Article

Comparison of the Performance of Artificial Neural Networks and Fuzzy Logic for Recognizing Different Partial Discharge Sources

Abdullahi Abubakar Mas'ud [1,*], Jorge Alfredo Ardila-Rey [2], Ricardo Albarracín [3] , Firdaus Muhammad-Sukki [4] and Nurul Aini Bani [5]

[1] Department of Electrical and Electronics Engineering, Jubail Industrial College, P.O. Box 10099, Jubail 31961, Saudi Arabia
[2] Department of Electrical Engineering, Universidad Técnica Federico Santa María, Santiago de Chile 8940000, Chile; jorge.ardila@usm.cl
[3] Departamento de Ingeniería Eléctrica, Electrónica, Automática y Física Aplicada, Escuela Técnica Superior de Ingeniería y Diseño Industrial, Universidad Politécnica de Madrid, Ronda de Valencia 3, 28012 Madrid, Spain; rasbarracin@gmail.com
[4] School of Engineering, Robert Gordon University, Aberdeen AB10 7GJ, UK; f.b.muhammad-sukki@rgu.ac.uk
[5] UTM Razak School of Engineering and Advanced Technology, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia; nurulaini.kl@utm.my
* Correspondence: masud_a@jic.edu.sa; Tel.: +966-538-138-814

Abstract: This paper compared the capabilities of the artificial neural network (ANN) and the fuzzy logic (FL) approaches for recognizing and discriminating partial discharge (PD) fault classes. The training and testing parameters for the ANN and FL comprise statistical fingerprints from different phase-amplitude-number (ϕ-q-n) measurements. Two PD fault classes considered are internal discharges in voids and surface discharges. In the void class, there are single voids, serial voids and parallel voids in polyethylene terephthalate (PET), while the surface discharge class comprises four different surface discharge arrangements on pressboard in oil at different voltages and angular positioning of the ground electrode on the respective pressboards. Previously, the ANN and FL have been investigated for PD classification, but there is no work reported in the literature that compares their performance, specifically when applied for real time PD detection problem. As expected, both the ANN and FL can recognize PD defect classes, but the results show that the ANN appears to be more robust as compared to the FL, but these conclusions required to be further investigated with complex PD examples. Finally, both the ANN and FL were assessed as practical PD classification. Despite of the limitations of the ANN, it is concluded that the ANN is better suited for practical PD recognition because of its ability to provide accurate recognition values and the severity level of PD defects.

Keywords: partial discharge; fuzzy logic; artificial neural network; polyethylene terephthalate

1. Introduction

One technique for examining failures in the insulation of high-voltage (HV) equipment is through the monitoring evaluation of partial discharges (PDs). PDs are well known electrical discharge phenomena that occur within the insulation system of HV power apparatus [1,2]. These discharges represent low energy degradation phenomena taking place in regions where the insulation dielectric strength is very low compared to the other materials. PDs can occur irrespective of the sources that produce them due to mechanical, thermal, electrical and environmental stresses, and once present,

they become principal mechanism for insulation degradation and in the long run can cause complete failure of the system, leading to potentially significant capital cost and financial consequences [3–5]. The nature and characteristics of PD have so far been established and evaluated [2,5], but what remains of interest nowadays is identifying enhanced techniques that can effectively identify and discriminate different PD sources and the noise associated with them such as voids in electrical insulation materials, surface discharges and corona [6–8]. This is important in order to provide a reliable assessment of the HV insulation condition and the nature of the PD fault [9,10].

During the early days, PD recognition was carried out by visual examination on a power frequency ellipse [11]. In this case, the positive half of the AC cycle ellipse starts at the 0° marker and goes clockwise around the ellipse to 180° while the negative half starts at 180° and goes around the lower half of the ellipse [11]. The position of the PD pulse and its magnitude are displayed on the ellipse. One drawback of this technique is that it can be complicated to implement due to interference from electrical equipment noise and radio frequency interference [9,11]. Afterwards, the availability of computer-assisted devices provides automatic recognition process and this led to the application of phased-resolved partial discharge (PRPD) pattern observations for evaluating the PD [8,12]. The typical trend in the PRPD patterns is to record the PD signals as a 3-D array in function of (ϕ-q-n) where ϕ is the phase angle, q is the amplitude and n is the number of discharges [12]. These patterns can be captured over different power cycles thereby creating new opportunities for an improved evaluation of the PD faults. It was then established that different PD sources produce different PRPD patterns [11,13,14]. Thereafter, the goal has identified an efficient technique for recognizing PD patterns using the expert systems. This comprises the artificial neural network (ANN) [11,14–17], FL [18,19], wavelet analysis [20,21], and support vector machines [22] among others. It is interesting to note that these techniques recorded recognition performance up to 90% for a number of cases of PD sources.

Based on the aforementioned techniques, the first task in PD recognition is selecting PD patterns, which produce excellent discriminating capabilities and where the PRPD patterns have been established. Secondly, feature extraction is carried out using statistical tools. These tools provide a well-defined evaluation of the PRPD patterns [11,13]. The work reported in references [9,23] shows that the application of many statistical parameters provides good recognition capability using the ANN, together with development and evaluation of the performance of difference combinations of the statistical fingerprints.

It is evident that a number of papers have reported the recognition and discrimination of PD patterns, but little work has been done to investigate the capability of these pattern recognition tools to recognize single or multiple identical PD sources and their degradation levels at various locations within the insulation system. This is vital in order to find out the variation in their insulation characteristics. Following this approach, Abubakar Mas'ud [3] reported that PD source positions relative to each other and to the ground have an effect on the PRPD patterns. Therefore, this paper investigates and compares the robustness of the ANN and the FL in recognizing and discriminating different PD fault positions within different HV insulation systems. Previously, ANN and FL have been compared in recognizing different dataset in other disciplines, but there appears to be no literature reports that compare their performance regarding PD classification. Ben Salah and Ouali [24] compared the ANN and FL techniques for maximum power point tracking (MPPT) of PV systems and their results shows that the FL deliver up to 7% more power than that one obtained by applying ANN. A similar comparison is carried out by Albedin et al. in [25], where the MPPT is obtained with ANN and FL algorithms under variable conditions. Their results show that the FL controller appears to be easier to implement for MPPT than the ANN, but its major setback is a slow transient performance and fluctuations. However, the ANN shows better performance and was able to show accurate estimate of maximum power generation under various conditions. Another research presented in [26] compares the ANN and FL performance for predicting a number of day's compressive strength of concretes containing low-lime and high-lime ashes. The results show that both, the ANN and FL, demonstrates strong potential for recognition, with none outperforming the other.

In this paper, ANN and FL are considered because research findings indicate that both are techniques already integrated for commercial services in PD recognition [27]. PD faults of interest include single and multiple voids in PET and surface discharges on pressboard in oil commonly used as insulation for oil-immersed transformers.

Voids are the most common PD sources in underground cable insulation and slot insulations in electrical machines while surface discharge are mostly associated with the oil insulation of transformers [5]. In this paper, multiple voids are considered because HV insulation may contain more than one void. Therefore, it is important to understand the discharge patterns from different void arrangements to establish whether there exist variations in the PD patterns in terms of uniqueness and statistics. In particular, to investigate whether any pattern recognition tool can capture slight variations of similar PD fault types.

2. Experimental Set-Up

2.1. The PD Detection System

In this research, the PD measurement circuit employed is in accordance with the IEC 60270 standard [28]. The HV AC test system is made up of a HV transformer, over voltage protection, current/voltage regulating devices, an AC measurement capacitor (C_1) of 1 nF, partial discharge free up to 100 kV, used as a low impedance path for PD pulses, a measurement impedance (Z) used for the HV AC signal measurement, and the PD acquisition system. A schematic diagram is shown in Figure 1.
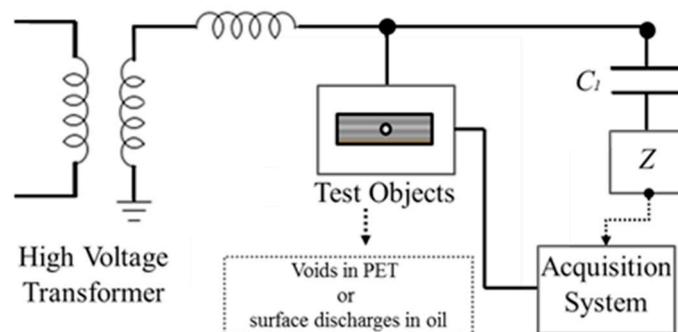

Figure 1. Diagram of the HV PD detection set-up.

The PD detection system offers capabilities such as calibration settings, data capture control and saving path and the basic function control. It can produce power cycle synchronized PRPD patterns in real time. It also has functionalities for data logging patterns. This is important for this investigation because it is needed to capture and store PD data over long stressing period to monitor the different levels of degradation that may be affected by environmental factors such as temperature, pressure etc. Since the PD measurement gives the amplitude in volts, there is a need for calibration to obtain the actual value in Coulombs. Therefore, prior to the real experimentation and testing of the standard PD geometry, the traditional calibration technique was employed to establish a scale factor relating the response of the measuring system to the level of the discharge pulse PD injector calibration pulse of amplitude levels 500 pC, 2000 pC and 10,000 pC are applied. To calibrate the PD, a known PD pulse is injected into the HV system when it is not energized and the amplitude level in mV is monitored using the PD detection system. To determine the calibration factor for the experiments, the peak amplitude level of the PD in mV corresponds to the injected pulse in pC.

