August 2014

Air-Gap Partial Discharge Development Stage Recognition for Power Transformer Insulation Monitoring Considering Effect of Cavity Size

Zhenze Long
University of Wisconsin-Milwaukee

AIR-GAP PARTIAL DISCHARGE DEVELOPMENT STAGE RECOGNITION FOR POWER TRANSFORMER INSULATION MONITORING CONSIDERING EFFECT OF CAVITY SIZE

by

Zhenze Long

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Engineering

at

The University of Wisconsin-Milwaukee

August 2014

ABSTARACT
AIR-GAP PARTIAL DISCHARGE DEVELOPMENT STAGE RECOGNITION FOR POWER TRANSFORMER INSULATION MONITORING CONSIDERING EFFECT OF CAVITY SIZE

by

Zhenze Long

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Professor David C. Yu

Oil-paper insulation system is commonly used for power transformer internal insulation. Partial discharge (PD) is one of the main reasons for aging and disruption of the insulation system. Air-gap PD occurs in gas-filled cavity in transformer oil-paper insulation and is an extremely common and serious defect type. For air-gap PD analysis, most experiments were conducted through the standard air-gap discharge model recommended by CIGRE. Some work has been done to diagnose air-gap PD severity. However, the effect of cavity size on PD activity has not been emphasized yet. My thesis systematically discusses the effect of cavity size on air-gap PD activity through experiments. And pattern recognition classifier is a critical part in PD diagnosis. Artificial neural network and support vector machine are commonly used nowadays and show some good results in site application. To enhance PD diagnosis accuracy is still a main task. In this work, Random Forests is first time introduced in PD diagnosis.

Experiments show that large cavity PD possesses lower inception field, higher charge magnitude, higher inception phase. PD happening in large cavity is more harmful than that happening in small cavity. Besides, during Air-gap PD development process, charge magnitude variation of large and small cavity model both presents concave curve shape with respect to time and discharge phase slowly expends. For small cavity model, when air-gap PD comes to the last stage, positive PD even can expand to the negative half cycle and vice versa. And through clustering, the PD development stage for large and small cavity model are both divided into three stages, initial discharge stage, weak discharge stage and pre-breakdown stage. For air-gap PD development stage identification, total accuracy of random forests classifier is 93.15%, showing a better performance than RBF neural network.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my major advisor, Professor David C. Yu, for his patient guidance and useful critiques on my thesis. I also highly appreciate that Professor Yu built up the communicational bridge between universities in China and University of Wisconsin-Milwaukee, which gave opportunities to many college students like me to come to United States for further study and better educational environment. Besides, I want to thank my parents, Zhou Long and Yi Zhu, for their unselfish love. I would never have this chance to study abroad without their full support.

Chapter 1 Introduction

1.1 Background

Power transformers are important apparatuses in electrical power system and need very large investments. They are usually very reliable, with a 25-35 year design life [1, 2]. However as transformers age, their internal condition degrades, which increases the risk of failure. Transformer failures due to dielectric problems are reported as high as 75% [3, 4].

1.1.1 The harm and types of partial discharge (PD)

Oil-paper insulation system is commonly used for power transformer internal insulation. Partial discharge (PD) is one of the main reasons for aging and disruption of the insulation system. PD is a weak discharge event which does not bridge the electrodes within an electrical insulation system under high field stress. PD normally occurs within some defects which are formed in long-term operation, such as cavities, cracks, joints and electrode burrs, in the insulation system. It does not cause direct breakdown of the insulation immediately because the surrounding insulation is strong enough to avoid a complete breakdown of the material. However, long-term PD can lead to continuous deterioration of the oil-paper insulation and consequently cause breakdown [5].

Cumulative PD brings about the aging and degradation of the oil-paper insulation through the following ways. ① Bombardment by charged particles. Electrons and positive ions generated by PD bombard the insulation with relatively high kinetic energy, which is more than $10eV$. Normally, the chemical bonds, $C-H$ and $C=H$, in oil-paper insulation have energy of $3.5eV$ and $6.2\,\mathrm{eV}$ respectively. They are possibly ruptured by the energetic particles. Molecular structure of the insulation destroyed, large quantity of tiny holes are generated on the insulation surface, eventually leading to the formation of discharge channels in the pressboard. ② Radiation effect. PD generates visible light, ultraviolet ray, X-ray and γ-ray. High-energy X-ray and γ-ray also lead to rupture of the chemical bonds. ③ Thermal effect. Temperature at PD point could be high enough to carbonize the pressboard. ④ Corrosion by discharge byproducts. Due to PD, N_2, H_2O and O_2 become HNO_3 and O_3, which cause the chemical deterioration of the material [2].

According to different PD locations and mechanism, PD is divided into four main types, air-gap (internal) PD, surface PD, corona PD and bubble PD. Air-gap PD occurs in gas-filled cavity between winding and pressboard or within pressboard. Though air-gap PD magnitude is relatively low, it can do serious harm to the solid insulation (pressboard) that can't be self-recovery [6]. Surface PD, which also can also increase the risk of failure, happens along the interface between pressboard and oil when the pressboard is affected with damp and contaminated by copper sulphide or other impurity [7]. Corona PD happens at conductor burrs and joints which are

naked in the oil. It doesn't damage the solid insulation and is less harmful than Air-gap and surface PD. Bubble PD happens in the mobile gas bubble in the oil. It is not stable and can be eliminated by oil filtering [8].

Failures in power transformer due to insulation breakdown can lead to expensive maintenance or replacement cost and reduce the reliability of power supply. Therefore, it is very necessary to monitor insulation condition and diagnose fault type as well as severity, then reasonable and proactive maintenance strategy can be implemented to reduce failure rate and prolong life span of power transformer.

1.1.2 PD measurement and diagnosis techniques

Generally speaking, the term "monitoring" describes a basic parameter measurement with threshold alarms. The term "diagnosis" indicates the addition of sophisticated analysis, such as an expert system capable of providing an assessment of equipment condition and suggested actions [2]. There are three main kinds of PD measurement methods, Dissolved gas analysis (DGA), Acoustic test and Electrical test.

(1) DGA detects the small quantities of gases which come from oil-paper insulation decomposition by abnormal electrical or thermal stresses. DGA is not affected by the strong electromagnetic interference in substations and is taken for on-line monitoring by most utilities. However, it just can distinguish faults such as arcing, overheating and PD but can't recognize the PD type.

