# PARTIAL DISCHARGE CLASSIFICATION ON XLPE CABLE JOINTS UNDER DIFFERENT NOISE LEVELS USING ARTIFICIAL INTELLIGENCE TECHNIQUES

**WONG JEE KEEN RAYMOND**



**FACULTY OF ENGINEERING**
**UNIVERSITY OF MALAYA**
**KUALA LUMPUR**

**2016**

# PARTIAL DISCHARGE CLASSIFICATION ON XLPE CABLE JOINTS UNDER DIFFERENT NOISE LEVELS USING ARTIFICIAL INTELLIGENCE TECHNIQUES

**WONG JEE KEEN RAYMOND**

**THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**FACULTY OF ENGINEERING**
**UNIVERSITY OF MALAYA**
**KUALA LUMPUR**

**2016**

**UNIVERSITY OF MALAYA**

**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: **Wong Jee Keen Raymond**

Registration/Matric No: **KHA130001**

Name of Degree: **Doctor of Philosophy**

Title of Thesis: **PARTIAL DISCHARGE CLASSIFICATION ON XLPE CABLE JOINTS UNDER DIFFERENT NOISE LEVELS USING ARTIFICIAL INTELLIGENCE TECHNIQUES**

Field of Study: **High Voltage**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

## ABSTRACT

Cross linked polyethylene (XLPE) cables are widely used in power industries due to their good electrical and mechanical properties. Cable joints are the weakest point in the XLPE cables and most susceptible to insulation failures. Any cable joint insulation breakdown may cause a huge loss to power companies. Therefore, it is vital to diagnose the insulation quality to detect early signs of insulation failure. Partial discharge (PD) measurement is a vital tool for assessing the insulation quality at cable joints. Since the past, there have been many pattern recognition methods to classify PD, where each method has its own strengths and weaknesses. Although many works have been done on PD pattern recognition, it is usually performed in a noise-free environment. Also, works on PD pattern recognition are mostly done on lab fabricated insulators, where works using actual cable joints are less likely to be found in literature. Therefore, in this work, classification of real cable joint defect types using partial discharge measurement under noisy environment was performed. Five cross-linked polyethylene (XLPE) cable joints with artificially created defects were prepared based on the defects commonly encountered on site. A novel high noise tolerance principal component analysis (PCA)-based feature extraction was proposed and compared against conventional input features such as statistical features and fractal features. These input features were used to train the classifiers to classify each PD defect type. Classifications were performed using three different artificial intelligence classifiers, which include Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Support Vector Machine (SVM). The performance of each classifier and feature extraction method was evaluated. It was found that SVM and ANN performed well while ANFIS classification accuracy was the weakest. As for input features, the proposed PCA features displayed highest noise tolerance with the least performance degradation compared to other input features.

## ABSTRAK

Kabel XLPE digunakan secara meluas dalam industri tenaga kerana ia mempunyai ciri-ciri mekanikal dan elektrikal yang bagus. Cantuman kabel merupakan tempat yang sering mengalami kegagalan kerana ia bahagian yang paling lemah dalam kebel. Sebarang kerosakan cantuman kebel akan mengakibatkan kerugian besar. Oleh itu, ia amat penting untuk mengetahui kuality bahan penebat dalam cantuman kabel. Penyukatan pelepasa separa (PD) adalah process yang penting untuk mengetahui kuality penebat dalam cantuman kebel. Sejak dahulu, banyak kerja pengenalan corak untuk mengklasifikasi PD telah dilakukan dan setiap teknik mempunyai kekuatan dan kelemahan sendiri. Walaupun banyak kerja telah dilakukan dalan pengenalan corak PD, ia biasanya dilakukan dengan penebat buatan dalam keadaan sunyi. Lagipun kerja tentang pengenalan corak PD pada cantuman kabel jarang dijumpai. Oleh itu, dalam kerja ini, klasifikasi PD dalam cantuman kabel yang mempunyai kerosakan tiruan dilakukan dalam keadaan bunyi bising. Lima cantuman kabel jenis XLPE yang mempunyai kerosakan tiruan telah disediakan berdasarkan kerosakan yang biasa dialami dalam situasi praktical. Penyukatan PD dilakukan pada setiap cantuman kabel dan tiga tanda pengenalan yang dipetik daripada corak PD dalan keadaan bunyi bising tiruan. Satu teknik pengekstrakan ciri yang bertoleransi bunyi tinggi yang novel berdasarkan ciri analisis komponent utama (PCA) telah dicadangkan dan dibanding dengan ciri pengekstrakan tradisi seperti ciri statistical dan ciri fraktal. Tanda pengenalan digunakan untuk melatihkan pengelas untuk mengelaskan setiap jenis kerosakan PD. Tiga pengelas yang digunakan ialah pengelasan rangkaian tiruan (ANN), adaptif neuro-fuzzy system kesimpulan (ANFIS) dan vektor sokongan mesin (SVM). Prestasi setiap pengelas dan pengekstrakan ciri telah dinilai. SVM dan ANN amat memuaskan tetapi ANFIS adalah yang paling lemah. Bagi ciri-ciri input, ciri-ciri PCA dicadangkan menunjukan toleransi bunyi yang lebih tinggi dengan kemerosotan prestasi yang paling sikit berbanding dengan ciri-ciri lain.

## ACKNOWLEDGEMENTS

In successful completion of this work, I would like to express my deepest gratitude to my supervisors, Dr. Hazlee Azil Illias and Dr. Ab Halim Abu Bakar for their excellent supervision and guidance throughout my time in this university. They have provided me with useful advice and shared their experience and wisdom without hesitation.

I would also like to thank my family and friends who have provided me with their unconditional love, support and encouragement.

I express my appreciation to the UM High Voltage Laboratory staffs for being helpful throughout my research work.

Finally, I wish to thank the University of Malaya and Malaysian Ministry of Higher Education (MOHE) for funding this research work through MOHE HIR (H-16001-D00048).

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF SYMBOLS AND ABBREVIATIONS

| | | |
|---|---|---|
| AC | : | Alternating current |
| ANFIS | | Adaptive neuro-fuzzy inference system |
| ANN | : | Artificial neural network |
| APNN | : | Adaptive PNN |
| BPNN | : | Back propagation neural network |
| CLARA | : | Clustering large applications |
| CMAC | : | Cerebellar model articulation controller |
| CNN | : | Cascaded neural network |
| CPN | : | Counter propagation neural network |
| CPNN | : | Composite PNN |
| CWT | : | Continuous wavelet transform |
| dB | : | Daubechies |
| DC | : | Direct current |
| DEM | : | Dynamically weighted ensemble |
| DWT | : | Discreet wavelet transform |
| ENN | : | Ensemble neural network |
| EXNN | : | Extension neural network |
| FCM | : | Fuzzy c-means |
| FCPN | : | Forward CPN |
| FOLS | : | Forward orthogonal least square |
| GILBS | : | Gas insulated load break switches |
| GIS | : | Gas insulated substation |
| GUI | : | Graphical user interface |
| HRPNN | : | Heteroscedastic PNN |

| | | |
|---|---|---|
| HV | : | High voltage |
| IEC | : | International electrotechnical commission |
| IEEE | : | Institute of Electrical and Electronics Engineers |
| kV | : | Kilo Volt |
| LDPE | : | Low density polyethylene |
| pC | : | Pico coulomb |
| PCA | : | Principal component analysis |
| PD | : | Partial discharge |
| PDEV | : | Partial discharge extinction voltage |
| PDF | : | Probability density function |
| PDIV | : | Partial discharge inception voltage |
| PE | : | Polyethylene |
| PNN | : | Probabilistic neural network |
| PRPD | : | Phase resolved partial discharge |
| PSO | : | Particle swarm optimization |
| RBNN | : | Radial basis neural network |
| RBPNN | : | Radial basis probabilistic neural network |
| RF | : | Radio frequency |
| RHRPNN | : | Robust heteroscedastic PNN |
| RST | : | Rough set theory |
| SOM | : | Self-organizing map |
| SVM | : | Support vector machines |
| USB | : | Universal serial bus |
| XLPE | : | Cross linked polyethylene |

# CHAPTER 1: INTRODUCTION

## 1.1 Introduction

Electrical insulation is a significant part in all high voltage power equipment. Failure in any part of the power system will be detrimental to power generation and transmission companies. Hence, it is extremely important to check the insulation quality frequently. Failure analysis reveals that insulation failure is the root cause for more than half of high voltage equipment damage (Tian, Lewin, & Davies, 2002). Cross-linked polyethylene (XLPE) possesses excellent mechanical, thermal and electrical attributes and has been widely used as insulation of high voltage cables (Cho, Shim, & Kim, 1998). For XLPE cables, cable joints are widely recognized as dielectrically weak points due to the introduction of insulation discontinuity and the manmade nature of their construction (Hunter, Lewin, Hao, Walton, & Michel, 2013). PD measurement is a nondestructive technique which has received global acceptance as an effective diagnostic tool with the capability to assess and monitor insulation systems for its integrity during manufacture and while in service (Satish & Zaengl, 1994).

Partial discharge (PD) is a type of breakdown that does not fully connect the electrodes. This can lead to serious insulation damage and considerably reduce the life span of high voltage equipment (Angrisani, Daponte, Lupò, Petrarca, & Vitelli, 2000). PD occurs if the local electric field is greater than the threshold value, causing a partial breakdown of the surrounding medium (Satish & Gururaj, 1993a). PD has a transient nature and is characterized by pulsating currents with a duration of several nanoseconds to few microseconds (Karthikeyan, Gopal, & Venkatesh, 2006). PD charge magnitude is not always proportional to the damage caused since PDs of tiny magnitude may rapidly lead to electrical tree growth, especially for high voltage cables (Z. Hao, Blackburn, Phung, & Sen, 2007b). Therefore, it is very cost effective if PD activity can be detected and

quantified in its early stage in order for replacement to be scheduled at a suitable time (Ma, Zhou, & Kemp, 2002b).

The IEC has a specific limit of PD for all power equipment. When insulation failure occurs, it is replaced with no information of the type of PD since current commercial PD detector does not provide any information on the PD source (Karthikeyan et al., 2006). PD development became active since 1980 due to revision of IEC 60270 (Gulski, 1995). PD classification is of interest because of the relationship between the PD activity and the dielectric materials aging process. Since each defect has a unique deterioration behavior, it is important to recognize the relationship between the PD patterns and the defect type in order to determine the insulation quality. PD pattern recognition is crucial in determining substantial risk of an imminent insulation breakdown where the current component may require servicing and replacement or not (Hoof, Freisleben, & Patsch, 1997). Many works have been performed on PD classification in various power system equipment, such as gas insulated switchgears and substations (Feng-Chang, Hong-Chan, & Cheng-Chien, 2013; Hamilton & Pearson, 1997; Liping, Ju, & Yilu, 2015; Meijer, Gulski, & Smit, 1998; Ziomek, Reformat, & Kuffel, 2000), power cables (Casals-Torrens, Gonz ález-Parada, & Bosch-Tous, 2012; Hunter et al., 2013) and transformers (Ibrahim, Sharkawy, Salama, & Bartnikas, 2012; Ke, Jinzhong, Shuqi, Ruijin, et al., 2015). Commonly used classifiers include neural networks (Feng-Chang et al., 2013; Mehrdad Majidi & Oskuoee, 2015), fuzzy logic (Contin, Cavallini, Montanari, Pasini, & Puletti, 2002; Salama & Bartnikas, 2000) and support vector machines (Jing & Hou, 2015; Khan et al., 2014).

PD has a group of unique discriminatory attributes which serves as an identification markers, allowing them to be recognized. In order to perform PD classification, it is necessary to choose which discriminatory features to be extracted and which feature

extraction method to be used (Hui, Chan, Saha, & Ekanayake, 2013). The purpose of feature extraction is to extract meaningful input feature from the unprocessed PD data to represent the PD pattern associated with a specific defect (Karthikeyan, Gopal, & Vimala, 2005). These extracted features are used as input of the classifier during the training process. Feature extraction also helps to reduce the size of raw PD data for quicker and simpler handling. PD classification requires some sort of data reduction method, such as reducing the matrix size. This is due to unprocessed PD data are too huge to be used as input to the classifiers as it will drastically increase the training time and cripple the performance of the classifier (Kranz, 1993; Yu & Song, 2003). Identification markers that are too huge in size will create redundancy and increase training time without improvement in classification accuracy. However, identification markers that are too small will cause unsatisfactory classification rate. Hence, ideal identification markers should be small in size and accurately represent the PD pattern produced by the defect.

## 1.2 Problem Statement

PD pulses have small magnitude but their long term effects will cause inevitable insulation breakdown. The interpretation of PD patterns is able to expose the cause of PD and measure the insulation performance of power equipment (Kai, Kexiong, Fuqi, & Chengqi, 2002). Although many researches had been done on PD classification in the past, they are based on self-fabricated insulation material from the laboratory instead of actual cable joints. Apart from that, most of PD classification works were performed in lab environment and under noise-free environment. However, in reality, on site PD measurement suffers from lower detection sensitivity due to the interference of external noises (Suzuki & Endoh, 1992). PD measurement often faces interference caused by radio transmissions, power electronic components, random noise from switching, lightning, arcing, harmonics and interferences from ground connections (Satish & Nazneen, 2003).

A lot of research works have been performed on denoising PD signals and it has improved over the years but a perfect and universal denoising standard has yet to be achieved. Most previous research works used artificial noise generated by adding evenly distributed random number to phase and charge or including random numbers with various standard deviation and zero mean instead of actual noise encountered on site. A more detailed research is needed for how different noise levels affect classification accuracy.

## 1.3 Research Objectives

This research aims to improve classification accuracy of PD under noisy conditions. The key objectives of this research are:

1. To perform measurement of partial discharge (PD) from artificially-prepared cable joint defects
2. To propose feature extractions using Principal Component Analysis (PCA) from PD data and artificial intelligence classifiers to identify PD types from cable joint defects
3. To compare the performance of the proposed principal component analysis (PCA) based input feature against traditional input features under noisy condition using different classifiers
4. To compare the performance of ANFIS classifier against SVM and ANN under noisy conditions

## 1.4 Research Methodology

This work focuses on obtaining PD patterns from five types of commonly encountered defect in actual XLPE cable joints and performing classification to identify the defect. Feature extractions were performed to obtain useful input features, which serve as identification marker for the cable joint defect. These input features were used as the

training input for three different classifiers. Different noise levels with increasing pulse count and magnitude were added to the PD signals to observe which input feature and classifier has the higher noise tolerance. The classifiers were trained using noise-free PD signals but tested with noisy PD signals. Noise levels of variable pulse count and noise level with variable amplitude were tested. The noise source was obtained from ground interference during raining which is not a randomly generated noise. This work also investigated a relatively new classifier, adaptive neuro-fuzzy inference system (ANFIS) and its feasibility as a PD classifier.

## 1.5 Thesis Contributions

The main contributions of this thesis are:

1. A new principal component analysis (PCA)-based input feature that has high noise tolerance compared to traditional input features used for PD classification has been proposed
2. A more realistic PD classification experiment by using actual cable joint samples has been performed
3. A better representation of noise contamination effect on PD classification has been proposed by using actual noise pattern from ground interferences with varying pulse count and charge magnitude

## 1.6 Thesis Organization

This thesis is divided into five chapters. Chapter 1 consists of an introduction, which introduces the background, problem statement, research objectives and research methodology of this work.

Chapter 2 deals with literature review while overviewing the previous related studies of this research area.

Chapter 3 presents the steps taken to achieve the objectives of this study. This includes preparation of the XLPE cable joints, PD measurement setup, acquisition of noisy signals and implementation of feature extraction and classifiers.

In Chapter 4, all measurement and classification results that have been performed in this work are detailed in a systematic manner. This chapter also deals with the comparison between different types of input feature and classifier under different noise levels.

Finally, Chapter 5 presents the conclusions and the future work recommended for this research work.

# CHAPTER 2: PD MEASUREMENT AND CLASSIFICATION

## 2.1 Introduction

This chapter begins with information about Cross-linked polyethylene (XLPE) cable joints. It also provides more information about partial discharge in terms of PD pattern representation, PD measurement methods, PD detection methods and PD denoising techniques. Detailed review of previous works related to PD classification was presented.

## 2.2 XLPE Cable Joints

Cross-linked polyethylene (XLPE) has well-balanced electrical, thermal and mechanical properties and it has been used for insulation of electrical cables (Cho et al., 1998). In order to connect different XLPE cables, cable joints are required. Installing a cable joint is a very delicate task and requires highly skilled professional cable jointers.

There are two main types of cable joints used in the power industry; heat shrink and cold shrink cable joints. Cold shrink and heat shrink cable terminations require very different installation methods. Heat shrink termination requires multiple tools such as gas torches and igniters. Heat shrinking requires very precise skill as uneven shrinking will produce inconsistent installation. Overheating while performing heat shrinking will produce scorch and burnt damage to the XLPE cable.

Cold shrinking cable joints do not require heat source and are more reliable and less susceptible to human error. It is also easier to install without any special training and tools. The cold shrink tube is an open ended rubber sleeve made of rubber elastomers that has been factory expanded and assembled onto a supporting removable plastic core. Cold shrink tubing shrinks upon removal of the supporting core and forms the cable joint connecting the two XLPE cable ends.

Cold shrink technology is well suited to outdoor applications as well as to medium and high voltage cable splices, cable joints and cable terminations. Generally, cold shrink tubing is more flexible than heat shrink tubing as it moves with the cable when it goes through thermal expansions and contractions.

Two main cautions must be made when connecting two high voltage cables. First, the conducting layer at the outer side of both cables must be terminated without causing a field concentration. This is similar with the construction of a cable terminal. The second issue is to avoid a void or any free space between the area of the cable insulation and the cable conductors (Frederik H. Kreuger, 1991). A bi-manchet as shown in Figure 2.1 can be used to overcome these problems.

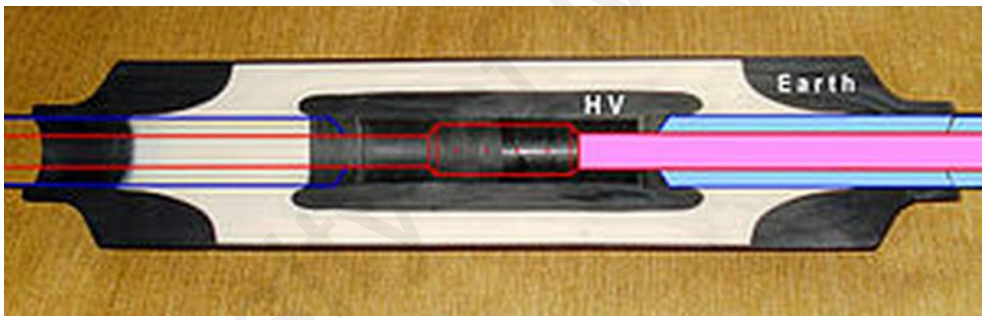



**Figure 2.1:** Bi-mancet at the XLPE cable joint

(Frederik H. Kreuger, 1991)

In Figure 2.1, the red color represents the conductor of the HV cable and the blue color is the cable insulation. The black sections are the semiconductor. The outer layer distributes the electric field while the inner layer diverts the field as seen in Figure 2.2. The equipotential lines were directed from the cable core to the outer section of the bi-mancet on both sides of the device.

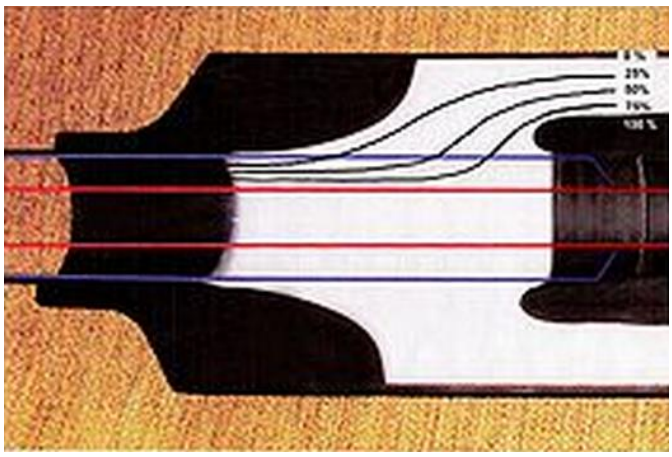


**Figure 2.2:** Field distribution in the cable joint

(Frederik H. Kreuger, 1991)

## 2.3 PD Representation

Time-resolved and phase-resolved data are the two major types of PD pattern representations currently used for PD related research. Phase-resolved PD (PRPD) data are acquired based on the AC test voltage waveform. The phase angle of the voltage is split into a specific amount of sections while the AC test voltage is maintained at a constant level. A PD detector is required to capture the individual PD signal and quantifies all pulses based on the phase angle occurrence ($\varphi$), charge magnitude ($q$) and the number of PD ($n$) over a predetermined time duration (Sahoo, Salama, & Bartnikas, 2005). The relevant phase and amplitude pulse numbers of PRPD patterns are usually stored in matrix format for computational purposes (Contin, Montanari, & Ferraro, 2000). These data are commonly known as $\varphi$-$q$-$n$ or PRPD patterns (Ardila-Rey, Martinez-Tarifa, & Robles, 2015).

Time-resolved data pattern has interesting benefits because individual pulse shape can be observed and there is correlation between the PD signal shape and the nature of the insulation defect, which provides aging information of the insulation system (F. H. Kreuger, Gulski, & Krivda, 1993). The measurement procedure of time-resolved patterns

typically requires less expensive measurement system compared to phase-resolved measurements.

Phase-resolved data are more widely used in PD classification research because it is able to represent the physical process at the PD location since individual PD pulse has a solid relation with the PRPD patterns (Karthikeyan, Gopal, & Venkatesh, 2008).

## 2.4 Online vs. Offline PD Measurement

PD measurement consists of two main approaches, which are off-line and on-line PD detections. Off-line methods refer to tests where the equipment under test is de-energized from normal operation and energized by an external voltage source ("IEEE Guide for Partial Discharge Testing of Shielded Power Cable Systems in a Field Environment," 2007). On-line tests are performed at operating voltage, thus the exact behavior of PD can be obtained and evaluated. The offline PD testing has several advantages, such as able to determine PD extinction voltage (PDEV) and PD inception voltage (PDIV) since measurement voltage can be controlled. The advantages of on-line PD testing are PD characteristics measurable at different load conditions and tests can be carried out without causing power outage.

Currently, only offline methods have been standardized in IEC 60270. On-line methods remain unconventional with no benchmark that one can make comparison with. However, online tests have become increasingly popular in cable PD detection in recent years. Both online and offline PD tests are complimentary of each other. By combining both methods, a more valid result of cable condition can be obtained.

Majority of the research related to PD classification uses offline detection method since it is convenient to conduct in a laboratory environment. In this work, offline measurement was used as it is more practical and feasible to be conducted in the

laboratory. Using a defective cable joint in an online grid is not only difficult but highly dangerous.

## 2.5 PD Detection Methods

The standard PD detection systems rely on electrical voltage or current pulse detection. Electrical pulse detection equipment is commercially available and can be installed in high voltage (HV) laboratory (F. H. Kreuger et al., 1993). Pulse detection method has its own advantages and disadvantages (Casals-Torrens et al., 2012). The advantages of this method are high demand for commercial applications, ability to detect PD level, auto calibration supported when double sensors are used, able to analyze PD direction for cable accessories and ability to work together with PD localization systems. However, the disadvantages of this method are expensive measurement equipment and challenging to install in the field, susceptibility to radio frequency (RF) interference when lack shielding, shield construction limits inductive methods and low coupling capacity reduces capacitive method's sensitivity.

Acoustic detection uses acoustic sensors to measure pressure fluctuations on the insulation surface. This offers an interesting measurement technique for PD detection (Casals-Torrens et al., 2012). The advantages of this method are electromagnetic noise immunity, non-destructive and non-intrusive, high sensitivity sensor, frequency spectrum has high range, sensor installation is not affected by shielding construction, robust mechanical strength, excellent electrical resistivity and more cost effective compared to other sensors. However, the disadvantages of this method are signal attenuation, measurement sensitivity is affected by temperature, cannot detect PD level, highly complex calibration required and limited capability when handling equipment with air insulation.

Although the acoustic detection has gained some popularity and quite a number of research had been done using acoustic methods (Al-geelani, Piah, Adzis, & Algeelani, 2013; Boya, Ruiz-Llata, Posada, & Garcia-Souto, 2015; Casals-Torrens et al., 2012; Danouj, Tahan, & David, 2013; Kundu, Kishore, & Sinha, 2012), electrical pulse detection is still the common industrial method for PD measurement. In this work, electrical pulse detection was used because it conforms to the IEC60270 PD measurement standard.

## 2.6 PD Denoising

Ideally, by analyzing the specific combination of PD phase distribution, pulse magnitude and pulse count with time, PD patterns can be classified (Ma, Zhou, & Kemp, 2002a). However, during PD measurement, a difficulty encountered is caused by external noise interference, which degrades the PD measurement detection sensitivity. Major interferences faced during PD measurements are caused by discrete spectral interferences, stochastic pulse shaped interferences and periodic pulse shaped interferences (Satish & Nazneen, 2003).

One way to easily disregard the presence of noise is to set a threshold to ignore signals that are less than 10% of the maximum discharge amplitude (Gulski, 1993). However, this method is not suitable for application which requires good accuracy. This is because big threshold level ignores actual PD pulses that have small magnitudes while small threshold level mistakenly detects huge amount of noise as PD pulse (Allahbakhshi & Akbari, 2011; Shim, Soraghan, & Siew, 2001).

The Mean Square Error method was used in (Sriram, Nitin, Prabhu, & Bastiaans, 2005) to compare the performance of 28 different types of denoising methods. It was found that wavelet based denoising has the best results. The fundamentals workings of wavelet transform can be found in (Bentley & McDonnell, 1994; Chul-Hwan & Aggarwal, 2000).

Steps to select an optimal mother wavelet and setting automated thresholding rule are documented in (Ma et al., 2002a, 2002b). The *J* criterion (Jin et al., 2006) can also be used to determine the optimum mother wavelet. For PD denoising, Daubechies (dB) was the most popular mother wavelet choice since the Daubechies wavelet and a single PD pulse shape have high similarities. Numerous research works have also used wavelet transform for denoising purposes, especially the Daubechies wavelet, which is capable of detecting high frequency, fast decaying, short duration and low amplitude signals (Angrisani et al., 2000; Lalitha & Satish, 2000; Satish & Nazneen, 2003; Shim et al., 2001).

After an optimal mother wavelet is determined, the maximum number of decomposition can be calculated using the $J_{max}$ formula (Zhou, Zhou, & Kemp, 2005). Discreet wavelet transform (DWT) is preferred over continuous wavelet transform (CWT) since CWT is much more difficult to compute and produces a lot of irrelevant data (Satish & Nazneen, 2003). Successful denoising is achieved if it has low amplitude reduction, minimum pulse shape distortion and high signal to noise ratio (Satish & Nazneen, 2003).

PD denoising are usually done offline because online PD denoising is much more challenging due to nonzero wavelet coefficients that are higher than the PD coefficients. A new method for online PD denoising is to raise the voltage to slightly below PD inception voltage (PDIV) to record noise level of the measurement system (Z. Hao, Blackburn, Phung, & Sen, 2007a). A threshold value is calculated using the recorded noise level. Lastly, wavelet transform de-noises the PD signal using the calculated threshold level.

Notable recent advancement in wavelet denoising is the introduction of second generation wavelet transform (Xiaodi, Chengke, Hepburn, Guobin, & Michel, 2007) and

complex wavelet transform (Xu, Tang, & Sun, 2007). Second generation wavelet transform differs from DWT by providing interpretation of a fully spatial domain of the transform compared to the original frequency domain-based constructions. Complex wavelet transform is shift invariant and possesses greater directional selectivity while filtering multidimensional signals.

## 2.7 Previous PD classification works

There are numerous research work related to PD classification that had been done in the past. This section summarizes them into 5 main categories according to the classification method used. They are ANN, Fuzzy logic related techniques, SVM, combination of two or more classifiers and other miscellaneous classifiers.

### 2.7.1 Artificial Neural Network (ANN)-Based Classifiers

The earliest PD classification work was reported in (Suzuki & Endoh, 1992). The test sample used was a 66 kV XLPE cable with artificial defect and it was tested at 38 kV. The phase vs. charge plot was reduced into a 20x30 pixel image to be used as the input feature. The classifier used was a three layer feed forward back propagation neural network. The neural network has three outputs, which are “no PD” for safe PD level, “Warning” for moderate PD level and “Alarm” for high PD level. The neural network was trained using 30 input data and able to achieve 90% classification rate.

Back propagation neural network was used in (Mazroua, Salama, & Bartnikas, 1993) to classify PD pattern caused by two different cavity sizes and two different cavity shapes. The input features used in this work include rise time, fall time, apparent charge, area and width of the PD pulse. The test samples used were acrylic disk with different artificially created cavity shape and sizes. 30 measurements were taken from each sample and 20 of them were used for training and 10 for testing purposes. It was found that the neural

network was able to differentiate different PD patterns caused by different cavity shapes with 100% accuracy, but only 60% average accuracy in distinguishing cavity sizes.

In (Satish & Gururaj, 1993a), the PD pattern observed on the oscilloscope was captured using a camera and processed using a custom boundary tracing algorithm to generate input features. A multilayer back propagation neural network was used to classify PD patterns. A total of 21 test samples with different artificial defects were prepared such as laminar cavity, sharp points in various conditions, gas bubbles, cavities etc. However, due to equipment limitation, the pattern information cannot be shown in a single image capture, thus all their samples were split into two main categories. The classifier was trained to classify PD patterns into the 2 categories. The classifier managed to achieve accuracy of 85.5% and 82.25% for both category.

ANN was used in (Satish & Zaengl, 1994) to classify PD patterns. Test samples used in their work were 7 different flat cylindrical void specimens. Feature extraction was performed on a phase resolved PD (PRPD) pattern. The PRPD pattern was reduced to a smaller dimension of 16 (amplitude) x 32 (phase) to be used as the input feature. The total average classification accuracy was 79%. A random noise was added into the PD pattern to observe the implication. Random noise was created by adding random number to non-zero values in the pulse count axis of the PRPD pattern. The classification accuracy of the contaminated sample reduced to 73.85% with 5% noise and 42.2% with 10% noise.

A research work done in (Cachin & Wiesmann, 1995) attempted to classify PD pattern. The PD patterns used in their work were generated using stochastic computer simulations. 8 classes of PD test pattern were used where 6 of them simulated void discharge and 2 of them simulated corona discharge. A total of 25 data set were prepared for each class. A 3D PRPD pattern that had been discretized to 256x256 was used as input feature to the classifier. They also discretized a 2D PRPD pattern into 8x8 points for comparison. ANN

was used to classify the PD patterns and the combined recognition rate of all samples was 79%. By normalizing the charge scale of the PRPD pattern and using contours, they managed to separate superimposed patterns that has minor overlapping.

Research in (Mazroua, Bartnikas, & Salama, 1995) used ANN to classify cavity defect sizes and electric tree defect aging times by using the PD patterns measured from both defect. Input features such as peak amplitude, rise time, fall time, width and area of the PD pulse were extracted from the PD signal. The neural network was able to distinguish perfectly between cavity defect and electric tree defect. Average accuracy of 94% was achieved when classifying different cavity sizes while 91% was achieved when classifying different aging time for electric tree.