2.2. PD Test Samples

One of the targets of this investigation is to produce appropriate physical PD models of voids in PET, surface discharges in oil and then produce relevant PRPD patterns. These patterns will be

processed using statistical tools for feeding as input to the ANNs and FL. The models for investigation are shown in Figure 2 and are described as follows:
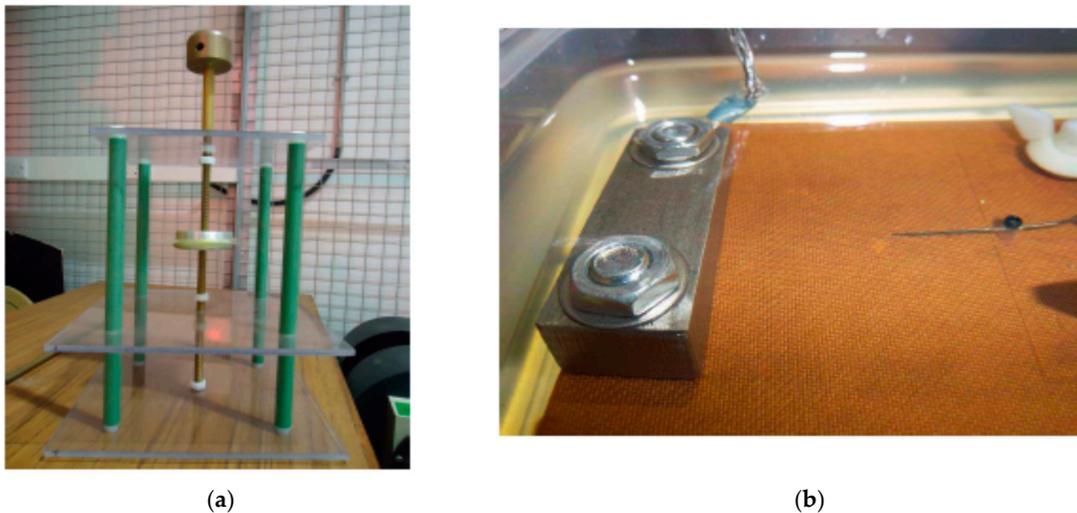

(**a**)

(**b**)

Figure 2. Photograph of the test set-up for (**a**) voids; and (**b**) surface discharge in oil.

(a) Single void (void 1): In this test object, nine PET sheets were sandwiched between the HV electrode and the ground. A single cylindrical void of $D = 0.6$ mm and $t = 50$ μm is located at the centre of the PET layers as shown in Figure 3a. The inception voltage for void 1 is approximately 2.82 kV. PD data was captured over 250 cycles from the start of the experiment through a 7 h continuous degradation period. However, to guarantee that the experimental set-up is discharge free prior to the samples with voids being investigated, PET layers without any voids were initially stressed to confirm the validity of discharge free PET environments [3]. Optimal external pressure is maintained on all samples to ensure an ideal specimen without any small voids on the edges and any other twist that can create a void.

(b) Parallel void (void 2): This test arrangement is similar to void 1, except that two voids with the same dimension (i.e., $D = 0.6$ mm and $t = 50$ μm) are created adjacent to each other in a horizontal form as shown in Figure 3b. The inception voltage for void 2 is approximately 3.2 kV.

(c) Serial void (void 3): This test arrangement is also similar to void 1, except that two voids with the same sizes (i.e., $D = 0.6$ mm and $t = 50$ μm) are created adjacent to each other in a vertical form as shown in Figure 3c. The inception voltage is approximately 2.82 kV. It can be seen that the inception voltage for void 2 is greater than that of void 1 and void 3 due to the influence of the electric field on the different void arrangements.

(d) Surface discharge in oil 1 (surf 1): Surface discharge along an oil-pressboard interface is investigated by means of an experimental test arrangement as shown in Figure 2b. In this test arrangement, a needle was placed at an angle of 10° to a pressboard surface, with the needle tip at a predetermined distance (d) = 25 mm from an earth electrode also placed on the pressboard in oil. The applied voltage is 18.5 kV. In order to have good representative of PD patterns for evaluation were captured continuously from onset up to 7 h of continuous stress.

(e) Surface discharge in oil 2 (surf 2): Test arrangement is similar to surf 1 except that $d = 45$ mm.

(f) Surface discharge in oil 3 (surf 3): This test arrangements is also identical to surf 1 except that $d = 25$ mm and applied voltage of 30 kV.

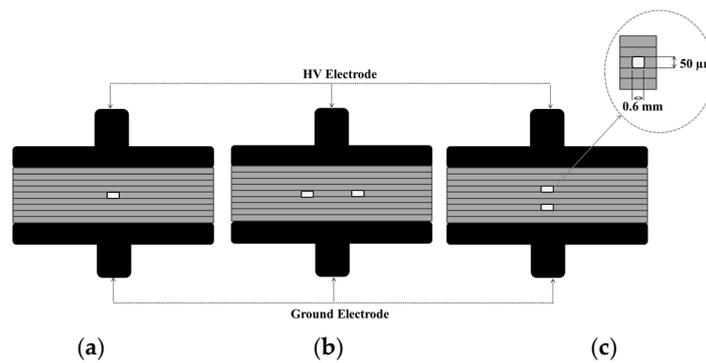(g) Surface discharge in oil 4 (surf 4): This test object is similar to surf 3 except that $d = 45$ mm.

Figure 3. Different arrangement of void in PET: (**a**) void 1; (**b**) void 2; and (**c**) void 3.

Different surface discharge samples were investigated in order to understand and quantify the variations that may be seen in the PRDP patterns from various surface discharge arrangements in oil when the ANN and fuzzy logic approaches were applied.

3. PD Pattern Recognition Tools

3.1. Artificial Neural Networks

ANNs are mathematical models that emulate the way humans classify patterns, learn tasks and solve problems [29,30]. The structure of an ANN consists of the input layer, middle layer (i.e., hidden layer) and output layer. The number of outputs of an ANN is directly related to the number of classes to be distinguished. Each layer of the ANN consists of one or more neurons, which compute the sum of the incoming signals, and passes it to a non-linear squashing function, e.g., hyperbolic function or sigmoid function, to give an output [29]. These neurons in the ANN are connected to each other by synapses which weights are assigned [29,30].

An example of the learning process in an ANN applied to the identification of PD is shown in Figure 4. In this context, the input vectors are the statistical fingerprints derived from different PRPD patterns and the outputs are a combination of 0 s and 1 s to differentiate the PD. The ANN can have many layers and normally has sigmoid functions in the hidden layer. The sigmoid function is the most commonly used function in the construction of the ANN because of its asymptotic properties and can normally squash input parameter to a range 0 to 1 [30]. There is no certain criterion for selecting the number of neurons in the hidden layer, but enough neurons are needed to obtain a very good performance [17]. Generally, in designing and training the ANN, certain considerations have to be taken to be able to get the best result, i.e., by choosing the number and types of neurons in the hidden layer and finding the best solution to avoid local minima in the error space [29]. Local minima are a sudden termination of the training error curve resulting from instability of the ANN [29].

In the majority of ANNs including PD, at the beginning of training, the weights are randomly chosen within the range of −0.1 to 0.1, while the bias values are threshold values initially chosen to be 1 and can changed depending on the training pattern [11]. They are then adjusted based on the difference between the output value and the target value, according to a certain training algorithms [11]. Basically, the inputs are fed into the ANN and the weights continuously adjusted until the error predefined (output value) reaches the minimum acceptable value. Among the training algorithms for the ANN, the back-propagation (BP) is the most widely used for PD classification and recognition rate as high as 90% have been recorded for some unseen PD samples [4,11,15]. The BP is simpler to implement and more efficient in determining the gradients in ANNs. The BP algorithm is a kind of supervised learning [29]. There are basically four steps in implementing the BP ANN [31]. These include the feed-forward propagation, BP to the output layer, BP to the hidden layer and weight updates [32]. The BP algorithm is typically used for updating weights and biases of ANN and allows reducing the

mean square error (MSE). The adjustment results, for training and testing phases, have been chosen according with the knowledge acquired from previous works, such as [3,14].

After training enough samples and continuously amending the connection weights, the final weight values and threshold values in the neurons are obtained to indicate the correct information. The neuron refers to an information-processing element important for the function of the ANN. The inputs are connected to the neuron and multiplied by certain weights. There is also an adder for summing the input signals weighted by the respective synapse. Next is the activation function that limits the amplitude of the output of the neuron, normalized to either [0, 1] or [1, −1] at the output. In the literature, Gulski and Krivda [11] have compared the ANN using the BP for identify and discriminate different arrangements of two electrode PD samples. They were able to demonstrate that with the BP ANN, recognition rate of PD reached as high as 100%.