⑵ Ultrasonic test detects the acoustic signals ranging from 100 to 300 kHz generated by PD. It can precisely tell if there is PD happening. If multiple sensors are mounted on the transformer tank, PD source can be located according to the arrival time of the pulses at the sensors. Nevertheless, ultrasonic test can't recognize the PD type or severity, either.

⑶ Electrical test is subdivided into two categories, pulsating current method and ultra high frequency (UHF) method, based on different detection frequency band. More PD information can be obtained by comparing electrically detected pulsating signals and power frequency signal for PD type and severity diagnosis. Pulsating current method which can be implemented discharge quantity calibration according to IEC 60270 is widely used for insulation condition assessment. The diagnosis system and procedures are shown below.

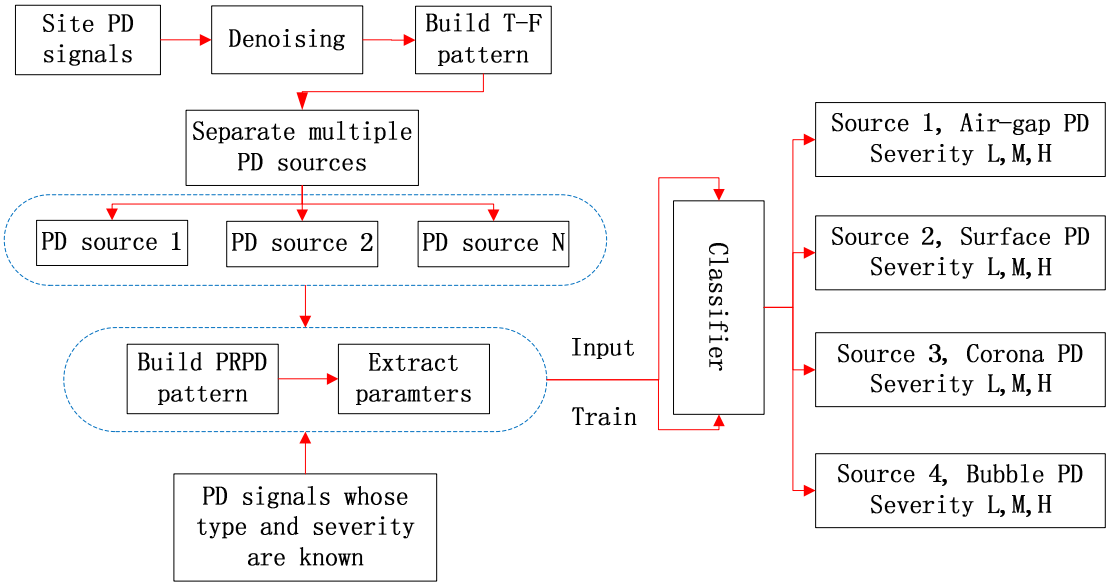


Figure 1. 1 PD diagnosis system

① Collect different PD type's signals at different development stage from experiment and site operation, extract parameters from Phase Resolved Partial Discharge (PRPD) pattern to be feature information database that is the training data for pattern recognition classifier.

② Collect field fault signal, denoise PD signal due to strong electromagnetic interference in substations.

③ Build T-F pattern to separate multiple PD sources, then extract parameters from each PD's PRPD pattern.

④ used the trained classifier to determine PD's type and severity.

1.2 Thesis objectives and outline

Large amount of work has been done to deal with PD type and severity diagnosis. For air-gap PD analysis, most researchers conducted experiments using the standard

cylindrical air-gap discharge model recommended by CIGRE [9]. However, the effect of cavity size, which could strongly affect the diagnosis accuracy of PD severity, has not been valued yet. Only Lutz and Schifani studied the inception field in different height's cavity in [10-12] and Hazlee studied the effect of cavity height on inception field and charge magnitude in [13]. This thesis systematically discusses the effect of cavity size on air-gap PD activity through experiments, and then distinguishes development characteristic difference between large cavity model and small cavity model.

Pattern recognition classifier is another critical part in PD diagnosis. Linear discriminant analysis, artificial neural network and support vector machine are commonly used nowadays and show some good results in site application. To enhance PD diagnosis accuracy is still the main task. In this thesis, random forests classifier that is originally proposed by Leo Breiman in [14], which has not been introduced in PD diagnosis before, is implemented to identify large cavity and small cavity PD's development stage.

This thesis is divided into 5 chapters. Chapter 1 contains the background of PD diagnosis and objectives of this work. Chapter 2 explains the mechanism of air-gap PD, introduces the procedure of air-gap PD experiment considering different cavity size and interpreters the effect of cavity size on PD activity through experiment results. In chapter 3, the development characteristics of large cavity and small cavity

PD are detailed. Furthermore, 27 dimensional statistical parameters are extracted from PRPD pattern and then parameter dimensions are reduced through kernel principle component analysis (KPCA). Chapter 4 presents the clustering results of large and small cavity PD development stage. In addition, RBF neural network and random forests are used to identify the development stage. Chapter 5 is the conclusions of this work.

Chapter 2 Effect of cavity size on PD activity

2.1 Mechanism of air-gap PD in oil-paper insulation

Air-gap PD occurs in the gas-filled cavity in transformer oil-paper insulation and is an extremely common and serious defect type [6, 15]. The gas-filled cavity defect may be created during manufacturing, installation or long-term operation process. Moisture in the pressboard can be evaporated when oil temperature is high enough, causing delamination in the pressboard [16]. Figure 2. 1 indicates a basic diagram of air-gap defect in solid dielectric.

