 70 ^{0}C. Table 24 shows low recognition rates for such conditions. Hence, it would be recommended to include training data from different temperatures to have a robust AE recognition system.

Table 24: Recognition rates when training equally at 23 ^{0}C and 70 ^{0}C and testing at 50 ^{0}C

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	17	66	64
QDA	13	57	69
SVM	34	57	59
KNN with Euclidean distance	46	65	67
KNN with cityblock distance	46	64	69
KNN with cosine distance	49	60	61
KNN with correlation distance	50	56	60

5.2.4 PD Type Recognition with Barrier Inclusion

The core of the transformer can act as a barrier between the AE waves emitted from a PD source and the AE sensor. In this this subsection, the effect of inserting a barrier in the line of sight between the PD source and the AE sensor is examined. Only surface and void discharges have data "with" and "without" a barrier. All PD data are measured at conditions as in Table 25.

Table 25: Measurement conditions for the barrier experiments

Condition	Comment
PD location	9
Sensor Location	**a**
Oil temperature	23 ^{0}C
Training samples per class	170
Testing samples per class	100

Table 26 presents the recognition rates when the barrier is inserted in the line-of-sight of the PD source. The results in Table 26 indicate high recognition rate, yet they are a bit lower when there is no barrier inserted as shown in Table 27. This could be justified by the fact that adding the barrier increases the acoustic field complexity inside the tank with extra wave reflections and diffractions, which caused the signals to lose some of their characteristics. The barrier effect might be better visualized by

comparing the scatter of the first three principle components of the data measured "with" and "without" a barrier presented in Figure 46 and Figure 47 respectively.

Table 26: Recognition rates with barrier inserted in the line-of-site of the PD source

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	83	94	87
QDA	84	93	87
SVM	88	94	87
KNN with Euclidean distance	84	88	89
KNN with cityblock distance	82	87	89
KNN with cosine distance	59	88	90
KNN with correlation distance	51	89	90

Table 27: Recognition rates without a barrier inserted in the line-of-site of the PD source

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	74	96	86
QDA	72	96	97
SVM	75	93	95
KNN with Euclidean distance	85	95	94
KNN with cityblock distance	85	95	92
KNN with cosine distance	75	95	95
KNN with correlation distance	50	94	94

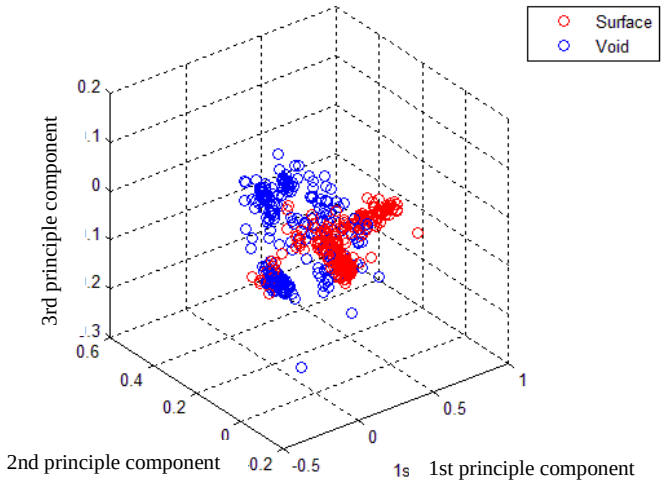


Figure 46: First three principle components scatter of data measured without a barrier

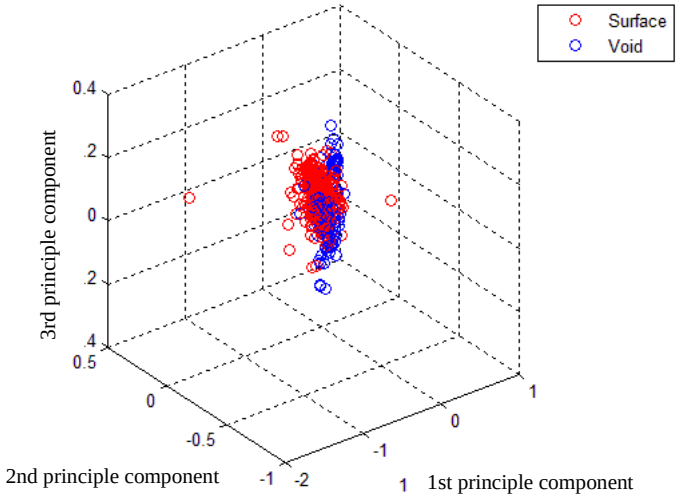


Figure 47: First three principle components scatter of data measured with a barrier

The data scatter corresponding to the AE sample measured “without” a barrier shows more distinguished distribution compared to the scatter of the data corresponding to the AE sample measured “with” a barrier.