Similar to Gulski, Abubakar Mas'ud [3] and Abubakar Mas'ud et al. [14] have tried in several attempts to recognize a number of PD patterns of corona and surface discharges using the ANN. Statistical metrics were applied as input to the ANN and recognition rate up to 90% was recorded for some PD fault geometries.
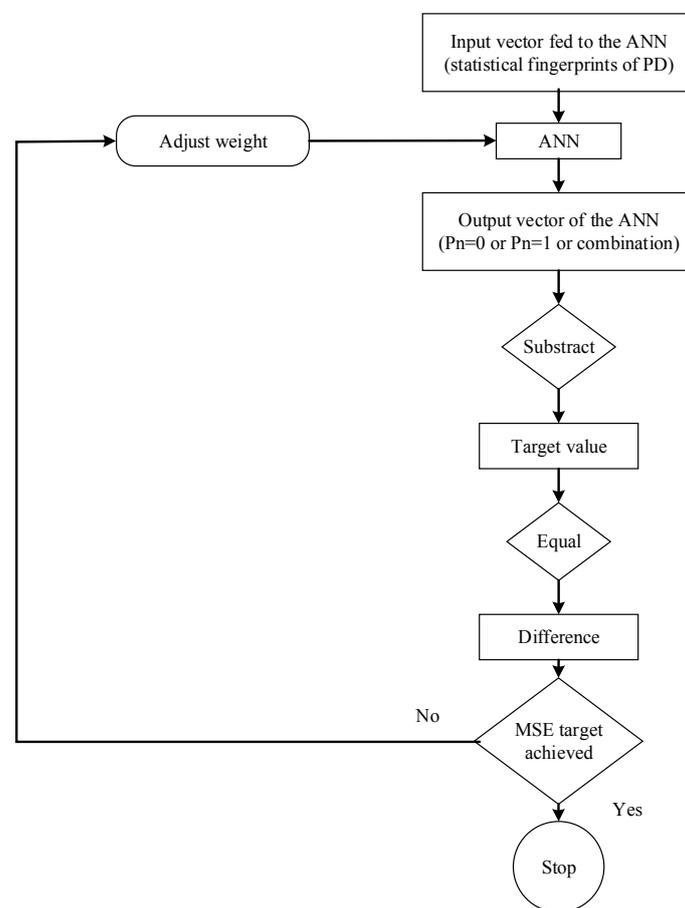

Figure 4. A learning algorithm in an artificial neural network (ANN) model.

3.2. Fuzzy Logic

The FL system is a knowledge-based or rule based system [33]. Generally, it is simply a nonlinear mapping of an input fingerprint set to a scalar output fingerprint [34]. The main component of the fuzzy system is a knowledge base made up of fuzzy IF-THEN rules. These IF-THEN statements are distinguished by certain membership functions (MF) that represent the degree of which a member

belongs to a set as follows: MF = 1 perfect membership; MF = 0 no membership, and MF = (0, 1) partial membership [35].

Figure 5a shows an example of MF, where the classes are numbers that are negative large (NL), negative medium (NM), negative small (NS), close to zero, positive small (PS), positive medium (PM) and positive large (PL). μ refers to the membership in a particular class or set. A FL comprises of four fundamental parts [36]: fuzzifier, rules, inference engine, and defuzzifier, as shown in Figure 5b. The fuzzy logic algorithm is as follows:

(a)　Fuzzifier: In this case, crisp input set is collected and then converted to fuzzy set using certain linguistic parameters, terms and membership functions. There are basically three types of fuzzifier namely singleton fuzzifier, Gaussian fuzzifier and trapezoidal (or triangular fuzzifier) [33]. In the singleton fuzzifier, the inputs are converted into fuzzy singletons. It simplifiers the fuzzy computations but cannot eliminate noise [33]. The Gaussian and trapezoidal are commonly used for pattern recognition because they can suppress noise and simplify fuzzy calculations based certain membership functions [34,35]. In this paper, trapezoidal fuzzifier will be applied to recognize PD.

(b)　The inference is carried out based on certain fuzzy rules. In the inference stage, a number of fuzzy operators (OR, AND, NOT) are applied in the IF part of the rule. This is important in order to define input and output behaviours and choosing the minimum number of variables that are applied to the fuzzy logic machine.

(c)　Finally, the defuzzification step, where the corresponding fuzzy output is mapped to a crisp output based on the membership functions. Many types of defuzzification exist which include the maximum defuzzification technique, centre of gravity and the bisector method. The maximum defuzzification technique is the most common and it selects the output with the maximum membership function [33,36].
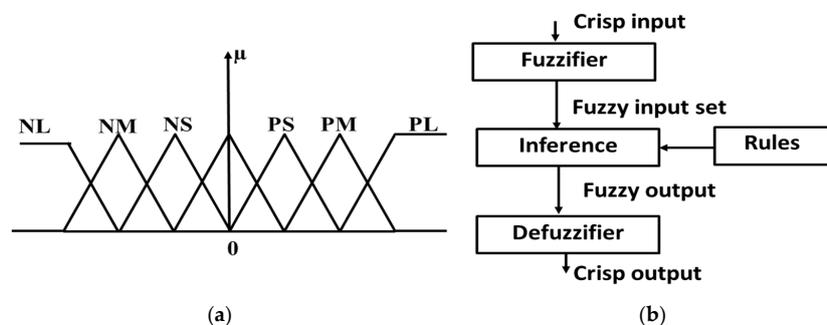

(a)　　　　　　　　　　　　(b)

Figure 5. (**a**) Membership Functions for different classes (N = Negative, P = Positive, L = Large, M = Medium, S = Small, NL = Negative Large, NM = Negative Medium, NS =Negative Small, PS = Positive Small, PM = Positive Medium, PL =Positive Large) (Adapted from [34]); and (**b**) a fuzzy inference system.

Previously, few papers have reported successful classification of PD using FL applied to classify PD [37–39], with the most recent one being the fuzzy art maps [39]. Gopal et al. [37] used different PRPD patterns (internal void, surface discharge, oil corona and corona) to represent the crisp number set of the fuzzy logic. Different membership functions were analysed (e.g., trapezoidal, triangular and bell shape) in order to obtain the optimum generalization and classification. Secondly, the corresponding values of q, ϕ and n represent the crisp number sub-set. Membership functions are defined for each subset and fuzzy rules are applied in order to classify the PD e.g., if ϕ is very low and n is very high and q is medium then the PD represent an internal void, i.e., $\phi U_q U_n$ = void. The PD logic classification chart as defined in the literature is summarized in Figure 6a. Prior to classification, the low, high and medium values for each subset are defined based on the understanding of the parameters of each

PD fault. Chen et al. [38] investigated the application of the statistical parameters for fuzzy logic classification. The optimum statistical fingerprints with higher discriminating capability of PD patterns were initially determined by comparing the 25%, 50% and 75% quartiles of the different PD faults.
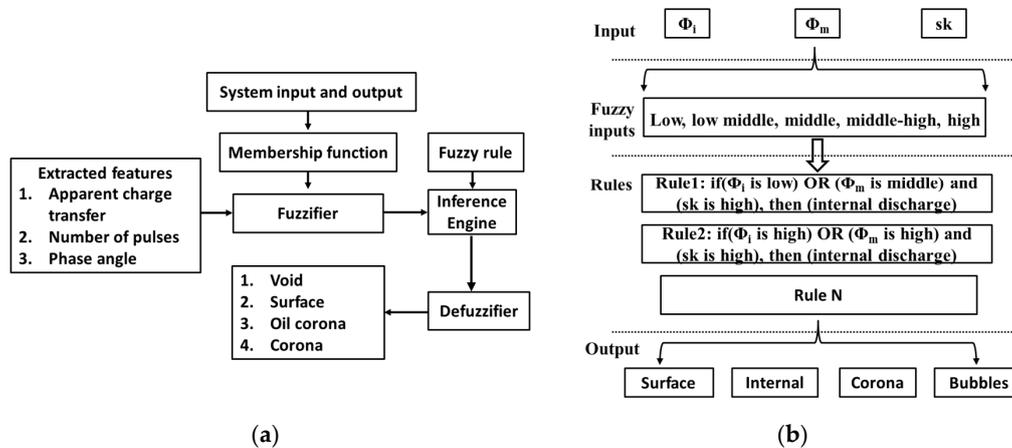

Figure 6. A fuzzy inference system to classify PD patterns as used in the literature (**a**) adapted from [35]; and (**b**) adapted from [36].

The four artificially created PD faults investigated include internal discharge, surface discharge, corona and discharges from bubbles in oil. Afterwards, Φ_i (inception phase), Φ_m (mean phase) and *sk* (skewness) were determined to be the most discriminating markers for fuzzy logic identification. Figure 6b shows the fuzzy logic classification tool as adopted by Chen. The PD markers (fingerprints) were fuzzified into five linguistic attributes i.e., low, low medium, medium, medium-high and high. Fuzzy rules for PD inference were determined based on the statistical parameters confidence limits in order to obtain the correct diagnosis of the PD fault. In this paper, similar algorithms as implemented by Chen et al. [35] will be developed for PD classification using the FL.