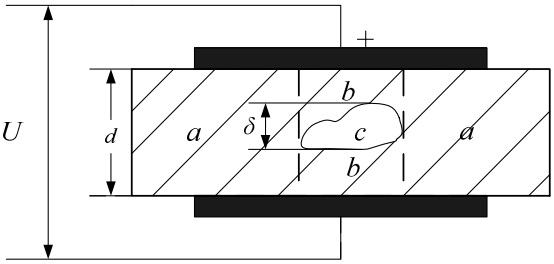


Figure 2. 1 Diagram of a gas-filled cavity within insulation pressboard

According to the principle of current continuity,

$$\dot{U}_c\dot{Y}_c = \dot{U}_b\dot{Y}_b \tag{2.1}$$

where U_c is the voltage on the cavity, U_b is the voltage on the insulation paperboard, Y_c , Y_b are the admittance of gas and pressboard respectively. γ_c, conductivity of gas and λ_b, conductivity of paper are both less than $10^{-11}(\Omega \bullet m)^{-1}$, which can be neglected in power frequency electric field. So the voltage ratio is modified in Equation 2.2.

$$\frac{u_c}{u_b}=\left|\frac{\dot{U}_c}{\dot{U}_b}\right|=\sqrt{\frac{\gamma_b^2+(\omega C_b)^2}{\gamma_c^2+(\omega C_c)^2}}=\frac{\omega C_b}{\omega C_c}=\frac{\varepsilon_b\delta}{\varepsilon_c(d-\delta)} \tag{2.2}$$

Where ε_b , ε_c are the permittivity of the pressboard and gas.

$$\frac{E_c}{E_b}=\frac{u_C/\delta}{u_b/(d-\delta)}=\frac{\varepsilon_b}{\varepsilon_c} \tag{2.3}$$

Since the permittivity of the gas is less than the permittivity of the insulation pressboard, the electric field in the cavity is higher than that in the paper. And the breakdown strength of gas is much lower than the paper's. When the electric field is sufficiently high, PD firstly occurs in the gas-filled cavity. The discharge starts from one end of the cavity surface, only bridging the cavity and does not connect the whole insulation between electrodes.

If PD occurs in cavity, two conditions must be satisfied. One is that there are enough initial free electrons, the other is sufficient electric field to keep electron avalanche happening. The electrons in the gas cavity are accelerated by the applied electric field, interacting with neutral gas molecules during the movement. If the kinetic energy of accelerated electron is high enough, it will collide with any gas molecule, resulting in releasing a new electron and positive ion. This process is called collision ionization. The new free electron would also be accelerated to hit other gas molecule, producing another electron. As the ionization repeats, large amount of electrons, which flow to anode, will be generated. The repetition of gas ionization is called as electron avalanche, making the gas property from a non-conducting to a conducting condition in a very short period of time.

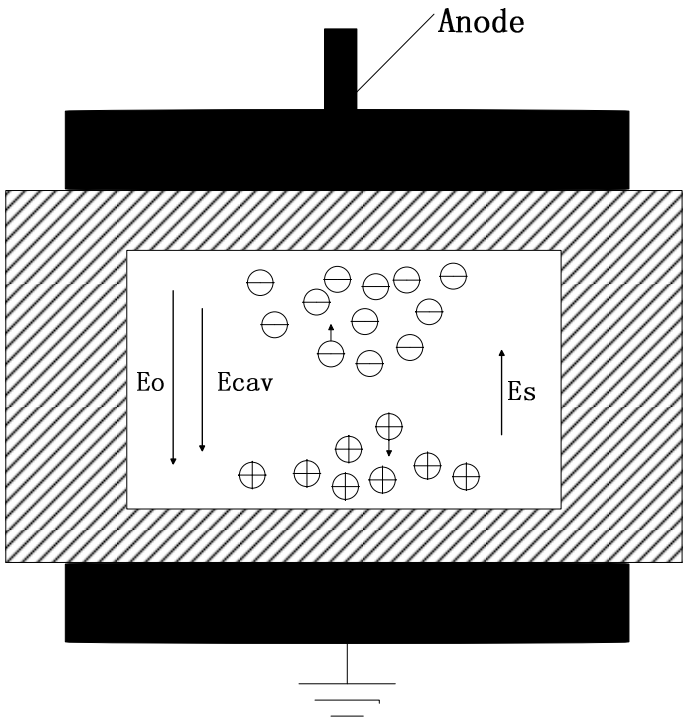


Figure 2. 2 Schematic diagram of PD event

As electron avalanche develops, the electric field in the cavity decreases. Figure 2.2 illustrates this event. When discharge occurs, the electrons and positive ions, generated by ionization, flow to the other end of cavity surface. So after a PD, the original electric field in the cavity, E_0 is offset partially by the opposite field, E_s which is due to charge accumulation on the cavity surface. When the resultant field in the cavity, E_{cav} drops to less than the breakdown strength of gas, free electrons lose their energy, the streamer channel collapses and discharge stops.

2.2 Air-gap PD experiments considering different cavity sizes

2.2.1 Air-gap PD model

The air-gap discharge models of different size were manufactured according to CIGRE Method II and ASTM-D149-09 in [9, 17]. Configuration of the basic model is shown in Figure 2. 3. A cylindrical cavity was made by a ring of pressboard embedded between two pressboards. In order to study the effect of cavity diameter and height on PD activity, five types of configurations are designed. When diameter of cavity is set to $40\,mm$, the heights are 0.5, 1.0 and $2.0\,mm$. And when height is set to $0.5\,mm$, the diameters are 10, 20 and $40\,mm$. The insulating glue is employed to seal the cavity in order to avoid the oil from penetrating into. The electrode system is a pair of plane-plane electrode arrangement. The upper plane electrode, with a diameter of $60\,mm$ and a height of $5\,mm$, was energized during the experiment. The under one, with same diameter and a height of $10\,mm$, was the grounding electrode. All the electrodes are made of brass and the insulating papers are Kraft pressboards. Besides, all the pressboards were fully dried and polished smooth before oil impregnation. The oil index satisfied the IEC 60296 and ASTM D 3487-09 standards.

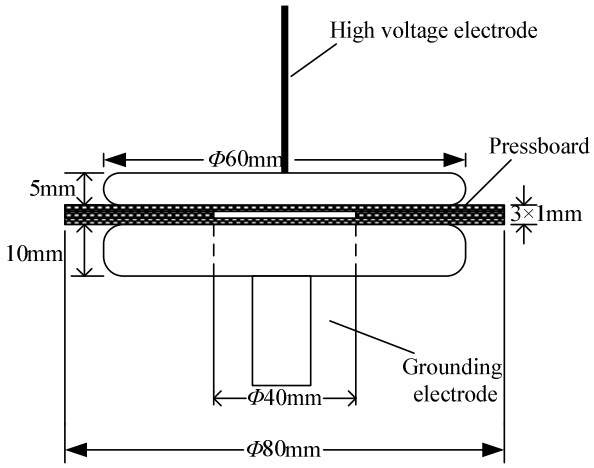


Figure 2. 3 Air-gap PD model

2.2.2 PD experiment setup

In order to simulate the real environment in power transformer, a simulative oil tank where the air-gap discharge model is in was designed, shown in Figure 2. 4. To ensure an even distribution of both temperature and dissolved gases in the oil tank, a circulating pump and oil loop was used for the oil circulation. A temperature sensor was installed in the oil tank and the heaters were placed in a large incubator. During the test, the temperature of the oil was set to 60 °C, which was similar to a typical temperature of power transformers in service.