A new type of ANN was proposed in (Hong, Fang, & Hilder, 1996) which is known as modular neural network. Modular neural network is basically 5 individual networks that are connected in series. Each sub network classifies one defect against the rest and passes the remaining test samples into the following sub network until all test samples are classified. The 5 test samples used in their work consist of high voltage buses with different defect caused by different size of needle point diameter. The discharge pulse count, maximum and average discharge amplitude within a 10 degree phase window were used as input feature to the classifier. The single neural network achieved 88.31% while the modular neural network achieved 93.6% accuracy. They later extended their work to show that modular neural network has higher training speed compared to singular neural network (Tao & Fang, 2001).

Envelope extraction is a new feature extraction proposed in (Hamilton & Pearson, 1997). Envelope extraction is an image processing technique based on mathematical morphology. A 420 kV GIS $SF_6$ chamber was used as the test sample with 3 types of defects created using aluminum particles, needle tip of different length and diameter and

aluminum wire of different length. ANN was used to classify the measured PD patterns into 3 groups. The overall classification rate achieved was 94.6%.

Counter propagation neural network was proposed in (Hoof et al., 1997). Test samples consist of low density polyethylene (LDPE) and plexi glass, six types of defects were created, which are electrode bound cavity, point to dielectric gap in air, surface discharge in air, corona discharge in air, electrical treeing in polyethylene and stochastic discharge. The $\Delta u_n(\Delta u_{n-1})$ pattern which was derived from the PD pulse sequence was used as input feature, $u_n$ is the magnitude of the PD pulse sequence. Classification of 100% was obtained using the counter propagation neural network. Counter propagation neural network (CPN) consists of partly self-organizing maps (SOM) that are combined with an outstar structure. CPN is used for estimating a function that is characterized by a group of desired pairs of input output and their function of inverse. For PD classification, the forward variant of the CPN (FCPN) was utilized. A FCPN is depicted in Figure 2.3 (Hoof et al., 1997). In this figure, part of the Kohonen self-organizing map is merged with the outstar structure (Grossberg, 1969). It works as a lookup table which can compare a pattern with the prototypes encoded in input-to-hidden weights and chooses the most identical one. Then, the results are encoded in the hidden-to-output weights. The previous weights are trained by unsupervised competitive learning while the latter weights are trained by supervised learning. To increase the classification efficiency, the network structure is dynamically modified by using a vigilant structure. The complete working of CPN can be found in (Hecht-Nielsen, 1988; Hoof et al., 1997).

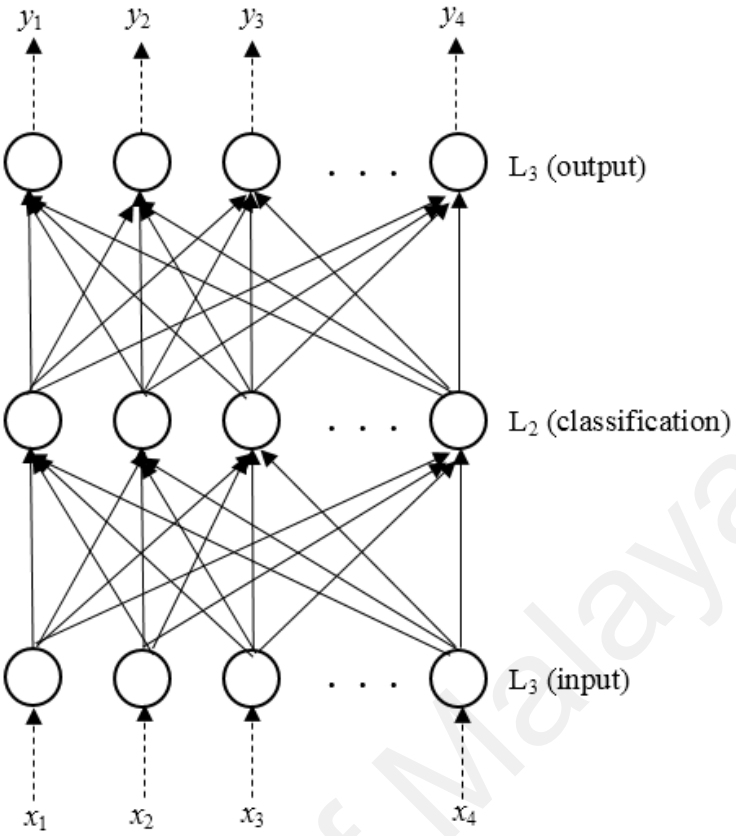



**Figure 2.3:** Architecture of the FCPN

(Hoof et al., 1997)

A new feature extraction was proposed in (Lalitha & Satish, 1998), where the PRPD patterns were compressed using quad tree partitioning fractal image compression. Then, the average pulse count and phase and magnitude spreads were calculated from the compressed pattern and used as the input feature. PD data used for classification are single point corona in air, cavity discharges, surface discharges in air and multiple corona in air, which were obtained from their previous work in (Krivda, Gulski, Satish, & Zaengl, 1995). Using ANN as the classifier, 20% of the data achieved 100% classification accuracy while the remaining 80% of test sample had classification accuracy of 75%.

ANN was also used in (Candela, Mirelli, & Schifani, 2000) for PD classification. Epoxy sheet was used as test samples and different defects were created such as surface discharge, spherical void, discharge between metallic and dielectric surface. Fractal

dimension, lacunarity, skewness, kurtosis and Weibull parameters such as α and β were used as input feature. Classification accuracy of each samples ranges from 87% to 98%.

Radial basis neural network was used in (Lalitha & Satish, 2000) to classify PD patterns obtained from single point corona in air, surface discharge and cavity discharge. They overlapped different types of PD to show that multi-resolution signal decomposition method was able to separate the PD patterns as long as the overlapping was not too severe. For the three discharge types, classification accuracy of more than 88% was achieved.

Extending their previous work in (Salama & Bartnikas, 2000), Salama and Bartnikas used the autoregressive time-series analysis to determine the number of input features to be used for the cascaded neural network (CNN) (Salama & Bartnikas, 2002). The CNN was able to determine small variation in different cavity sizes better than normal neural network. This is due to CNN uses an indexed feature that sets up the highest vital input to the next stage of CNN that is influential in producing a classified attribute output from the CNN. Figure 2.4 shows an example of CNN with double outputs.

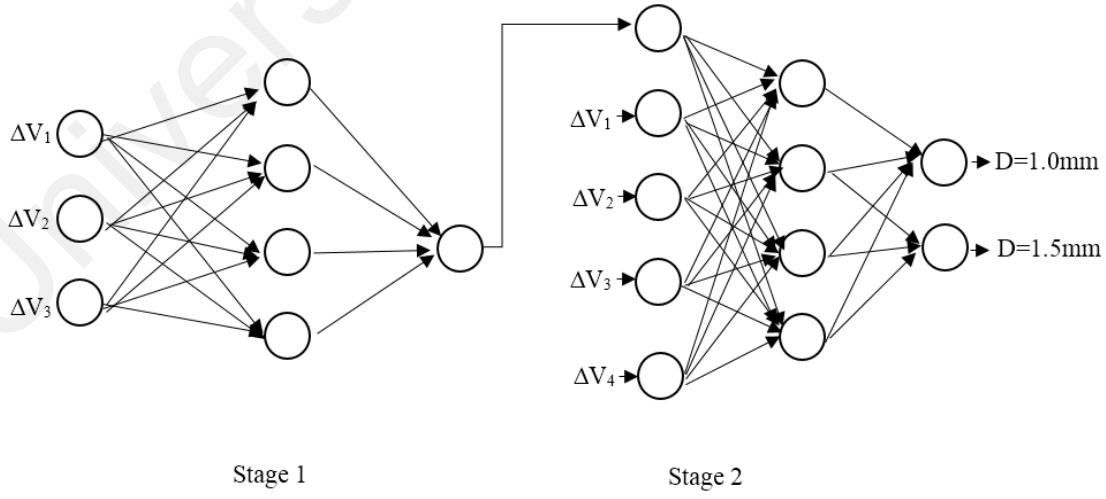



**Figure 2.4:** CNN with double outputs

(Salama & Bartnikas, 2002)

For this CNN system, the output of the first stage that is indexed $I_d$, is created to represent unity output for a 1.5 mm void and zero for a 1.0 mm void. The indexed signal $I_d$ is passed to the second stage along with $\Delta V$ and $\Delta Q_m$ inputs as one of the input. After training, CNN can successfully distinguish different cavity sizes efficiently using three inputs with 6 neurons in the hidden layer.

Kai et al. used stator bar as test samples and created artificial defects such as cavity discharge, surface corona, tiny cavity discharge and slot discharge (Kai et al., 2002). Back propagation neural network was used to test several different input features. The 6 types of input features used and their average classification accuracy are 95.75% for tabulated data, 92.25% for surface fitting parameters, 92.63% for moment features, 89.25% for statistical features, 75.67% for fractal features and 90.75% for combination of fractal features and barycenter coordinates.

ANN was used in (C. Chang et al., 2005) to classify different PD patterns in gas insulated substation (GIS). Artificial defects created included corona, particle on enclosure, particles on bus conductor and particle on spacer. Fast Fourier transform and discreet Wavelet transform were used to extract four types of input features from the PD wave shape, which are simple count ratio, mean within partition variation, between partition migration count and a combination of all previous features. By using all input features, no misclassification occurs. Extending their work C. S. Chang et al. used local discriminant features obtained using wavelet packet transform to be used as input feature (C. S. Chang et al., 2005). ANN was used to differentiate PD signals and PD signals combined with corona. The error rate was 1%.

A new variant of ANN known as probabilistic neural network (PNN) was used in (Karthikeyan et al., 2005) for PD classification. Perspex material was used to create three types of defects, void, corona in air and corona in oil. Statistical features were used as

input features. A total of 55 PD reading were measured from the test samples. It was found that 100% accuracy was achieved if larger test samples were used. PNN network is a network formulated by the probability density function (PDF). It is a network that mainly relies on competitive learning by using a winner takes all style and the fundamental theory based upon the multivariate probability (Karthikeyan et al., 2008). It offers a general key to classification of PD pattern using a statistical approach known as Bayesian classifiers. PNN also utilizes the Parzen Estimators created to build the PDF needed by Bayes theory (Karthikeyan et al., 2005; Su, Chia, Chen, & Chen, 2014). The architecture of a standard PNN is shown in Figure 2.5.

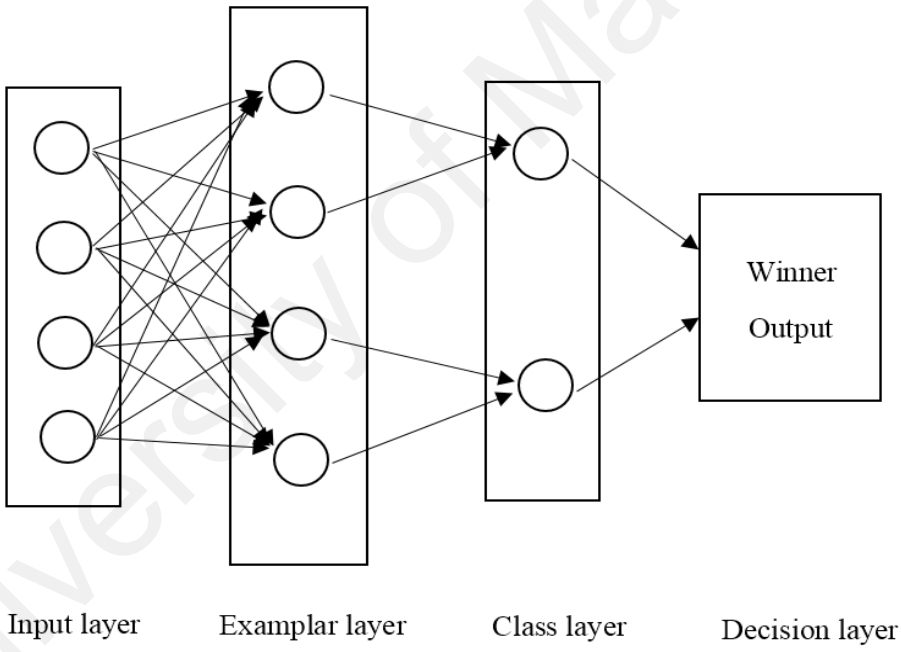



**Figure 2.5:** Architecture of PNN

(Karthikeyan et al., 2005)

There will be an exemplar node (unit) or a single pattern of example for each training. Every pattern unit forms a dot product of the example given for classification and weight vector, where the weights entering a node came from a given example. A principal advantage of the PNN paradigm is that it is usually much quicker than the usual back-propagation network, where the incremental time adaption of back propagation is a

considerable fraction of the total computation time (Specht, 1988, 1990). PNN has several variations such as Adaptive PNN (APNN), composite PNN (CPNN), Radial Basis PNN (RBPNN), heteroscedastic PNN (HRPNN) and robust heteroscedastic PNN (RHRPNN). APNN utilizes the alteration in the smoothing parameter or the variance parameter, which requires different variance parameter of each class depending on the calculation of the average distance and Euclidean distance computation (Karthikeyan et al., 2008). CPNN uses multifarious inputs to obtain a precise and reliable decision on the classification group (Karthikeyan et al., 2008). RBPNN is trained using the forward orthogonal least square method (FOLS). The combination of FOLS-RBPNN is capable of handling a range of complexities such as ill-conditioned and large dataset training, overlapped multiple discharge sources with ease and varying the applied voltage (Venkatesh & Gopal, 2011a). HRPNN is basically similar to the original PNN with few big alterations in the exemplar layer and the method of vector selection during training at the output layer. RHRPNN is a modified version of HRPNN by including a statistical tool that is powerful for improved interval approximation known as the Jack-knife method (Venkatesh & Gopal, 2011b).

Extension neural network (EXNN) was used in (Mang-Hui, 2005; Mang-Hui & Chih-Yung, 2005). High voltage current transformer was used as test samples with artificial defects such as corona, low voltage coil PD and high voltage coil PD. The mean value of total discharge magnitude, maximum discharge magnitude and mean value of the maximum discharge magnitude were extracted to be used as the input features. Extension neural network was compared with back propagation neural network, where both network achieved 100% accuracy under noise-free condition. Noisy data were created by adding uniformly distributed 30% random error data. Extension neural network retained 92% classification accuracy while back propagation neural network fell to 80%. The extension neural network is a new type of neural network, which essentially combines extension theory with neural network by using extension distance instead of Euclidean distance to

check similarities between test data (Wang, Tseng, Chen, & Chao, 2009). Extension theory provides an innovative distance measurement for PD recognition while the neural network helps to entrench the parallel computing features and learning abilities. The extension neural network is able to process the clustering difficulties of a range of feature values with continuous input, supervised learning and descriptive output. Due to the lack of complicated training process, extension neural network is capable of quick adaptive process for huge quantity of training data and defined the lower and upper bounds straight from the training patterns. Figure 2.6 shows the extension neural network structure which consists of the output and input layers.

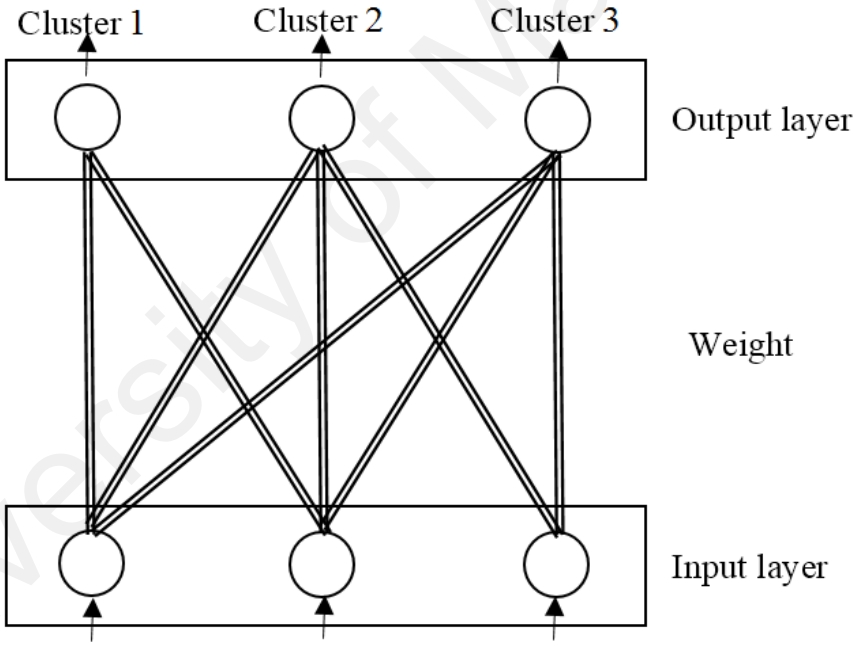



**Figure 2.6:** Structure of the EXNN

(Mang-Hui, 2005; Mang-Hui & Chih-Yung, 2005)

There are 2 connection values known as weights between every output node and every input node. They represents the feature upper bound and lower bound of the classical domain feature. One node exists at the output layer of every pattern of the prototype and only single non-zero output node to designate the prototype pattern, which is closest to

the input vector. The inner workings of extension neural network is described in (H.-C. Chen, Gu, & Wang, 2012).

The effectiveness of fractal features and statistical features as input features was tested in (Jian, Caixin, Grzybowski, & Taylor, 2006). Five types of PD patterns were obtained from a plane to electrode systems such as dielectric surface discharges in air, single point corona discharges, single point discharges, dielectric surface discharges in oil and discharges in cavity electrode system. A back propagation neural network was used as a classifier. It was found that fractal features achieved 84.26%, statistical features achieved 88.44% and a combination of fractal and statistical achieved 93.64% classification accuracy.

A research work done in (Jin et al., 2006) used ultra-high frequency (UHF) measurement of PD signals and wavelet packet transform was used to extract input features from the PD signals. GIS was used as test samples with artificial defects such as corona, particle on surface, particle on conductor and free particle on enclosure. ANN was used as a classifier and achieved 98.3% classification accuracy for 4 test subjects. This work was further expanded in (Karthikeyan et al., 2008) to compare different variant of probabilistic neural network.

ANN was used in (Boczar, Borucki, Cichon, & Zmarzly, 2009) to classify PD patterns measured using acoustic method. 8 types of artificial defects were created, which include point to point discharge in oil, point to point in oil with gas bubbles, point to plane in oil, surface discharge, multipoint surface discharge, multipoint plane discharge in oil, multipoint plane discharge in oil with gas bubbles and moving particles discharge in oil. Input features were extracted using short time Fourier transform and power spectrum density method. A classification accuracy of up to 90% were achieved.

Probabilistic neural network was used in (Evagorou et al., 2010) to classify PD patterns measured from lab created artificial PD sources. Input features used were the wavelet coefficients at various scales extracted using wavelet packet transform. Classification accuracy of 97.49%, 91.9%, 100% and 99.8% were obtained for corona in air, floating discharge in oil, internal discharge in oil and surface discharge in air.

A type of ANN known as radial basis probabilistic neural network (RBPNN) was used in (Venkatesh & Gopal, 2011a) to classify 4 types of PD sources; void, air corona, oil corona and void with air corona. Forward Orthogonal Least Square algorithm was used for feature extraction. RBPNN managed to achieve classification accuracy between 80% and 90% while PNN achieved classification accuracy of 50% to 70%. Robust Heteroscedastic Probabilistic Neural Network was later used in (Venkatesh & Gopal, 2011b), which achieved classification accuracy more than 90%.

A research work done in (Al-geelani, Piah, & Shaddad, 2012) used radial basis neural network to classify acoustic signals caused by surface discharge. The test samples used were high voltage glass insulators. Wavelet transform was performed on the measured signals and statistical features were calculated from the wavelet coefficients to be used as input feature. It was found that no misclassification occurred.

ANN was used in (F. C. Gu, Chang, Chen, Kuo, & Hsu, 2012) to classify PD patterns measured from XLPE cable joints. The 3 defects created at the cable joints were short insulation, long insulation and knife defect. PDs were measured using acoustic method and processed using Hilbert Huang transform. Fractal features were extracted from the Hilbert energy levels and PRPD pattern. Classification accuracy was 100% for both fractal features under noise-free condition. However, when 30% noise was applied, classification accuracy of ANN when using fractal features from Hilbert energy levels was 76.7% while fractal features from PRPD pattern was 68.3%. The same test sample

and input feature was used again in (F.-C. Gu, Chang, Chen, & Kuo, 2012), but extension method and ANN was used as classifier. Both classifiers achieved 100% accuracy under noise-free condition. When 30% noise was introduced, extension method and ANN achieved 81.67% and 73.33% accuracy. Similar input feature and classifier were again used to classify defects from GIS (Feng-Chang et al., 2013) and high voltage transformer (Hung-Cheng, 2013).

Particle swarm optimization (PSO) was proposed in (Al-geelani et al., 2013) to optimize the input features to increase classification accuracy. PD signals were measured using acoustic method from glass insulators contaminated by three levels of salt water concentration. Wavelet transform was used as feature extraction. By using PSO, classification accuracy of radial basis neural network managed to increase from 96.98% to 100%. Similar work with more detailed description was done in (Al-geelani, M. Piah, & Bashir, 2015).

RF antenna was used to measure PD in a research work done in (Shurrab, El-Hag, Assaleh, Ghunem, & Jayaram, 2013). Silicon rubber surface was used as a test subject with artificial surface discharge and corona from positive and negative end. Statistical and wavelet analysis features were used as input features to an ANN classifier. The classification accuracy was 96%.

A modified ANN known as ensemble neural network (ENN) was proposed in (Abubakar Mas'ud, Stewart, & McMeekin, 2014) for PD classification. 4 types of artificial defects were created in laboratory, which includes corona in air, surface discharge, single void and electrode bound cavity. Statistical features were used as input feature and classification accuracy of 95% was obtained. The ENN technique is based on training a number of neural network models with statistical parameters from PD patterns and combining their predictions. ENN improves upon the generalization performance of

a single neural network by simply training many NNs and combining their component predictions. These models include the naive classifier, basic ensemble network, a generalized ensemble network and dynamically weighted ensemble (DEM) network. Among various ENN output aggregation systems, DEM has been shown to outperform other techniques when applied to a number of dataset categories. DEM determines the output weight of each of the individual NNs and provides the best performance at any instant of evaluation. The weight is proportional to the certainty of the individual neural network predictions and this certainty evaluates how close the output is to any known target value. An example of ENN model is shown in Figure 2.7, where statistical parameters are used as the input (Abubakar Mas'ud et al., 2014).

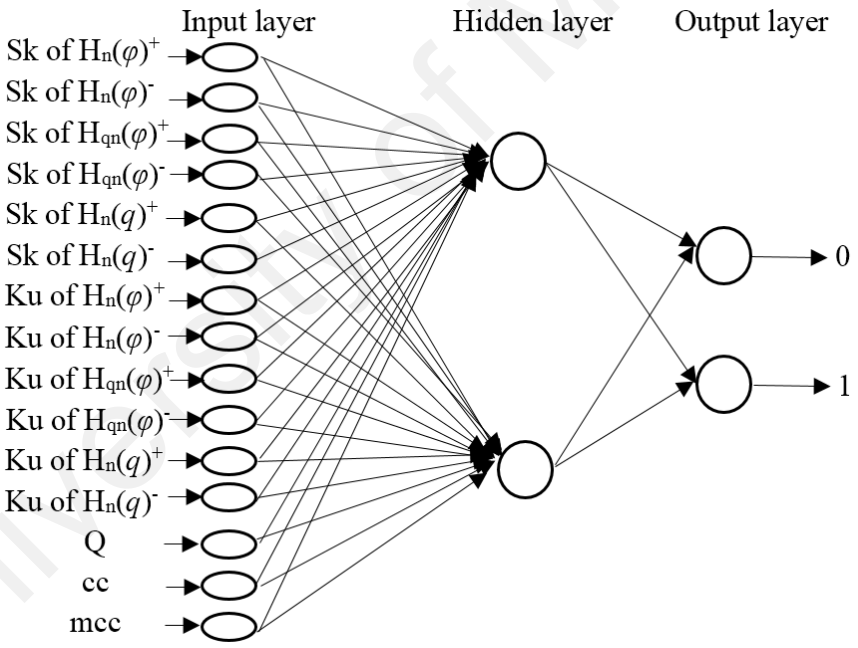



**Figure 2.7:** Structure of the ENN

(Abubakar Mas'ud et al., 2014)

The modular neural network proposed in (Hong et al., 1996; Tao & Fang, 2001) shares similar concept as ENN, where multiple individual neural networks are trained and their outputs combined.

A research work done in (Sinaga, Phung, & Blackburn, 2014) used UHF sensor to measure PD. Wavelet transform was used to denoise the PD signals. 3 types of PD patterns were measured from defects created in the laboratory such as void, floating metal, mixture of void and floating metal. Statistical features and wavelet features were used as input feature to an ANN classifier and the classification accuracy achieved was 95.3% when denoising was used and 75.3% with no denoising.

A combination of phase resolved PD data and time resolved PD data was proposed in (Liping et al., 2015) to be used as input feature. A technique called Dempster Shafer evidence theory was used as a data fusion technique. GIS was used as a test sample with the following defects, protrusion, surface contamination, void in spacer and free particles. ANN was used as a classifier, which achieved classification accuracy of 82.94% when using time resolved PD data, 92% when using phase resolved PD data and 97.25% when using the fused data.

### 2.7.2 Fuzzy Logic-Based Classifiers

A fuzzy expert system was used in (Salama & Bartnikas, 2000) to classify different void sizes of 1 mm,1.5 mm and 2 mm based on the PD data measured. Input features such as pulse width apparent charge, pulse fall time, pulse rise time and area under pulse were extracted from the PD signal to be used as input feature to the classifier. The classification accuracy was not specified but it was said to be inferior compared to previous attempts with ANN. However, it was mentioned that fuzzy expert system is much easier to be implemented than neural network.

Fuzzy C mean classifier was used in (Contin et al., 2002) to separate overlapping PD signals. The equivalent time length and equivalent bandwidth were calculated from the PD signal to be used as input feature. Weibull parameters from the pulse height

distribution were also used as input feature. Turbo generator stator bars and induction motor coil were used as test samples. Artificial defects created include internal discharge from micro voids, electrical noise, corona and surface delamination discharge. Their work shows that separation of PD signals is possible, provided that both signals are not fully overlapped.

Further extending their work from (T. K. Abdel-Galil, Hegazy, Salama, & Bartnikas, 2005) , (T. K. Abdel-Galil, R. M. Sharkawy, M. M. A. Salama, & R. Bartnikas, 2005) used the same apparent charge and applied voltage as input features to the adaptive neuro-fuzzy inference system (ANFIS) and achieved classification accuracy of 82.2%. Since then, there have been several works which used ANFIS for PD classification (Canxin et al., 2009; Chalashkanov, Kolev, Dodd, & Fothergill, 2008; Fard, Akbari, Shojaee, Mirzaei, & Naderi, 2010; Sinaga, Phung, & Blackburn, 2010).

Neuro fuzzy network was used in (Mazzetti et al., 2006) to classify PD patterns in heat shrink joints and terminations of XLPE cables. 4 types of PD pattern were obtained from a cable joint with burned semiconductor and cable termination with cut in dielectric layer, particle insertion and asymmetrical shrinking. Statistical parameters were used as input feature and the classification accuracy of 80% to 99% was achieved depending on the test voltage. Their work was extended in (Rizzi, Mascioli, Baldini, Mazzetti, & Bartnikas, 2009) by using fuzzy min max classifier optimized by genetic algorithm. The same test samples and input features were used and there was no misclassification.

### 2.7.3 Support Vector Machines-Based Classifiers

Support vector machine (SVM) was used in (Sharkawy, Mangoubi, Abdel-Galil, Salama, & Bartnikas, 2007) for PD classification. Test samples used were cylindrical glass specimen holder with two parallel plane aluminum electrodes and four types of PD

pattern were obtained. Higher order moments and entropy of the estimated density function were used as input feature. Different combinations of moments were tested and the best classification accuracy of 98% was obtained by using 6 moments as input feature.

SVM was used in (Si, Li, Yuan, & Li, 2008) to classify PD patterns obtained from internal discharge, surface discharge and corona. Input features were extracted from the pulse wave shapes using the equivalent time-frequency method. PD was measured at DC voltage instead of AC voltage. The classification accuracy was not specified but the proposed system was able to distinguish between different PD patterns.

SVM was used in (L. Hao & Lewin, 2010) to classify PD patterns from artificial defects such as corona with remote earth, surface discharge in air and internal discharge in oil. The phase charge average was used as input feature. Out of 60 samples, only 2 were misclassified.

SVM was used as a classifier in (Hunter et al., 2013) to classify PD measured from PILC cables joints with the following defects; spike on the ferrule, void on the top, spike on the ferrule and void in the crutch. Denoising was done by ignoring PD signal below 25pC. A combination of statistical features and features extracted from wavelet transform were used as input features. The classification accuracy achieved was 91.1%.

SVM was used in (Khan et al., 2014) to classify PD patterns measured from GIS. Artificial defects were created by placing metal particles at different location at a Perspex spacer in the GIS. Statistical features were used as input features and classification accuracy of 94% was achieved.

A research work done in (Xiaoxing, Song, Na, Ju, & Wei, 2014) made a GIS simulator in the laboratory and obtained 4 PD readings from defects such as metal needle, free moving particles, fixed metal particles and air gap defect. A mixture of statistical features

and features extracted via chaos theory was used as input features. SVM was used as a classifier and achieved classification accuracy of 98%.

### 2.7.4 Combination of Different Classifiers

Three types of classifiers were used in (Gulski & Krivda, 1993) for PD classifications, back propagation neural network, Kohonen self-organizing map (SOM) and learning vector quantization. Statistical features such as skewness, kurtosis, cross correlation factor and discharge factor were used as input feature to train the classifiers. Similar to previous work, noise suppression of 10% was implemented and a GIS was used as a test sample. Defects that can be categorized by the classifiers include cavity, surface, corona and treeing. Many different combinations of test and training were performed. When only two PD sources were classified, all classifiers were able to achieve 100% classification. However, when more samples are tested simultaneously, classification rate decrease significantly to only 30%. When all samples were tested simultaneously, back propagation neural network outperformed the other classifiers.

A work was done in (Mazroua, Bartnikas, & Salama, 1994) to classify PD patterns on artificial cylindrical cavities of different sizes using nearest neighbor classifier, learning vector quantization and ANN. When performing classification between 1 mm and 2 mm, 1 mm and 3 mm cavity size, all classifiers used achieved similar accuracy of 97.5% and 100%. However, when classifying smaller cavity sizes of 1 mm and 1.5 mm, learning vector quantization has higher accuracy of 77.5% compared to the other two classifiers, which has 72.5% accuracy. Their works concluded that all three classifiers have equivalent performance for most conditions, but learning vector quantization has the upper hand when attempting to distinguish smaller cavity sizes.

A research work done in (Dey, Chatterjee, Chakravorti, & Munshi, 2010) used cross wavelet transform as feature extraction method. The test samples used were acrylic resin with four types of voids, narrow void in contact with the electrode, narrow void away from the electrode, large void in contact with the electrode and large void away from the electrode. Rough set theory (RST), ANN and fuzzy c-means (FCM) were used as classifiers. Their test was repeated with and without noise. RST, ANN and FCM achieved classification accuracy of 87.5%, 83.3%, 85.4% with noise and 91.6%, 93.7%, 91.6% without noise respectively.