A further analysis is carried out by training with data “without” a barrier and testing with data “with” a barrier. Table 28 shows significant drop in the recognition rate for this case.

Table 28: Recognition rates when training without a barrier and testing with barrier inserted in the line-of-site of the PD source

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	63	62	48
QDA	65	59	63
SVM	60	58	56
KNN with Euclidean distance	68	63	51
KNN with cityblock distance	67	62	50
KNN with cosine distance	59	64	54
KNN with correlation distance	50	76	54

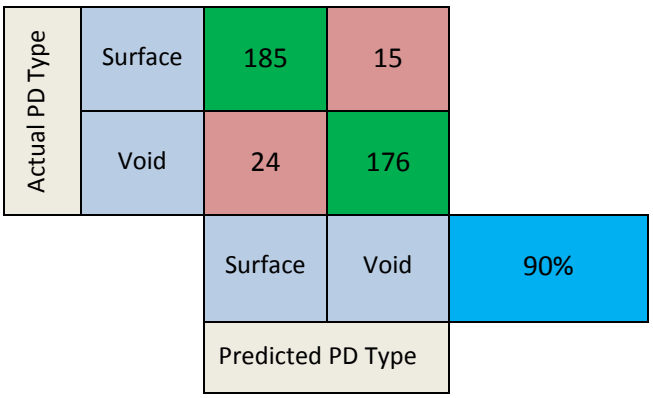
As previously discussed, minimizing the recognition error can be achieved by training from data taken "with" and "without" a barrier as shown in Table 29.

Table 29: Recognition rates when training and testing equally from data with and without a barrier inserted in the line-of-site of the PD source

Classifier/Feature	DFT	Wavelet	PCA
Polynomial (quadratic)	76	94	88
QDA	78	92	87
SVM	77	93	87
KNN with Euclidean distance	84	90	89
KNN with cityblock distance	83	90	89
KNN with cosine distance	68	90	90
KNN with correlation distance	50	92	90

The confusion matrix of PCA followed by KNN with distance correlation in Table 29 is depicted in Table 30.

Table 30: Confusion matrix (training and testing with equally from data with and without a barrier inserted in the line-of-site of the PD source using PCA followed by KNN with distance correlation)



Based on the presented analysis, it would be recommended to train the system with data “with and without a barrier” to have a complete system. A multiple sensor system could help in minimizing the barrier effect because a PD source could be seen “with” or “without” a barrier with reference to the different locations of the installed sensors.

Chapter 6: Conclusions and Recommendations

This thesis aimed not only to achieve high recognition rate for different simulated PD types using AE signals, but also to consider more AE practical measurement conditions. The measurement conditions include having aged insulation material (oil/paper), a tank size of 1×1×0.5 m dimensions, and a high surrounding noise level. In addition, other practical condition effects on the recognition rate are examined like changing of PD location, sensor location, oil temperature, and having a barrier in the line-of-sight between the PD source and the AE sensor. Four common types of PDs are considered for the classification problem; surface discharge, PD from a sharp point to ground plane, PD from semi parallel plates, and PD from an air void in the insulation.

The results indicate the effectiveness of using PCA as a feature extractor in general with KNN as a classifier. In addition, the utilization of high frequency AE sensors, 100-450 kHz, proved to provide good detection for the different PD sources. Furthermore, the results show the effect of each simulated practical condition on the recognition rate as follows:

1- Effect of PD location: PD location has an impact on the recognition rates since PD at different locations includes different amounts of interferences. However, by including training data from different PD locations, the recognition error is minimized.

2- Effect of AE sensor location: the results indicate that the recognition rates are less affected by the sensor depth location. However, it is recommended to keep the sensor away from any edges to avoid signal reflections and interferences.

3- Effect of oil temperature: a great impact can be noticed on the recognition rates when training the data at one temperature and testing it from another temperature. PD’s at higher temperatures tend to initiate at lower applied voltage levels. In addition, AE waves in oil have different propagation

speeds at different oil temperatures, which have a direct relationship with the acoustic impedance (Z) of the oil. Therefore, data taken at different temperatures must be considered in the training phase to achieve a high recognition rate.

4- Effect of barrier insertion: a higher applied voltage was required to be able to detect AE signals. In addition, the barrier causes different data distribution compared to data collected when no barrier is inserted, which causes misclassification when training with data at one condition and testing with data at the other condition. However, by training with data taken “with” and “without” a barrier, high recognition rates are achieved.