4. PRPD Patterns and Statistical Parameters for the ANN and FL

4.1. PRPD Patterns for PD Faults

The PRPD represented as *ϕ-q-n* plots for all void arrangements at the initial degradation stage are shown in Figure 7, while that for the surface discharges are shown in Figure 8. From Figure 8, it is obvious that there may be statistical variability among the three void arrangements. In terms of the *ϕ-q-n* distributions, void 1 and void 3 appear to be similar, but are different from void 2. The number of discharges occurring at larger amplitudes appears to be higher in the void 2 situation when compared to void-3. However, this depends on the distance between the two voids in the void 2 arrangement [3]. When PD occurs, the closely spaced void 2 arrangement affect the field in each of the voids due to electrostatic interaction, while more distanced spaced voids in void 2 arrangement have little influence on the electric field in the voids themselves [40]. Generally, it is expected that when the number of voids is increased, the number of PD per cycle increases. Theoretically, PD patterns from different arrangement of voids in any HV insulation may vary due to change in temperature, air pressure within the void and fabrication tolerance of the PET sheets, which may affect the PD discharge amplitude levels [3]. However, for a practical cable (e.g., cross-linked polyethylene (XLPE)), the partial discharge inception voltage (PDIV) may vary from these experiments because of certain cable parameters considered, i.e., insulation medium, geometry (conductor radius), local conditions (distance of void to conductor, rated voltage, temperature, internal pressure), level of degradation (void size) and their travelling paths up to reach the sensors [41].

For tests involving a point to ground on pressboard in oil (Figure 8a–d), the applied voltage, gap distance and angular positioning of the point discharge source play a significant part in the PD repetition rate and oil-pressboard degradation process. Although the PRPD pattern shown illustrates only the initial degradation, it is clear that the higher the applied voltage, the higher the discharge intensity. From the experiment, it was observed that distinguishing aspect between 25 mm and 45 mm point-ground gaps on pressboard in oil is the spread of the tracking marks on the pressboard surface due to sustained intense PD. This result is an indication of variability in the PD patterns and statistical fingerprints defining the surface discharge in oil.
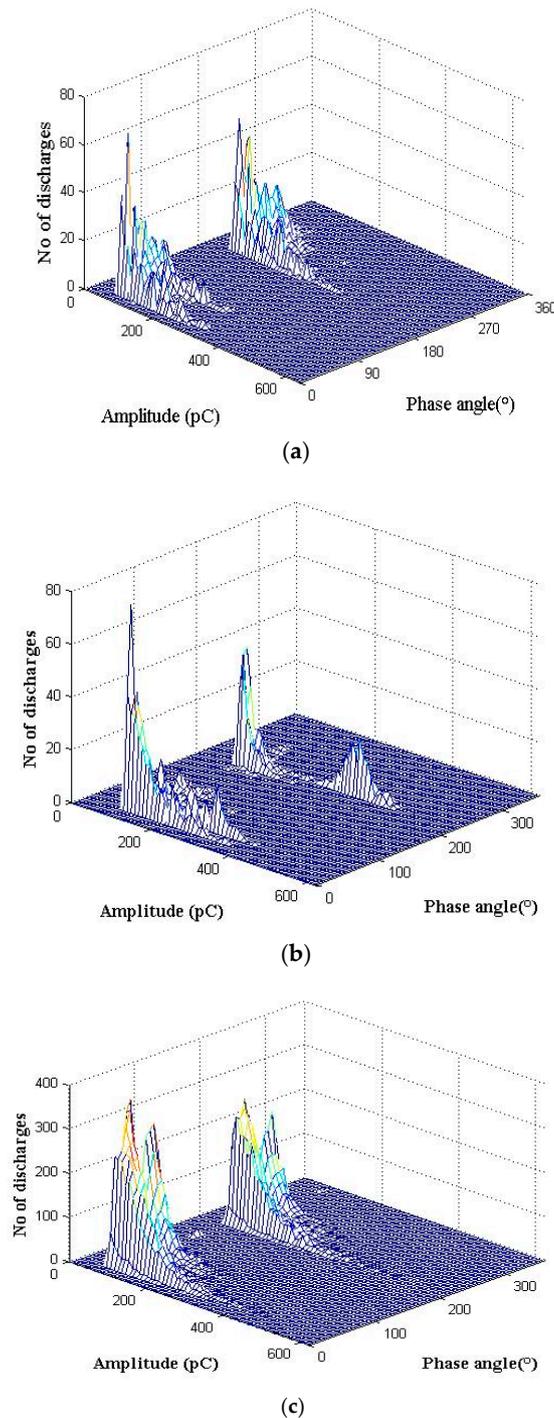
(**a**)

(**b**)

(**c**)

Figure 7. Sample PRPD pattern for: (**a**) void 1; (**b**) void 2; and (**c**) void 3.
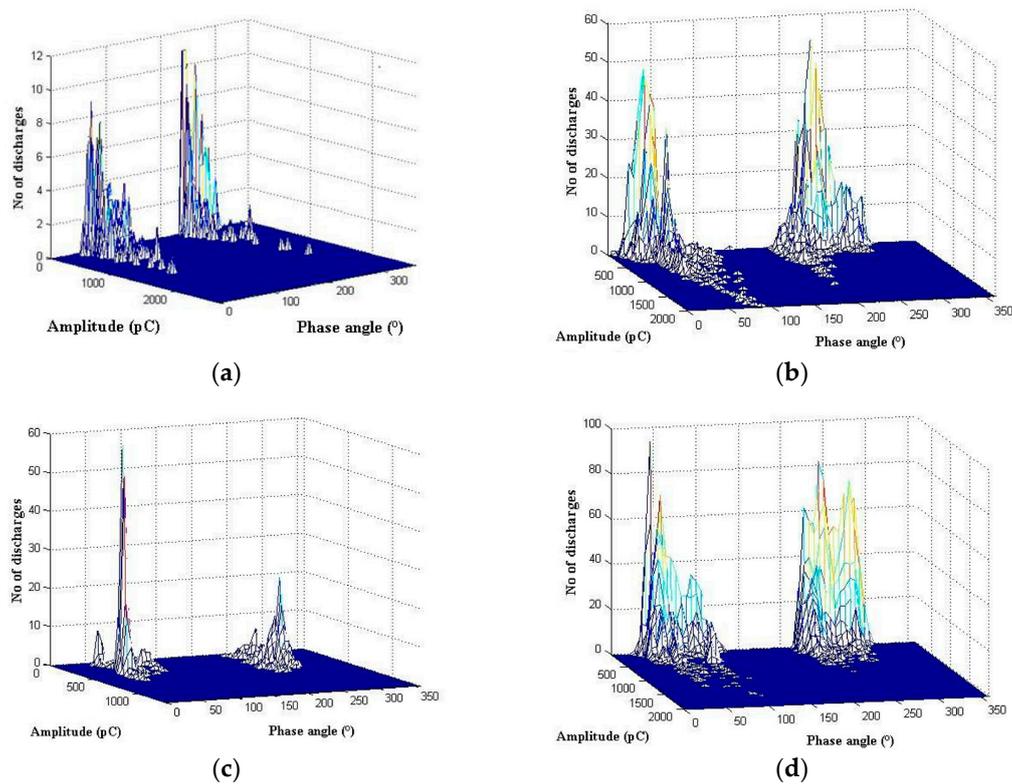
Figure 8. *ϕ-q-n* pattern for surface discharges in oil (**a**) *v* = 18.5 kV, *d* = 45 mm; (**b**) *v* = 30 kV, *d* = 45 mm, in oil; (**c**) *v* = 18.5 kV, *d* = 25 mm and (**d**) *v* = 30 kV, *d* = 25 mm.

4.2. Statistical Fingerprints

For this investigation, numerous PRPD pattern fingerprint structures are obtained so as to determine the appropriate input parameter vector to which the ANN and FL can learn from. Since PD events that transpire in the dielectric material are essentially complex random phenomena that demonstrates considerable statistical inconsistency in characteristics, i.e., pulse magnitude, appearance and time of event. The PD behaviour can also be either continuous or random process, which reveals that statistical measures (operators) should be the most reliable choice for analysis. In this paper, in order to simplify the pattern recognition process, the *ϕ-q-n* plots are obtained and transformed into 2-D plots. It is complex to analyse the *ϕ-q-n* plots in the 3D and therefore 2D plots are commonly used. The 2D plots chosen for this investigation are the phase-number histograms ($H_n(\phi)$), amplitude-number histograms ($H_n(q)$) and average amplitude-number histograms ($H_{qn}(\phi)$). These distributions are evaluated in both the positive and negative half power of the AC cycle. Example of the *ϕ-q-n* (2D) derived pattern is shown in Figure 9.
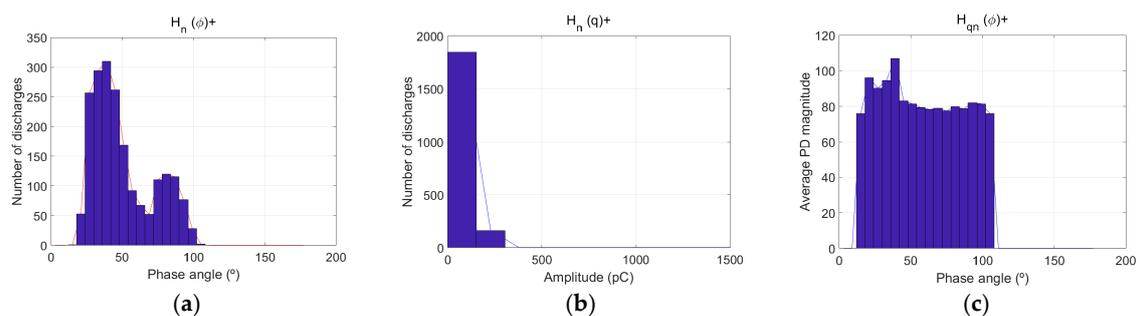

Figure 9. Example of 2D (**a**) $H_n(\phi)$; (**b**) $H_n(q)$ and (**c**) $H_{qn}(\phi)$ derived plots for single void.