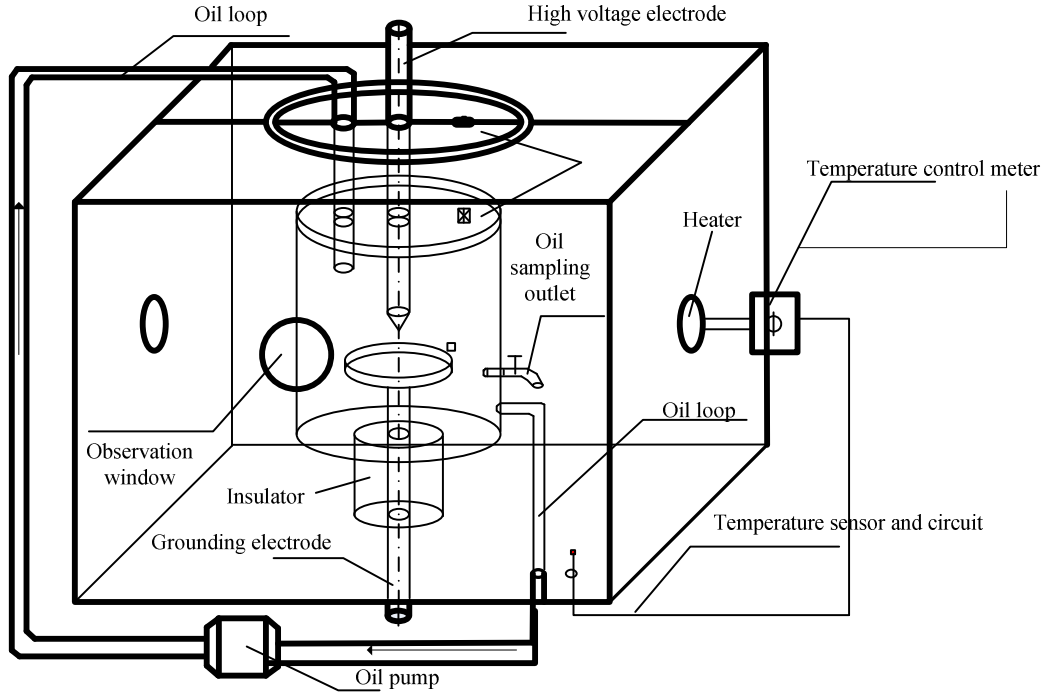
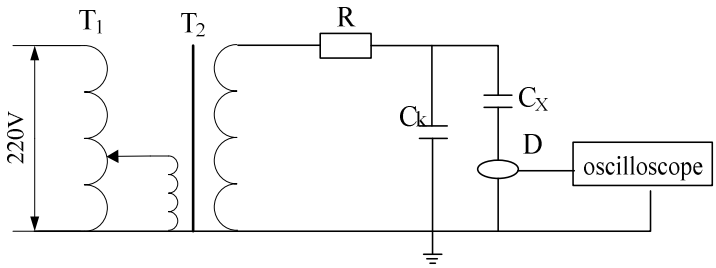


Figure 2. 4 Simulative oil tank

PD signals were detected according to the pulsating current method based on the standard of IEC 60270. A discharge-free AC voltage transformer (60 kV / 60 kVA) was applied to energize samples with a power frequency of 50 Hz. The coupling capacitors (1000 pF) facilitated the passage of the high-frequency current impulses. PD signals were sampled by a digital oscilloscope (Lecroy Wavepro 7100). The sampling frequency was set to 10 MS/s during the test. The measurement circuit sketch is shown in Figure 2. 5 and the whole PD test system is shown in Figure 2. 6.



T_1: Voltage regulator； T_2：High voltage testing transformer； C_K：Coupling capacitance；C_X：PD model;：
R：Protective resistance; D: Current transducer

Figure 2. 5 Measurement Circuit of Partial Discharge

Figure 2. 6 PD test system

2.2.3 Charge magnitude correction

According to IEC 60270, PD pulse voltage magnitude detected by oscilloscope can be corrected to charge magnitude through the correction circuit shown in Figure 2. 7. The artificial simulation branch consists of a voltage source, U_0 in series with a capacitance, C_0. When a given charge magnitude, $q_0 = U_0 C_0$ is injected into test sample, C_x, a pulse voltage magnitude, U can be recorded from the oscilloscope. Then U_0 is changed for several times and charge injection is repeated again and

again. Relationship between charge magnitude and pulse voltage magnitude is obtained by curve fitting shown in Figure 2. 8. In terms of the function, $q_0 = 0.748U + 1.7078$, PD charge magnitude can be estimated.