This work can be improved in the future by improving the capability of the high voltage laboratory with new equipment and higher voltage levels. Furthermore, more PD types and practical conditions can be added like having oil circulation. In addition, the oil tank can be upgraded by inserting real oil-paper insulated windings. Moreover, multiple AE sensors can be interfaced at different locations around the tank, which can minimize the PD location and barrier challenge. Finally, this research was based on receiving AE waves without knowing the shape of the original waves at the PD source. A future study can be conducted to simulate the propagation of an AE from the PD source until it reaches the sensor mounted on the tank’s wall under the different measurement conditions like PD location, oil temperature, and barrier insertion. With such simulations, the behavior of AE waves in a transformer environment can be investigated further, which could help in choosing more discriminant features.

Bibliography

[1] Matthew Sedano, H. Brown, and P. Richard, "Electricity transmission," National Conference of State Legislatures, A Primer ISBN 1-58024-352-5, 2004.

[2] Office of Electricity Delivery and Energy Reliability, "Large power transformers and the U.S. electric grid," U.S. Department of Energy, June 2012.

[3] Bart Tichelman, "Transformer asset management - a new paradigm," in *Energy Pulse*, 2004, April 16.

[4] William H. Bartley P.E., "Analysis of transformer failures," in *International Association of Engineering Insurers*, Stockholm, 2003.

[5] Cheng-Chien Kuo and Horng-Lin Shieh, "Artifical classification system of aging period based on insulation status of transformer," in *8th International Conference on Machine Learning and Cybernetics*, Baoding, July 2009, pp. 3310-3315.

[6] J. Rubio-Serrano et al., "Electro-acoustic detection, identification and location of partial discharge sources in oil-paper insulation systems," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no. 5, October 2012.

[7] Wojciech Sikorski and Waldemar Ziomek, "Detection, recognition and location of partial discharge sources using acoustic emission method," in *Acoustic Emission*, 2012, ch. 3, pp. 49-74.

[8] T. Boczar, "Identification of a specific type of PD from acoustic emission frequency spectra," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 8, no. 4, pp. 598-606, 2001.

[9] W. Sikorski and K. Siodla, "Identification of partial discharge sources in high voltage insulating systems using acoustic emission method," in *14th International Symposium on High Voltage Engineering*, Beijing, China, 2005.

[10] T. Boczar, S. Borucki, A. Cichon, and D. Zmarzly, "Application possibilities of artificial neural networks for recognizing partial discharges measured by the acoustic emission method," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 16, no. 1, pp. 214-223, February 2009.

[11] S. R. Hardie and P. S. Bodger, "A new technique for detecting partial discharges within an on-line power transformer subjected to interference," in *Electrical Engineers' Association (EEA) of NZ Conference*, Auckland, New Zealand, 2006.

[12] A. Swedan, A. H. El-Hag, and K. Assaleh, "Acoustic detection of partial discharge using signal processing and pattern recognition techniques," *Insight–Non-Destructive Testing And Condition Monitoring*, vol. 54, no. 12, December 2012.

[13] M.L Chai, Y.H.M. Thayoob, P.S. Ghosh, A.Z. Sha'ameri, and M.A. Talib, "Identification of different types of partial discharge sources from acoustic emission signals in the time-frequency representation," in *First International Power and Energy Coference PECon*, Putrajaya, Malaysia, 2006, pp. 581-586.

[14] W. Chen, X. Chen, S. Peng, and J Li, "Canonical correlation between partial discharges and gas formation in transformer oil paper insulation," *Energies*, vol. 5, no. 4, pp. 1081-1097, April 2012.

[15] T. Jiang, J. Li, Y. Zheng, and C. Sun, "Improved bagging algorithm for pattern recognition in UHF signals of partial discharges," *Energies*, vol. 4, no. 7, pp. 1087-1101, July 2011.

[16] Yangchun Cheng, E. Gockenbach, C. Eichler, and Chengrong Li, "The partial discharge phenomena on the surface of oil impregnated paper with parallel electric field," in *Electrical Insulation and Dielectric Phenomena*, West Lafayette, IN, Oct. 2010, pp. 1-4.

[17] ea technology. Case study partial discharge of cb spouts. [Online]. Available: http://www.eatechnology.com/analytical/power-forensics/casestudypartialdischargeofcbspouts

[18] "Solid dielectric breakdown," American University of Sharjah, Sharjah, Lecture notes 2012.

[19] F. Massingue, S. Meijer, Pantelis D. Agoris, J.J. Smit, and Jose Lopez-Roldan, "Partial discharge pattern analysis of modeled insulation defects in transformer insulation," in *IEEE International Symposium on Electrical Insulation*, 2006.

[20] Baharudin, George Chen, and Fauzan, "Partial discharge modelling based on a cylindrical model in solid dielectrics," in *International Conference on Condition Monitoring and Diagnosis*, Beijing, 2008.