Previously, encouraging performance was recorded by means of a number of aforementioned statistical tools as training and testing set for the ANN [11]. In agreement with the literature this paper also utilizes 15 statistical fingerprints from the *φ-q-n* patterns. These include: skewness (sk) and kurtosis (ku) of the $H_n(\phi)+$, $H_n(\phi)-$, $H_{qn}(\phi)+$, $H_{qn}(\phi)-$, $H_n(q)+$ and $H_n(q)-$ distributions. The mathematical representations for these statistical fingerprints is shown in Table 1, where μ is the average value, σ is the standard deviation, m represents the size of the data and P_j is the probability of the discrete value x_j and y_j as the case may be. Q_S^+ and Q_S^- represent the sum of discharge amplitudes in both the *+ve* and negative half power cycles. Similarly N_S^+ and N_S^- represent the number of discharges in both the *+ve* and *−ve* half power cycle.

The $H_n(\phi)+$, $H_n(\phi)-$ distributions represent pulse count distribution (positive and negative half cycle) in phase while the $H_{qn}(\phi)+$, $H_n(q)-$ plots are the mean pulse height distribution (+*ve* and −*ve* half cycle) in phase. Similarly, the $H_n(q)+$ and $H_n(q)-$ are the pulse amplitude distribution (+*ve* and −*ve* half cycle) in amplitude.

The discharge factor (*Q*), cross-correlation (cc) and the modified cross-correlation (mcc) also form part of the statistical variables as input to the ANN and FL. *sk* measures the asymmetry of each of the 2D PRPD derived distributions in comparison to the normal distribution. *ku* measures the sharpness or evenness of the distributions. The cross-correlation (cc) determines the variation in shape between any two $H_{qn}(\phi)$ plots in the positive and negative half power cycles. *Q* determines the difference in the average discharge level of any two $H_{qn}(\phi)$ plots.

Figures 10 and 11 shows the statistical variability of selected 2D plots for the voids and surface discharge over a 7 h degradation period. From Figure 10, it is clear that *sk* ($H_{qn}(\phi)+$) and *Q* for single and serial voids follow identical patterns which correlates with their PD mechanism as described in Section 4.1. However there is no clear variability between the surface discharges in Figure 11, clearly showing that there is no clear pattern for surface discharges degradation over different gap distances or applied voltage, what is more, basically little *φ-q-n* pattern variation may exist between them.

Table 1. Mathematical expressions of statistical operators.

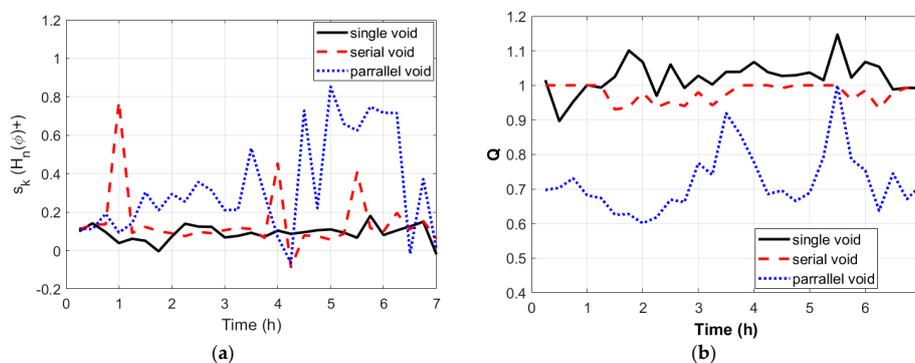| Statistical Operator	Mathematical Equation
Skewness	$sk = \dfrac{\sum (x_j-\mu)^3 P_j}{\sigma^3}$
Kurtosis	$ku = \dfrac{\sum (x_j-\mu)^4 P_j}{\sigma^4}$
Discharge Factor	$Q = \dfrac{Q_S^-/N_S^-}{Q_S^+/N_S^+}$
Cross-correlation (cc)	$cc = \dfrac{\sum x_j y_j - \frac{\sum x_j \sum y_j}{n}}{\sqrt{\left[\sum x_j^2 - \frac{(\sum x_j)^2}{n}\right]\left[\sum y_j^2 - \frac{(\sum y_j)^2}{n}\right]}}$
Modified cross-correlation	$Q\ cc$

Figure 10. Statistical variation of single void (void 1), serial void (void 2) and parallel void (void 3) (**a**) *sk* ($H_{qn}(\phi)+$) and (**b**) *Q*.
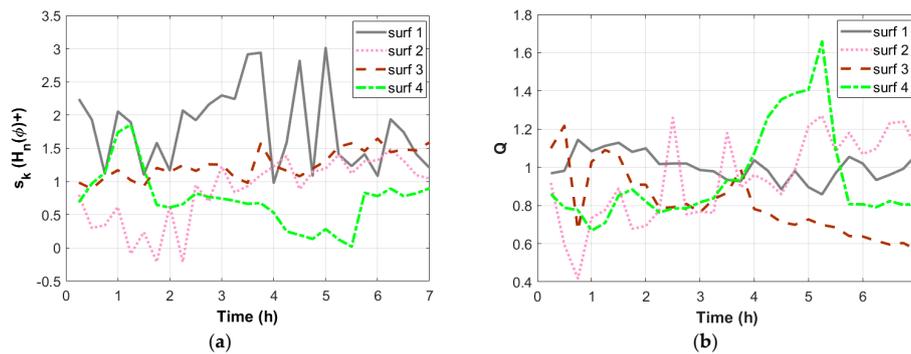
Figure 11. Statistical variation of surf 1, surf 2 and surf 3 and surf 4 (**a**) *sk* ($H_n(\phi)$+) and (**b**) Q.

4.3. Results and Analysis

4.3.1. ANN results

In this section, attention is drawn to the behaviour of the ANN in classifying and discriminating PD patterns of: (a) different void arrangement in PET, and (b) different surface discharge arrangements in oil. The training and testing fingerprints comprises statistical operators acquired from the $H_{qn}(\phi)$, $H_n(q)$ and $H_n(\phi)$ distributions in both the positive and negative voltage half cycles. For the ANN, two strategies are employed. Firstly, training the ANN (with either single, serial or parallel voids distributions) and then testing undertaken with the others. The same procedure of training and testing for voids is applied for the different surface discharges on pressboard in oil. For each PD defect, the input fingerprints for the ANN comprised of 28 samples derived from the *sk* and *ku* of the ($H_n(\phi)$+, $H_n(\phi)$−, $H_{qn}(\phi)$+ and $H_{qn}(\phi)$−), Q, cc and mcc. The choice of samples depends on the complexity of the data. For less complex data, few samples are enough. But for more complex data, such as in this case, large samples are needed to obtain better approximation. The output parameters are chosen to be [0, 1], [1, 0] and [1, 1] for the void classification while [0, 0], [0, 1], [1, 0] and [1, 1] are selected for the surface discharges on pressboard in oil. Out of the 28 samples, eight are applied as testing dataset and the remaining are the training dataset.

To determine the performance of the ANN for each training and testing strategy, up to 100 iterations of ANN results are obtained starting with different initial states (weights and biases). After 100 iterations, the result became stable and therefore adding more iteration is meaningless. Much iteration is needed for this investigation in order to obtain the confidence limits recognition rate of the ANN for a more reliable PD diagnosis. In deciding the configuration of ANN for each strategy, the hidden layer and the learning rate are varied, though the momentum rate remains the same. The momentum rate is applied in order to speed up the training process, thereby smoothen up weight updating and offering some resistance to irregular weight variations. The learning rate is usually set between value of 0 and 1. After several trial and errors with several layers, 10 hidden layers, a momentum rate of 0.6 and learning rate of 0.05 are determined to be the optimal for all the strategies. For this investigation, no validation samples were applied due to small data examples and training is said to be completed when the learning cycle reaches the maximum acceptable value or when the mean square error reaches 10^{-3}.