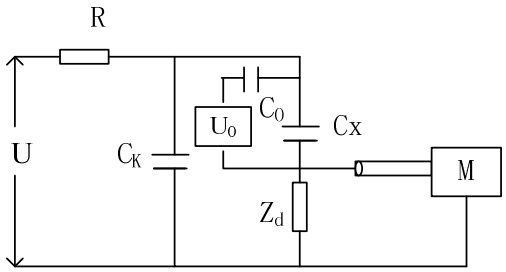


Figure 2. 7 Correction circuit of series connection

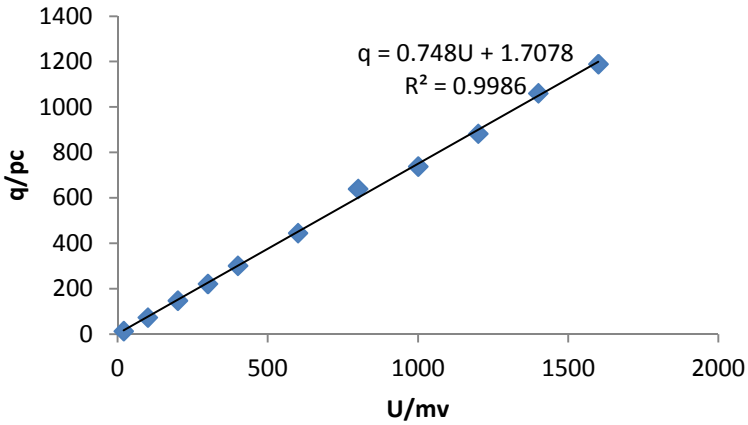


Figure 2. 8 Relationship between charge magnitude and pulse voltage magnitude

2.3 Effect of cavity size on PD inception field

PD Inception voltage, V_{inc} for each cavity configuration was tested for ten times, and PD inception field, E_{inc}, the minimum field that is required for a PD to occur, was calculated with the help of Ansys Maxwell. The field in each cavity configuration is

almost uniform. Figure 2. 9 shows the field for the cavity whose diameter is 40 millimeters and height is 1.0 millimeters. Mean value of experiment results is shown in Table 2. 1 and Table 2. 2. It was found that larger cavity possesses lower E_{inc}.

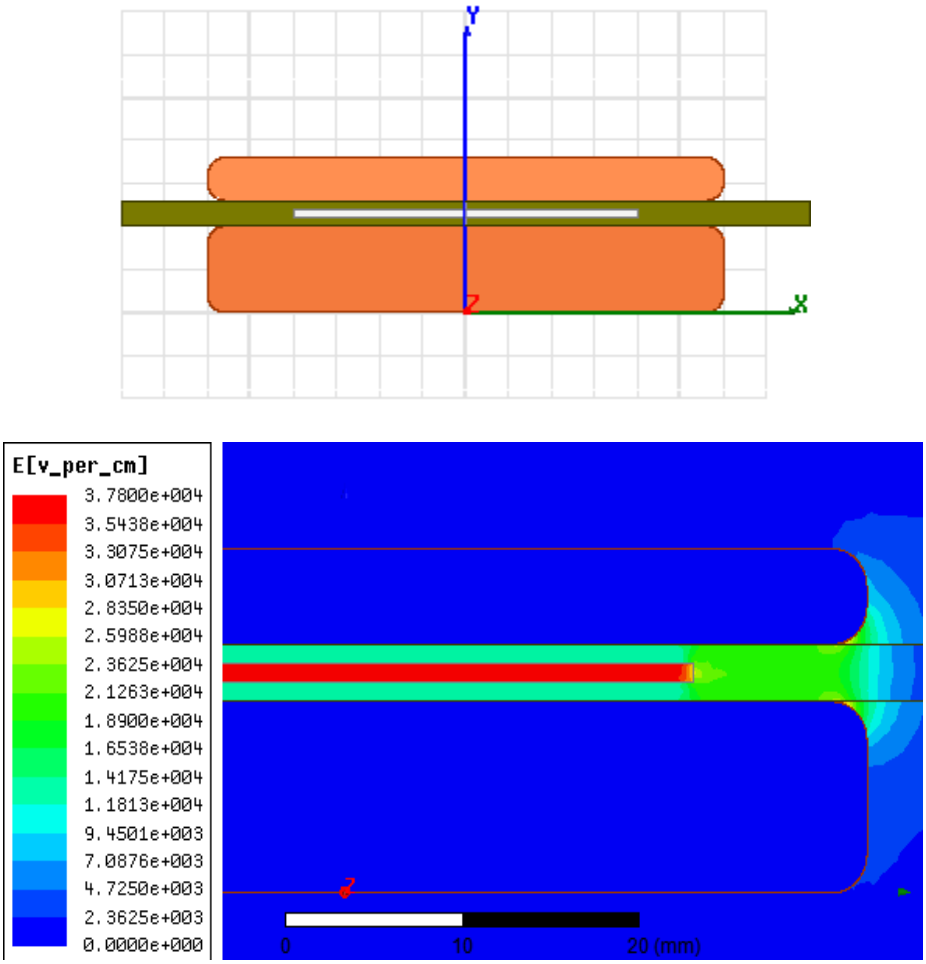


Figure 2. 9 Simulation of inception field, d=40mm, h=1mm

Table 2. 1 Inception voltage and field for cavity of different height, d=40mm

Cavity height (mm)	0.5	1.0	2.0
Mean inception voltage (kV)	5.9	6.3	6.6
Mean inception field (kV/cm)	47.63	37.79	28.28

Table 2. 2 Inception voltage and field for cavity of different diameter, h=0.5mm

Cavity diameter (mm)	10	20	40
Mean inception voltage (kV)	6.5	6.1	5.9
Mean inception field (kV/cm)	52.75	49.96	47.63

In other published research, for spherical cavity, E_{inc} has been defined as

$$E_{inc} = (\frac{E}{p})_{cr} p[1 + \frac{B}{(ph)^n}] \tag{2.4}$$

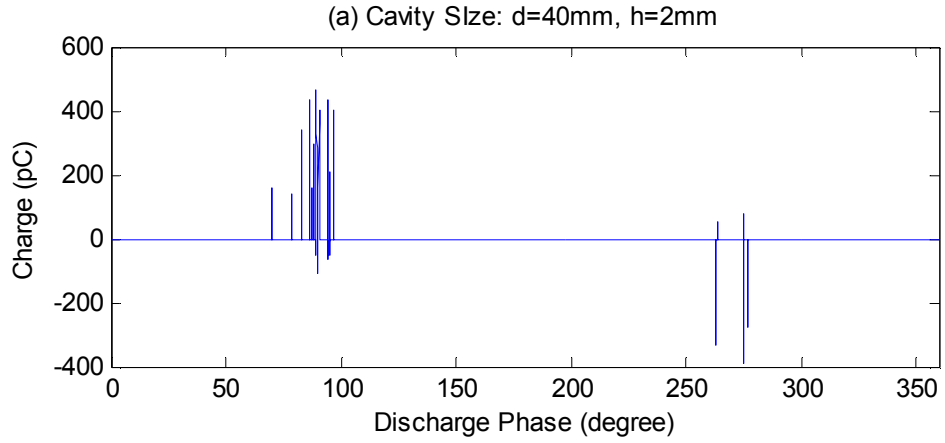
where $(E/p)_{cr}$, B and n are parameters associated with ionization processes in the air gap, p is the pressure in the cavity and h is the cavity height. For air, $(E/p)_{cr} = 24.2 VPa^{-1}m^{-1}$, $n = 0.5$ and $B = 8.6\,\mathrm{Pa}^{0.5}\,m^{0.5}$. This equation indicates that spherical cavity with higher diameter has lower E_{inc}. Results in Table 2. 1 and Table 2. 2 show good agreement with the equation. The effect of cavity height on E_{inc} can be explained by that as cavity height increases, lower electric field is needed for electrons to reach the kinetic energy level to trigger electron avalanche. For longer diameter, E_{inc} decreases slightly. Because the formation of streamer is a random event, to some degree, larger air gap area increases the probability of the formation of discharge channel, reducing E_{inc}.