A comparison between SVM, ANN and self-organizing maps (SOM) was done in (Lai, Phung, & Blackburn, 2010). PD sources were obtained from 4 samples, which are corona discharge from two needle with different sharpness, surface discharge from cable termination and surface discharge from Perspex block. Statistical features and principal component analysis (PCA) features were used as input features. It was proven that SVM has the highest classification accuracy (98.63%), followed by ANN (96.89%) and lastly SOM (96.04%).

Extension method was used in (H. C. Chen, 2012) to classify PD patterns. Fractal features and mean discharge values were used as input feature. XLPE cable was used as test samples with artificial defects such as semiconductor layer defect, insulation layer defect, cable termination defect and a good cable. Artificial white noise up to 30% of the mean discharge value was added to the test data. Classification accuracy of extension method was compared with ANN and K-means method. Under noise-free condition, extension method and ANN achieved 100% while K-means method achieved 88.75%. When 30% noise was added, classification accuracy of extension method, ANN and K-means method fell to 80%, 69.9% and 60.7% respectively.

A research done in (Darabad, Vakilian, Phung, & Blackburn, 2013) has tested multiple input features and classifier pairs. PDs were measured from discharge sources such as air corona, discharge in oil, surface discharge, single cavity and noise. K nearest neighbor, regression tree and ANN were used as classifiers. Input features consist of statistical features, texture features, fast Fourier transform features, Cepstral features and linear discriminant analysis features. Most combination achieved satisfactory classification accuracy between 90% and 100%.

A research work was done in (Hui et al., 2013) to classify 5 types of PD sources such as corona, discharge in oil, internal discharge, surface discharge and discharge due to floating particles. Feature extractions used include stochastic neighbor embedding, PCA and discrete wavelet transform. SVM and ANN were used as the classifiers. It was shown all classifiers and input features achieved classification accuracy of above 90% and SVM outperformed ANN in all input features used.

A research work done in (Su et al., 2014) attempted to classify PD measured from gas insulated load break switches (GILBS). 3 artificial defects were made in the GILBS such as high voltage bushing pollution, metal particles in cavity and floating electrode. Wavelet transform was used as a denoising method. Statistical features were used as input to probabilistic neural network and fuzzy C mean classifier. Classification accuracy of 87.6% and 85.6% was achieved respectively.

Research work done in (Harbaji, Shaban, & El-Hag, 2015) measured PD by using acoustic emission method. The test samples used were pressboard with artificial defects such as surface discharge, corona, void and PD from semi parallel planes. Discreet wavelet transform and PCA were used to extract features directly from the acoustic signal. SVM and K nearest neighbor were used as classifiers. Classification accuracy of up to 90% was achieved for all classifiers.

A new feature extraction was proposed in (Ke, JinZhong, Shuqi, Fei, et al., 2015) for PD classification. They used a method known as two directional modified fuzzy weighted two dimensional linear discriminant analysis for feature extraction. Artificial PD source created in the laboratory include cavity discharge inside pressboard, surface discharge in oil, corona in oil and air. Classification accuracy of 93% and 96% was achieved by the fuzzy C means and SVM classifier. Similar work was done in (Ke, Jinzhong, Shuqi, Ruijin, et al., 2015) but it mainly explores separation of overlapping PD signals using S transform and affinity propagation clustering.

A total of 17 PD sources were created in (M. Majidi, Fadali, Etezadi-Amoli, & Oskuoee, 2015; Mehrdad Majidi & Oskuoee, 2015) for PD classification. The PD sources were categorized into 4 groups, which are corona, surface discharge and three different void sizes (1 mm, 1.5 mm and 2 mm) with 5 different number of voids (1 to 5 voids). Primal dual interior point and basis pursuit denoising algorithm were used for feature extraction method. Sparse representation classifier and ANN were used as the classifier. The classification accuracy of each group varies from 80% to 95%.

### 2.7.5 Other Classifiers

A work in (Gulski, 1993) performed PD classification on polyethylene (PE) cables and 400kV gas insulated substation system (GIS). Artificial defects created at the PE cables include sharp electrode via stainless needle, cavity discharge via oval cavity under the semiconductor layer and treeing at cavity via sharp needle tip at flat cavity. For GIS, artificial defects include internal discharges via air pockets, floating part discharge via contact spring and corona discharge via splinter inside the GIS. Statistical features such as skewness, kurtosis and number of peaks were used as input features. In order to overcome noise interference, a 10% noise suppression was implemented, meaning only signals greater than 10% of the maximum PD amplitude was taken into consideration. It

was mentioned that 1 to 2 minutes of PD data is sufficient to determine the statistical parameters of the sample. Classification was done by comparing the statistical parameters of known defect models to the statistical parameters of the test samples. Samples that have statistical parameters within 90% similarity were considered to be in the same group. Using this classification method, recognition rate varies from 13% to 100%, the average recognition rate is 49.5% for all samples.

In a work done in (F. H. Kreuger et al., 1993), skewness, kurtosis, number of peak, discharge factor and modified cross correlation factor were used as input features. For classification, two techniques were used, which are the recognition rate and the centour score. Recognition rate is the same comparison method which was used in (Gulski, 1993) and centour score is defined as the percentage rank of the data that are further from the center of a known population that varies from 0 to 100%. It is not the probability of how the test sample is similar to a standard sample, but the best possible estimation of it. The test samples used are 12 dielectric sheets with different defects such as corona, surface discharge, treeing in air and oil surroundings. It was found that the centour score and recognition rate technique have similar classification capabilities and is able to achieve up to 80% accuracy.

Hidden Markov models was used in (Satish & Gururaj, 1993b) for PD classification. A camera was used to capture the PD signal pattern. The signal pattern was split into 36 points long horizontally and each point was scaled to the range or 0 to 1 vertically. This signal was used as input feature and a total of 30 data set were collected from each sample. This work used the same data set as in (Satish & Gururaj, 1993a) but the test samples were split into 3 groups and overall recognition rate of 84% was achieved.

In a work done in (Gulski & Kridva, 1995), PD patterns were used to classify different insulation based on the aging duration. Test samples used include polyethylene with flat

cavity, current transformer and epoxy insulator. The test samples were subjected to accelerated aging by applying high voltage for extended period and feature extraction was performed at different aging stages. The classification aims to classify which stage of accelerated aging the insulator is based on the PD patterns measured. Statistical features were extracted from the PRPD distribution to be used as input feature. Classification was done by using statistical evaluation of mutual comparison of all measured PD data. The polyethylene sample was used to determine how many distinguishable aging state are possible. Their work managed to achieve average accuracy of 84% for 4 classes in current transformer and 95.6% for 6 classes in the epoxy insulator.

A team of researchers in (Hucker & Kranz, 1995) used PD classification on GIS. GIS with two different gases were tested, which are $SF_6$ and air insulation. 8 types of artificial defects were created in the $SF_6$ and 7 types in air insulation. The L2-Euclidean distance was use as a classifier. Average classification accuracy of 66.4% was achieved for classifying 15 types of defects in both $SF_6$ and air by using input feature obtained from electrical PD detection. When input feature from acoustic PD detection was used, average classification accuracy of 69.5% was achieved for classification of 8 types of defects in $SF_6$ GIS.

A new type of input feature was introduced in (Krivda et al., 1995) for PD classification purposes. They introduced fractal features, which consist of the fractal dimension and lacunarity. 7 types of PD patterns were measured from different defect such as single point corona in air, multi-point corona in air, surface discharge in air, single point corona in oil, air bubbles in oil, dielectric bounded cavity and background noise. It was shown that fractal dimension and lacunarity obtained from the PRPD pattern formed different cluster for different defect. No classification was performed as this work mainly

suggested that fractal features can be used as an input feature for PD classification. More details of fractal features can be found in (Satish & Zaengl, 1995).

A research work done in (Rahman, Arora, & Srivastava, 2000) used a combination of dielectric sheet, glass and needle and plane electrode to create 4 types of PD patterns, which are glow corona, streamer corona, surface discharge, internal discharge, sharp point and multi sharp point discharge. Four types of texture analysis algorithms were used for feature extraction based on the PD pattern image, gray-level difference histogram, spatial gray-level dependence method, gray-level run-length method and power spectrum method. Principal component analysis was used to reduce the dimension of the input feature before using minimum distance classifier for classification purposes. Classification accuracy up to 95% was achieved by using the first three principal components.

A research work done in (Ziomek et al., 2000) used genetic algorithms to assist PD classification. PD patterns were obtained from GIS with artificial defects such as voids in spacers, protrusions on electrodes, moving metallic particles and floating electrodes. Statistical features were obtained from the PRPD pattern and used as input feature to the Bayes classifier combined with genetic algorithm. The amount of PD data obtained is not the same for all test samples. Defects with higher amount of data points manage to achieve classification accuracy of 96% but samples with lower data points has only 20% to 60% accuracy.

SOM was used in (Yu & Song, 2003) to classify PD patterns on stator windings. Four artificial defects were prepared such as normal PD, internal void discharge, slot discharge and end winding discharge. Input features used consist of Weibull, statistical and fractal features. Their investigations showed that adding fractal features does not improve SOM classification accuracy. Weibull features are usable but does not give good accuracy

alone. The best result was achieved by using four Weibull features and eight statistical features.

Hidden Markov Model was used in (T. Abdel-Galil, Hegazy, Salama, & Bartnikas, 2004) for classification of cavity size based on PD pattern. The applied charge transfer characteristic and applied voltage were used as input feature. The classifier used was compared with neural networks and the classification accuracy was slightly higher, 95.5% compared to 93%.

Inductive inference algorithm was used in (T. Abdel-Galil, R. M. Sharkawy, M. M. A. Salama, & R. Bartnikas, 2005) as a classifier for PD patterns. High density polyethylene (HDPE) with different cylindrical cavities size was used as test samples. Pulse shape analysis such as rise time, fall time, width and area were used as input feature for classification purposes. The method has accuracy comparable to neural network. The average accuracy achieved by the inductive inference algorithm is 86.45% for 4 samples tested. In a similar work, (T. K. Abdel-Galil, Y. G. Hegazy, et al., 2005) used the apparent charged and applied voltage as the input feature to the fast match-based vector quantization to classify 3 different void sizes (1.0 mm, 1.5 mm and 2.0 mm). The classification accuracy achieved was 93%.

Adaptive resonance theory was used in (Karthikeyan et al., 2006) to classify PD patterns for void, surface discharge, corona and oil corona. Statistical features were used as input features and the classification accuracy was compared with ANN. A classification accuracy of above 90% was achieved.

Cerebellar model articulation controller (CMAC) was proposed in (H.-C. Chen & Gu, 2012) for PD classification. Epoxy samples were used as test samples with 5 types of lab created defects. Fractal features and mean discharge were used as input features. The

CMAC method was compared with ANN and K-means classifiers. Under noise-free condition, CMAC, ANN and K-means classifiers achieved 96%, 100% and 82% while under 30% noise, their accuracy reduced to 94%, 83.8% and 67.5% respectively. Extension method was used in (H.-C. Chen et al., 2012) to classify the same defect as their previous work using the same input features. Under noise-free condition, extension method, ANN and K-means classifier managed to achieve 100%, 100% and 82%. Under 30% noise condition, the accuracy fell to 93%, 83.8% and 67.5%.

A research work done in (Jian, Tianyan, Harrison, & Grzybowski, 2012) used Linear discriminant analysis for classifying PD patterns measured from 6 artificial PD sources which are corona in oil, surface discharge in oil, gas cavity, oil cavity, floating discharge in oil and paper fiber discharge. Wavelet packet decomposition was used to extract input features. It was proven that linear discriminant analysis was able to achieve classification accuracy comparable to ANN and SVM.

Five optical sensors were used in (Biswas, Dey, Chatterjee, & Chakravorti, 2013) for PD measurement in a custom made measurement setup. An acrylic disk with different void size was used as the test sample. Wavelet transform and inductive inference algorithms were used as feature extraction. A rough set theory based decision support system was used as the classifier and classification accuracy of 92% was achieved.

A research work in (Perpiñán, Sánchez-Urán, Álvarez, Ortego, & Garnacho, 2013) has attempted to cluster PD signals using CLARA (clustering large applications) algorithm. PD signals were measured using acoustic method. A large range of input features were used such as wavelet variances, energy range, maximum location, damping factors and frequency. PD signals were measured from 4 different points of a 1.5km length of XLPE power cable. The CLARA algorithm managed to create 4 clusters with reasonable success.

## 2.8 Chapter Summary

Numerous works have been done related to PD classification. However, there are still room for improvement. Previous works on PD classification were mostly done using lab fabricated material such as polyethylene, acrylic, Perspex and glass. Very few research works were done using actual XLPE cable joints. For research works that used XLPE cable joints as the test subject, actual noise was not used to test the noise contamination effect on the PD classifiers.

Noise contamination is another factor that can be improved. Previous works used software simulated white noise. For example, adding evenly distributed random number to phase and charge of PD data, adding white noise with zero mean and fluctuating power, including random numbers with various standard deviation and zero mean and merging randomly distributed noise that are within 10 % to 30 % of the test data.

ANFIS is a classifier reported to be used once in 2005 for PD classification. Since then, they have been several works which used ANFIS to classify void shapes based on PD patterns and to classify PD patterns caused by void, surface, floating metal and corona discharge. However, all these works only used statistical features and no comparison was done with other types of classifier. Apart of that, only lab fabricated materials were used as test samples instead of actual power system components. Therefore, there is no record of how capable ANFIS is compared to tried and tested classifiers such as ANN and SVM.

To address shortcomings of the current issues, this work used actual XLPE cable joints and actual noise caused by ground interference. A better analysis of how noise levels affect the classification accuracy can be done by using actual noise with varying pulse counts and charge magnitudes to contaminate the test data prior to classification. A novel PCA based feature extraction, which has high noise resistant was tested against widely

used methods, such as statistical features and fractal features. The feasibility of ANFIS as a classifier was also tested by comparing the classification accuracy against ANN and SVM.

# CHAPTER 3: METHODOLOGY

## 3.1 Introduction

This chapter describes the methodology that has been used in the work. The cable joint samples with artificial defect preparation, PD measurement setup and implementation of PD feature extraction, PD classification and the effect of noise contamination are described in detail. A general methodology flow chart of PD classification used in this work is shown in Figure 3.1. The cable joints were prepared and PD measurement was performed. Feature extraction was used to obtain input feature to train the classifier. The training process was repeated until good accuracy was achieved.

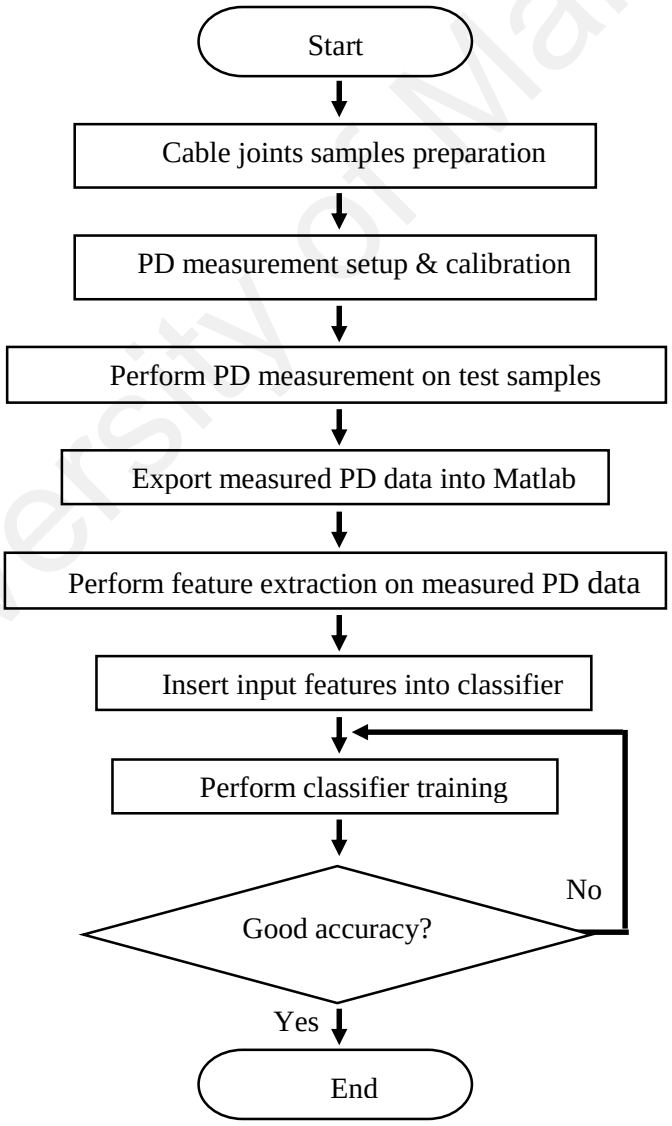



**Figure 3.1:** Methodology flow chart

Although the rated operating voltage for the XLPE cables is 11 kV, measurement was conducted at 9 kV. This is because the cable joints used contain artificially created defects. Using higher applied voltage will greatly increase the chance of insulation breakdown, which will permanently damage the cable, rendering it unusable. Each cable was energized at 9 kV and allowed to stabilize for 1 hour before PD measurement was taken. Each PD measurement has a duration of 60 seconds and the time gap between each measurement was 15 seconds. A total of 100 measurements were done on each sample. This results in a grand total of 500 PD patterns captured for all 5 test samples.

## 3.2 Cable Joint Samples Preparation

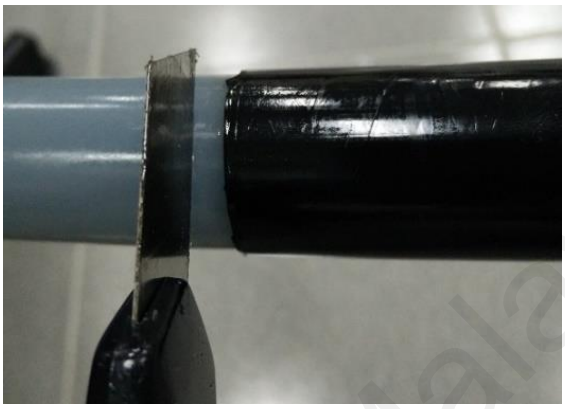
Six 11 kV XLPE cables of 3 meter length with a cable joint in the middle were prepared in this work. Five of them contain artificially created defect and one without any defect at the cable joint. The good cable joint was used as a benchmark for checking noise contamination. The cable joint defects created were based on commonly encountered defects in power industries (Mazzetti et al., 2006; Yun et al., 2010). The list of cable joint samples is shown in Table 3.1.

**Table 3.1:** Cable joint samples that have been prepared

| Cable Joint | Defect type |
|---|---|
| C1 | Insulation incision defect |
| C2 | Axial direction shift defect |
| C3 | Semiconductor layer tip defect |
| C4 | Metal particle on XLPE defect |
| C5 | Semiconductor layer air gap defect |
| C6 | Without defect |

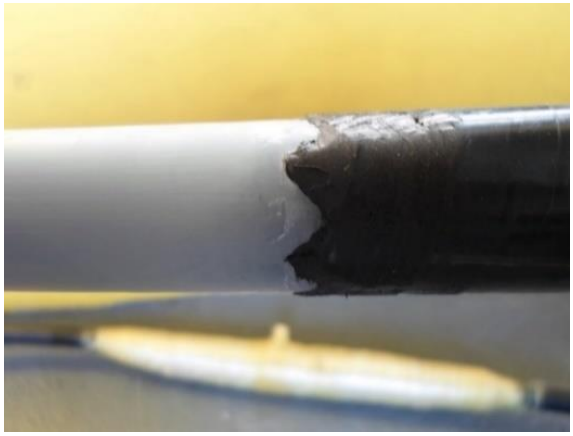
Insulation incision defect was created by making a shallow cut at the XLPE layer using a sharp knife. Axial direction shift defect was made by inserting the cable at an angle which was off centre. Semiconductor layer tip defect was created by creating rough edges at the semiconductor tip. Metal particle on XLPE defect was made by sprinkling metal

particles on the XLPE layer. Finally, semiconductor layer air gap defect was made by using insulation tape to wrap a layer of air around the semiconductor edge. All defects were made at the XLPE cable ends before the cable joints were installed. Pictures of the defects are shown in Figure 3.2.



(a)

(b)



(c)

(d)

(e)

**Figure 3.2:** Defects created; (a) Insulation incision defect, (b) Axial direction shift defect, (c) Semiconductor layer tip defect, (d) Metal particle on XLPE defect and (e) Semiconductor layer air gap defect

## 3.3 Experimental Setup

### 3.3.1 PD Measurement Equipment

In this research work, the Mtronix MPD 600 manufactured by Omicron was used for PD measurement. The block diagram of the PD measurement system is shown in Figure 3.3. The test setup consists of a PD-free step-up transformer, which acts as a HV source, measuring capacitor, which measures the operating voltage, a test object, a coupling capacitor and a coupling device, which acts as an equivalent RLC circuit, a USB controller and a PD detector connected to a PC. The PC was used to configure the PD detector settings and save the measured data.

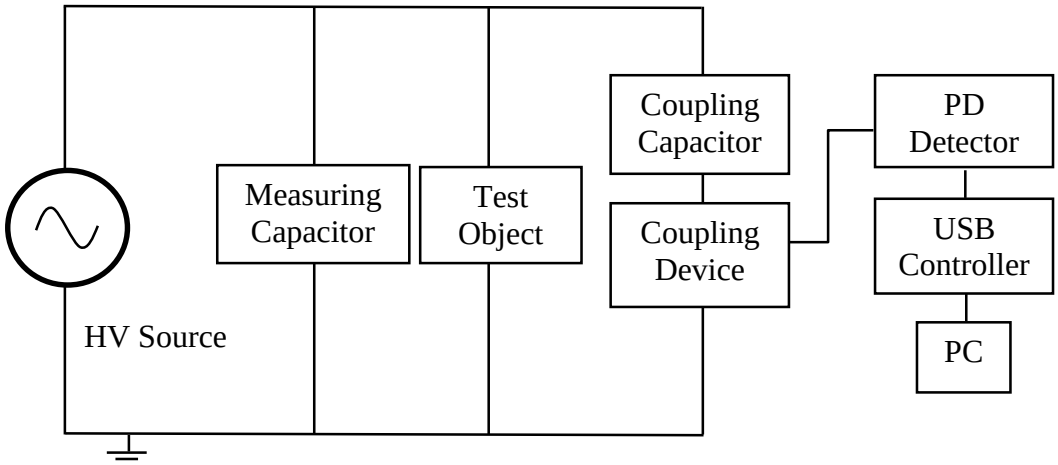



**Figure 3.3:** Setup of PD measurement under AC voltage

The Mtronix MPD 600 system is a digital system, which is excellent for both on-site and laboratory PD measurement. It has an operating frequency range of 0.1 Hz to 2.5 Hz. It supports the detection of PD signal with center frequency from 0 up to 32 MHz with a time resolution of less than 2 ns. These specifications enable it to have highly precise pulse detection.

The MPD 600 measures PD by measuring the change in current pulses between the two terminals of the test object. When an electrical discharge occurs in the test object, the coupling capacitor will transfer charges to the test object to stabilize the voltage drop across the test object. This will create a current pulse in a nanosecond range to flow in the circuit, which leads to the generation of a voltage pulse across the coupling device. The total number of charged transferred is the apparent charge. This charge magnitude is influenced by the total number of induced dipole moments of the real charge created by the abrupt variation of the capacitance of the test subject and their interactions with the system's electrodes (Boggs, 1990). In the event of PD occurrence, the measuring impedance or the coupling device will detect the short duration voltage pulse.

The USB controller handles data transfer between the PC and PD detector. The PD detector sends the measurement results to the USB controller through fiber optic cables.

The benefit of using fiber optic cables is that they provide absolute galvanic isolation between the USB controller and PD detector. The exclusion of ground loops lessens interference and boost signal to noise ratio, which will improve the system sensitivity. Once the data was received by the USB controller, it transfers the data to the laptop via USB cable. The user is then able to analyze and store the measurement data.

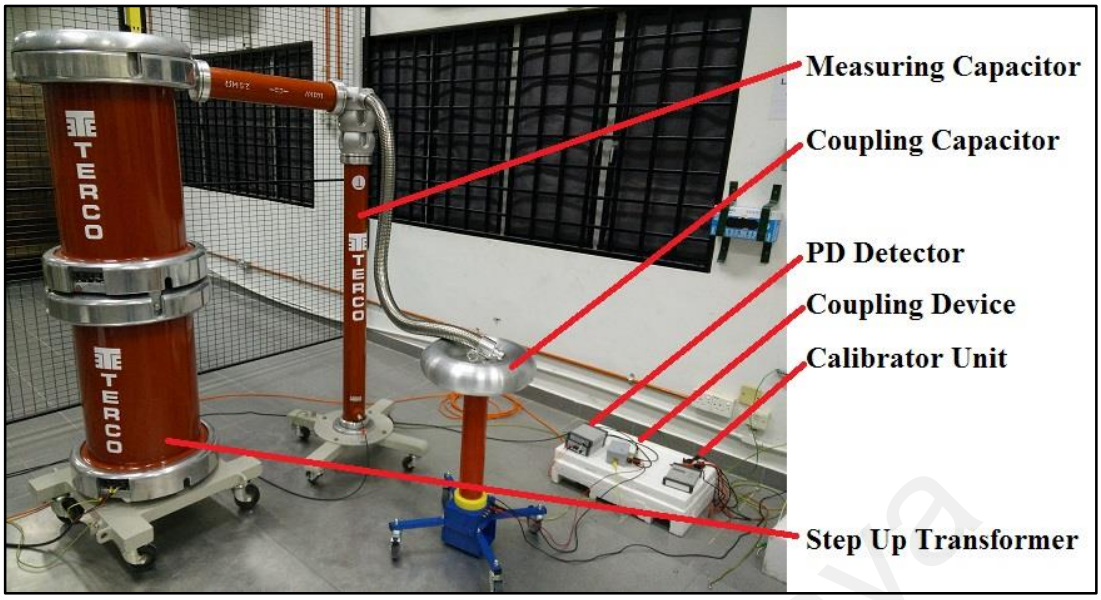
### 3.3.2 PD Detector Calibration

Calibration was required each time the system was rebooted. The MPD 600 system can be calibrated digitally using the CAL542 charge calibrator unit, which has a charge range of 1 to 100 pC, where 1 mV equals to 1 pC. Calibration was done by connecting the CAL542 to the circuit. The CAL542 will inject a specific charge value into the electrodes to be detected by the PD detector. From the Mtronix software graphical user interface (GUI), the same value was set as the target and the software will compute and perform the calibration. Once completed, the calibrator was removed from the system.

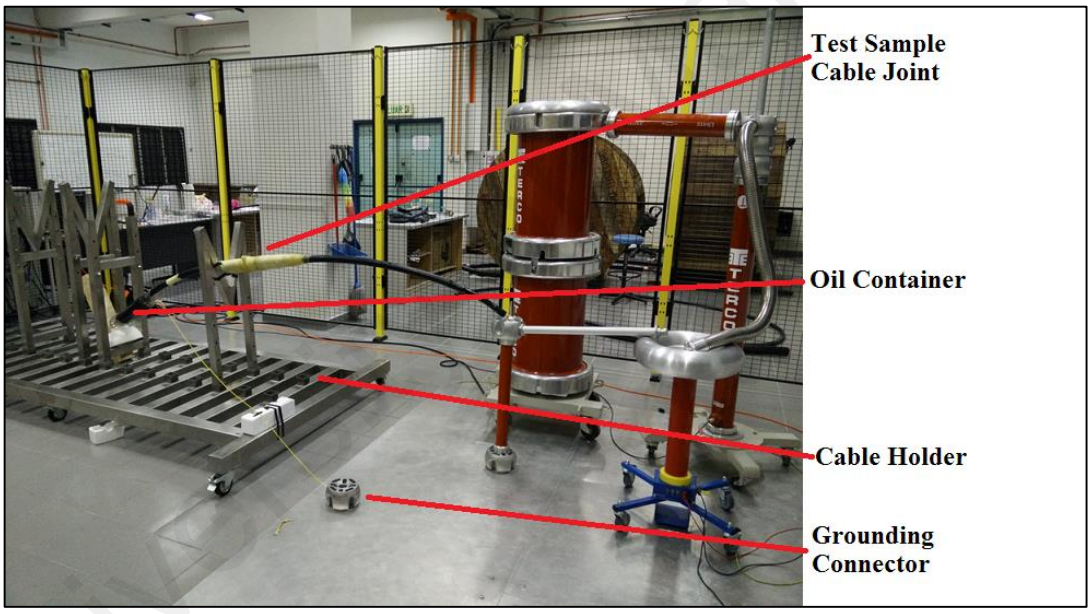
The equipment also allows calibration for the applied voltage amplitude. This was done similarly to the charge calibration. A known voltage amplitude was applied and the voltage value was set via the Mtronix software GUI. The high voltage source can be turned off after the calibration to make sure the setup was PD from the voltage source.

### 3.3.3 PD Measurement Hardware Setup

The actual PD measurement setup and the test sample setup are shown in Figure 3.4. Ideally, the PD detector should only detect PD activity from the test object. However, in real scenario, it is common for it to detect surface discharges from the connectors and grounding interference. The grounding cable was especially sensitive and must be carefully isolated from other conducting materials. This will cause it to detect random noisy signals even without any applied external voltage.




(a)




(b)

**Figure 3.4:** (a) PD measurement setup and (b) Test sample connection

An example of noisy PD signals detected when the ground wiring was not isolated properly can be seen in Figure 3.5. To get rid of these unwanted noises, careful handling of the grounding wire and other precaution was taken. Silicon rubber was used as an insulating material to prevent the grounding wire from touching other conductive materials. The cable holder was laced with thick insulator to support the cable joint. Both ends of the termination were covered with insulating tape. The cable end which contains

the earthing braid was soaked in silicon oil suspended in an oil container. Improvements made to the test object setup can be seen in Figures 3.6.