Figure 12 shows examples of varying recognition efficiencies when any of the void samples (single, serial or parallel) is applied for training the ANN and other void samples utilized for testing. The result clearly indicates that for ANN trained with the similar PD statistical sample, higher average recognition rate with lower variance has been observed, while for the ANN trained with one void sample and test with another sample, higher variance of recognition rates with lower variance is observed. Similarly, results for training and testing with different surface discharge samples are shown in Figure 13. To further evaluate Figures 12 and 13, average recognition rates of different training and testing permutations are shown in Table 2. The results show that patterns of single and serial void are

approximately similar with recognition efficiency of up to 98%, which is different from a parallel void. This result clearly correlates with the statistical variability of the voids (Figure 10) and the respective PD patterns (Figure 7). From this result, it can be established that when training the ANN with any of the three void arrangements considered in this paper and testing with others, high recognition probability is attained up 70%, which is enough to reveal that they all belong to the same defect, though with different geometry or arrangement. This result further demonstrates, though the single and the parallel void have similar PD characteristics such as the inception voltage, PD magnitude among others but they vary in terms of the PRPD distributions.
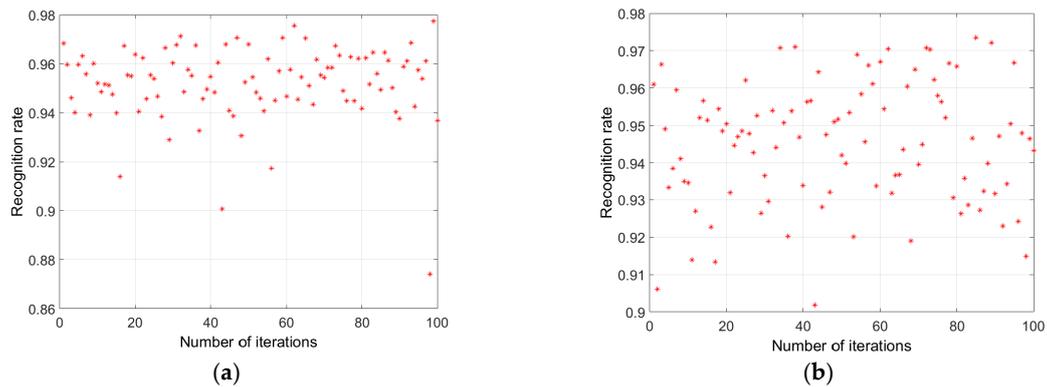

Figure 12. ANN Recognition performance when trained with (**a**) void 3 and testing with the same void 3; and (**b**) void 3 and testing carried out with void 1.
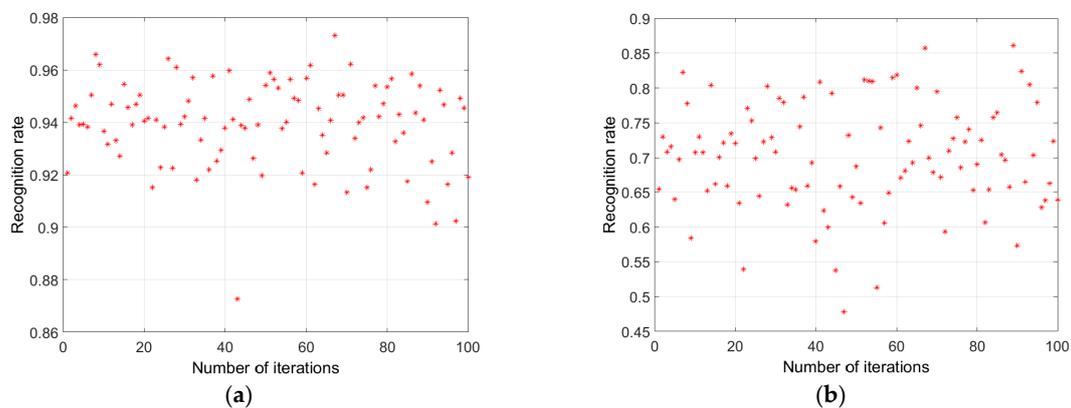

Figure 13. ANN Recognition performance when trained with (**a**) surf 1 and testing with surf 1; and (**b**) surf 1 and testing carried out with the same surf 1.

Table 2. Comparison of the mean recognition efficiency for the ANN for the different void arrangements and different surface discharge arrangements.

Training Data	Testing Data and the Corresponding Recognition Rate						
	Void 1	Void 2	Void 3	Surf 1	Surf 2	Surf 3	Surf 4
Void 1	98%	75%	96%				
Void 2	65%	97%	69%				
Void 3	94%	70%	96%				
Surf 1				95%	74%	82%	74%
Surf 2				78%	96%	79%	75%
Surf 3				80%	77%	95%	76%
Surf 4				73%	78%	78%	97%

Similarly, for the ANN applied for training with one surface discharge sample and testing with the same surface discharge, higher recognition performance is recorded up 97% in some instance. But when training any surface discharge sample and testing with another surface discharge sample, 73% appears to be the minimum recognition performance recorded. In contrast to the voids, the surface discharges appear to be different from each other in terms of the PD patterns and their degradation. The results indicate that the needle angular position, gap distance and applied voltage may play a role in determining the PD degradation pattern. This investigation implies that depending on training or testing with any PD source, the ANN is capable to capturing even slight changes in PD patterns of voids or surface discharge on pressboard in oil captured over long stressing period and this has laid the foundation for further investigation that specific ANN can be developed to recognized PD patterns from the similar arrangements.

4.3.2. FL results

In this section, the FL capability of being able to discriminate different void arrangement in PET and different surface discharge arrangements in oil will be investigated. However, due to the statistical and distribution similarity of some of the PD patterns as used in [42], it is necessary to scale down the input parameters for the FL to the ones that demonstrates higher discriminating potential. Box plots showing different quartiles (25%, 50% and 75%) will be applied to show the statistical distinction of the PD patterns. To determine quartiles intervals for recognition, 20 samples are used for each PD fault geometry of voids and surface discharges. Any statistical fingerprint applied for FL classification must be able to demonstrate non-overlapping characteristics [42,43]. However, this may be difficult for all statistical features because of stochastic nature of the PD.

Furthermore, to ensure effective classification, box plots for both positive and negative half power cycles are evaluated. For the different voids, *sk* of the $H_{qn}(\phi)$, *Q* and cc appear to show clear discrimination for the PD void patterns and will therefore be applied as input to the FL (see Figure 14). This is expected because of the different discharge levels and distributions on both half of the AC power cycle in the parallel void as opposed to the single or serial void. Similarly, for the surface discharges, *sk* of $H_n(q)$, *sk*, $H_{qn}(\phi)$ and the cc appear to show better discrimination parameters (see Figure 15). Figure 16a,b show the developed fuzzy identification system containing the fuzzy rules implemented for the void and surface discharge classification respectively. Initially, the fuzzy inputs are classified to five linguistic elements, i.e., low, low-medium, medium, high-medium and high depending on the discrimination level of the statistical parameters. Logical operators of fuzzy intersection (AND) and fuzzy union (OR) are used to determine the fuzzy rules. Similar to the literature [38] and for the voids the fuzzy rules are defined as follows:

a.   If (*Q* is high) ∪ (cc is high) ∩ (*sk* ($H_{qn}(\phi)$ is low)- Then (void 1)
b.   If (*Q* is low) ∪ (cc is low) ∩ (*sk* ($H_{qn}(\phi)$ is medium), Then (void 2)
c.   If (*Q* is high-medium) ∪ (cc is high) ∩ (*sk* ($H_{qn}(\phi)$ is low)- Then (void 3)

The fuzzy rules for surface discharge classification are as follows:

a.   If (*sk* ($H_n(q)$ is low) ∩ (cc is medium) ∪ (*sk* ($H_{qn}(\phi)$ is medium)- Then (surf 1)
b.   If (*sk* ($H_n(q)$ is medium) ∩ (cc is low) ∪ (*sk* ($H_{qn}(\phi)$ is low)- Then (surf 2)
c.   If (*sk* ($H_n(q)$ is medium) ∩ (cc is low) ∩ (*sk* ($H_{qn}(\phi)$ is high)- Then (surf 3)
d.   If (*sk* ($H_n(q)$ is medium) ∩ (cc is medium) (*sk* ($H_{qn}(\phi)$ is medium)- Then (surf 4)

In order to recognize the PD samples for voids and surface discharges, eight samples for each defect are used and recognition is decided when majority of the samples show particular trend. The recognition is decided as either 0 or 1 based on the membership functions of the fuzzy logic (see Figure 16). Table 3 shows the recognition performance of the FL as applied to recognize PD faults of similar geometry. The results clearly demonstrated FL capability to categorize the PD faults scenarios, depending on testing with different *φ-q-n* samples.

Table 3. Comparison of the mean recognition efficiency for the FL for the different void arrangements and different surface discharge arrangements.
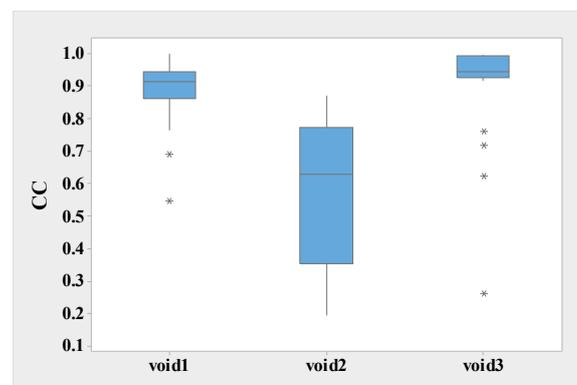
Training Data	Testing Data and the Corresponding Recognition Rate						
	Void 1	Void 2	Void 3	Surf 1	Surf 2	Surf 3	Surf 4
Void 1	100%	0%	0%				
Void 2	0%	100%	0%				
Void 3	0%	0%	100%				
Surf 1				100%	0%	0%	0%
Surf 2				0%	100%	0%	0%
Surf 3				0%	0%	100%	0%
Surf 4				0%	0%	0%	100%
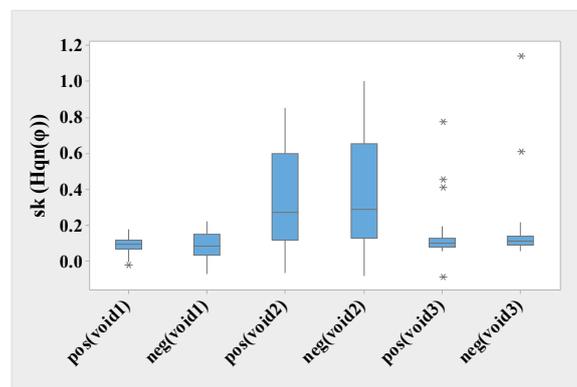
(**a**)
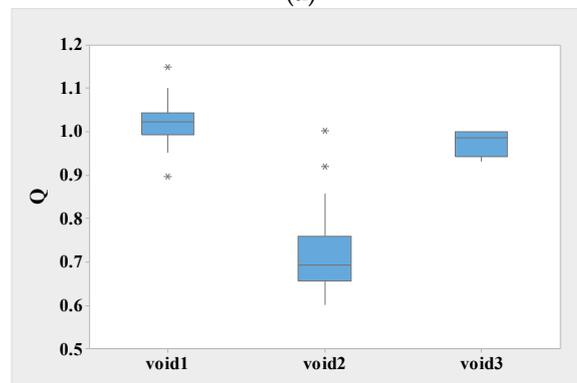
(**b**)

(**c**)

Figure 14. Comparison of quartiles for different statistical parameters for voids: (**a**) *sk* of the pulse height distributions; (**b**) *Q* and (**c**) cc.