2.4 Effect of cavity size on PD charge magnitude

During the PD experiment of different cavity size, two hundred power frequency cycles' discharge signals, at the inception voltage, were acquired for each cavity

configuration. Discharge signal of a cycle is shown in Figure 2. 10. Figure 2. 10 (a) is the discharge signal of the cavity whose diameter is 40 millimeters and height is 2 millimeters, and Figure 2. 10 (b) is for the cavity 10 millimeters diametric and 0.5 millimeter high. Obviously, the PD magnitude of large-sized cavity is much greater than that of small size.

Table 2. 3 and Table 2. 4 indicate the detailed experiment results, which come from the average values of two hundred cycles' discharge. As diameter stay constant, the charge magnitude increases rapidly along with the increased cavity height. When height remains the same, the magnitude of different diameter cavity almost stays unchanged. It was found that charge magnitude mainly depends on the height rather than diameter.



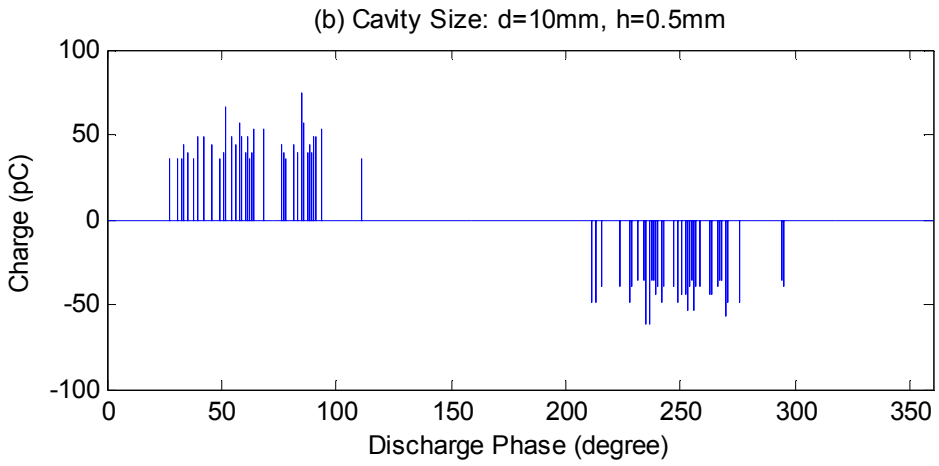


Figure 2. 10 Discharge pulse of a power frequency cycle

Table 2. 3 Charge magnitude for cavity of different height, d=40mm

Cavity height (mm)	0.5	1.0	2.0
Maximum PD magnitude (pC)	98	416	624.
Mean PD magnitude (pC)	47	187	303

Table 2. 4 Charge magnitude for cavity of different diameter, h=0.5mm

Cavity diameter (mm)	10	20	40
Maximum PD magnitude (pC)	84	79	98
Mean PD magnitude (pC)	55	43	47

As shown in Figure 2. 9, electric field in cavity is relatively uniform. The charge magnitude of a single pulse can be represented as Equation 2.5.

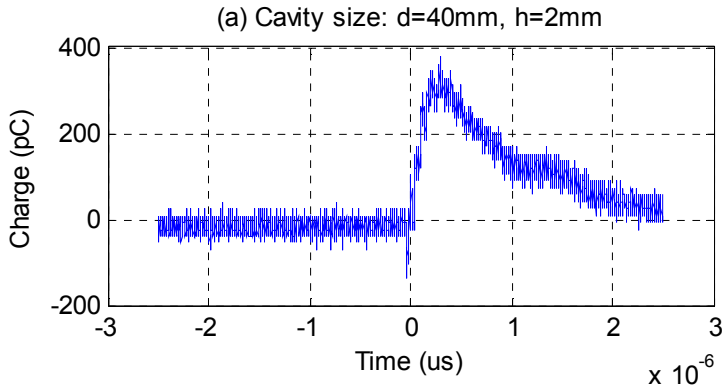
$$Q = NqvSt \tag{2.5}$$

where N is total quantity of electrons generated by avalanche, q is electron charge, v is the electron velocity through a specific streamer area, S and t stands for the single discharge duration time.

In uniform field, the number of electrons due to avalanche, which starts from cathode and then ends on anode, can be calculated by Equation 2.6.

$$N = n_0 e^{\alpha h} \tag{2.6}$$

where n_0 is the number of initial electron, α represents the first Thompson ionization coefficient. As cavity becomes higher, total number of electrons generated by avalanche certainly increases. And according to Table 2. 2, electrons in higher cavity are accelerated by higher potential difference, so the electron velocity in high cavity is greater than that in a low one. Besides, Figure 2. 11 shows the single discharge waveform of different height's cavity. Both discharge duration time, t is almost equal to 2.5 μs, which has no effect on the distinction of Q for each cavity configuration. Therefore, the higher cavity has larger charge magnitude, electrons in higher cavity are accelerated by higher potential difference owing to more electrons generated and electrons of higher speed.