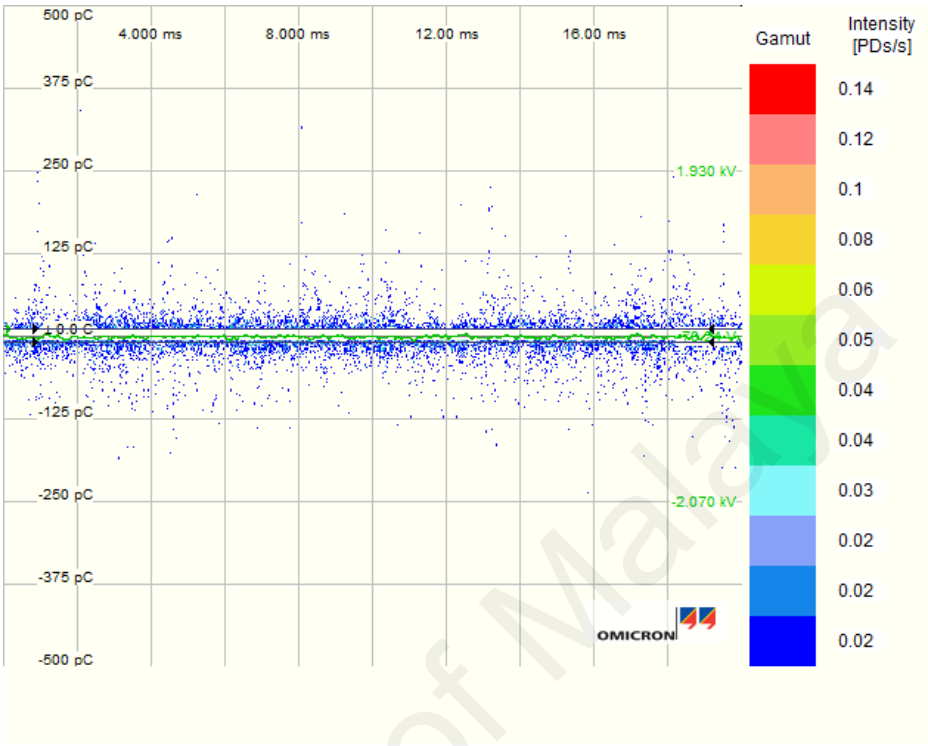
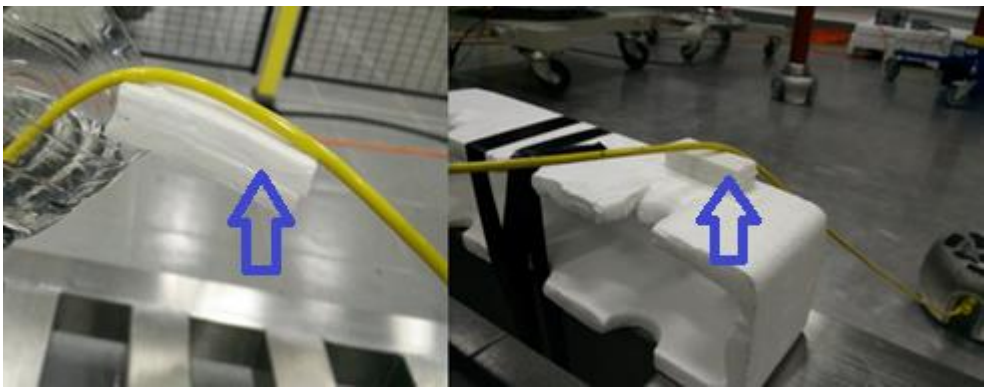
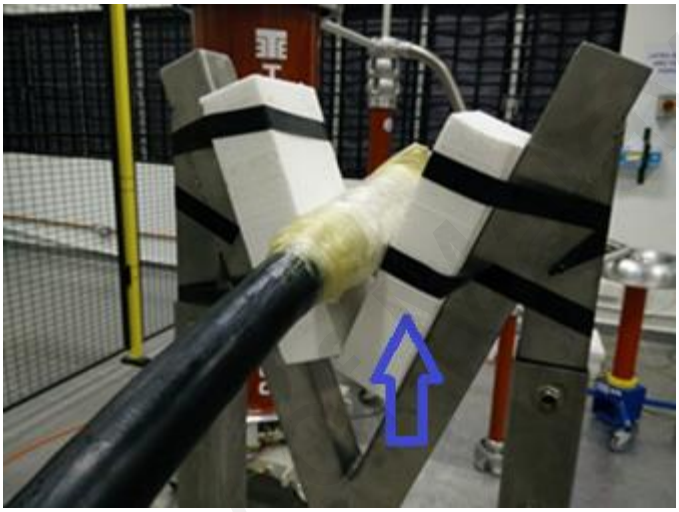



**Figure 3.5:** Noise detected at 0 kV

With these steps taken, no signal was detected at 0 kV with no interference from the connectors as shown in Figure 3.7. When the good cable joint was connected and energized, it remained PD free up to 9 kV. At 9 kV, the average PD magnitude detected was less than 10 pC, which was very small. This shows that the entire measurement setup system was properly isolated and not contributing any unwanted noise.



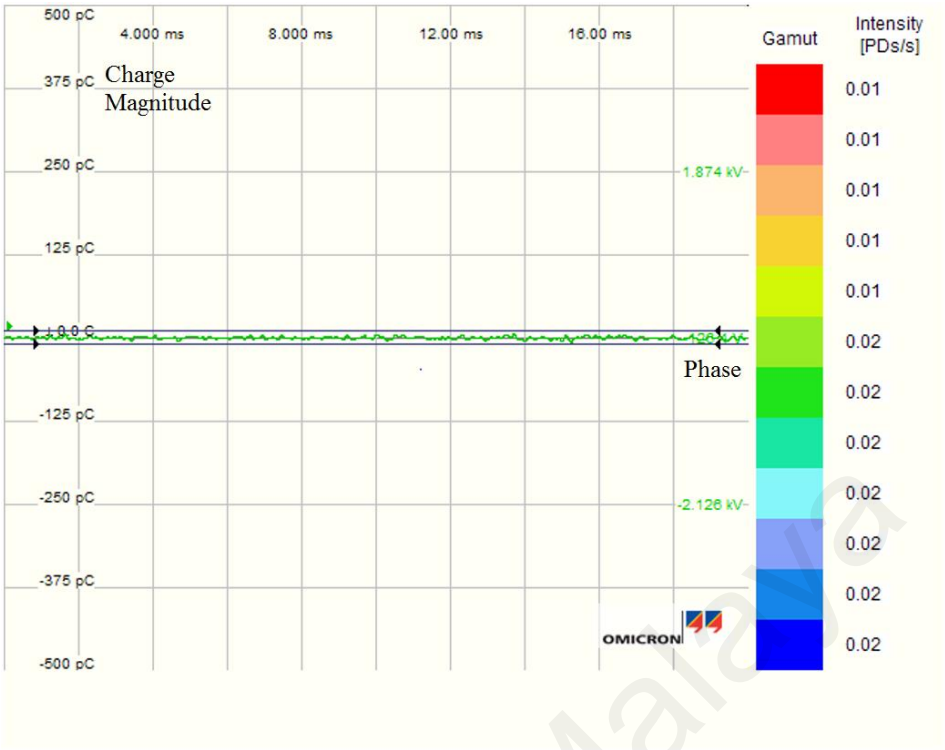
(a)



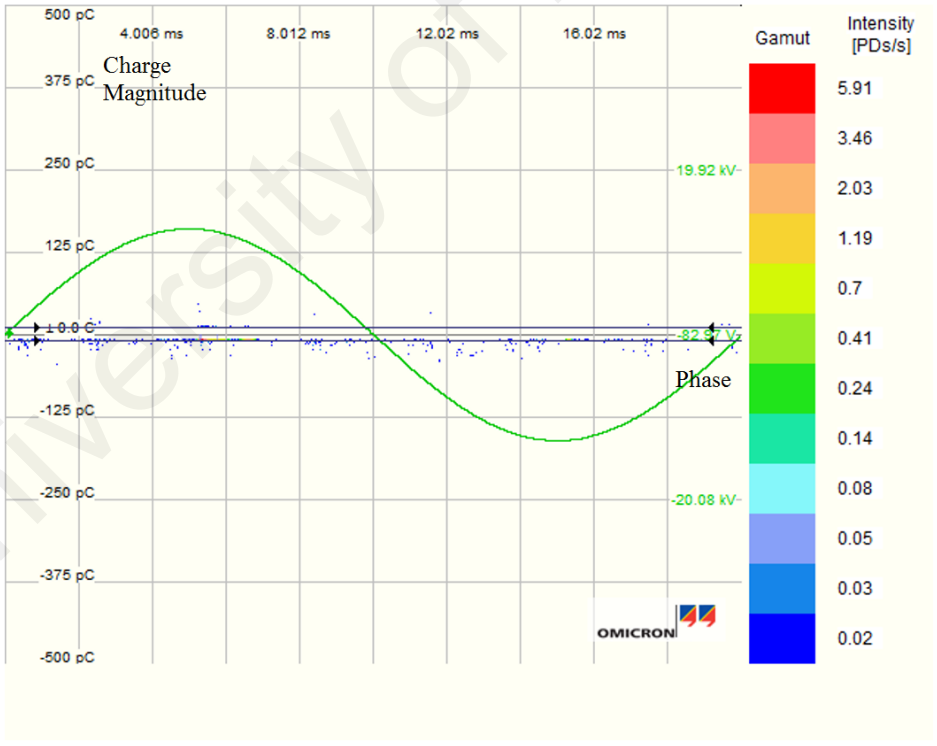
(b)

(c)

**Figure 3.6:** Improvement made; (a) Silicon rubber supporting grounding wire, (b) Cable holder laced with thick insulator and (c) Earthing braid soaked in silicon oil




(a)




(b)

**Figure 3.7:** PD reading of good cable joint at (a) 0 kV and (b) 9 kV

### 3.3.4 PD Measurement Software Setup

The Mtronix software GUI was used to control the PD detector settings and measurement parameters. As mentioned in previous section, charge and voltage calibration were performed using this software. For this work, the calibration value used was 100 pC. All settings used can be seen in the screenshot of the software GUI as shown in Figure 3.8.

One of the most important functions of this software is to export recorded PD data. There are three ways the software can export the measured data. First, it can save the current PRPD pattern and pulse shape as a png image. Second, it can export the last PD pulse signal into a text file containing the time and voltage level of the last recorded pulse. The last method is to export the measured phase, voltage and charge data as binary files. These binary files were imported into Matlab workspace using a Matlab code.

## 3.4 Noise Signal Acquisition

After calibration, the PD detector would not detect any reading when no voltage source was applied, hence the PRPD pattern display would be a complete horizontal line. However, in the event of rain, the PD detector will detect noisy signals even though when no voltage was applied. Since no changes have been made to the test object and equipment, these noisy signals must be originated from ground interference. These noisy signals were recorded and used to test the noise tolerance of each PD classifiers and input features. The noisy signals were stored in PRPD format with phase and charge information. In order to contaminate a sample of PD data, the phase and charge information of the noisy signals and test sample of PD data were combined to form a new contaminated PD data. Figure 3.9 shows how noise signals were overlapped onto PD data to contaminate it. Feature extractions were performed on the contaminated PD data for noise analysis.

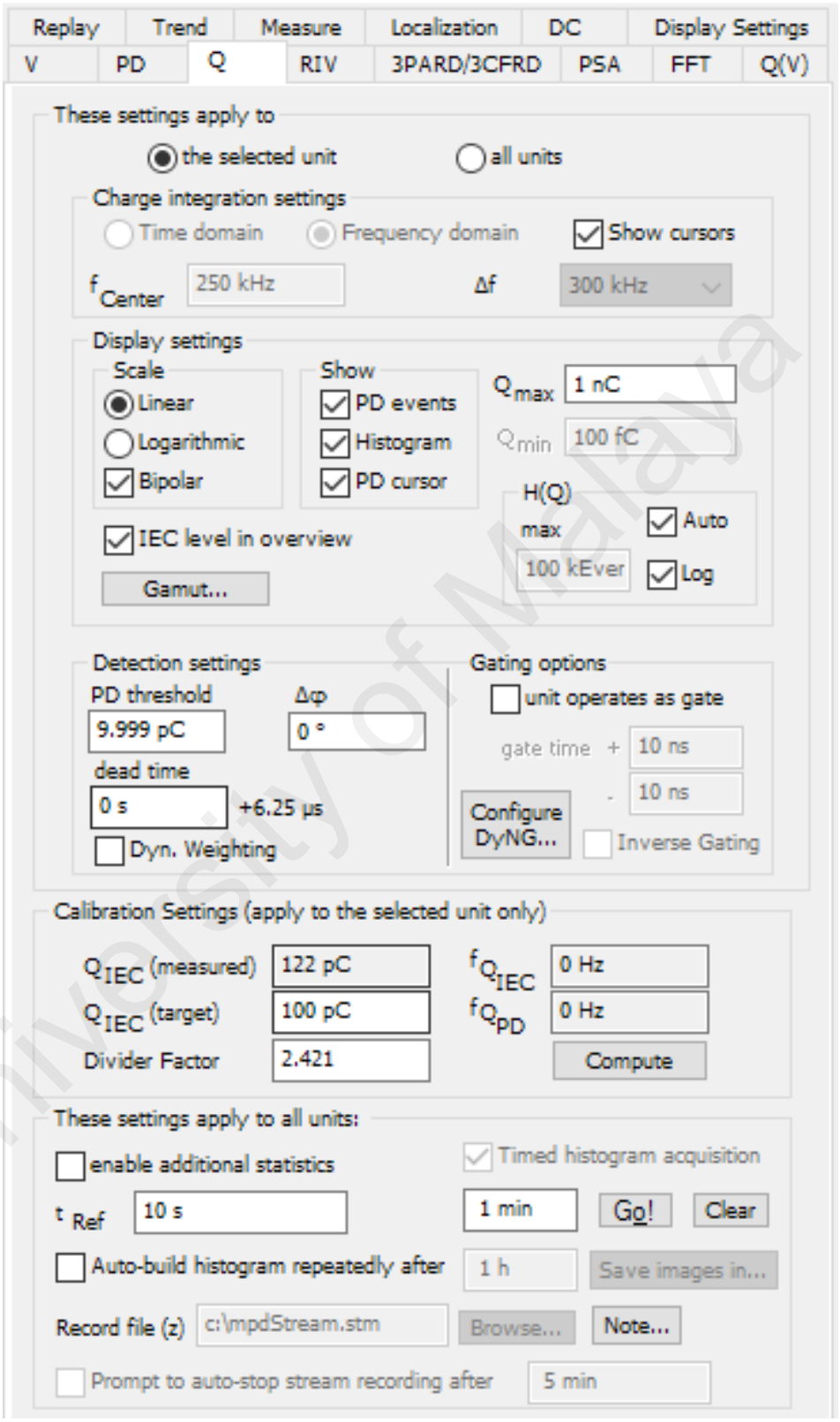
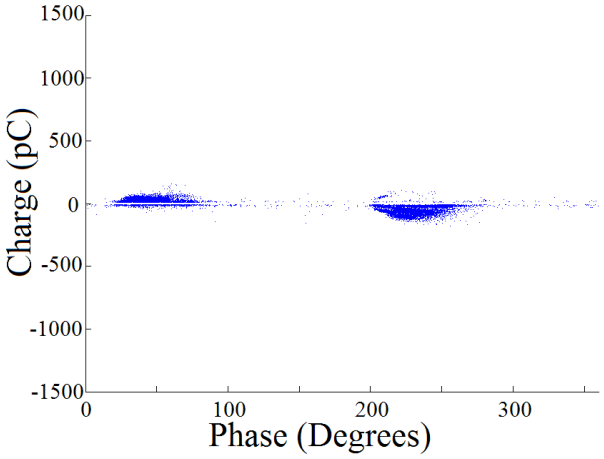
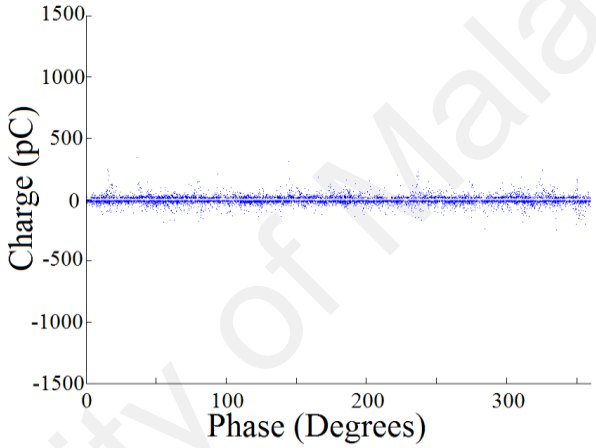


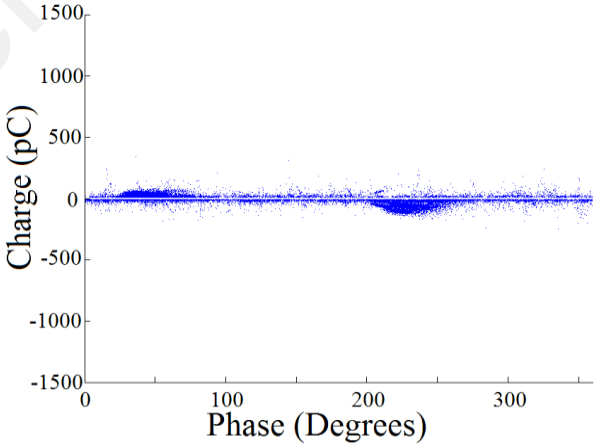

**Figure 3.8:** Mtronix software GUI




(a)




(b)




(c)

**Figure 3.9:** (a) Uncontaminated PD pattern, (b) Noise PD pattern and (c) Contaminated PD pattern

## 3.5 PD Feature Extraction Methods

The purpose of feature extraction is to obtain relevant input features from PD data to represent PD characteristics associated with a particu1ar defect. Apart from collecting relevant input features, feature extraction also tries to decrease the dimension of the original data for easier processing (Kranz, 1993; Yu & Song, 2003). Raw data contained in each of the stored patterns might be too enormous for direct handling. A form of data reduction such as reducing the matrix size of the data is normally required (Satish & Zaengl, 1994). This section explains the extracted input features used in this work, which includes statistical features, fractal features and the proposed PCA features.

### 3.5.1 Statistical features

Statistical parameters were first used in (Gulski, 1993) and consist of skewness, kurtosis, mean and standard deviation. These are proven methods that had been used by many researchers ever since.

The mean pulse height distribution, $H_{qn}(\varphi)$ is used to represent the average PD magnitude vs. the phase angle $\varphi$. The number of PD vs. phase angle $\varphi$ is represented by the pulse count distribution, $H_n(\varphi)$. Both of them can be split into two distributions, which are the positive and negative half cycle, the $H_{qn}^{+}(\varphi)$ and $H_n^{+}(\varphi)$ (from the positive half cycle of the voltage) and $H_{qn}^{-}(\varphi)$ and $H_n^{-}(\varphi)$(from the negative half cycle of the voltage). Statistical parameters were calculated from these distributions.

Skewness describes the asymmetry of the distribution with respect to normal distribution. Positive skewness represents asymmetric with larger left side, zero skewness represents symmetric and negative skewness shows asymmetric with larger right side (James & Phung, 1995).

Kurtosis describes the sharpness of the distribution with respect to a normal distribution. Zero kurtosis represents a normal distribution, positive kurtosis represents a sharp distribution and negative kurtosis represents a flat distribution (Gulski & Krivda, 1993).

Variance describes how much a group of numbers is spread out. Zero variance means all values are identical. The standard deviation is the square root of the variance. A very detailed mathematical description of skewness, kurtosis and cross correlation can be found in (F. H. Kreuger et al., 1993). The formulas of all mentioned statistical parameters are shown as follows,

$$Average\ (mean)\ \mu = \frac{\sum_{i=1}^{N} x_i f(x_i)}{\sum_{i=1}^{N} f(x_i)} \tag{1}$$

$$Variance: \sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2 f(x_i)}{\sum_{i=1}^{N} f(x_i)} \tag{2}$$

$$Skewness: S_k = \frac{\sum_{i=1}^{N} (x_i - \mu)^3 f(x_i)}{\sigma^3 \sum_{i=1}^{N} f(x_i)} \tag{3}$$

$$Kurtosis: K_u = \frac{\sum_{i=1}^{N} (x_i - \mu)^4 f(x_i)}{\sigma^4 \sum_{i=1}^{N} f(x_i)} - 3 \tag{4}$$

where $f(x_i)$ is the function of interest, $N$ is the size of the data and $x_i$ is discrete values of the distribution.

Weibull analysis offers a mathematical approach for the pulse height analysis patterns. The probability distribution of PD pulse rate ($F$) versus PD magnitude ($q$) can be expressed by the Weibull function (Contin et al., 2000; Yu & Song, 2003),

$$F(q; \alpha; \beta) = 1 - exp\left[-\left(\frac{q}{\alpha}\right)^{\beta}\right] \tag{5}$$

where $\alpha$ is the scale parameter and $\beta$ is the shape parameter. Each pulse height analysis curve is represented by the features $\alpha$ and $\beta$ and the PD pulse amplitude is represented by $q$. The features $\alpha+$, $\beta+$, $\alpha-$ and $\beta-$ are extracted from the negative and positive pulse height analysis curves. The pulse height analysis pattern is compressed using the Weibull method for digital analysis while retaining its relevant information. The values of $\alpha+$, $\beta+$, $\alpha-$, and $\beta-$ are then used as input to the classifier.

A Matlab code was written to import PD data and perform the calculations required to compute and save the statistical parameters. The flowchart of the Matlab code is shown in Figure 3.10.

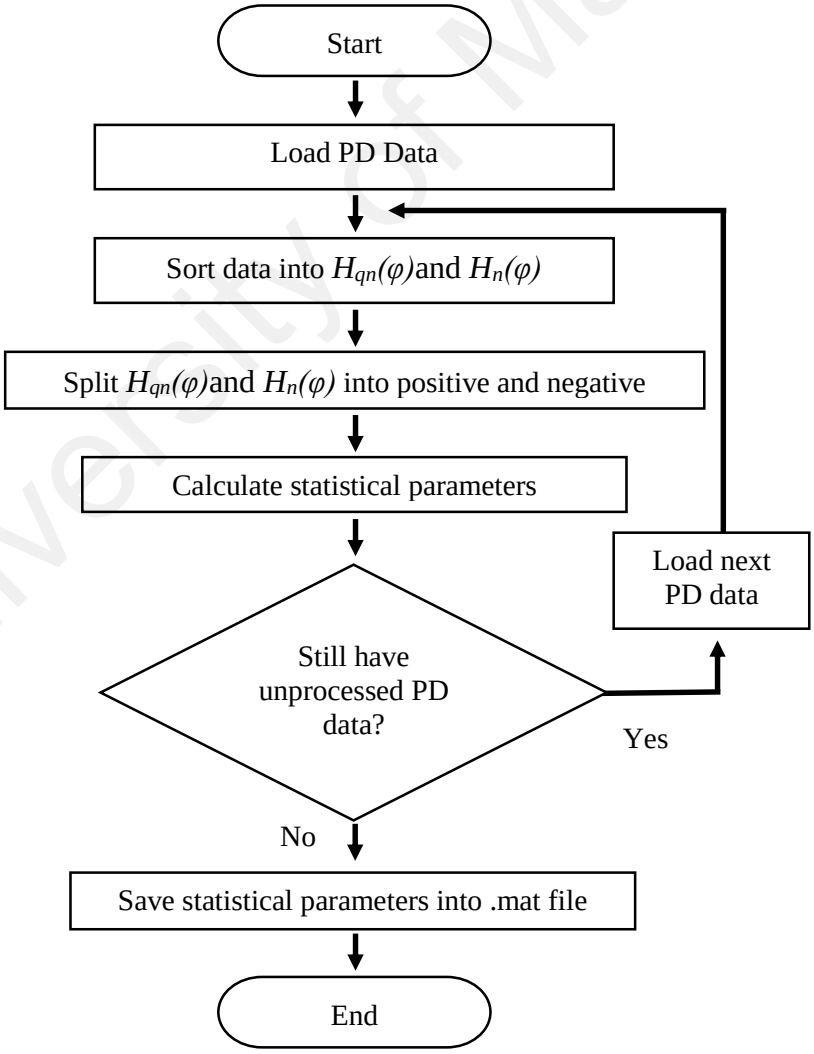



**Figure 3.10:** Flowchart of statistical parameters feature extraction

### 3.5.2 Fractal Features

Fractals are good for modeling complex shapes and natural phenomena, where existing mathematical techniques are found to be insufficient. Since PD can be considered a natural phenomenon that has complex surfaces and shapes, fractals can be used to model it. The usage of fractal features in PD classification is fascinating because it represents PRPD patterns directly (Krivda et al., 1995).

PRPD patterns can be represented using two fractal features, lacunarity ($\Lambda$) and fractal dimension ($D$) that are extracted using box counting method. Since $D$ is invariant to variation in scale, it can be utilized to gauge the roughness of the surface. $\Lambda$ is the denseness of the fractal surface. Both $D$ and $\Lambda$ are functions of the box size $L$. The number of boxes $N$, of side $L$ needed to cover a fractal set is shown in Equation (6), where $D$ is the fractal dimension set and $K$ is a constant (Satish & Zaengl, 1995):

$$N(L) = KL^{-D} \tag{6}$$

The lacunarity $\Lambda(L)$ relies on the second order statistics of $p(m,L)$. It can be defined after calculating $M(L)$ and $M^2(L)$. The formulas of $\Lambda(L)$, $M(L)$ and $M^2(L)$ are shown in Equations (7) to (9). Mathematical derivation of $D$ and $\Lambda$ can be found in (Kundu et al., 2012)

$$\Lambda(L) = \frac{M^2(L) - [M(L)]^2}{[M(L)]^2} \tag{7}$$

$$M(L) = \sum_{m=1}^{N} mp(m, L) \tag{8}$$

$$M^2(L) = \sum_{m=1}^{N} m^2 p(m, L) \tag{9}$$

In order to obtain the fractal dimension and lacunarity from the PD data to be used as input feature, a PRPD scatter plot has to be plot from the PD data. Matlab was used to

load the PD data, plot the scatter plot and saved it as a JPEG picture. ImageJ software was used to convert the JPEG picture into binary image. The fractal dimension and lacunarity was then computed from the binary image and saved into an Excel file. Matlab was then used to load the fractal dimension and lacunarity data from the Excel and stored it as .mat file to be used in the classifier.

ImageJ is a powerful image processing software with many advance features. For this work, it was used to convert scatter plot images into binary image (black and white) and compute the fractal dimension and lacunarity using the FracLac extension. A screenshot of ImageJ software with FracLac extension is shown in Figure 3.11. The conversion from JPEG to binary is shown in Figure 3.12. The settings used for the FracLac extension is shown in Figure 3.13 while a flowchart of this entire process is shown in Figure 3.14.

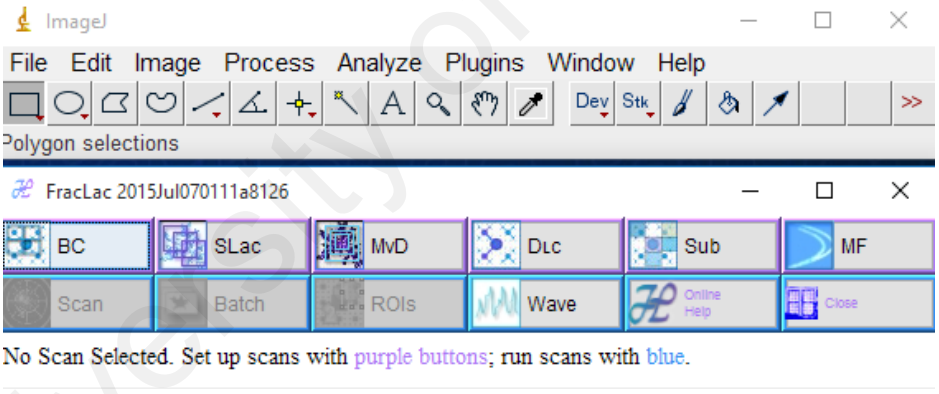
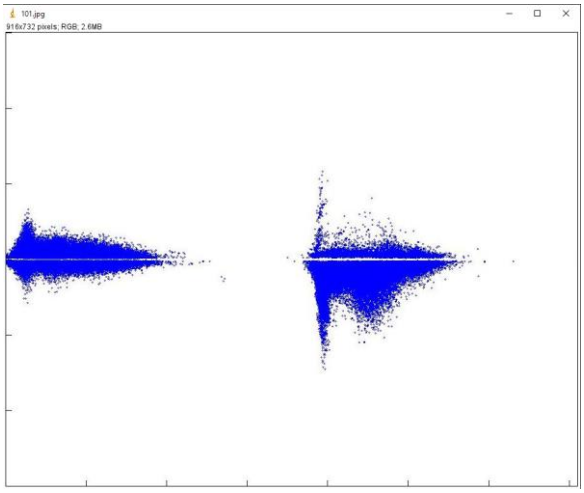
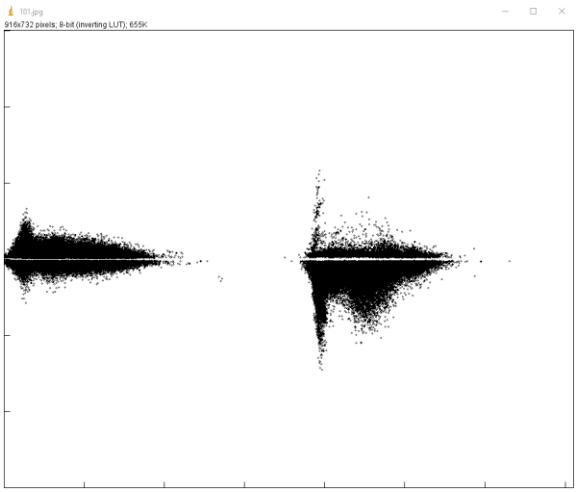



**Figure 3.11:** ImageJ software with FracLac extension



(a)



(b)