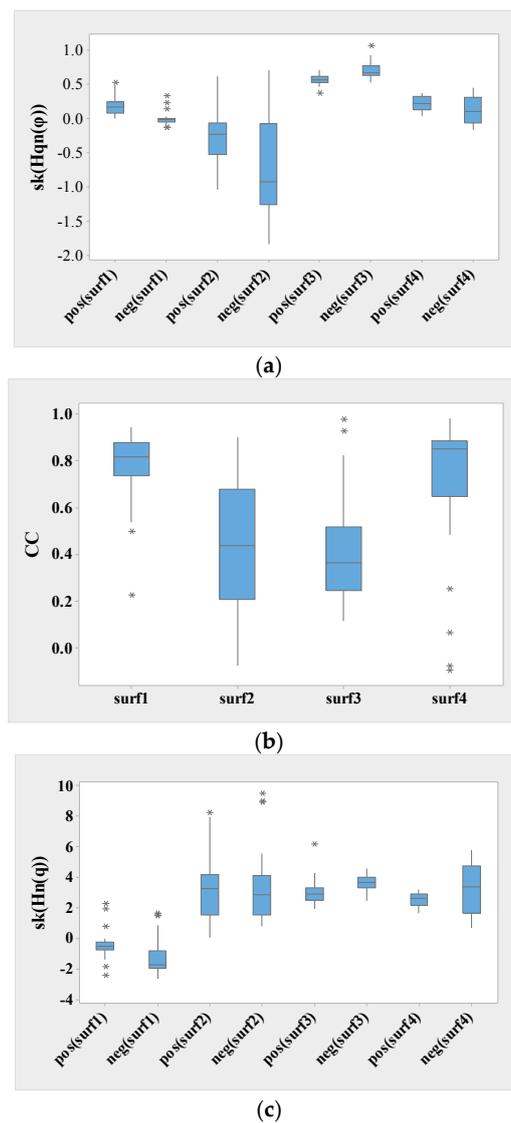(**a**)

(**b**)

(**c**)

Figure 15. Comparison of quartiles for different statistical parameters for surface discharges: (**a**) *sk* of the pulse height distribution; (**b**) cc, and (**c**) *sk* of the $H_n(q)$.
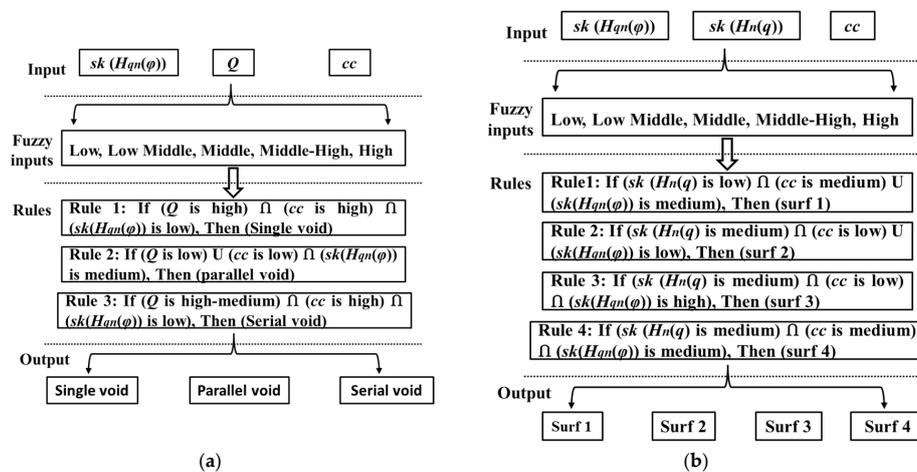

Figure 16. Structure of the Fuzzy classification system for (**a**) voids; and (**b**) surface discharges in oil.

4.3.3. ANN and FL results comparison

Generally, the ANN is well known as a nonlinear technique and easier to implement. It has the capability to give important conclusion useful complex data. On the other hand, the FL is more useful in interpreting uncertainties associated with data by interpreting data relationship using fuzzy rules. For complex problem solving, the ANN requires large data for correct interpretation while fuzzy does not [43].

In this paper, both the ANN and the FL have been applied to recognize similar PD faults of different arrangement within the HV insulation system. The ANN was able to capture even slight changes in the statistical features of the PD patterns of voids and surface discharges. One interesting thing about the ANN is its ability to provide recognition probability value of different defects in order to show the extent of similarity or variations of the PD defects. These have been achieved with few PD samples and through extensive evaluation of the ANN using different initial training conditions. Limited numbers of PD samples are used in this work because even in practice, PD data are commonly known to be few. The implication of the ANN results is that it is possible to recognize and discriminate PD faults of closely similar geometry depending on training or testing PD samples. Contrary to the ANN, the FL can provide recognition probability of the PD faults as either correct (1) or incorrect (0). However, it is difficult to understand the similarity of the PD defects and whether they are of the same geometrical arrangement or not.

One interesting thing about the FL is that it permits to develop fuzzy rules with suitable statistical parameters that can provide better discrimination of the PD faults of internal PD, (i.e., *sk* of the $H_{qn}(\phi)$, Q and cc) and surface discharges, (i.e., *sk* of $H_n(q)$, *sk*, $H_{qn}(\phi)$ and cc) which can be beneficial for practical implementation of this algorithm. In Table 4, it is included a brief qualitative comparison between ANN and FL techniques for PD pattern recognition.

Table 4. Qualitative comparison between ANN and FL for PD recognition.

ANN	FL
ANN is well-known as a non-linear technique and easier to be implemented	FL is more useful in interpreting uncertainties associated with data by interpreting data relationship using fuzzy rules
For complex problems, ANN requires large data for training to obtain correct result	FL do not require large data for correct interpretation
For PD recognition, ANN can provide recognition probability values of different defects in order to show the extent of similarity or variations of the PD defects	FL can provide recognition probability of the PD faults as either correct (1) or incorrect (0)
The ANN can recognize and discriminate PD faults of closely similar geometry depending on training or testing PD samples	With FL, it is difficult to understand the similarity of the PD defects and whether they are of the same geometrical arrangement or not

5. Conclusions

This paper has compared the ANN and FL for recognizing and discriminating a number of PD geometries of two classes, i.e., voids class and surface discharge class. Experiments were carried out and PD samples captured over a 7 h stressing period in order to obtain a fair representative of the PD samples for evaluation by the ANN and the FL techniques. Then, statistical features obtained from the pulse height and pulse count distributions were applied as input to the ANN and FL systems. The results clearly show that both the ANN and FL can recognize and discriminate the PD faults. The ANN algorithm can better provide the level of similarity or closeness between the PD fault samples. On the other hand, the FL permits to know the best statistical parameters for the discrimination of the void classes and for surface discharges, which is not inappropriate for the ANN because of the limited PD samples available. Although, the results show that the ANN provides a more robust classification but these conclusions have to be further investigated with more complex PD examples.

Based on this analysis, it is obvious that the ANN may be best suited for practical PD classification. As first step, it is important to firstly identify the kind of defect through online PD classification looking at its PRPD pattern. Then offline ANN PD testing tools already trained with either void class or surface discharge class of defects can be used to recognize in detail the PD fault of similar PD geometry. In training the offline PD classification tool, data must be available from actual HV installation either in service or from factory test. If such data is not available, then the possibility of using simulated PD samples may be investigated. If sufficient PD sample data can be assured, then the ANN can be used to behave as an experienced evaluator. However, the ANN may have some limitations such as excessive training and inability to obtain sufficient PD samples. On the other hand, the FL does not have these limitations but its inability to provide accurate recognition probability values may be a setback and may not provide the condition-monitoring engineer's correct information about the similarity of different PD faults. Though this work is purely based on laboratory data, future work will concentrate on comparing the capabilities of the ANN and FL for practical PD dataset in order to arrive at a more reliable conclusion.