[21] Hui Ma, J.C. Chan, T.K. Saha, and C. Ekanayake, "Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 20, no. 2, April 2013.

[22] M. Hikita, S. Okabe, H. Murase, and H. Okubo, "Cross-equipment evaluation of partial discharge measurement and diagnosis techniques in electric power apparatus for transmission and distribution," *IEEE Transactions on Dielectrics*

and Electrical Insulation, vol. 15, no. 2, pp. 505-518 , April 2008.

[23] Lynn Hamrick. (2010) Dissolved gas analysis for transformers. [Online]. Available: http://www.netaworld.org/sites/default/files/public/neta-journals/NWwtr09_Hamrick.pdf

[24] SIEMENS. (2013, October) SIEMENS Energy. [Online]. Available: http://www.energy.siemens.com/hq/en/services/power-transmission-distribution/transformer-test-laboratory/dissolved-gas-analysis.htm#content=Details

[25] S. Coenen, S. Tenbohlen, and S. M. Markalous, "Detection of partial discharges in power transformers using UHF PD measurements and acoustic measurements," , Germany.

[26] "Acoustic emission sensors," Vallen Systeme GmbH, Schaeftlarner, Specification 2012.

[27] E. Grossmann and K. Feser, "Improvement of sensitivity in online PD-measurements on transformers using acoustic emission techniques," in *12th International Symposium of High Voltage*, Bangalore, 2001.

[28] Copper Development Association Inc. (2013) [Online]. Available: http://www.copper.org/environment/sustainable-energy/transformers/education/trans_efficiency.html

[29] Erik Gregersen, Ed., *The Britannica guide to sound and light.*: Britannica Educational Publishing, 2011, ch. 1.

[30] Kaye & Laby Online. (2008) Tables of physical & chemical constants. [Online]. Available: http://www.kayelaby.npl.co.uk/general_physics/2_6/2_6_5.html

[31] Inc. Industrial Noise Control. (2010) Comparative examples of noise levels. [Online]. Available: http://www.industrialnoisecontrol.com/comparative-noise-examples.htm

[32] Alaa El Khatib, "Learning-based space-time adaptive processing," American University of Sharjah, Sharjah, Master Thesis 2013.

[33] P.S. Penev and L. Sirovich, "The global dimensionality of face space," in *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, 2000, pp. 264-270.

[34] "Dimensionality reduction," American University of Sharjah, Sharjah UAE, Lecture notes for pattern classification course 2012.

[35] Lindsay I Smith, "A tutorial on principal components analysis," Tutorial 2002.

[36] P.N. Belhumeur, J.P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 19, no. 7, pp. 711-719, July 1997.

[37] Edwin Chen. (2012, December) http://blog.echen.me. [Online]. Available: http://blog.echen.me/2011/03/14/pca-transpose-trick/

[38] Xindong Wu et al., "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1-37, January 2008.

[39] Gábor J. Székely, Maria L. Rizzo, and Nail K. Bakirov, "Measuring and testing dependence by correlation of distances," *Annals of Statistics*, vol. 35, no. 6, pp. 2769-2794, 2007.

[40] S. Santosh, R. Gupta, and A. B´ela, "Bayesian Quadratic Discriminant Analysis," *Journal of Machine Learning Research*, vol. 8, pp. 1277-1305, June 2007.

[41] J.A. Hunter et al., "Comparison of two partial discharge classification methods ," in *IEEE International Symposium on Electrical Insulation (ISEI)* , San Diego, CA, 2010, pp. 1-5.

[42] mathworks. (2013, December) mathworks.com. [Online]. Available: http://www.mathworks.com/help/stats/support-vector-machines-svm.html

[43] C.M.Bishop, *Neural networks for pattern recognition*. Oxford University press, 1995.

Vita

Mustafa Harbaji received his B.S. (cum laude) in Electrical Engineering from the American University of Sharjah (AUS) in fall 2011. During, his undergraduate studies he was a co-author for two conference papers in communication and energy management respectively and one journal paper under the title of "Renewable and Storage Energy Integration for Smart Grid Housing". In spring 2012, he was granted for a full graduate teaching assistantship to complete his M.S. in Electrical Engineering at AUS. During his M.S. period, he published and presented a conference paper based on the thesis preliminary results. He is now in the process of writing a transaction paper based on the thesis final results. In addition, he is working now in a joint research with Qatar University funded by National Priorities Research Program (NPRP), Qatar National Research Fund (QNRF) under the title of "Asset Management for Power Transformers in Smart Grid". He has been admitted to complete PhD studies in the area of the smart grid and distributed generation at Waterloo University, Canada, for the fall 2014 semester.