**Figure 3.12:** Conversion from (a) JPEG to (b) binary image

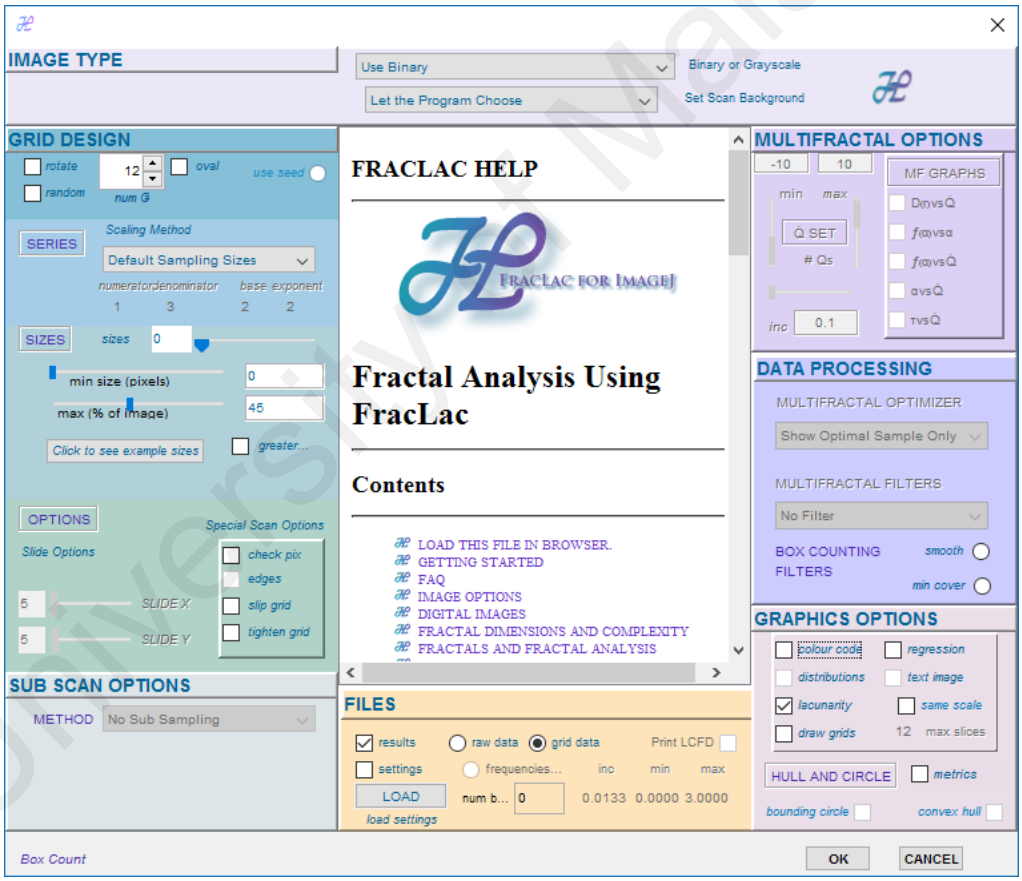



**Figure 3.13:** Settings used for the FracLac extension

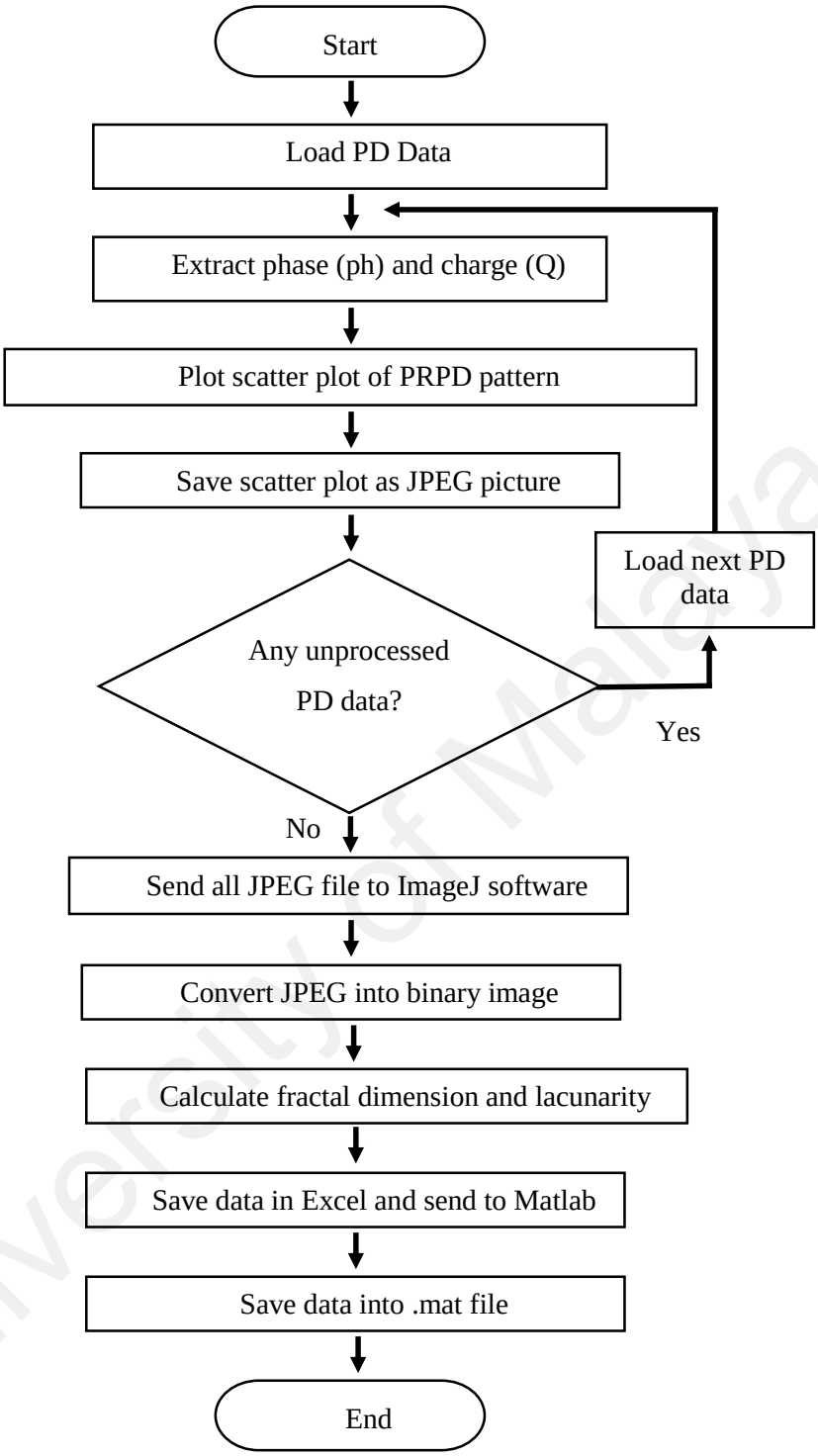



**Figure 3.14:** Flowchart of fractal dimension and lacunarity feature extraction

### 3.5.3 Principal Component Analysis (PCA) Features

PCA is a data reduction method which reserves data information in the reduced space with the least information loss. This is done by projecting data at a direction with the biggest variance at the lower dimension that will maximize the scatter of the projected

samples (Harbaji et al., 2015). This linear subspace is found by solving this Eigen problem,

$$cov(X)M=\lambda M \tag{10}$$

where *cov(X)* is the covariance matrix of the dataset *X*, *M* is a linear mapping created by the *d* principle eigenvectors of the covariance matrix and $\lambda$ is the *d* principal eigenvalues (Harbaji et al., 2015; Jing & Hou, 2015). This covariance matrix is able to determine which direction contains the most significant variance in the dataset, making PCA an effective tool for feature subset selection. The low-dimensional data $y_i$ of the data points $x_i$ are calculated using linear mapping $Y=XM$. The elements of *Y* will produce the feature sets (Hui et al., 2013). The best number of principal components to represent the data can be found using a scree plot. Scree plot is a graph of the eigenvalue magnitude vs. its number. The best number is chosen at a point where the graph has a sudden change in slope where the slope on its left side is much higher than the right side (Babnik, Aggarwal, & Moore, 2007). Each PCA component has a different weighting, thus normalization is not recommended to be done on data reduced using PCA. Normalization will alter the relative significance between the components leading to high error (Lai et al., 2010).

PCA was used as a data reduction method in previous PD classification works (Rahman et al., 2000). It was used to reduce the dimension of Fourier transformed signals (Babnik et al., 2007), simplify the acoustic emission signals (Harbaji et al., 2015) and reducing the dimension of energy levels for each wavelet decomposition levels (L. Hao et al., 2011). However, in this work, it is proposed that PCA is performed directly on the PRPD pattern. The original PRPD pattern was split twice into 6 groups. Firstly, the original PRPD pattern was split into 2 groups of positive and negative section of the charge magnitude. Secondly, the original PRPD pattern was split into 4 groups, which are the four phase quadrants (90 degrees for each quadrant). PCA was performed separately

on these 6 groups of data and the calculated PCA components were combined to form the input feature for the classifier.

Matlab was used to import PD data and split them into the 6 groups as mentioned above and perform the calculations required to compute the PCA features. The flowchart of the Matlab code is shown in Figure 3.15.

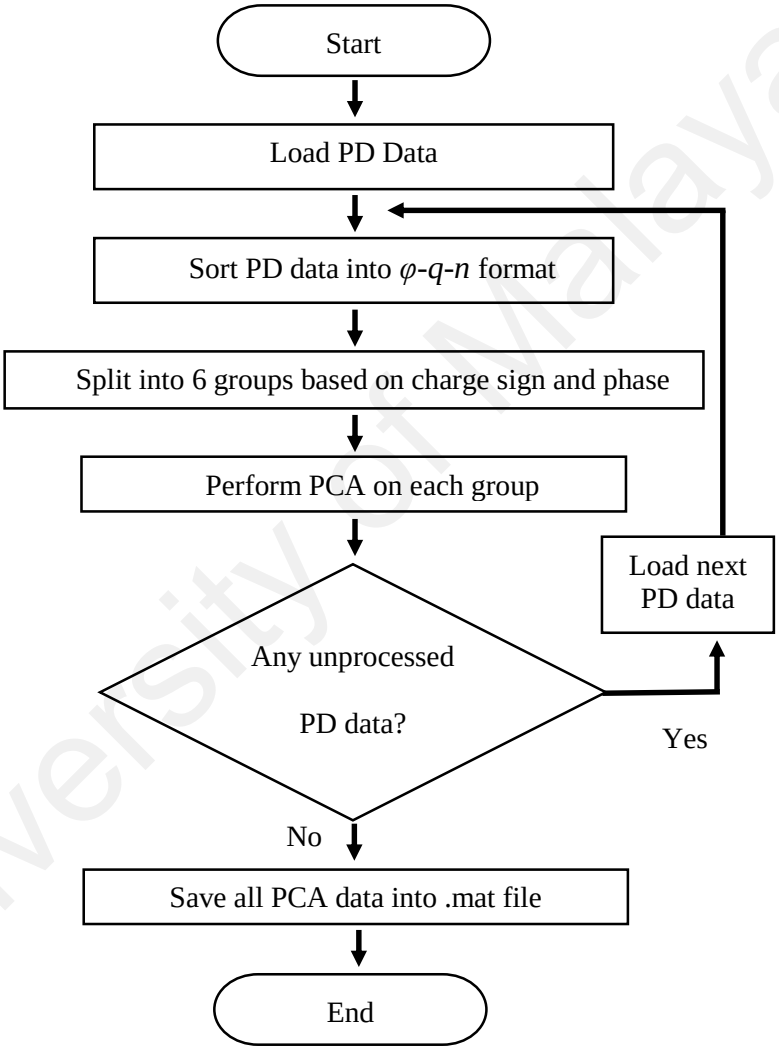



**Figure 3.15:** Flowchart of PCA feature extraction

After importing the phase (ph) and charge (Q) data, the number of times each pair repeats itself was counted and saved as a variable "n". The data are sorted into *φ-q-n* format, a three column wide Matlab array. This array was split into 6 groups according to the charge sign and phase quadrants. PCA was performed on these *φ-q-n* array and the

results were stored into a .mat file. The Matlab PCA function will return 2 variables, the score and the latent. Principal component scores are the representations of the input function in the principal component space. Rows of score correspond to observations and columns correspond to principal components. Latent is the vector containing the eigenvalues of the covariance matrix of X.

## 3.6 PD Classifiers

### 3.6.1 Artificial Neural Networks (ANN)

ANN is suitable for PD classification because it is insensitive to small input changes. ANN has the ability to continue making correct decisions even when the input presented is slightly different from the input used during training process. This is very important for PD classification where the discharge patterns are usually not the same (Mazroua et al., 1993).

The feed forward back propagation neural network (BPNN) is the most commonly used learning mode in ANN. It is a supervised learning network that is trained in a forward backward process. BPNN consists of one layer of input, a minimum of one hidden layer and one layer of output. It has been shown that with two hidden layers, any complex decision region can be generated (Satish & Zaengl, 1994). A typical model of BPNN is shown in Figure 3.16. Every layer is completely joined to the next layer. The primary function of the hidden layer is to obtain PD features from different sources and pass the information to the output layer. The amount of processing elements in the input layer relies on the amount of PD fingerprint data. The amount of processing element within the output layer is dependent on the number of defects to be classified (Gulski & Krivda, 1993). Details about the mathematical models and learning algorithm can be found in (Al-geelani et al., 2015; Gulski & Krivda, 1993; M. Majidi et al., 2015; Mazroua et al., 1995; Mazroua et al., 1993; Suzuki & Endoh, 1992). For PD classification purposes, at least

two input features are required to ensure convergence of the BPNN during training (Jin et al., 2006).

A three layer feed forward ANN was created in Matlab using the "patternnet" command. This ANN has 15 neurons in the hidden layer. This value was chosen as it gives the best accuracy with lowest training time. The "train" command was used to train the ANN. A block diagram of the created ANN can be seen in Figure 3.17.

In the previous section, it was mentioned that all input features were saved into a *.mat file. In order to use this input feature, the *.mat file was loaded into the ANN to be used as the input. The output was defined manually in binary form where each class has a value 1 at different columns. There were 5 different defects in this work, therefore 5 different groups of output were defined. Class 1 was defined as 00001, Class 2 as 00010, Class 3 as 00100, Class 4 as 01000 and lastly Class 5 as 10000.

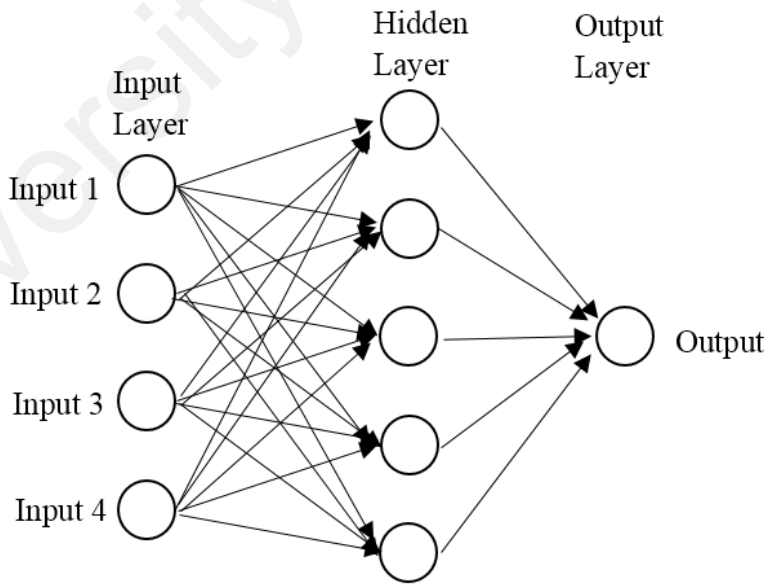



**Figure 3.16:** Typical model of BPNN with one hidden layer

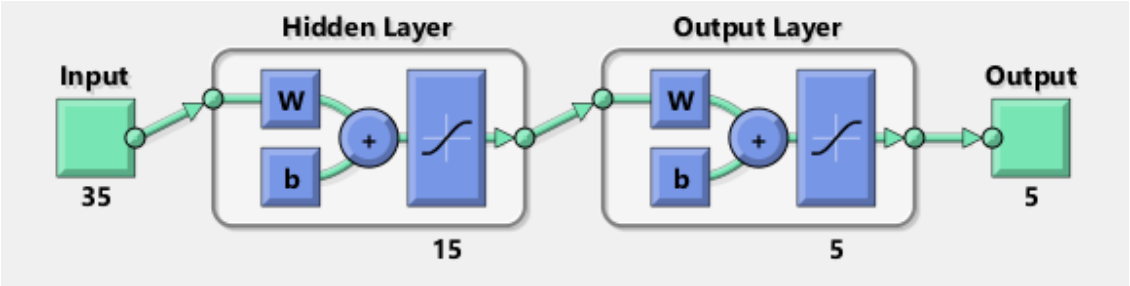



**Figure 3.17:** Multilayer feed forward ANN structure

### 3.6.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS uses both fuzzy system and neural network to determine the optimum fuzzy parameters. Neural network usage eradicates the need to choose the fuzzy parameters manually since it will be optimized by the neural network. The fuzzy system needs to be built by fuzzy logic before ANFIS is used to train the fuzzy scheme. ANFIS is a good tool to map PD input features to the defect class using If-Then rules created from the decision tree and the stipulated input output data matrices (T. K. Abdel-Galil, R. M. Sharkawy, et al., 2005). ANFIS is based on a fuzzy Sugeno model, where

Rule 1: If $x$ is $A_1$ and $y$ is $B_1$, then $f_1 = p_1x + q_1y + r_1$,

Rule 2: If $x$ is $A_2$ and $y$ is $B_2$, then $f_2 = p_2x + q_2y + r_2$.

The ANFIS architecture has five main layers (Sinaga et al., 2010). In the first layer, all nodes are adaptive nodes. The outputs of the first layer are the fuzzy membership grade of the inputs. The second layer contains fixed nodes that acts as the incoming signal multiplier. The output of this layer is the rules firing strength. The third layer also contains fixed nodes which focuses on normalizing the previous layer triggering strength. The forth layer contains adaptive nodes, which output is the product of the first order polynomial and the normalized firing strength. The firth layer contains one fixed node which sums all incoming signals. A general structure of ANFIS is shown in Figure 3.18

Rules fuzzification is performed by assigning fuzzy membership function (MF) to each condition in the premise part of the rules. Every input variable was normalized between zero and one in order to improve the training efficiency (Jang, 1994). Using these fuzzy rules, ANFIS was used to train, test and analyze the Sugeno type fuzzy inference system (Jang, 1993b). Each output rule acts as a linear combination of input variables and a constant. The final output will be each rule's output weighted average. These weights are automatically adjusted using the knowledge obtained during the training process.

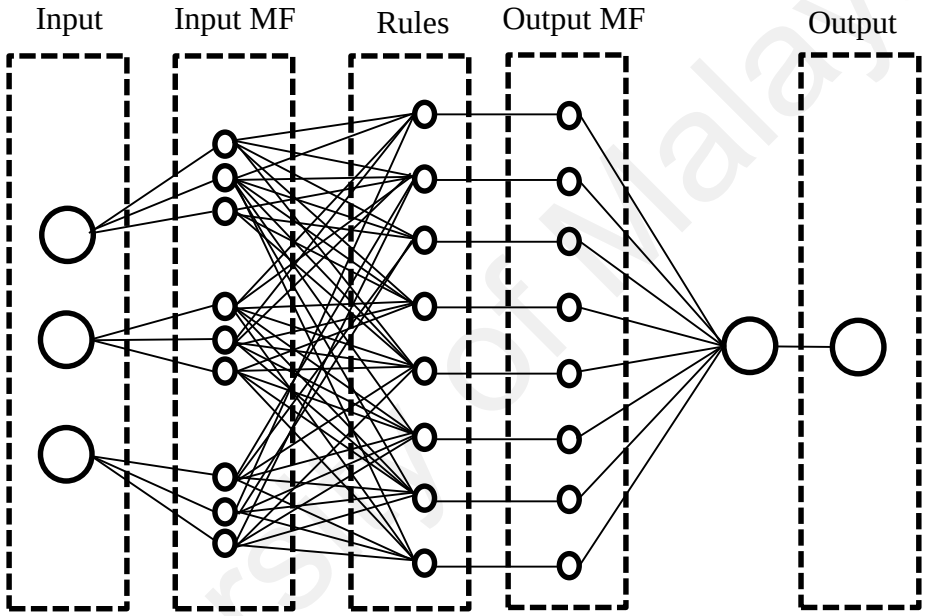



**Figure 3.18:** Structure of ANFIS

Matlab command "genfis2" was used to generate a Sugeno type fuzzy inference system using subtractive clustering. Genfis2 was used instead of genfis1 because the former was more suitable for large amount of data used in this work. The "anfis" command was used to train the classifier. Epoc and radii were 2 parameters that need to be defined prior to training. Epoc is the maximum number of times before training process was stopped. Radii is a vector that specifies a cluster center range of influence in each of the data dimensions, assuming the data falls within a unit hyperbox.

Using higher epoc value will increase accuracy and training time. However, there will be a saturation point where further increase of epoc will only increase training time without improving accuracy. As for radii, using larger value of radii will decrease training time and accuracy. After a vigorous testing, an optimum value of 20 epoc and 1 radii was chosen. In order to increase the efficiency of ANFIS classifier, it was recommended to normalize all input parameters to the value between 0 to 1. Normalization was done in Matlab by dividing all values in the input feature array with the largest value from the input feature array.

### 3.6.3 Support Vector Machine (SVM)

SVM is a machine learning algorithm that stems from statistical learning theory that can handle difficulties of complex pattern classification. Using a linear classification, SVM maps the data to a higher dimensional space (Mota, Rocha, Salles, & Vasconcelos, 2011). SVM can be modified to various domains and tasks using kernel methods based on the selection of base algorithm and kernel function. SVM is good for small amount of sample, high dimensionality and nonlinear problems in pattern recognition (L. Hao & Lewin, 2010).

SVM can be used to find functions from a set of labeled training data. Each data set can be represented by a vector whose dimension relies on the amount of patterns used. The function is either a regression or classification function (L. Hao & Lewin, 2010). SVM has fewer parameters to be modified, causing it to be less dependent on empirical procedures (Mota et al., 2011). SVM is based on several distinct concepts such as linear learning machines, kernel functions, feature spaces, optimization and statistical learning. These theories are merged to form the system of SVM learning. SVM was proven to outperform neural network in a numerous fields (L. Hao & Lewin, 2010).

SVM was initially intended to handle linearly separable cases. Unfortunately, not all practical problems are linearly separated. When dealing with non-linear problems, conventional SVM as a linear classifier will not function effectively. To overcome this problem, a technique known as kernel was presented to deal with non-linear problems using multiple linear classifier. According to the pattern recognition theory, a lower dimensional space and non-linear inseparable model are transformed into linear separable by mapping it nonlinearly into a higher dimensional feature space. Therefore, the usage of kernel method will avoid the curse of dimensionality (Jing & Hou, 2015).

SVM algorithm was initially intended for binary classification, which means they can only classify inputs into two classes (Guo & Wang, 2015). This is due to SVM uses a hyper plane to split data into two categories. In Matlab, the commands "svmtrain" and "svmclassify" can be used to perform simple binary classification. If more than two groups of classification are required, multi-level SVM is needed (Carrasco, López, & Maldonado, 2015; Liu, You, & Cao, 2006). Multi-level SVM is a one against all classifiers, where multiple binary SVM is performed. During multi-level SVM training, a category sample will be classified as one class while the other residual samples as other classes. For this research, 5 classes of outputs were needed to fit 5 types of artificially created cable joint defects. Therefore, multi svm classification was required (Khan et al., 2014).

Multilevel SVM was done by performing SVM multiple times. In the first cycle, all inputs were split into 2 categories, which are Class 1 and non Class 1. The remaining data which was non Class 1 was used as input to the next SVM, which splits the data into Class 2 and non Class 2. This process was repeated until only two classes remain. The summary of this process can be seen in Figure 3.19.

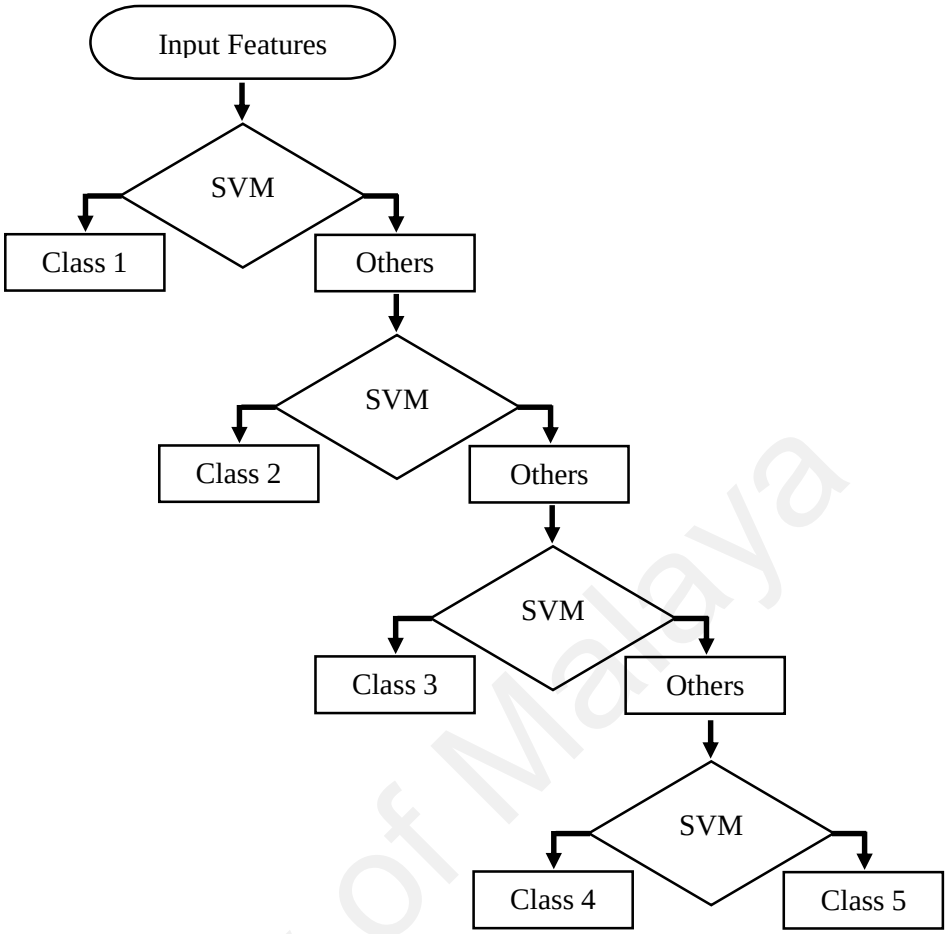



**Figure 3.19:** Multilevel SVM classifier

## 3.7 Chapter Summary

The procedures needed for PD classification have been explained in detail in this chapter. The hardware preparation such as XLPE cable joints sample preparation, measurement setup, equipment calibration and PD measurement have been presented. The software preparation such as feature extraction techniques (statistical, fractal and PCA features) and classifiers (ANN, ANFIS and SVM) used in this work has been elaborated as well. Comparison of classification results of each classifiers are shown in detail in Chapter 4. The classifiers were tested with contaminated noise signals with increasing pulse count and increasing charge magnitude to better represent the actual noise condition.

# CHAPTER 4: RESULTS & DISCUSSION

## 4.1 Introduction

The steps described in Chapter 3 such as PD measurement, feature extraction and classification have been successfully performed and the results of this research work are reported in this chapter. The PRPD patterns of each cable joint sample, sample data obtained from feature extractions and the classification accuracy of each classifier under different noise levels are presented in this chapter. A summary of the result analysis and discussion can be found at the end of this chapter.

## 4.2 Measured PRPD Patterns

The measured PRPD patterns of all five cable joint samples are shown in Figure 4.1. These measurements were performed after all noise and interference prevention methods described in Section 3.3.3 had been implemented. Therefore, all PD patterns seen are contributed by the defect within the XLPE cable joints and not from the surrounding environment or cable connectors. It can be seen that the PRPD pattern for each sample are different from each other.

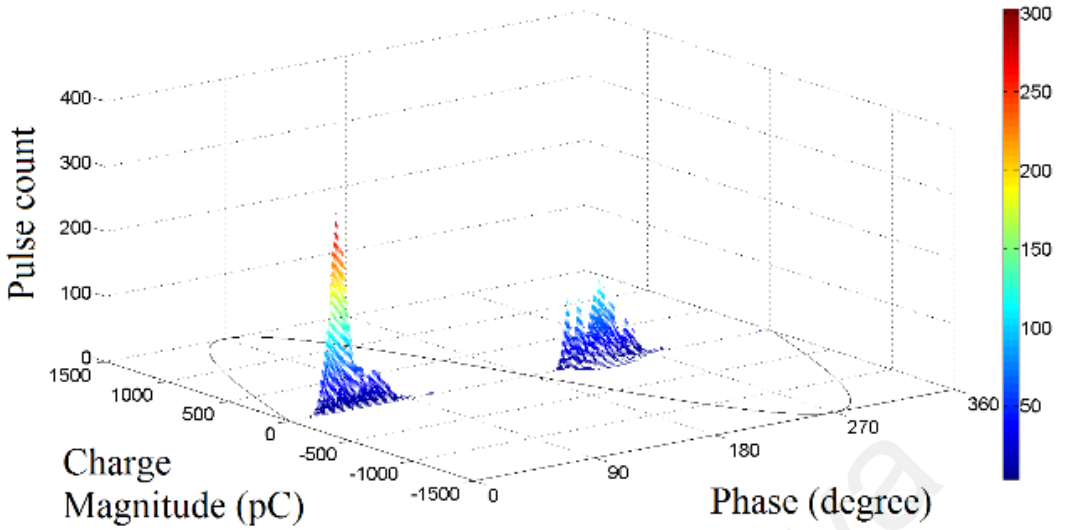
Based on visual inspection on the PRPD patterns, the insulation incision defect has two tall peaks at the end of both positive and negative cycles. The axial directional shift defect has more PD activities in the positive cycle, which accumulate at the first quadrant. It has a very sharp peak at around 80 degrees. The semiconductor layer tip defect has PD activities, which extend evenly between the positive and negative cycles. It has 5 noticeable peaks, 3 at the negative cycle and 2 at the positive cycle. The metal particle on XLPE defect has one main PD group at each positive and negative cycles and it has a prominent peak at 260 degrees. The semiconductor layer air gap defect has two main PD groups; one at the positive cycle and another at the negative cycle with a peak at 230 degrees. Two small clusters of PD with high charge magnitude but low pulse count can

also be seen, where the negatively charged PD spread out between 180 to 360 degrees while the positively charged PDs are distributed between 0 and 180 degrees. The details of PD average and maximum charge magnitude and the pulse count are shown in Table 4.1.

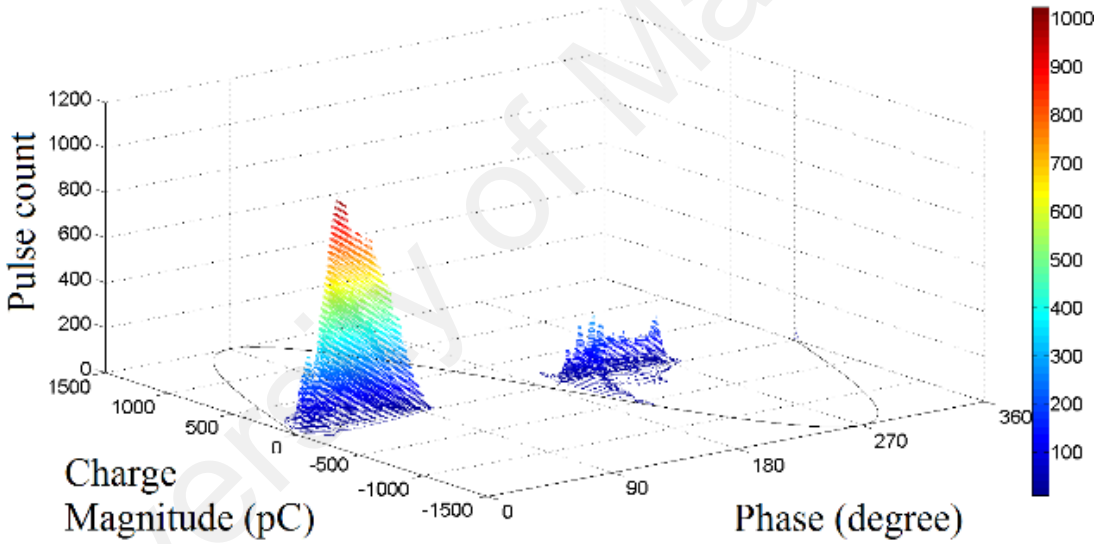
Although different defect types of cable joint have different PRPD patterns, classification of different defect types in the cable joint samples can be hardly done based on visual inspection on these PRPD patterns alone. Therefore, feature extractions from PD patterns and intelligent classifiers were used in this work to classify different defect types in the cable joint.