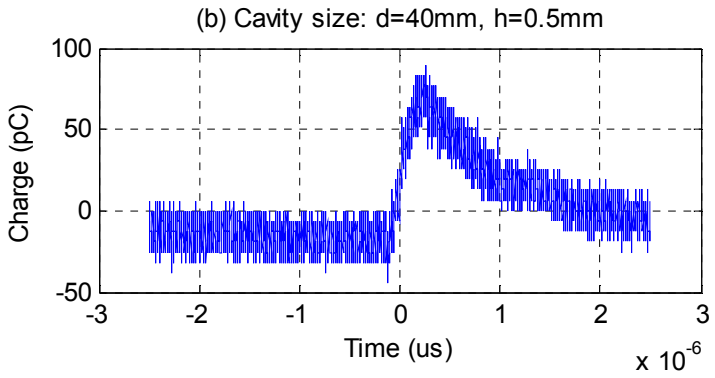


Figure 2. 11 Single discharge waveform

2.5 Effect of cavity size on PD inception phase

From Figure 2. 10, it is obvious to see both positive and negative discharge inception phase of large-sized cavity is much higher than that of small one. Detailed statistical results for two hundred power frequency cycles' discharge signal of each cavity configuration are shown in Table 2. 5 and Table 2. 6. Both larger height and diameter contribute to the higher PD inception phase.

Table 2. 5 PD inception phase for cavity of different height, d=40mm

Cavity height (mm)	0.5	1.0	2.0
Mean positive PD inception phase (degree)	50.10	69.33	76.08
Mean negative PD inception phase (degree)	231.12	245.38	252.51

Table 2. 6 PD inception phase for cavity of different diameter, h=0.5mm

Cavity diameter (mm)	10	20	40
Mean positive PD inception phase (degree)	26.21	42.67	50.10
Mean negative PD inception phase (degree)	208.43	218.38	231.12

When the inception field in the cavity, E_{inc} has been exceeded, the first free electron may not trigger avalanche immediately. Since a neutral gas molecule may absorb the electron and becoming a negative ion or the kinetic energy of electron is not enough to start ionization. Therefore, there should be a time delay before PD occurs. This delay is called average statistical time lag, t_{stat} and it is inversely proportional to the electron generation rate, P_e. Higher electron generation rate results in more collision opportunity for electron and gas molecule, reducing statistical time lag. Referring to Figure 2. 10, Table 2. 5 and Table 2. 6, smaller cavity, both height and diameter, will cause a higher electron generation rate.

Initial free electrons in the cavity may come from volume ionization, P_{vol} or surface emission, P_{surf}.

$$P_e = P_{vol} + P_{surf} \tag{2.7}$$

$$P_{surf}(t) = N_{sc}(t) v_o \exp[-\frac{\psi - \sqrt{q^3 \mid E(t) \mid /(4\pi\varepsilon_0)}}{KT}] \tag{2.8}$$

Volume ionization of initial electron generation is due to gas ionization by energetic photon and detachment of electrons from negative ions. Surface emission is an electron generation process where free electrons are emitted from the cavity surface under the influence of the amount of electrons trapped by the cavity wall $N_{sc}(t)$, detrapping work function ψ, electric field $E(t)$ and temperature T. Nevertheless, in a cavity that has never experienced PD, free surface electrons and energetic photons are extremely scarce. The generation of initial electrons is a highly random

event and where they exactly come from has not been decided yet.

However, after PD has occurred in the cavity, new electrons which start the next avalanche almost all come from the electrons which were left on the cavity wall by last discharge. The electron generation rate basically depends on the surface emission, P_{surf} rather than volume ionization, P_{vol}. The reason why smaller cavity has the property of higher surface emission rate is detailed as follows.

As indicated in equation 2.6, electron generation rate due to surface emission, P_{surf} is a function of the amount of electrons trapped by the cavity wall or on the cavity surface, $N_{sc}(t)$ and the detrapping work function, ψ. The electrons accumulated on the cavity duo to PD decay with time through three phenomenon, ① injection into electrode through insulation board conductivity [18], ② emission from traps at the following half power frequency cycle and ③ recombination with positive ions left by previous PD [13].

At the half power frequency cycle when PD occurs, after electrons hit the cavity surface, some of them can continue to go through the insulating board due to the material conductivity and the applied stress. From the point of equivalent circuit, there must be more paralleled resistances connecting cavity surface and electrode for the cavity with larger diameter. Higher conduction results in higher surface charge decay

rate. Therefore, the amount of charges left on the cavity surface, $N_{sc}(t)$ is lower for larger diameter. With regard to effect of cavity height, the work function, ψ plays a important role. Electrons within higher cavity hit insulating pressboard at higher speed, going into deeper traps of the pressboards. This leads to higher detrapping work function that electrons require more energy to be emitted from the traps to start another avalanche.

Hence, according to Equation 2.8, lower $N_{sc}(t)$ and higher ψ for larger cavity result in lower electron generation rate duo to surface emission, also longer statistical time lag and higher inception phase.

2.6 Summary

This chapter introduces the mechanism of air-gap PD as well as the setup of air-gap PD experiment considering different cavity size and emphasizes on interpreting the effect of cavity size on PD activity through experiment results.

1) Large cavity PD possesses lower inception field. Higher cavity height contributes more to the lower inception field than larger cavity diameter.
2) Large cavity PD possesses higher charge magnitude. Charge magnitude mainly depends on the cavity height.
3) Due to lower inception field and higher charge magnitude, PD happening large cavity is more harmful than that happening in small cavity.

4) Large cavity PD possesses higher inception phase. Both cavity height and diameter have impact on the inception phase.

Chapter 3 Air-gap PD development characteristics and feature information extraction

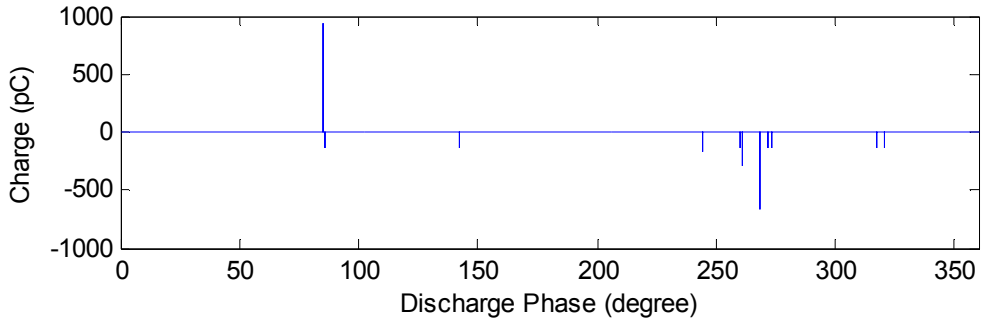
In order to evaluate the harmfulness of the detected air-gap PD, the PD development characteristic has been studied in some published research. However, due to the effect of cavity size on PD activity, previous work on PD development process needs to be researched further, making the severity diagnosis more accurate. This chapter first discusses about the PD development characteristic within a large cavity as well as a small one. Then feature information extraction based on statistical parameters of PRPD pattern is introduced.