Author Contributions: Abdullahi Abubakar Mas'ud carried out the experimental measurements and wrote the paper. Ricardo Albarracín and Jorge Alfredo Ardila-Rey proposed improvements for the analysis of the results and for the manuscript. Firdaus Muhammad-Sukki and Nurul Aini Bani carried out a proofreading on the article and provides suggestions for improvement.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kuffel, E.; Zaengl, W.; Kuffel, J. *High Voltage Engineering Fundamentals*, 2nd ed.; Newnes: Oxford, UK, 2000; pp. 613–651.
2. Kreuger, F.H. *Partial Discharge Detection in High Voltage Equipment*; Butterworths: London, UK, 1989.
3. Abubakar, M.A. The Application of Ensemble Neural Networks for Partial Discharge Pattern Recognition. Ph.D. Thesis, Glasgow Caledonian University, Glasgow, UK, 2013.
4. Hirata, A.; Nakata, S.; Kawasaki, Z. Toward automatic classification of partial discharge sources with neural network. *IEEE Trans. Power Deliv.* **2006**, *21*, 526–527. [CrossRef]
5. Gill, P. *Electrical Power Equipment Maintenance and Testing*, 2nd ed.; Power Engineering (Willis); Taylor & Francis: Boca Raton, FL, USA, 2008.
6. Cavallini, A.; Montanari, G.; Contin, A.; Pulletti, F. A new approach to the diagnosis of solid insulation systems based on PD signal inference. *IEEE Electr. Insul. Mag.* **2003**, *19*, 22–30. [CrossRef]
7. Ardila-Rey, J.A.; Martínez-Tarifa, J.M.; Robles, G.; Rojas-Moreno, M.V. Partial discharge and noise separation by means of spectral power clustering techniques. *IEEE Trans. Electron.* **2013**, *20*, 1436–1443. [CrossRef]
8. Albarracín, R.; Robles, G.; Martinez-Tarifa, J.M.; Ardila-Rey, J. Separation of sources in radio frequency measurements of partial discharges using time–power ratio maps. *ISA Trans.* **2015**, *58*, 389–397. [CrossRef] [PubMed]
9. Gulski, E.; Kreuger, F. Computer aided analysis of discharge patterns. *J. Appl. Phys.* **1990**, *23*, 1569–1575. [CrossRef]
10. Raymond, W.J.K.; Illias, H.A.; Mokhlis, H. Partial discharge classifications: Review of recent progress. *Measurement* **2015**, *68*, 164–181. [CrossRef]
11. Gulski, E.; Krivda, A. Neural Network as a tool for recognition of partial discharges. *IEEE Trans. Electr. Insul.* **1993**, *28*, 984–1001. [CrossRef]
12. Álvarez, F.; Garnacho, F.; Ortego, J.; Sánchez-Urán, M.Á. Application of HFCT and UHF Sensors in On-Line Partial Discharge Measurements for Insulation diagnosis of High Voltage Equipment. *Sensors* **2015**, *15*, 7360–7387. [CrossRef] [PubMed]
13. Gulski, E. Computer Aided Recognition Of Partial Discharges Using Statistical Tools. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1999.

14. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G. Application of an ensemble neural network for classifying partial discharge patterns. *Electr. Power Syst. Res.* **2014**, *110*, 154–162. [CrossRef]

15. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G. An investigative study into the sensitivity of different partial discharge *φ-q-n* pattern resolution sizes on statistical neural network pattern classification. *Measurement* **2016**, *92*, 497–507. [CrossRef]

16. Candella, R.; Mirelli, G.; Schifani, R. Partial discharge recognition by means of statistical and fractal parameters and a neural network. *IEEE Trans. Electr. Insul.* **2000**, *7*, 87–94. [CrossRef]

17. Mas'ud, A.A.; Albarracín, R.; Ardila-Rey, J.A.; Muhammad-Sukki, F.; Illias, H.A.; Bani, N.A.; Munir, A.B. Artificial Neural Network Application for Partial Discharge Recognition: Survey and Future Directions. *Energies* **2016**, *9*, 574. [CrossRef]

18. Abdel-Galil, T.; Sharkawy, R.; Salama, M. Partial discharge pattern classification using the fuzzy decision tree approach. *IEEE Trans. Instrum. Meas.* **2005**, *54*, 2258–2263. [CrossRef]

19. Wang, G.; Guo, Z.; Lei, Q. Online incremental learning for tool condition classification using modified Fuzzy ARTMAP network. *J. Intell. Manuf.* **2014**, *25*, 1403–1411. [CrossRef]

20. Ma, X.; Zhou, C.; Kemp, I. Interpretation of wavelet analysis and its application in partial discharge detection. *IEEE Trans. Dielectr. Electr. Insul.* **2002**, *9*, 446–457. [CrossRef]

21. Alisha, E.; Satish, L. Wavelet analysis for classification of multi-source PD patterns. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 40–47.

22. Robles, G.; Parrado-Hernández, E.; Ardila-Rey, J.; Martínez-Tarifa, J.M. Multiple partial discharge source discrimination with multiclass support vector machines. *Expert Syst. Appl.* **2016**, *55*, 417–428. [CrossRef]

23. Mazroua, A.; Bartnikas, R.; Salama, M. Partial discharge pattern recognition with neural networks using the multi-layer perceptron technique. *IEEE Trans. Electr. Insul.* **1995**, *28*, 1082–1089. [CrossRef]

24. Ben Salah, C.; Ouali, M. Comparison of Fuzzy Logic and Neural Network in Maximum Power Point Tracker for PV Systems. *Electr. Power Syst. Res.* **2011**, *81*, 43. [CrossRef]

25. Alabedin, A.M.Z.; El-Saadany, E.F.; Salama, M.M.A. Maximum Power Point Tracking for Photovoltaic Systems Using Fuzzy Logic and Artificial Neural Networks. In Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–29 July 2011; pp. 1–9.

26. Topçu, İ.B.; Sarıdemir, M. Prediction of Compressive Strength of Concrete Containing Fly Ash Using Artificial Neural Networks and Fuzzy Logic. *Comput. Mater. Sci.* **2008**, *41*, 305. [CrossRef]

27. James, R.E.; Su, Q. IET Power and Energy Series. In *Condition Assessment of High Voltage Insulation in Power System Equipment*; The Institution of Engineering and Technology: London, UK, 2008.

28. IEC-60270. *High-Voltage Test Techniques—Partial Discharge Measurements*, 3rd ed.; International Electrotechnical Commission (IEC): New Delhi, India, 2000.

29. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: Upper Saddle River, NJ, USA, 1998.

30. Priddy, K.; Keller, P. *Artificial Neural Network: An Introduction*; SPE: New York, NY, USA, 2005.

31. Cilimkovic, M. Neural Networks and Back Propagation Algorithm. Master's Thesis, Institute of Technology Blanchardstown, Dublin, Ireland, 2015.

32. Li, C.; Yu, C. Performance Evaluation of Public Non-Profit Hospitals Using a BP Artificial Neural Network: The Case of Hubei Province in China. *Int. J. Environ. Res. Public Health.* **2013**, *10*, 3619–3633. [CrossRef] [PubMed]

33. Wang, L.X. *A Course in Fuzzy Systems and Control*; Prentice-Hall: London, UK, 1997.

34. Mendel, J. Fuzzy logic systems for engineering: A tutorial. *Proc. IEEE* **1995**, *83*, 345–377. [CrossRef]

35. Knapp, R.B. Fuzzy Sets and Pattern Recognition. Available online: http://hci.sapp.org/lectures/knapp/fuzzy/fuzzy.pdf (accessed on 2 January 2016).

36. International Electrotechnical Commission (IEC). *Part 7: Fuzzy Control Programming*; International Electrotechnical Commission: London, UK, 1997.

37. Gopal, S.; Karthikeyan, B.; Kavitha, D. Partial Discharge pattern classification using fuzzy expert system. In Proceedings of the IEEE 8th International Conference on Solid Dielectrics (ICSD), Toulouse, France, 5–9 July 2004; pp. 500–503.

38. Chen, X.; Cavallini, A.; Montanari, G.C. Statistical analysis and fuzzy logic identification of partial discharge in paper-oil insulation system. In Proceedings of the IEEE 9th International Conference on the Properties and Applications of Dielectric Materials (ICPADM), Harbin, China, 19–23 July 2009; pp. 505–508.

39. Marungsri, B.; Boonpoke, S. Applications of simplified fuzzy ARTMAP to partial discharge classification and pattern recognition. *WSEAS Trans. Syst.* **2011**, *10*, 69–80.

40. Illias, H.; Chen, G.; Lewin, P. Partial discharge behaviour within two spherical cavities in dielectric material. In Proceedings of the IEEE Conference on Electrical Insulation and Dielectric Phenomena, Cancun, Mexico, 16–19 October 2011; pp. 456–459.

41. Alvarez, F.; Ortego, J.; Garnacho, F.; Sánchez-Urán, M.Á. A clustering technique for partial discharge and noise sources identification in power cables by means of waveform parameters. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 469–481. [CrossRef]

42. Cavallini, A.; Montanari, G.C.; Puletti, F. A fuzzy logic algorithm to detect electrical trees in polymeric insulation systems. *IEEE Trans. Dielectr. Electr. Insul.* **2005**, *12*, 1134–1144. [CrossRef]

43. Majidi, M.; Fadali, M.S.; Etezadi-Amoli, M.; Oskuoee, M. Partial discharge pattern recognition via sparse representation and ANN. *IEEE Trans. Dielectr. Electr. Insul.* **2015**, *22*, 1061–1070. [CrossRef]