**Table 4.1:** Details of PRPD patterns of each cable joint samples

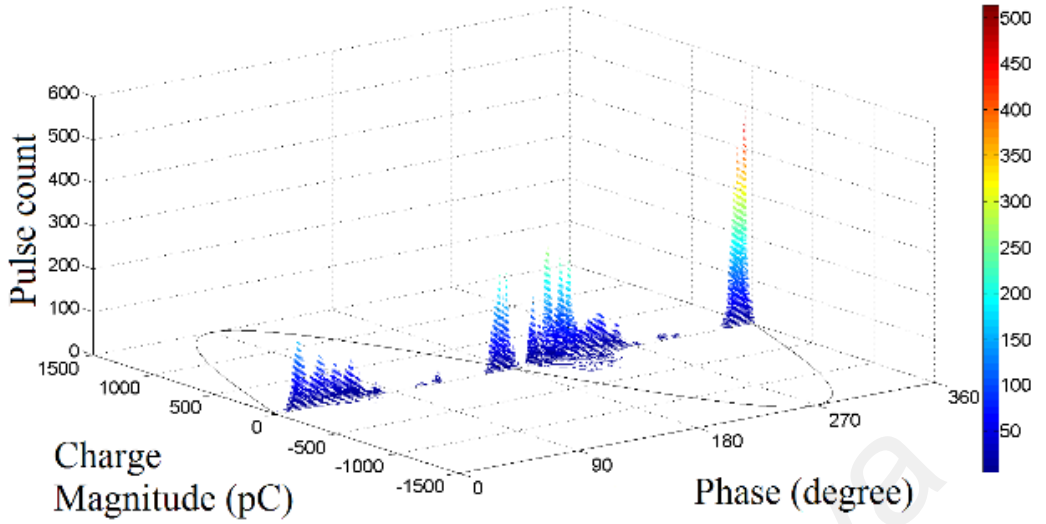
| Sample | Defect type | Mean charge (pC) | Maximum charge (pC) | Pulse count |
|---|---|---|---|---|
| C1 | Insulation incision defect | 50.5848 | 578.1893 | 25085.25 |
| C2 | Axial direction shift defect | 62.5773 | 1110.9720 | 130852.40 |
| C3 | Semiconductor layer tip defect | 47.4091 | 858.4207 | 25720.46 |
| C4 | Metal particle on XLPE defect | 32.5417 | 770.1085 | 22576.75 |
| C5 | Semiconductor layer air gap defect | 243.3954 | 1576.8620 | 13682.92 |



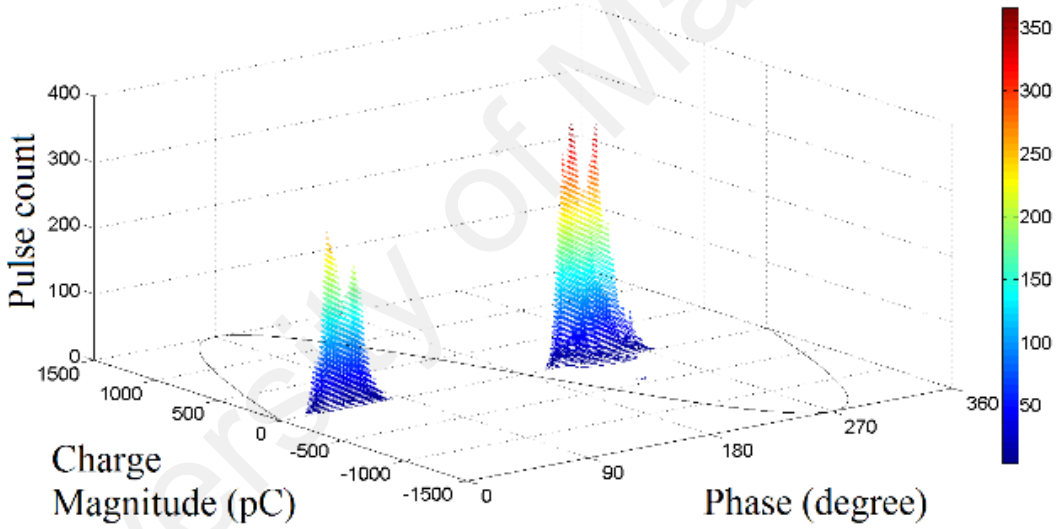

(a)



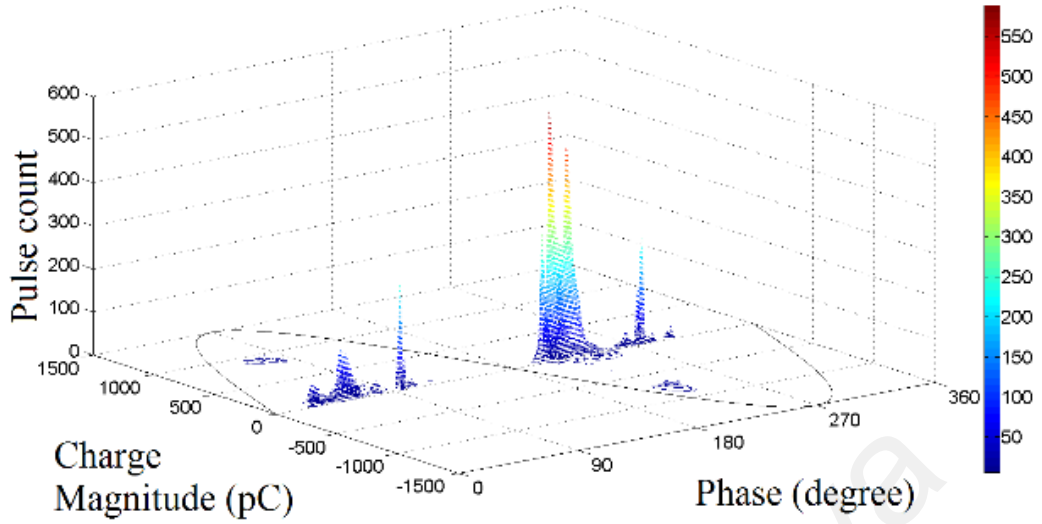

(b)




(c)




(d)




(e)

**Figure 4.1:** PRPD patterns from (a) Insulation incision defect, (b) axial direction shift defect, (c) semiconductor layer tip defect, (d) metal particle on XLPE defect and (e) semiconductor layer air gap defect

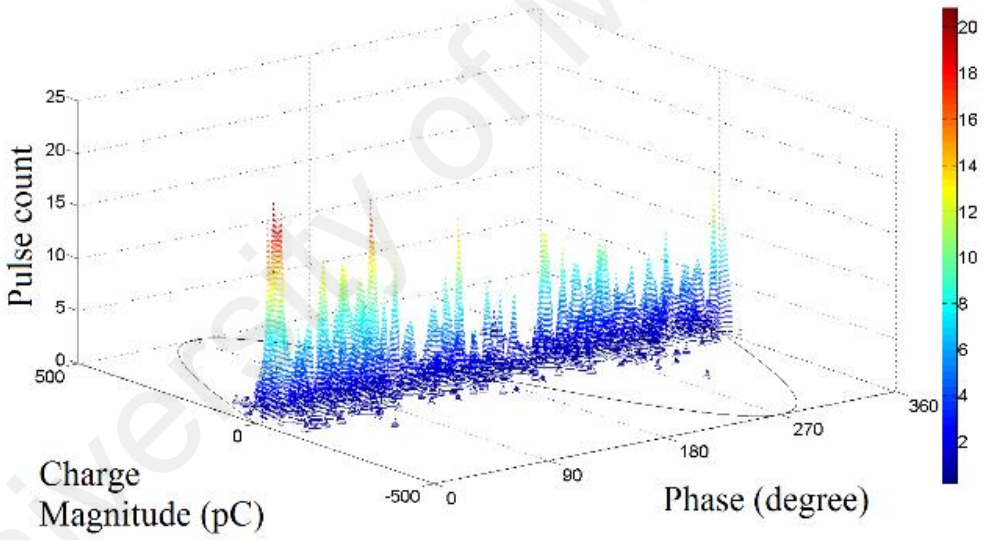
## 4.3 Measured Noise Patterns

In order to determine the classification accuracy of each feature extraction and intelligence classifier method under noisy environment, the classifiers were trained using uncontaminated inputs but tested with input signals that were contaminated with noise. Feature extractions were performed on the contaminated PD signals and used as the test input for each classifier method. The noise signals were recorded from ground interference during raining events.

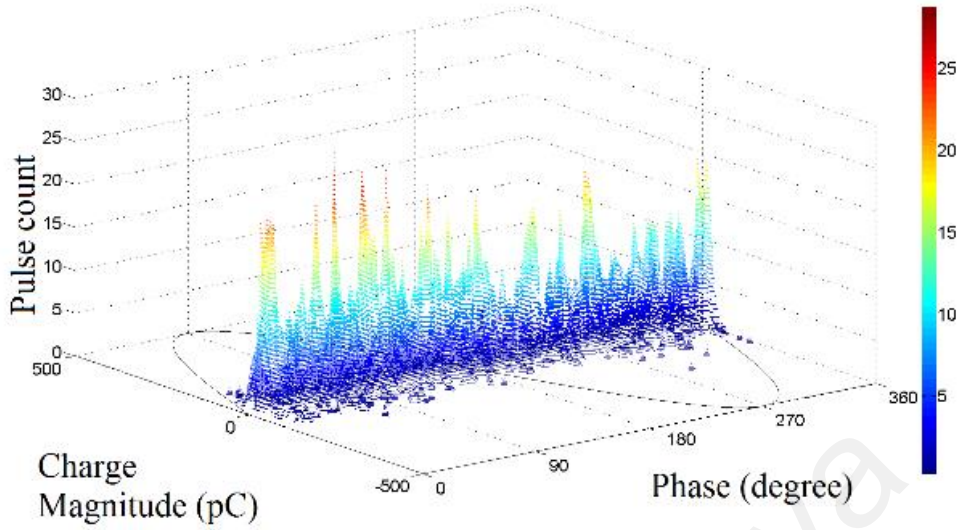
Four PRPD patterns of the recorded noisy signal are shown in Figure 4.2. It can be seen that the noise pattern occurs randomly at every phase and the number of PD activity increases as the duration of noise increases. In order to observe how the increased pulse count of noise affects the classification accuracy, different durations of the noise signal from 5 to 60 seconds were overlapped onto the measured PD data. This is to represent the scenario from a short period of noise duration to a persistent noise throughout the measurement. The details of the noisy data such as average and maximum charge magnitude are shown in Table 4.2.

**Table 4.2:** Details of noisy PD patterns with increasing pulse count

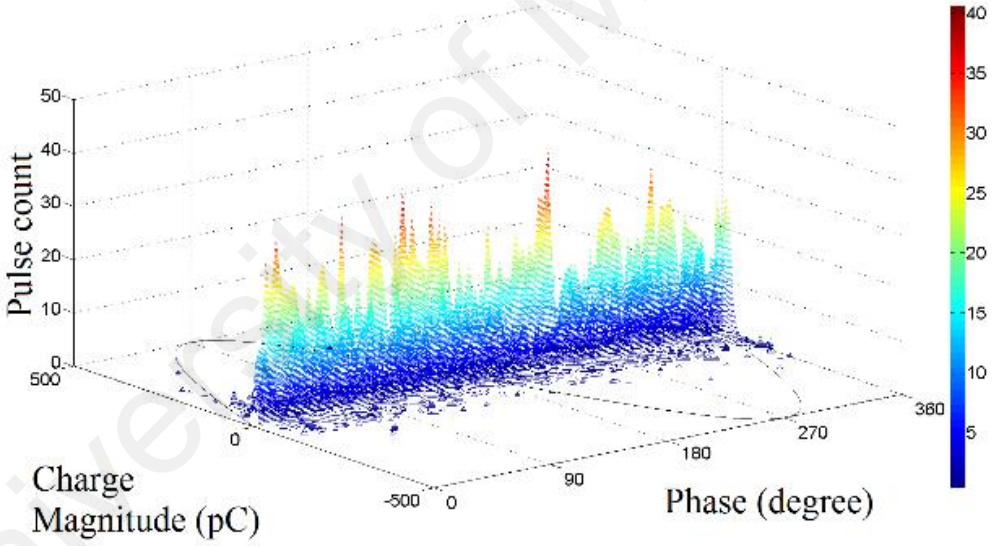
| Noise duration (s) | Mean charge (pC) | Maximum charge (pC) | Pulse count |
|---|---|---|---|
| 5 | 25.10 | 239.63 | 557 |
| 10 | 24.14 | 239.63 | 1101 |
| 15 | 24.00 | 239.63 | 2225 |
| 20 | 23.70 | 239.63 | 2739 |
| 25 | 23.98 | 239.63 | 3554 |
| 30 | 24.05 | 239.63 | 4265 |
| 35 | 24.47 | 312.58 | 5290 |
| 40 | 24.46 | 337.08 | 6310 |
| 45 | 24.36 | 337.08 | 7841 |
| 50 | 24.29 | 337.08 | 8412 |
| 55 | 23.93 | 337.08 | 9050 |
| 60 | 23.91 | 337.08 | 9181 |



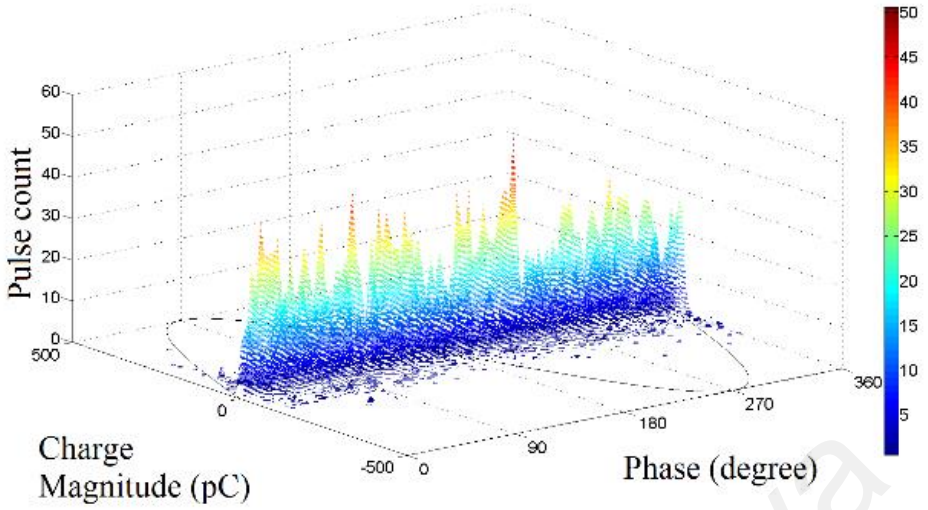

(a)




(b)




(c)




(d)

**Figure 4.2:** PRPD patterns of different noise duration; (a) 15 seconds, (b) 30 seconds, (c) 45 seconds and (d) 60 seconds

In order to investigate the effect of increasing noise magnitude on PD classification accuracy, noise with increasing charge magnitude is required. Since there is no way to control the noise magnitude from the measurement equipment physically, a noise sample with 25 pC average charge magnitude was recorded. This noisy signal was multiplied by 2, 4, 6, 8, 10 and 12 to achieve average charge magnitude of 50 pC, 100 pC, 150 pC, 200 pC, 250 pC and 300 pC. The pulse count in this noise sample remained constant while the charge magnitude was increased. The details of the noise with increasing charge magnitude are shown in Table 4.3. These noisy signals were used to investigate the performance of PD classifiers and input features under noisy condition with increasing average charge magnitude.

**Table 4.3:** Details of noisy PD patterns with increasing charge magnitude

| Mean charge (pC) | Maximum charge (pC) | Pulse count |
|---|---|---|
| 50 | 480 | 657 |
| 100 | 960 | 657 |
| 150 | 1440 | 657 |
| 200 | 1920 | 657 |
| 250 | 2400 | 657 |
| 300 | 2880 | 657 |

## 4.4 Sample Data of Input Features

A total of 500 sets of PD data were recorded in this work, where 100 PD data were taken from each cable joint sample. Due to space constrain, a data was randomly chosen from each cable joint sample to be presented in this report. Tables 4.4 to 4.6 show the sample data that were extracted such as statistical features, PCA features and fractal features.

**Table 4.4:** Extracted statistical features

| Sample | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| Mean $H_n^+(\varphi)$ | 41.2944 | 548.2333 | 73.8111 | 45.3944 | 35.6944 |
| Mean $H_n^-(\varphi)$ | 43.3833 | 293.5556 | 107.9278 | 108.9667 | 52.8833 |
| Stdev $H_n^+(\varphi)$ | 78.4387 | 690.5267 | 105.5642 | 83.9630 | 46.7474 |
| Stdev $H_n^-(\varphi)$ | 69.3773 | 384.3038 | 147.7483 | 184.3328 | 81.0928 |
| Skewness $H_n^+(\varphi)$ | 2.2340 | 0.7831 | 1.5270 | 1.7639 | 1.1117 |
| Skewness $H_n^-(\varphi)$ | 1.4984 | 0.9107 | 1.6990 | 1.5827 | 2.0038 |
| Kurtosis $H_n^+(\varphi)$ | 7.1230 | 1.9383 | 4.1013 | 4.7041 | 2.8598 |
| Kurtosis $H_n^-(\varphi)$ | 3.9166 | 2.3737 | 5.2176 | 4.1422 | 6.5342 |
| Mean $H_{qn}^+(\varphi)$ | 29.8984 | 60.3932 | 36.7744 | 24.8165 | 239.9170 |
| Mean $H_{qn}^-(\varphi)$ | -46.9201 | -70.9458 | -50.7191 | -33.0217 | -168.3107 |
| Stdev $H_{qn}^+(\varphi)$ | 14.8031 | 355.9565 | 60.3246 | 44.6759 | 390.1542 |
| Stdev $H_{qn}^-(\varphi)$ | 32.0612 | 231.3776 | 72.4628 | 47.5718 | 337.4364 |
| Skewness $H_{qn}^+(\varphi)$ | 1.3793 | 16.4857 | 5.3307 | 8.0583 | 1.5959 |
| Skewness $H_{qn}^-(\varphi)$ | -0.7037 | -25.4588 | -5.8096 | -6.9205 | -2.4466 |
| Kurtosis $H_{qn}^+(\varphi)$ | 8.2998 | 278.1250 | 42.6534 | 78.3225 | 3.9256 |
| Kurtosis $H_{qn}^-(\varphi)$ | 2.5962 | 722.1208 | 76.0322 | 68.9398 | 7.5861 |
| Weibull $\alpha+$ $H_{qn}^+(\varphi)$ | 33.8424 | 47.5005 | 49.1166 | 45.9543 | 54.6889 |
| Weibull $\beta+$ $H_{qn}^+(\varphi)$ | 2.1356 | 0.8096 | 1.0378 | 1.0411 | 0.8202 |
| Weibull $\alpha-$ $H_{qn}^-(\varphi)$ | 52.2874 | 67.0304 | 58.0309 | 55.3513 | 60.2303 |
| Weibull $\beta-$ $H_{qn}^-(\varphi)$ | 1.5167 | 0.9212 | 0.9392 | 0.9421 | 0.8430 |

**Table 4.5:** Extracted PCA features

| Sample | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| 1st Principal component | 15466.92 | 53812.19 | 28191.27 | 22645.52 | 24781.90 |
| | -10023.63 | -86101.03 | -18434.80 | -13376.07 | -36949.43 |
| | -5443.29 | 32288.85 | -9756.46 | -9269.45 | 12167.53 |
| 2nd Principal component | 396.12 | 23468.59 | -1368.24 | -513.11 | 9458.52 |
| | 1808.37 | 4266.60 | -5982.90 | -3987.69 | 2429.16 |
| | -2204.49 | -27735.19 | 7351.14 | 4500.80 | -11887.68 |
| Latent | 184664096 | 5675854466 | 614888988 | 388830821 | 1063725994 |
| | 4143434 | 669109687 | 45853246 | 18211088 | 118340647 |

**Table 4.6:** Extracted fractal features

| Sample | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| Fractal dimensions | 1.0194 | 1.1832 | 1.2625 | 1.1153 | 1.4403 |
| | 1.0216 | 1.2099 | 1.2707 | 1.1323 | 1.4284 |
| | 1.0165 | 1.2239 | 1.2939 | 1.1385 | 1.4472 |
| | 1.0312 | 1.1872 | 1.2455 | 1.0955 | 1.4237 |
| | 0.9945 | 1.1593 | 1.2234 | 1.0693 | 1.3844 |
| | 0.9807 | 1.1614 | 1.2184 | 1.0679 | 1.3809 |
| | 0.9745 | 1.1518 | 1.2207 | 1.0687 | 1.3734 |
| | 0.9800 | 1.1530 | 1.2161 | 1.0611 | 1.3819 |
| | 0.9801 | 1.1496 | 1.2143 | 1.0539 | 1.3809 |
| | 0.9793 | 1.1584 | 1.2269 | 1.0711 | 1.3822 |
| | 0.9774 | 1.1593 | 1.2245 | 1.0652 | 1.3759 |
| | 0.9754 | 1.1524 | 1.2124 | 1.0641 | 1.3757 |
| Lacunarity | 3.2030 | 2.8662 | 3.6734 | 4.3090 | 2.1689 |
| | 3.1826 | 2.7053 | 3.6556 | 4.2790 | 2.1232 |
| | 3.1259 | 2.6616 | 3.4191 | 4.1019 | 2.1006 |
| | 3.1382 | 2.7285 | 3.7433 | 4.4175 | 2.1653 |
| | 3.3055 | 2.9040 | 3.9493 | 4.7882 | 2.3764 |
| | 3.5426 | 2.8755 | 3.9130 | 4.7669 | 2.3790 |
| | 3.4176 | 2.8940 | 3.8376 | 4.6951 | 2.4244 |
| | 3.5118 | 2.9119 | 3.8289 | 4.7892 | 2.4159 |
| | 3.3419 | 2.8563 | 3.7805 | 4.5176 | 2.3105 |
| | 3.5261 | 2.9234 | 3.9009 | 4.7871 | 2.3782 |
| | 3.4471 | 2.8645 | 3.8581 | 4.7480 | 2.4381 |
| | 3.4664 | 2.7786 | 3.8659 | 4.7270 | 2.4359 |

## 4.5 Classification Results under Noise-Free Condition

After performing feature extraction, the statistical features, fractal features and PCA features were obtained and split into 7 groups of input feature. The statistical features were split into 3 groups; the first group consists of variance, skewness, kurtosis and mean (var, skew, kur, mean), the second group consists of Weibull parameters while the third group is the combination of the first two groups. Fractal features were also split into 3 groups, which are fractal dimensions, lacunarity and a combination of fractal dimensions and lacunarity. The last group of input feature was the proposed PCA features. These input features were used as the input for the classifiers to determine the classification accuracy.

In order to test the accuracy of the classifier, the hold-one-out technique was used, where all input data were randomly split into a 70:30 ratio, where 70% of the data were used for training and 30% of the remaining data were used for testing purposes. The 30% of the data are unique data that were not used in the training and will serve as a good benchmark of the classification accuracy. All classification tests were repeated 100 times (using a new randomly selected 70:30 ratio of training and test input) to get the average performance. The results were sorted in a table with information such as training time needed for 100 tests, the size of the input feature used, individual classification accuracy of each sample and total classification accuracy of the classifier.

### 4.5.1 ANN Classification Results

The classification accuracy of ANN when using statistical features, fractal features and PCA features is shown in Table 4.7.

**Table 4.7:** ANN classification results

| Classifier | Input type | Size | Time (s) | Classification accuracy (%) | | | | | **Total (%)** |
|---|---|---|---|---|---|---|---|---|---|
| | | | | C1 | C2 | C3 | C4 | C5 | |
| ANN | Var, skew, kur, mean | 16 | 95.52 | 92.89 | 90.19 | 88.18 | 88.03 | 94.29 | 90.71 |
| | Weibull features | 4 | 85.61 | 94.23 | 79.72 | 75.67 | 80.20 | 88.89 | 83.74 |
| | Statistical features | 20 | 94.25 | 94.71 | 89.89 | 88.91 | 89.65 | 94.67 | **91.57** |
| | Fractal dimensions | 12 | 80.63 | 76.49 | 76.42 | 52.23 | 58.3 | 91.65 | 71.02 |
| | Lacunarity | 12 | 97.07 | 79.83 | 83.41 | 91.95 | 80.34 | 91.50 | 85.41 |
| | Fractal features | 24 | 82.12 | 84.48 | 85.95 | 86.84 | 77.09 | 92.86 | **85.44** |
| | PCA features | 48 | 84.73 | 88.84 | 92.44 | 81.73 | 88.05 | 93.93 | **89.00** |

The classification accuracy of variance, skewness, kurtosis and mean (90.71%) is higher than Weibull features, which is 83.74%. When statistical and Weibull parameters were combined and used as one input, the total classification accuracy increases slightly to 91.57%. The training time needed for variance, skewness, kurtosis and mean is l0 seconds longer compared to Weibull features. However, when both groups of features were combined, the training time is improved slightly to 94.25 seconds.

It can be seen that lacunarity gives higher classification accuracy of 85.41% compared to fractal dimension at 71.02% but its training time is longer, which is 97.07 seconds compared to 80.63 seconds. However, when fractal dimension and lacunarity were used as the input feature, the ANN classifier was able to achieved higher classification accuracy of 85.44% and shorter training duration of 82.12 seconds compared to only using fractal dimensions.

Using PCA features, the classification accuracy of 89.00% was achieved. The training time required is 84.73 seconds. The training time of ANN is relatively constant in the range of 80 to 100 seconds regardless of the input feature size.

### 4.5.2 ANFIS Classification Results

The classification accuracy of ANFIS when using statistical features, fractal features and PCA features is shown in Table 4.8.

**Table 4.8:** ANFIS classification results

| Classifier | Input type | Size | Time (s) | Classification accuracy (%) | | | | | **Total (%)** |
|---|---|---|---|---|---|---|---|---|---|
| | | | | C1 | C2 | C3 | C4 | C5 | |
| ANFIS | Var, skew, kur, mean | 16 | 58.29 | 95.87 | 94.48 | 97.88 | 92.66 | 99.92 | 96.16 |
| | Weibull features | 4 | 13.48 | 89.37 | 29.43 | 90.66 | 86.77 | 88.79 | 77.00 |
| | Statistical features | 20 | 97.33 | 95.42 | 94.81 | 98.86 | 93.33 | 99.58 | **96.40** |
| | Fractal dimensions | 12 | 54.59 | 42.53 | 88.16 | 73.05 | 54.12 | 78.97 | 67.37 |
| | Lacunarity | 12 | 32.93 | 48.74 | 93.30 | 87.00 | 67.70 | 94.54 | 78.23 |
| | Fractal features | 24 | 198.02 | 64.80 | 98.44 | 85.73 | 67.86 | 99.98 | **83.36** |
| | PCA features | 48 | 429.16 | 49.95 | 72.91 | 76.9 | 56.49 | 77.43 | **66.74** |

Similar to ANN, the classification accuracy of ANFIS is higher when using variance, skewness, kurtosis and mean (96.16%) compared to Weibull features (77.00%). When using a combination of statistical and Weibull features, the total accuracy has a marginal increase from 96.16% to 96.40%. However, this causes the training time to increase significantly from 58.29 seconds to 97.33 seconds. ANFIS training duration increases when the feature size used is increased.

It was observed that lacunarity gives higher classification accuracy of 78.23% than fractal dimensions (67.37%) and its training time is also shorter, 32.93 seconds compared to 54.59 seconds. When fractal dimension and lacunarity were combined as one input feature, the ANFIS classifier was able to achieve higher classification accuracy of 83.36%. However, the training time increases significantly to 198.02 seconds.

For PCA features, ANFIS performed badly with only 66.74% accuracy. This is very low compared to statistical and fractal features. Apart of that, the training duration is 429.16 seconds, which is the longest training duration recorded. It is observed that ANFIS training duration is heavily affected by the feature size, where larger feature size will cause the training duration to increase.

### 4.5.3 SVM Classification Results

The classification accuracy of SVM when using statistical features, fractal features and PCA features is shown in Table 4.9.

**Table 4.9:** SVM classification results

| Classifier | Input type | Size | Time (s) | Classification accuracy (%) | | | | | **Total (%)** |
|---|---|---|---|---|---|---|---|---|---|
| | | | | C1 | C2 | C3 | C4 | C5 | |
| SVM | Var,skew,kur,mean | 16 | 21.33 | 99.87 | 99.34 | 98.07 | 86.38 | 99.98 | 94.47 |
| | Weibull features | 4 | 24.18 | 99.87 | 59.97 | 95.01 | 0 | 89.28 | 68.82 |
| | Statistical features | 20 | 20.05 | 99.86 | 99.1 | 98.18 | 96.2 | 100 | **98.67** |
| | Fractal dimensions | 12 | 54.07 | 94.47 | 94.48 | 68.19 | 16.61 | 95.81 | 73.91 |
| | Lacunarity | 12 | 56.07 | 92.47 | 100 | 84.86 | 75.54 | 98.44 | 90.26 |
| | Fractal features | 24 | 59.75 | 94.76 | 99.12 | 93.54 | 75.53 | 99.70 | **91.93** |
| | PCA features | 48 | 38.33 | 97.33 | 95.73 | 91.53 | 82.62 | 99.51 | **93.35** |

For SVM, the classification accuracy is higher when using variance, skewness, kurtosis and mean (94.47%) compared to Weibull parameters. Using Weibull parameters alone as input feature yields very low classification accuracy of 68.82%. When using a combination of statistical and Weibull features, the classification accuracy has a marginal increase to 98.67% without negatively affecting the training duration.

It was observed that lacunarity gives higher classification accuracy than fractal dimension, 90.26% over 73.91% with similar training time. When both fractal dimension and lacunarity were used as input feature, the ANN classifier was able to achieve a slightly higher classification accuracy of 91.93%.

For PCA features, the training duration is faster compared to fractal features despite the larger feature size. SVM achieved classification accuracy of 93.35% when using PCA features. It was observed that larger feature size does not affect the classification time of SVM, where it also has a very short overall training duration compared to ANN and ANFIS.

## 4.6 Classification Results under Noisy Condition

Based on observation from the classifiers under noise-free conditions, it is obvious that for each classifier, statistical features (combination of variance, skewness, kurtosis and mean and Weibull features) and fractal features (combination of fractal dimensions and lacunarity) performed better when they were combined instead of separated. Therefore, when investigating the effect of noise contamination on classification accuracy, only statistical features and fractal features were compared against PCA features.

Similar to previous tests, the test data were once again split into 70:30 ratio where 70% of the test data were used as training input and 30% of the remaining data were used as test input. The 70% training input were input features extracted from noise-free PD

signals, but the remaining 30% test data that were not used for training were overlapped with noise and this contaminated data were used for testing. Therefore, the classifiers were trained using clean input data and tested using contaminated data that were not used during training process. Each test was repeated for 100 times and the average classification accuracy is presented in the following section.

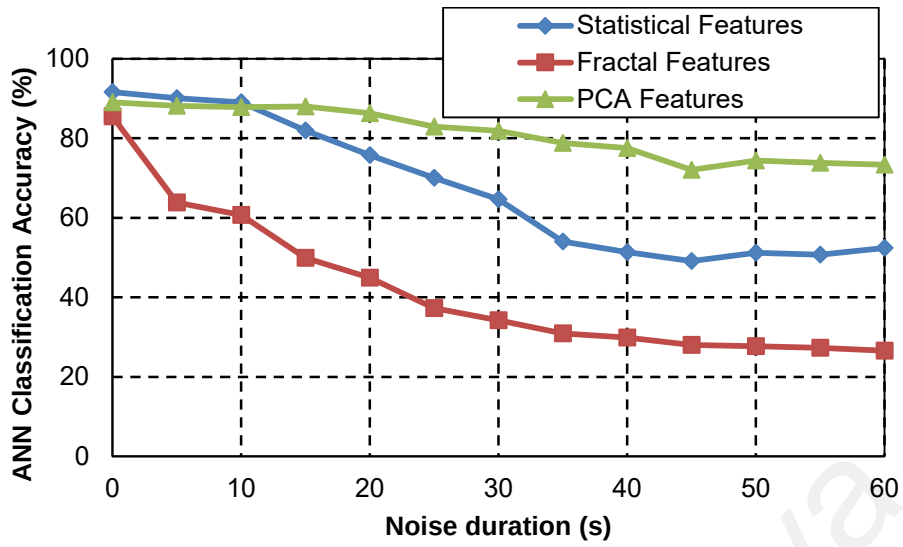
### 4.6.1 Noise with Increasing Pulse Count

The classification accuracy of ANN, ANFIS and SVM when using statistical, fractal and PCA features when using input features contaminated with different noise duration is shown in Table 4.10. The noisy signals used range from 5 seconds and have a 5 second increment up to 60 seconds. More PD pulses are contained in the noise sample when longer noise duration was applied. The amount of PD pulse in each noise duration has been shown in Table 4.2.

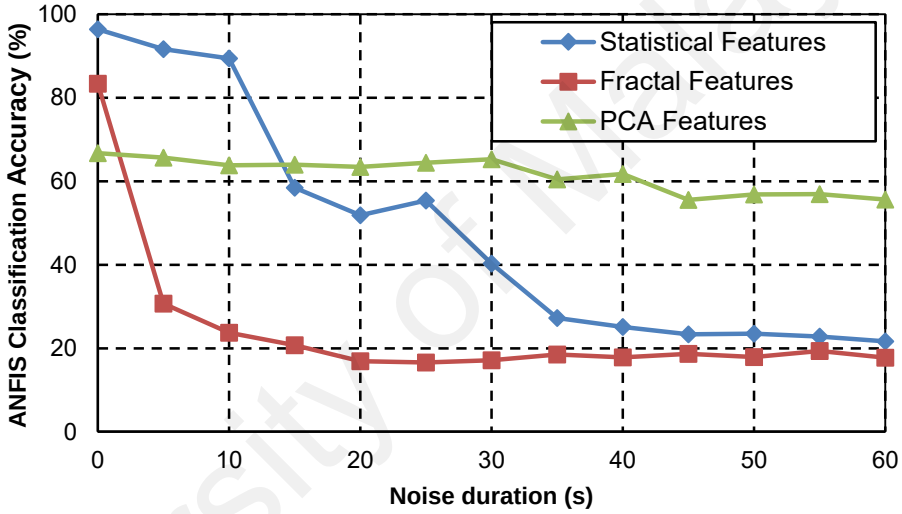
It can be seen that all classifiers and input feature combination experience a certain degree of decrement in the classification accuracy. However, the rate of decrement is not the same because each classifier and input feature combination has its own noise tolerance. To better observe the rate of accuracy reduction, a graph of classification accuracy against noise contamination of each classifier is shown in Figure 4.3.

**Table 4.10:** Classification accuracy of different classifier and input feature under varying duration of noise signal

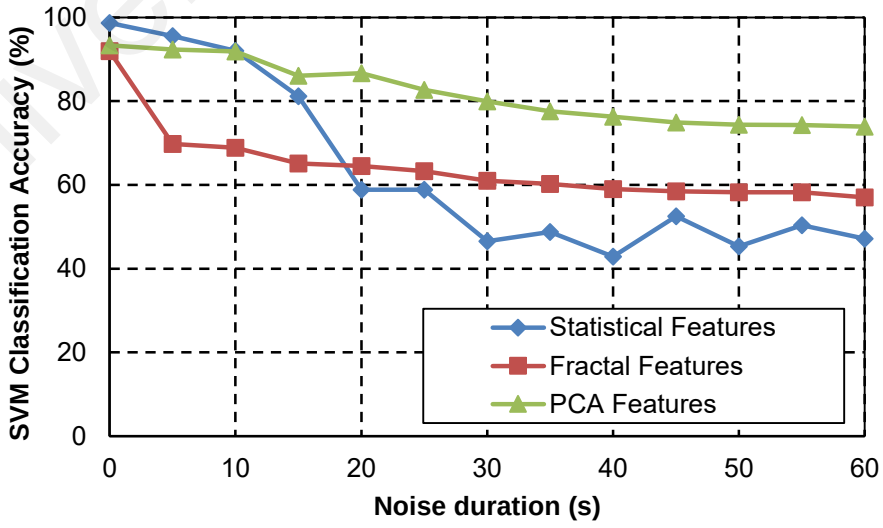
| Classifier | Input type | Duration of noisy signals (s) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 |
| ANN | Statistical | 90.1 | 89.0 | 82.0 | 75.7 | 70.0 | 64.6 | 54.0 | 51.3 | 49.1 | 51.2 | 50.7 | 52.4 |
| | Fractal | 63.8 | 60.7 | 49.9 | 44.9 | 37.3 | 34.2 | 30.9 | 29.9 | 28.0 | 27.7 | 27.3 | 26.5 |
| | PCA | 88.1 | 87.8 | 88.0 | 86.3 | 82.9 | 81.8 | 78.8 | 77.5 | 72.0 | 74.3 | 73.8 | 73.3 |
| ANFIS | Statistical | 91.6 | 89.4 | 58.4 | 51.9 | 55.4 | 40.3 | 27.2 | 25.1 | 23.3 | 23.5 | 22.8 | 21.7 |
| | Fractal | 30.7 | 23.7 | 20.7 | 16.9 | 16.6 | 17.1 | 18.5 | 17.8 | 18.7 | 17.9 | 19.4 | 13.3 |
| | PCA | 65.7 | 63.8 | 64.0 | 63.5 | 64.4 | 65.3 | 60.4 | 61.8 | 55.5 | 56.9 | 56.9 | 55.6 |
| SVM | Statistical | 95.6 | 92.0 | 81.2 | 58.9 | 58.9 | 46.5 | 48.8 | 42.9 | 52.5 | 45.3 | 50.4 | 47.1 |
| | Fractal | 69.8 | 68.9 | 65.1 | 64.5 | 63.3 | 61.0 | 60.3 | 59.0 | 58.5 | 58.3 | 58.2 | 57.1 |
| | PCA | 92.3 | 91.9 | 86.1 | 86.7 | 82.7 | 80.0 | 77.6 | 76.3 | 74.9 | 74.3 | 74.3 | 73.9 |




**(a)**




**(b)**




**(c)**

**Figure 4.3:** Noise tolerance against increasing pulse count for (a) ANN, (b) ANFIS and (c) SVM

For ANN, as the noise duration increases, the classification accuracy of statistical and fractal features decreases at a steeper rate compared to PCA feature. The classification accuracy of PCA feature manages to overtake the statistical and fractal features above 10 seconds noise duration due to its slower reduction rate. At the maximum of 60 seconds noise duration, the statistical, fractal and PCA features achieved classification accuracy of 52.4%, 26.5% and 73.3% respectively.

For ANFIS, as the noise duration increases, fractal features classification accuracy decreases sharply in the beginning but retains a very low classification rate. The noise tolerance of statistical feature is better than the fractal feature but is poorer than the PCA features. PCA features managed to achieve higher classification accuracy compared to statistical and fractal feature once the noise duration exceeded 15 seconds. At the maximum of 60 seconds noise duration, the statistical, fractal and PCA features achieved classification accuracy of 21.7%, 13.3% and 55.6% respectively.

For SVM, as the noise duration increases, the performance of statistical features reduces at a faster rate than fractal features. Thus, after the 20 seconds noise duration, statistical features performed worse than fractal features despite having higher classification accuracy under noise-free condition. The classification accuracy of PCA features decreases at a slower rate and becomes the input feature with the highest classification accuracy for SVM after 10 seconds noise duration. At the maximum of 60 seconds noise duration, the statistical, fractal and PCA features achieved classification accuracy of 47.1%, 57.1% and 73.9% respectively.

### 4.6.2 Noise with Increasing Charge Magnitude

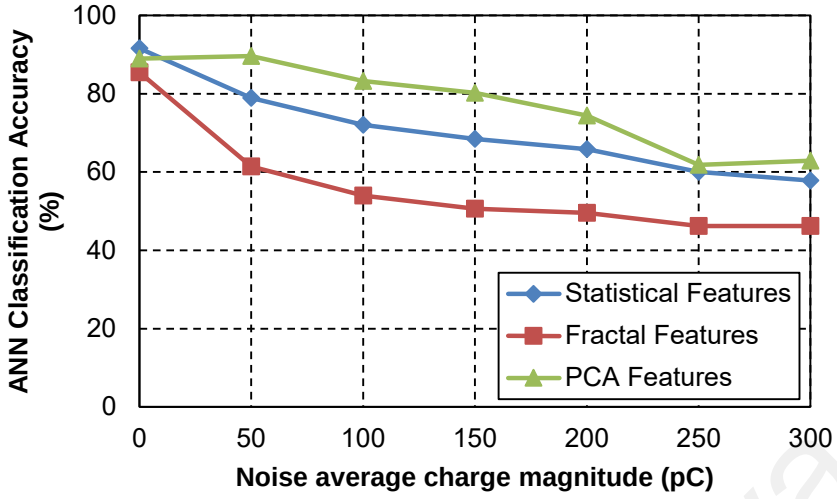
The classification accuracy of ANN, ANFIS and SVM when using statistical, fractal and PCA features with input data contaminated with noise of different charge magnitudes

is shown in Table 4.11. The noisy signals have an average noise charge magnitude from 50 pC to 300 pC. Detail specification of the noise signal has been shown in Table 4.3.

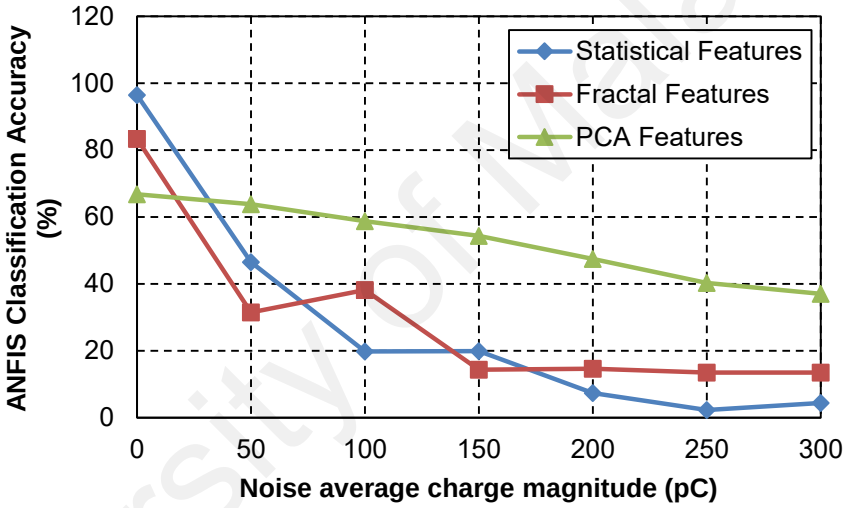
All classifiers and input feature combinations suffer classification accuracy reduction when the input features were contaminated with noise of different average charge magnitudes. All classifiers and input feature combinations experience a certain degree of deterioration in the classification accuracy. However, the rate of classification accuracy decrement is not the same as each classifier and input feature combination has its own noise tolerance. A graph of classification accuracy against noise contamination is shown in Figure 4.4 for a clearer view of the noise tolerance of each classifier and input feature combination.

**Table 4.11:** Classification accuracy of different classifier and input feature under varying magnitude of noise signal

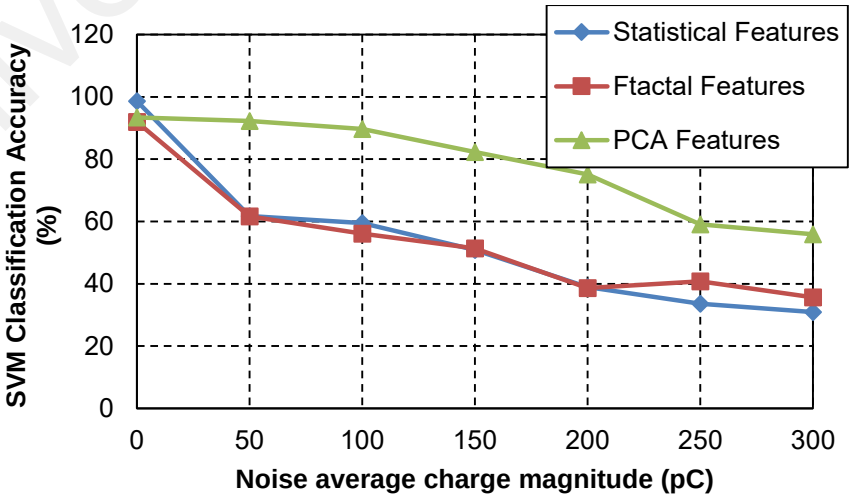
| Classifier | Input type | Average noise charge magnitude (pC) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0 | 50 | 100 | 150 | 200 | 250 | 300 |
| ANN | Statistical | 91.57 | 78.94 | 72.04 | 68.45 | 65.82 | 60.06 | 57.83 |
| | Fractal | 85.44 | 61.45 | 53.95 | 50.60 | 49.59 | 46.18 | 46.18 |
| | PCA | 89.00 | 89.61 | 83.28 | 80.24 | 74.45 | 61.85 | 62.86 |
| ANFIS | Statistical | 96.40 | 46.46 | 19.81 | 19.90 | 7.33 | 2.25 | 4.39 |
| | Fractal | 83.36 | 31.45 | 38.14 | 14.30 | 14.63 | 13.49 | 13.49 |
| | PCA | 66.74 | 63.83 | 58.74 | 54.31 | 47.48 | 40.26 | 37.04 |
| SVM | Statistical | 98.67 | 61.74 | 59.46 | 50.90 | 38.96 | 33.61 | 30.92 |
| | Fractal | 91.93 | 61.70 | 56.09 | 51.34 | 38.64 | 40.80 | 35.64 |
| | PCA | 93.35 | 92.29 | 89.74 | 82.29 | 75.06 | 59.08 | 55.92 |




(a)




(b)




(c)

**Figure 4.4:** Noise tolerance against increasing charge magnitude for (a) ANN (b) ANFIS and (c) SVM

For ANN, as the noise charge magnitude increases, the classification accuracy of statistical and fractal features decreases faster compared to PCA feature. PCA feature manages to overtake the statistical and fractal features in classification accuracy once a noise of 50 pC was introduced due to its slower reduction rate. At the maximum of 300 pC charge magnitude, the statistical, fractal and PCA features achieved classification accuracy of 57.83%, 46.18% and 62.86% respectively.

For ANFIS, as the noise charge magnitude increases, the classification accuracy of statistical and fractal features fell sharply in the beginning and fell to a very low classification rate of less than 20%. PCA features experienced a slower deterioration rate compared to statistical and fractal features. At the maximum of 300 pC charge magnitude, the statistical, fractal and PCA features achieved classification accuracy of 4.39%, 13.49% and 37.04% respectively.

For SVM, as the noise charge magnitude increases, the classification accuracy of statistical and fractal features experienced a very similar reduction rate. PCA features experienced a much slower deterioration rate compared to statistical and fractal features. At the maximum of 300 pC charge magnitude, the statistical, fractal and PCA features achieved classification accuracy of 30.92%, 35.64% and 55.92% respectively.

## 4.7 Discussion

Under noise-free condition, statistical features give the best classification accuracy for all three classifiers. The performance is noticeably better compared to fractal and PCA features, which performed differently when being used in different classifiers. For ANN and SVM, PCA features performed better than fractal features while in ANFIS, fractal features performed better than PCA features.

ANFIS performance is on par with ANN for statistical and fractal features. However, it performed very weak when using PCA features, achieving only 66.74% classification accuracy under noise-free condition. This is due to ANFIS requires normalizing the input data during the training process to improve its efficiency (Jang, 1993a). PCA components contain a different weighting; hence normalization will change the relative significance between each components, causing higher error rate in ANFIS (Lai et al., 2010).

In terms of training speed, SVM has the fastest training speed. Regardless of the input feature sized used for training, SVM always finishes 100 training cycles in less than 1 minute. ANN is slightly slower than SVM for all input features used. ANFIS training speed varies according to feature size, where it takes longer time to train when the input feature size is larger. When using input feature with a dimension of less than 20, ANFIS is as fast as ANN and SVM, but once input feature with larger dimension was used, the training time increased almost linearly. The effect of increasing feature size on the training duration for all classifiers is shown in Figure 4.5. It is seen that SVM and ANN training speed is not affected by the size of the input feature and remains relatively consistent when the feature size is increased. ANFIS, on the other hand, experienced an increased training duration when the input feature size was increased.

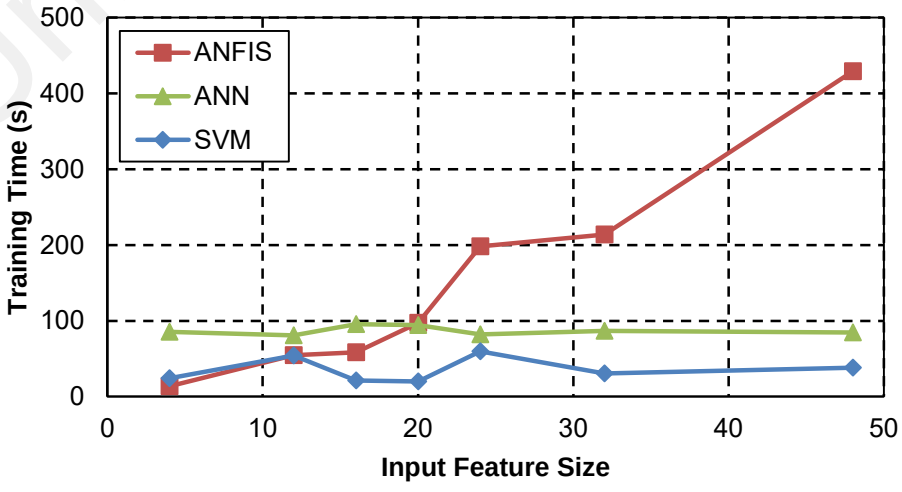



**Figure 4.5:** Training time vs. input feature size for PD classifiers

Under noisy conditions, the performance of each classifier and input feature combination are different from each other. In previous works, noise analysis was performed using random noise generated between 10% and 30% of the PD data. Some of these noises were added into the input feature itself, while others were added into the PRPD pattern. In order to better represent the effect of actual noise contamination, real noise signals from ground interference were used in this work. The details of measured PD and noise signals used have been shown in Table 4.1 to Table 4.3. Each PD reading measured from all 5 test samples has different average charge magnitudes. Test samples 1 to 4 have an average charge magnitude of around 50 pC while test sample 5 is around 250 pC. In the first noise analysis, the noise average charge magnitude was 25 pC, which is approximately 50% of test samples 1 to 4 and 30% of sample 5. This is very high amount of noise contamination.

In the second part of the noise contamination analysis, noise of different average charge magnitude was used, where the average charge magnitude was multiplied mathematically to observe the effect of increasing noise amplitude. The range of average charge magnitude tested were from 50 pC to 300 pC. This is 6 times the average charge magnitude of samples 1 to 4 and 120% of sample 5. Therefore, the deterioration is more severe in the second noise analysis. This provides a sufficient range of noise to observe the actual effect of noise contamination.

Table 4.12 shows the reduction of classification accuracy under the maximum noise duration. The highest classification accuracy is during noise free condition and the worst classification accuracy is during the maximum noise duration. The first trend that can be observed is that PCA features show highest noise tolerance compared to statistical and fractal features in all three classifiers. PCA features does not have the best classification accuracy under noise-free condition in all three classifiers. However, once noise

contamination was introduced, the other input features experience huge classification accuracy reduction and PCA features have the best classification accuracy. Fractal features experienced the most reduction in classification accuracy in ANN and ANFIS. ANN and ANFIS are better in coping against contaminated statistical features compared to fractal features. On the other hand, SVM is better in coping against contaminated fractal features compared to statistical features. For all classifiers, the proposed PCA feature has the highest noise tolerance with least percentage reduction of classification accuracy under noisy conditions.

**Table 4.12:** Percent reduction of classification accuracy under maximum noise duration

| Classifier | Input type | Highest classification accuracy (%) | Worst classification accuracy (%) | Reduction (%) |
|---|---|---|---|---|
| ANN | Statistical | 91.57 | 52.4 | 42.78 |
| | Fractal | 85.44 | 26.5 | 68.98 |
| | PCA | 89.00 | 73.3 | **17.64** |
| ANFIS | Statistical | 96.40 | 21.7 | 77.49 |
| | Fractal | 83.36 | 13.3 | 84.05 |
| | PCA | 66.74 | 55.6 | **16.69** |
| SVM | Statistical | 98.67 | 47.1 | 52.27 |
| | Fractal | 91.93 | 57.1 | 37.89 |
| | PCA | 93.35 | 73.9 | **20.84** |

Table 4.13 shows the reduction of classification accuracy under the maximum average charge magnitude. It can be seen that noise contamination with increasing charge magnitude has a more profound effect on classification accuracy compared to noise contamination with increasing pulse count. This is expected due to the high magnitude noisy signals used for testing. All three classifiers and input features experience a greater reduction rate of classification accuracy compared to Table 4.12. ANFIS classification accuracy fell to an unusable level for statistical and fractal features, where it no longer recognizes the input features under such high noise contamination. ANN and SVM

experienced a more linear deterioration and is noticeably better than ANFIS. Similar to Table 4.12, SVM is better at recognizing PD patterns under noisy conditions when using fractal features compared to statistical features while ANN performed better when using statistical features compared to fractal features. The proposed PCA feature once again managed to display the highest noise tolerance by having the least reduction compared to other input features under harsh noise contamination.

**Table 4.13:** Percent reduction of classification accuracy under maximum noise charge amplitude

| Classifier | Input type | Highest classification accuracy (%) | Worst classification accuracy (%) | Reduction (%) |
|---|---|---|---|---|
| ANN | Statistical | 91.57 | 57.83 | 36.85 |
| | Fractal | 85.44 | 46.18 | 45.95 |
| | PCA | 89.00 | 62.86 | **29.37** |
| ANFIS | Statistical | 96.40 | 4.39 | 95.45 |
| | Fractal | 83.36 | 13.49 | 83.82 |
| | PCA | 66.74 | 37.04 | **44.50** |
| SVM | Statistical | 98.67 | 30.92 | 68.66 |
| | Fractal | 91.93 | 35.64 | 61.23 |
| | PCA | 93.35 | 55.92 | **40.10** |

The proposed PCA features exhibited high noise tolerance properties and managed to achieve reasonably good classification accuracy when used with ANN and SVM classifiers under different noise levels. This is due to changes to the original PD data due to noise are minimized while transforming the PD data from a higher dimension to a lower dimension in the PCA process. Thus, these result in classification accuracy using PCA features and intelligent classifiers to be less affected by different durations of noise signals compared to statistical and fractal features. Previous works only used PCA as a data reduction method. However, in this work, it is shown that PCA features performed directly on the PRPD data is very beneficial with its high noise tolerance properties.

ANFIS has slower training speed compared to ANN and SVM and the training is heavily affected by input feature size. Under noise-free condition, it performed similar to ANN. However, under noise contamination, its classification accuracy fell greatly and experienced a sharper decline in classification accuracy compared to ANN and SVM. It is less efficient in handling input feature with huge variation. Furthermore, ANFIS is incompatible with the proposed high noise tolerance PCA features due to the normalization process. Therefore, it is concluded that ANFIS is not suitable for PD classification with noise contamination.

## 4.8 Chapter Summary

Measurement results from a PD detector have been displayed in this chapter. The details of the noise signals used and PD pattern measured from the test samples have been presented. The classification accuracy of all input features and classifiers has been elaborated. From the results, it is clear that the proposed PCA input features managed to display very good noise tolerance in all tested noisy conditions. The feasibility of ANFIS as a PD classifier has been evaluated and it is not recommended as a PD classifier due to its slow training time and inability to recognize PD pattern accurately under noisy conditions.

# CHAPTER 5: CONCLUSION & FUTURE WORK

## 5.1 Conclusion

Based on the findings from the literature review, PD classification under noise-free condition is well established with many works proposing numerous input features and classifier. They were able to achieve more than 90% classification accuracy when using lab fabricated materials or artificial PD source. However, the real challenge is to replicate that results in noise contaminated scenario as noise interference is a real problem in practical PD measurement. Therefore, this research work has successfully filled the gap in PD classification researches, where actual XLPE cable joints were used as a test sample and PD classification was done under noisy conditions using actual noise obtained from ground interference. A novel high noise tolerance PCA feature has also been successfully proposed.

Five XLPE cable joints with artificial defects have been successfully prepared and PD measurement was successfully performed after noise isolation precautions have been taken. Using the PD measurement data, three feature extraction methods, which are statistical features, fractal features and the proposed PCA features have been performed. Three different artificial intelligence based classifiers, which are ANN, ANFIS and SVM were used to classify the PD pattern based on the extracted input feature.

After analyzing the performance of each classifier and input feature, the tests were repeated with contaminated signals. Under noise-free condition, all PD classifiers and input features performed very well with the exception of ANFIS and PCA features. This is due to ANFIS requires normalizing the input during training process for efficiency purpose. This will lead to higher error because normalization will cause PCA feature to lose its relative significance.

Two different noisy conditions were tested, noise with increasing duration or pulse count and noise with increasing average charge magnitude. Under noisy conditions, all PD classifiers and input features suffer from degradation in classification accuracy. However, the proposed PCA features displayed the highest noise tolerance, suffering only a minimal 20% reduction under the most severe noise contamination scenario while other input features suffered must more severe reduction. PD classifiers were also tested against noise with increasing average charge amplitude, which pushed the PD classifiers to the limit as the charge magnitudes are higher than the average PD pulse of the test samples. Once again, the proposed PCA features emerged as the input feature with the highest classification accuracy.

For PD classifier performance, ANFIS is weaker than ANN and SVM. Under noise-free condition, ANFIS performed equally well as ANN. However, when noise contamination was introduced, ANFIS classification drops drastically as it is unable to cope with huge variation of input feature that have been contaminated. ANFIS also had very long training time when larger input features were used. Apart from that, ANFIS was incompatible with the proposed PCA input features, making it highly susceptible to noise contamination.

In noise-free condition, SVM performed better than ANN and it also had faster training speed. Under noisy conditions, it was found that SVM was better at adapting to noise with high pulse count while ANN was better at adapting to noise with high average charge magnitude. SVM experienced less performance reduction when contaminated with noise with increasing noise pulse count compared to ANN. On the contrary, ANN experienced less performance reduction using all three input features when contaminated with noise with increasing average charge magnitude compared to SVM.

In conclusion, all objectives of this research work have been successfully achieved. PD classification of actual XLPE cables has been performed under noisy conditions and the proposed PCA features were found to have high noise tolerance and may benefit power industries in practical PD classification and condition monitoring.

## 5.2 Future work

Future work that can be performed are:

1. Design a PD classifier that is capable of recognizing overlapping PRPD patterns from two different defects. This is due to there are more than one defect that may exist in a test object. Hence, a PD classifier that is able to identify multiple defect sources will be extremely useful. It will also make the PD recognition system more versatile with multiple defect recognition capabilities.
2. Use a more effective input feature that has higher accuracy and invariant to noise contamination. The current method has much higher noise tolerance compared to traditional input features. However, the classification accuracy under noisy conditions is lower compared to classification accuracy under noise-free conditions. Hence, it would be good to bridge the gap of the input feature performance between under noise-free and noisy conditions.

## REFERENCES


Abdel-Galil, T., Hegazy, Y. G., Salama, M. M. A., & Bartnikas, R. (2004). Partial discharge pulse pattern recognition using Hidden Markov Models. *IEEE Transactions on Dielectrics and Electrical Insulation, 11*(4), 715-723. doi:10.1109/TDEI.2004.1324361

Abdel-Galil, T., Sharkawy, R. M., Salama, M. M. A., & Bartnikas, R. (2005). Partial discharge pulse pattern recognition using an inductive inference algorithm. *IEEE Transactions on Dielectrics and Electrical Insulation, 12*(2), 320-327. doi:10.1109/TDEI.2005.1430400

Abdel-Galil, T. K., Hegazy, Y. G., Salama, M. M. A., & Bartnikas, R. (2005). Fast match-based vector quantization partial discharge pulse pattern recognition. *IEEE Transactions on Instrumentation and Measurement, 54*(1), 3-9. doi:10.1109/TIM.2004.839762

Abdel-Galil, T. K., Sharkawy, R. M., Salama, M. M. A., & Bartnikas, R. (2005). Partial discharge pattern classification using the fuzzy decision tree approach. *IEEE Transactions on Instrumentation and Measurement, 54*(6), 2258-2263. doi:10.1109/TIM.2005.858143

Abubakar Mas'ud, A., Stewart, B. G., & McMeekin, S. G. (2014). Application of an ensemble neural network for classifying partial discharge patterns. *Electric Power Systems Research, 110*(0), 154-162. doi:http://dx.doi.org/10.1016/j.epsr.2014.01.010

Al-geelani, N. A., M. Piah, M. A., & Bashir, N. (2015). A review on hybrid wavelet regrouping particle swarm optimization neural networks for characterization of partial discharge acoustic signals. *Renewable and Sustainable Energy Reviews, 45*, 20-35. doi:http://dx.doi.org/10.1016/j.rser.2015.01.047

Al-geelani, N. A., Piah, M. A. M., Adzis, Z., & Algeelani, M. A. (2013). Hybrid regrouping PSO based wavelet neural networks for characterization of acoustic signals due to surface discharges on H.V. glass insulators. *Applied Soft Computing, 13*(12), 4622-4632. doi:http://dx.doi.org/10.1016/j.asoc.2013.07.011

Al-geelani, N. A., Piah, M. A. M., & Shaddad, R. Q. (2012). Characterization of acoustic signals due to surface discharges on H.V. glass insulators using wavelet radial basis function neural networks. *Applied Soft Computing, 12*(4), 1239-1246. doi:http://dx.doi.org/10.1016/j.asoc.2011.12.018

Allahbakhshi, M., & Akbari, A. (2011). A method for discriminating original pulses in online partial discharge measurement. *Measurement, 44*(1), 148-158. doi:http://dx.doi.org/10.1016/j.measurement.2010.09.036

Angrisani, L., Daponte, P., Lupò, G., Petrarca, C., & Vitelli, M. (2000). Analysis of ultrawide-band detected partial discharges by means of a multiresolution digital signal-processing method. *Measurement, 27*(3), 207-221. doi:http://dx.doi.org/10.1016/S0263-2241(99)00067-6

Ardila-Rey, J. A., Martinez-Tarifa, J. M., & Robles, G. (2015). Automatic selection of frequency bands for the power ratios separation technique in partial discharge measurements: part II, PD source recognition and applications. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(4), 2293-2301. doi:10.1109/TDEI.2015.004822

Babnik, T., Aggarwal, R., & Moore, P. (2007). Data mining on a transformer partial discharge data using the self-organizing map. *IEEE Transactions on Dielectrics and Electrical Insulation, 14*(2), 444-452. doi:10.1109/TDEI.2007.344626

Bentley, P. M., & McDonnell, J. T. E. (1994). Wavelet transforms: an introduction. *Electronics & Communication Engineering Journal, 6*(4), 175-186. doi:10.1049/ecej:19940401

Biswas, S., Dey, D., Chatterjee, B., & Chakravorti, S. (2013). An approach based on rough set theory for identification of single and multiple partial discharge source. *International Journal of Electrical Power & Energy Systems, 46*(0), 163-174. doi:http://dx.doi.org/10.1016/j.ijepes.2012.10.050

Boczar, T., Borucki, S., Cichon, A., & Zmarzly, D. (2009). Application possibilities of artificial neural networks for recognizing partial discharges measured by the acoustic emission method. *IEEE Transactions on Dielectrics and Electrical Insulation, 16*(1), 214-223. doi:10.1109/TDEI.2009.4784570

Boggs, S. A. (1990). Partial discharge. II. Detection sensitivity. *Electrical Insulation Magazine, IEEE, 6*(5), 35-42. doi:10.1109/57.63081

Boya, C., Ruiz-Llata, M., Posada, J., & Garcia-Souto, J. A. (2015). Identification of multiple partial discharge sources using acoustic emission technique and blind source separation. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(3), 1663-1673. doi:10.1109/TDEI.2015.7116363

Cachin, C., & Wiesmann, H. J. (1995). PD recognition with knowledge-based preprocessing and neural networks. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(4), 578-589. doi:10.1109/94.407023

Candela, R., Mirelli, G., & Schifani, R. (2000). PD recognition by means of statistical and fractal parameters and a neural network. *IEEE Transactions on Dielectrics and Electrical Insulation, 7*(1), 87-94. doi:10.1109/94.839345

Canxin, G., Li, Z., Yong, Q., Chengjun, H., Hui, W., Linpeng, Y., & Xiuchen, J. (2009, 20-22 Nov. 2009). *Application of adaptive neuro fuzzy inference system to the partial discharge pattern recognition.* Paper presented at the Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on.

Carrasco, M., López, J., & Maldonado, S. (2015). A multi-class SVM approach based on the l1-norm minimization of the distances between the reduced convex hulls. *Pattern Recognition, 48*(5), 1598-1607. doi:http://dx.doi.org/10.1016/j.patcog.2014.12.006

Casals-Torrens, P., Gonz áez-Parada, A., & Bosch-Tous, R. (2012). Online PD detection on high voltage underground power cables by acoustic emission. *Procedia Engineering, 35,* 22-30. doi:10.1016/j.proeng.2012.04.161

Chalashkanov, N., Kolev, N., Dodd, S., & Fothergill, J. C. (2008, 26-29 Oct. 2008). *PD Pattern Recognition Using ANFIS.* Paper presented at the Electrical Insulation and Dielectric Phenomena, 2008. CEIDP 2008. Annual Report Conference on.

Chang, C., Chang, C. S., Jin, J., Hoshino, T., Hanai, M., & Kobayashi, N. (2005). Source classification of partial discharge for gas insulated substation using waveshape pattern recognition. *IEEE Transactions on Dielectrics and Electrical Insulation, 12*(2), 374-386. doi:10.1109/TDEI.2005.1430405

Chang, C. S., Jin, J., Chang, C., Hoshino, T., Hanai, M., & Kobayashi, N. (2005). Separation of corona using wavelet packet transform and neural network for detection of partial discharge in gas-insulated substations. *IEEE Transactions on Power Delivery, 20*(2), 1363-1369. doi:10.1109/TPWRD.2004.839187

Chen, H.-C., & Gu, F.-C. (2012). Pattern recognition with cerebellar model articulation controller and fractal features on partial discharges. *Expert Systems with Applications, 39*(7), 6575-6584. doi:10.1016/j.eswa.2011.12.044

Chen, H.-C., Gu, F.-C., & Wang, M.-H. (2012). A novel extension neural network based partial discharge pattern recognition method for high-voltage power apparatus. *Expert Systems with Applications, 39*(3), 3423-3431. doi:10.1016/j.eswa.2011.09.030

Chen, H. C. (2012). Fractal features-based pattern recognition of partial discharge in XLPE power cables using extension method. *Generation, Transmission & Distribution, IET, 6*(11), 1096-1103. doi:10.1049/iet-gtd.2012.0080

Cho, Y.-S., Shim, M.-J., & Kim, S.-W. (1998). Electrical tree initiation mechanism of artificial defects filled XLPE. *Materials Chemistry and Physics, 56*(1), 87-90. doi:http://dx.doi.org/10.1016/S0254-0584(98)00160-6

Chul-Hwan, K., & Aggarwal, R. (2000). Wavelet transforms in power systems. I. General introduction to the wavelet transforms. *Power Engineering Journal, 14*(2), 81-87. doi:10.1049/pe:20000210

Contin, A., Cavallini, A., Montanari, G. C., Pasini, G., & Puletti, F. (2002). Digital detection and fuzzy classification of partial discharge signals. *IEEE Transactions on Dielectrics and Electrical Insulation, 9*(3), 335-348. doi:10.1109/TDEI.2002.1007695

Contin, A., Montanari, G. C., & Ferraro, C. (2000). PD source recognition by Weibull processing of pulse height distributions. *IEEE Transactions on Dielectrics and Electrical Insulation, 7*(1), 48-58. doi:10.1109/94.839341

Danouj, B., Tahan, S. A., & David, E. (2013). Using a new generation of piezoelectric sensors for partial discharge detection. *Measurement, 46*(1), 660-666. doi:http://dx.doi.org/10.1016/j.measurement.2012.09.005

Darabad, V. P., Vakilian, M., Phung, B. T., & Blackburn, T. R. (2013). An efficient diagnosis method for data mining on single PD pulses of transformer insulation defect models. *IEEE Transactions on Dielectrics and Electrical Insulation, 20*(6), 2061-2072. doi:10.1109/TDEI.2013.6678854

Dey, D., Chatterjee, B., Chakravorti, S., & Munshi, S. (2010). Cross-wavelet transform as a new paradigm for feature extraction from noisy partial discharge pulses. *IEEE Transactions on Dielectrics and Electrical Insulation, 17*(1), 157-166. doi:10.1109/tdei.2010.5412014

Evagorou, D., Kyprianou, A., Lewin, P. L., Stavrou, A., Efthymiou, V., Metaxas, A. C., & Georghiou, G. E. (2010). Feature extraction of partial discharge signals using the wavelet packet transform and classification with a probabilistic neural network. *Science, Measurement & Technology, IET, 4*(3), 177-192. doi:10.1049/iet-smt.2009.0023

Fard, M. A., Akbari, A., Shojaee, R., Mirzaei, H. R., & Naderi, P. (2010, 4-9 July 2010). *Partial discharge defects classification using neuro-fuzzy inference system.* Paper presented at the Solid Dielectrics (ICSD), 2010 10th IEEE International Conference on.

Feng-Chang, G., Hong-Chan, C., & Cheng-Chien, K. (2013). Gas-insulated switchgear PD signal analysis based on Hilbert-Huang transform with fractal parameters enhancement. *IEEE Transactions on Dielectrics and Electrical Insulation, 20*(4), 1049-1055. doi:10.1109/TDEI.2013.6571416

Grossberg, S. (1969). Embedding fields: A theory of learning with physiological implications. *J. Mathematical Psychology, 6*, 209-239.

Gu, F.-C., Chang, H.-C., Chen, F.-H., & Kuo, C.-C. (2012). Partial discharge pattern recognition of power cable joints using extension method with fractal feature enhancement. *Expert Systems with Applications, 39*(3), 2804-2812. doi:10.1016/j.eswa.2011.08.140

Gu, F. C., Chang, H. C., Chen, F. H., Kuo, C. C., & Hsu, C. H. (2012). Application of the Hilbert-Huang transform with fractal feature enhancement on partial discharge recognition of power cable joints. *Science, Measurement & Technology, IET, 6*(6), 440-448. doi:10.1049/iet-smt.2011.0213

Gulski, E. (1993). Computer-aided measurement of partial discharges in HV equipment. *IEEE Transactions on Electrical Insulation, 28*(6), 969-983. doi:10.1109/14.249370

Gulski, E. (1995). Digital analysis of partial discharges. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(5), 822-837. doi:10.1109/94.469977

Gulski, E., & Kridva, A. (1995). Influence of aging on classification of PD in HV components. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(4), 676-684. doi:10.1109/94.407032

Gulski, E., & Krivda, A. (1993). Neural networks as a tool for recognition of partial discharges. *IEEE Transactions on Electrical Insulation, 28*(6), 984-1001. doi:10.1109/14.249372

Guo, H., & Wang, W. (2015). An active learning-based SVM multi-class classification model. *Pattern Recognition, 48*(5), 1577-1597. doi:http://dx.doi.org/10.1016/j.patcog.2014.12.009

Hamilton, D. J., & Pearson, J. S. (1997). Classification of partial discharge sources in gas-insulated substations using novel preprocessing strategies. *Science, Measurement and Technology, IEE Proceedings -, 144*(1), 17-24. doi:10.1049/ip-smt:19970862

Hao, L., & Lewin, P. L. (2010). Partial discharge source discrimination using a support vector machine. *IEEE Transactions on Dielectrics and Electrical Insulation, 17*(1), 189-197. doi:10.1109/TDEI.2010.5412017

Hao, L., Lewin, P. L., Hunter, J. A., Swaffield, D. J., Contin, A., Walton, C., & Michel, M. (2011). Discrimination of multiple PD sources using wavelet decomposition and principal component analysis. *IEEE Transactions on Dielectrics and Electrical Insulation, 18*(5), 1702-1711. doi:10.1109/TDEI.2011.6032842

Hao, Z., Blackburn, T. R., Phung, B. T., & Sen, D. (2007a). A novel wavelet transform technique for on-line partial discharge measurements. 1. WT de-noising algorithm. *IEEE Transactions on Dielectrics and Electrical Insulation, 14*(1), 3-14. doi:10.1109/tdei.2007.302864

Hao, Z., Blackburn, T. R., Phung, B. T., & Sen, D. (2007b). A novel wavelet transform technique for on-line partial discharge measurements. 2. On-site noise rejection application. *IEEE Transactions on Dielectrics and Electrical Insulation, 14*(1), 15-22. doi:10.1109/tdei.2007.302865

Harbaji, M., Shaban, K., & El-Hag, A. (2015). Classification of common partial discharge types in oil-paper insulation system using acoustic signals. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(3), 1674-1683. doi:10.1109/TDEI.2015.7116364

Hecht-Nielsen, R. (1988). Application of counter propagation networks. *Neural Networks, 1*, 131-139.

Hong, T., Fang, M. T. C., & Hilder, D. (1996). PD classification by a modular neural network based on task decomposition. *IEEE Transactions on Dielectrics and Electrical Insulation, 3*(2), 207-212. doi:10.1109/94.486772

Hoof, M., Freisleben, B., & Patsch, R. (1997). PD source identification with novel discharge parameters using counterpropagation neural networks. *IEEE Transactions on Dielectrics and Electrical Insulation, 4*(1), 17-32. doi:10.1109/94.590861

Hucker, T., & Kranz, H. G. (1995). New approach in partial discharge diagnosis and pattern recognition. *Science, Measurement and Technology, IEE Proceedings -, 142*(1), 89-94. doi:10.1049/ip-smt:19951431

Hui, M., Chan, J. C., Saha, T. K., & Ekanayake, C. (2013). Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources. *IEEE Transactions on Dielectrics and Electrical Insulation, 20*(2), 468-478. doi:10.1109/TDEI.2013.6508749

Hung-Cheng, C. (2013). Partial discharge identification system for highvoltage power transformers using fractal featurebased extension method. *Science, Measurement & Technology, IET, 7*(2), 77-84. doi:10.1049/iet-smt.2012.0078

Hunter, J. A., Lewin, P. L., Hao, L., Walton, C., & Michel, M. (2013). Autonomous classification of PD sources within three-phase 11 kV PILC cables. *IEEE Transactions on Dielectrics and Electrical Insulation, 20*(6), 2117-2124. doi:10.1109/TDEI.2013.6678860

Ibrahim, K., Sharkawy, R. M., Salama, M. M. A., & Bartnikas, R. (2012). Realization of partial discharge signals in transformer oils utilizing advanced computational techniques. *IEEE Transactions on Dielectrics and Electrical Insulation, 19*(6), 1971-1981. doi:10.1109/TDEI.2012.6396955

IEEE Guide for Partial Discharge Testing of Shielded Power Cable Systems in a Field Environment. (2007). *IEEE Std 400.3-2006*, 1-44. doi:10.1109/IEEESTD.2007.305045

James, R. E., & Phung, B. T. (1995). Development of computer-based measurements and their application to PD pattern analysis. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(5), 838-856. doi:10.1109/94.469978

Jang, J. S. R. (1993a). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics, 23*(3), 665-685. doi:10.1109/21.256541

Jang, J. S. R. (1993b). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on, 23*(3), 665-685. doi:10.1109/21.256541

Jang, J. S. R. (1994, 26-29 Jun 1994). *Structure determination in fuzzy modeling: a fuzzy CART approach.* Paper presented at the Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on.

Jian, L., Caixin, S., Grzybowski, S., & Taylor, C. D. (2006). Partial discharge image recognition using a new group of features. *IEEE Transactions on Dielectrics and Electrical Insulation, 13*(6), 1245-1253. doi:10.1109/TDEI.2006.258196

Jian, L., Tianyan, J., Harrison, R. F., & Grzybowski, S. (2012). Recognition of ultra high frequency partial discharge signals using multi-scale features. *IEEE Transactions on Dielectrics and Electrical Insulation, 19*(4), 1412-1420. doi:10.1109/TDEI.2012.6260018

Jin, J., Chang, C. S., Chang, C., Hoshino, T., Hanai, M., & Kobayashi, N. (2006). Classification of partial discharge events in gas-insulated substations using

wavelet packet transform and neural network approaches. *Science, Measurement and Technology, IEE Proceedings, 153*(2), 55-63. doi:10.1049/ip-smt:20045036

Jing, C., & Hou, J. (2015). SVM and PCA based fault classification approaches for complicated industrial process. *Neurocomputing, 167*, 636-642. doi:http://dx.doi.org/10.1016/j.neucom.2015.03.082

Kai, G., Kexiong, T., Fuqi, L., & Chengqi, W. (2002). PD pattern recognition for stator bar models with six kinds of characteristic vectors using BP network. *IEEE Transactions on Dielectrics and Electrical Insulation, 9*(3), 381-389. doi:10.1109/TDEI.2002.1007700

Karthikeyan, B., Gopal, S., & Venkatesh, S. (2006). ART 2—an unsupervised neural network for PD pattern recognition and classification. *Expert Systems with Applications, 31*(2), 345-350. doi:http://dx.doi.org/10.1016/j.eswa.2005.09.029

Karthikeyan, B., Gopal, S., & Venkatesh, S. (2008). Partial discharge pattern classification using composite versions of probabilistic neural network inference engine. *Expert Systems with Applications, 34*(3), 1938-1947. doi:http://dx.doi.org/10.1016/j.eswa.2007.02.005

Karthikeyan, B., Gopal, S., & Vimala, M. (2005). Conception of complex probabilistic neural network system for classification of partial discharge patterns using multifarious inputs. *Expert Systems with Applications, 29*(4), 953-963. doi:http://dx.doi.org/10.1016/j.eswa.2005.06.014

Ke, W., JinZhong, L., Shuqi, Z., Fei, G., Huanchao, C., Rui, L., . . . Grzybowski, S. (2015). A new image-oriented feature extraction method for partial discharges. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(2), 1015-1024. doi:10.1109/TDEI.2015.7076803

Ke, W., Jinzhong, L., Shuqi, Z., Ruijin, L., Feifei, W., Lijun, Y., . . . Jiaming, Y. (2015). A hybrid algorithm based on s transform and affinity propagation clustering for separation of two simultaneously artificial partial discharge sources. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(2), 1042-1060. doi:10.1109/TDEI.2015.7076806

Khan, Y., Khan, A. A., Budiman, F. N., Beroual, A., Malik, N. H., & Al-Arainy, A. A. (2014). Partial discharge pattern analysis using support vector machine to estimate size and position of metallic particle adhering to spacer in GIS. *Electric Power Systems Research, 116*, 391-398. doi:http://dx.doi.org/10.1016/j.epsr.2014.07.001

Kranz, H. G. (1993). Diagnosis of partial discharge signals using neural networks and minimum distance classification. *IEEE Transactions on Electrical Insulation, 28*(6), 1016-1024. doi:10.1109/14.249375

Kreuger, F. H. (1991). *Industrial High Voltage* (Vol. 1): Delft University Press.

Kreuger, F. H., Gulski, E., & Krivda, A. (1993). Classification of partial discharges. *IEEE Transactions on Electrical Insulation, 28*(6), 917-931. doi:10.1109/14.249365

Krivda, A., Gulski, E., Satish, L., & Zaengl, W. S. (1995). The use of fractal features for recognition of 3-D discharge patterns. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(5), 889-892. doi:10.1109/94.469983

Kundu, P., Kishore, N. K., & Sinha, A. K. (2012). Identification of two simultaneous partial discharge sources in an oil-pressboard insulation system using acoustic emission techniques. *Applied Acoustics, 73*(4), 395-401. doi:http://dx.doi.org/10.1016/j.apacoust.2011.11.004

Lai, K. X., Phung, B. T., & Blackburn, T. R. (2010). Application of data mining on partial discharge part I: predictive modelling classification. *IEEE Transactions on Dielectrics and Electrical Insulation, 17*(3), 846-854. doi:10.1109/TDEI.2010.5492258

Lalitha, E. M., & Satish, L. (1998). Fractal image compression for classification of PD sources. *IEEE Transactions on Dielectrics and Electrical Insulation, 5*(4), 550-557. doi:10.1109/94.708272

Lalitha, E. M., & Satish, L. (2000). Wavelet analysis for classification of multi-source PD patterns. *IEEE Transactions on Dielectrics and Electrical Insulation, 7*(1), 40-47. doi:10.1109/94.839339

Liping, L., Ju, T., & Yilu, L. (2015). Partial discharge recognition in gas insulated switchgear based on multi-information fusion. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(2), 1080-1087. doi:10.1109/TDEI.2015.7076809

Liu, Y., You, Z., & Cao, L. (2006). A novel and quick SVM-based multi-class classifier. *Pattern Recognition, 39*(11), 2258-2264. doi:http://dx.doi.org/10.1016/j.patcog.2006.05.034

Ma, X., Zhou, C., & Kemp, I. J. (2002a). Automated wavelet selection and thresholding for PD detection. *Electrical Insulation Magazine, IEEE, 18*(2), 37-45. doi:10.1109/57.995398

Ma, X., Zhou, C., & Kemp, I. J. (2002b). Interpretation of wavelet analysis and its application in partial discharge detection. *IEEE Transactions on Dielectrics and Electrical Insulation, 9*(3), 446-457. doi:10.1109/TDEI.2002.1007709

Majidi, M., Fadali, M. S., Etezadi-Amoli, M., & Oskuoee, M. (2015). Partial discharge pattern recognition via sparse representation and ANN. *IEEE Transactions on Dielectrics and Electrical Insulation, 22*(2), 1061-1070. doi:10.1109/TDEI.2015.7076807

Majidi, M., & Oskuoee, M. (2015). Improving pattern recognition accuracy of partial discharges by new data preprocessing methods. *Electric Power Systems Research, 119*, 100-110. doi:http://dx.doi.org/10.1016/j.epsr.2014.09.014

Mang-Hui, W. (2005). Partial discharge pattern recognition of current transformers using an ENN. *IEEE Transactions on Power Delivery, 20*(3), 1984-1990. doi:10.1109/TPWRD.2005.848441

Mang-Hui, W., & Chih-Yung, H. (2005). Application of extension theory to PD pattern recognition in high-voltage current transformers. *IEEE Transactions on Power Delivery, 20*(3), 1939-1946. doi:10.1109/TPWRD.2005.848673

Mazroua, A. A., Bartnikas, R., & Salama, M. M. A. (1994). Discrimination between PD pulse shapes using different neural network paradigms. *IEEE Transactions on Dielectrics and Electrical Insulation, 1*(6), 1119-1131. doi:10.1109/94.368651

Mazroua, A. A., Bartnikas, R., & Salama, M. M. A. (1995). Neural network system using the multi-layer perceptron technique for the recognition of PD pulse shapes due to cavities and electrical trees. *IEEE Transactions on Power Delivery, 10*(1), 92-96. doi:10.1109/61.368411

Mazroua, A. A., Salama, M. M. A., & Bartnikas, R. (1993). PD pattern recognition with neural networks using the multilayer perceptron technique. *IEEE Transactions on Electrical Insulation, 28*(6), 1082-1089. doi:10.1109/14.249382

Mazzetti, C., Mascioli, F. M. F., Baldini, F., Panella, M., Risica, R., & Bartnikas, R. (2006). Partial discharge pattern recognition by neuro-fuzzy networks in heat-shrinkable joints and terminations of XLPE insulated distribution cables. *IEEE Transactions on Power Delivery, 21*(3), 1035-1044. doi:10.1109/tpwrd.2006.875861

Meijer, S., Gulski, E., & Smit, J. J. (1998). Pattern analysis of partial discharges in SF6 GIS. *IEEE Transactions on Dielectrics and Electrical Insulation, 5*(6), 830-842. doi:10.1109/94.740764

Mota, H. d. O., Rocha, L. C. D. d., Salles, T. C. d. M., & Vasconcelos, F. H. (2011). Partial discharge signal denoising with spatially adaptive wavelet thresholding and support vector machines. *Electric Power Systems Research, 81*(2), 644-659. doi:http://dx.doi.org/10.1016/j.epsr.2010.10.030

Perpiñán, O., Sánchez-Urán, M. A., Álvarez, F., Ortego, J., & Garnacho, F. (2013). Signal analysis and feature generation for pattern identification of partial discharges in high-voltage equipment. *Electric Power Systems Research, 95*, 56-65. doi:10.1016/j.epsr.2012.08.016

Rahman, M. K. A., Arora, R., & Srivastava, S. C. (2000). Partial discharge classification using principal component transformation. *Science, Measurement and Technology, IEE Proceedings -, 147*(1), 7-13. doi:10.1049/ip-smt:20000074

Rizzi, A., Mascioli, F. M. F., Baldini, F., Mazzetti, C., & Bartnikas, R. (2009). Genetic Optimization of a PD Diagnostic System for Cable Accessories. *IEEE Transactions on Power Delivery, 24*(3), 1728-1738. doi:10.1109/tpwrd.2009.2016826

Sahoo, N. C., Salama, M. M. A., & Bartnikas, R. (2005). Trends in partial discharge pattern classification: a survey. *IEEE Transactions on Dielectrics and Electrical Insulation, 12*(2), 248-264. doi:10.1109/tdei.2005.1430395

Salama, M. M. A., & Bartnikas, R. (2000). Fuzzy logic applied to PD pattern classification. *IEEE Transactions on Dielectrics and Electrical Insulation, 7*(1), 118-123. doi:10.1109/94.839349

Salama, M. M. A., & Bartnikas, R. (2002). Determination of neural-network topology for partial discharge pulse pattern recognition. *IEEE Transactions on Neural Networks, 13*(2), 446-456. doi:10.1109/72.991430

Satish, L., & Gururaj, B. I. (1993a). Partial discharge pattern classification using multilayer neural networks. *Science, Measurement and Technology, IEE Proceedings A, 140*(4), 323-330.

Satish, L., & Gururaj, B. I. (1993b). Use of hidden Markov models for partial discharge pattern classification. *IEEE Transactions on Electrical Insulation, 28*(2), 172-182. doi:10.1109/14.212242

Satish, L., & Nazneen, B. (2003). Wavelet-based denoising of partial discharge signals buried in excessive noise and interference. *IEEE Transactions on Dielectrics and Electrical Insulation, 10*(2), 354-367. doi:10.1109/TDEI.2003.1194122

Satish, L., & Zaengl, W. S. (1994). Artificial neural networks for recognition of 3-d partial discharge patterns. *IEEE Transactions on Dielectrics and Electrical Insulation, 1*(2), 265-275. doi:10.1109/94.300259

Satish, L., & Zaengl, W. S. (1995). Can fractal features be used for recognizing 3-d partial discharge patterns. *IEEE Transactions on Dielectrics and Electrical Insulation, 2*(3), 352-359. doi:10.1109/94.395421

Sharkawy, R. M., Mangoubi, R. S., Abdel-Galil, T., Salama, M. M. A., & Bartnikas, R. (2007). SVM classification of contaminating particles in liquid dielectrics using higher order statistics of electrical and acoustic PD measurements. *IEEE Transactions on Dielectrics and Electrical Insulation, 14*(3), 669-678. doi:10.1109/tdei.2007.369530

Shim, I., Soraghan, J. J., & Siew, W. H. (2001). Detection of PD utilizing digital signal processing methods. Part 3: Open-loop noise reduction. *Electrical Insulation Magazine, IEEE, 17*(1), 6-13. doi:10.1109/57.901611

Shurrab, I., El-Hag, A., Assaleh, K., Ghunem, R., & Jayaram, S. (2013). RF-based monitoring and classification of partial discharge on wet silicone rubber surface. *IEEE Transactions on Dielectrics and Electrical Insulation, 20*(6), 2188-2194. doi:10.1109/TDEI.2013.6678869

Si, W., Li, J., Yuan, P., & Li, Y. (2008). Digital detection, grouping and classification of partial discharge signals at DC voltage. *IEEE Transactions on Dielectrics and Electrical Insulation, 15*(6), 1663-1674. doi:10.1109/TDEI.2008.4712671

Sinaga, H. H., Phung, B. T., & Blackburn, T. R. (2010, 27-29 Oct. 2010). *Neuro fuzzy recognition of ultra-high frequency partial discharges in transformers.* Paper presented at the IPEC, 2010 Conference Proceedings.

Sinaga, H. H., Phung, B. T., & Blackburn, T. R. (2014). Recognition of single and multiple partial discharge sources in transformers based on ultra-high frequency signals. *Generation, Transmission & Distribution, IET, 8*(1), 160-169. doi:10.1049/iet-gtd.2013.0131

Specht, D. F. (1988, 24-27 July 1988). *Probabilistic neural networks for classification, mapping, or associative memory.* Paper presented at the Neural Networks, 1988., IEEE International Conference on.

Specht, D. F. (1990). Probabilistic neural networks and the polynomial Adaline as complementary techniques for classification. *IEEE Transactions on Neural Networks, 1*(1), 111-121. doi:10.1109/72.80210

Sriram, S., Nitin, S., Prabhu, K. M. M., & Bastiaans, M. J. (2005). Signal denoising techniques for partial discharge measurements. *IEEE Transactions on Dielectrics and Electrical Insulation, 12*(6), 1182-1191. doi:10.1109/TDEI.2005.1561798

Su, M.-S., Chia, C.-C., Chen, C.-Y., & Chen, J.-F. (2014). Classification of partial discharge events in GILBS using probabilistic neural networks and the fuzzy c-means clustering approach. *International Journal of Electrical Power & Energy Systems, 61*, 173-179. doi:http://dx.doi.org/10.1016/j.ijepes.2014.03.054

Suzuki, H., & Endoh, T. (1992). Pattern recognition of partial discharge in XLPE cables using a neural network. *IEEE Transactions on Electrical Insulation, 27*(3), 543-549. doi:10.1109/14.142717

Tao, H., & Fang, M. T. C. (2001). Detection and classification of partial discharge using a feature decomposition-based modular neural network. *IEEE Transactions on Instrumentation and Measurement, 50*(5), 1349-1354. doi:10.1109/19.963209

Tian, Y., Lewin, P. L., & Davies, A. E. (2002). Comparison of on-line partial discharge detection methods for HV cable joints. *IEEE Transactions on Dielectrics and Electrical Insulation, 9*(4), 604-615. doi:10.1109/TDEI.2002.1024439

Venkatesh, S., & Gopal, S. (2011a). Orthogonal least square center selection technique – A robust scheme for multiple source Partial Discharge pattern recognition using Radial Basis Probabilistic Neural Network. *Expert Systems with Applications, 38*(7), 8978-8989. doi:http://dx.doi.org/10.1016/j.eswa.2011.01.115

Venkatesh, S., & Gopal, S. (2011b). Robust Heteroscedastic Probabilistic Neural Network for multiple source partial discharge pattern recognition – Significance of outliers on classification capability. *Expert Systems with Applications, 38*(9), 11501-11514. doi:http://dx.doi.org/10.1016/j.eswa.2011.03.026

Wang, M.-H., Tseng, Y.-F., Chen, H.-C., & Chao, K.-H. (2009). A novel clustering algorithm based on the extension theory and genetic algorithm. *Expert Systems with Applications, 36*(4), 8269-8276. doi:http://dx.doi.org/10.1016/j.eswa.2008.10.010

Xiaodi, S., Chengke, Z., Hepburn, D. M., Guobin, Z., & Michel, M. (2007). Second generation wavelet transform for data denoising in PD measurement. *IEEE*

*Transactions on Dielectrics and Electrical Insulation, 14*(6), 1531-1537. doi:10.1109/TDEI.2007.4401237

Xiaoxing, Z., Song, X., Na, S., Ju, T., & Wei, L. (2014). GIS partial discharge pattern recognition based on the chaos theory. *IEEE Transactions on Dielectrics and Electrical Insulation, 21*(2), 783-790. doi:10.1109/TDEI.2013.004020

Xu, Z., Tang, J., & Sun, C. (2007). Application of Complex Wavelet Transform to Suppress White Noise in GIS UHF PD Signals. *IEEE Transactions on Power Delivery, 22*(3), 1498-1504. doi:10.1109/tpwrd.2007.899767

Yu, H., & Song, Y. H. (2003). Using improved self-organizing map for partial discharge diagnosis of large turbogenerators. *IEEE Transactions on Energy Conversion, 18*(3), 392-399. doi:10.1109/TEC.2003.815834

Yun, J., Hong, M., Junhua, L., Yu, L., Hai, L., Xiaojuan, J., . . . Wenjie, L. (2010, 13-16 Sept. 2010). *Partial discharge pattern Characteristic of MV Cable joints with artificial defect.* Paper presented at the Electricity Distribution (CICED), 2010 China International Conference on.

Zhou, X., Zhou, C., & Kemp, I. J. (2005). An improved methodology for application of wavelet transform to partial discharge measurement denoising. *IEEE Transactions on Dielectrics and Electrical Insulation, 12*(3), 586-594. doi:10.1109/TDEI.2005.1453464

Ziomek, W., Reformat, M., & Kuffel, E. (2000). Application of genetic algorithms to pattern recognition of defects in GIS. *IEEE Transactions on Dielectrics and Electrical Insulation, 7*(2), 161-168. doi:10.1109/94.841804

## LIST OF PUBLICATIONS AND PAPERS PRESENTED

List of ISI paper publications:

1. W.J.K. Raymond, H. A. Illias, A. H. A. Bakar, and H. Mokhlis, "Partial Discharge Classifications: Review Of Recent Progress," *Measurement*, Volume 68, pp. 164-181, February 2015 (ISI-Cited Publication).

2. W.J.K. Raymond, H. A. Illias, A. H. A. Bakar, and H. Mokhlis, "Classification of Partial Discharge Measured under Different Levels of Noise Contamination," Plos One, under review since February 2016.

3. W.J.K. Raymond, H. A. Illias, A. H. A. Bakar, and H. Mokhlis, "A High Noise Tolerance Feature Extraction for Partial Discharge Classification in XLPE Cable Joints," IEEE Transactions on Dielectrics and Electrical Insulation, accepted in August 2016.

List of conference paper publications:

1. W.J.K. Raymond, H. A. Illias, H. Mokhlis, A. H. A. Bakar, " Investigation Of Partial Discharge Severity At Xlpe Cable Without Termination," *IEEE International Conference on Power and Energy*, pp. 13-16, 1-3 Dec. 2014, Kuching, Malaysia.