3.1 PD development characteristic

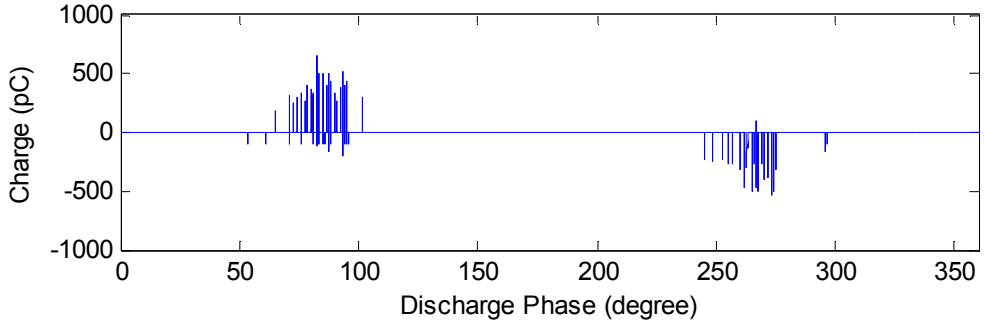
The insulation degradation due to air-gap PD is a long-term process, which could continue more than five years until breakdown depending on the field in the cavity[19]. Under laboratory conditions, the testing voltage should be higher than the inception voltage to accelerate the process of insulation deterioration, making whole system breakdown occurs at a reasonable time. According to the previous experiment experience, the testing voltage for large cavity, whose diameter is $40mm$ and height is $2mm$, is set to 1.4 times as much as the inception voltage and the voltage for small one, $10mm$ wide and $0.5mm$ high, is double inception voltage. Discharge signals of four hundred power frequency cycles were acquired every half an hour until breakdown.

3.1.1 Large cavity model

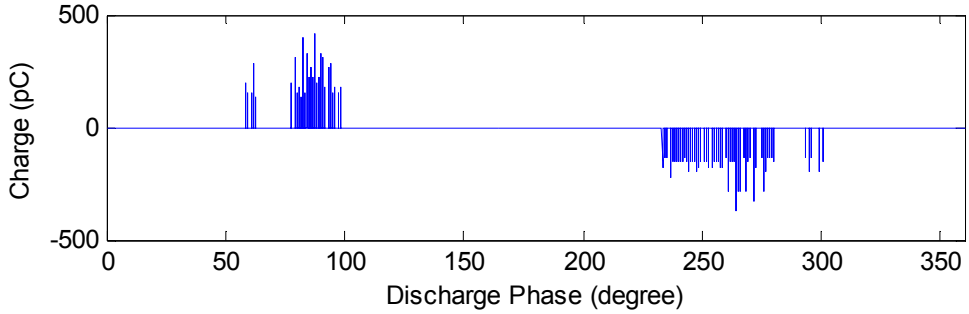
Large cavity model was energized at $8.5kV$. It took twenty-eight hours to the final breakdown. Discharge signals at different time are shown in Figure 3. 1. At first, PD magnitude was relatively high and the discharges mainly concentrated on 90 and 270 degree where the peak values of power frequency cycle are. After 5^{th} hour, the magnitude started declining but inception phase extended to around 60 degree. PD magnitude stayed about $400pC$ for a long time. At 19^{th} hour, PD suddenly became extremely fierce, the magnitude increasing sharply and inception phase expanding much. Then discharge continued developing and the whole system was broken down at 28^{th} hour.



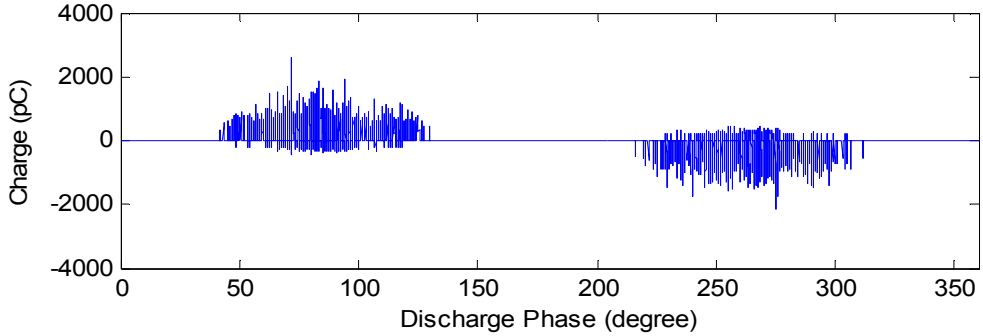
(a) 0.5^{th} hour



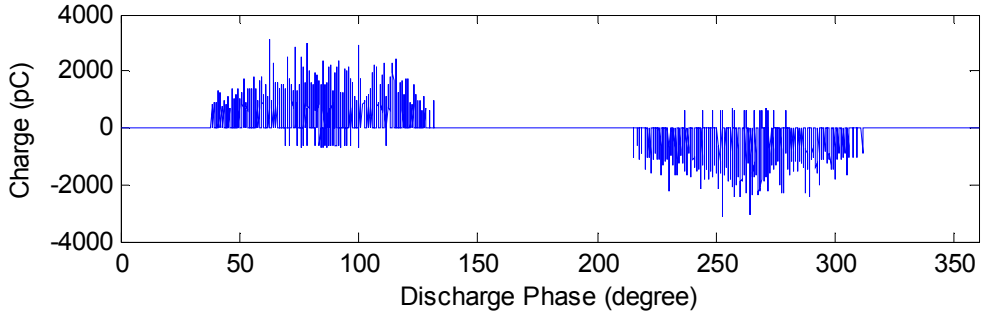
(b) 6th hour



(c) 14th hour



(d) 22nd hour



(e) 27th hour

Figure 3. 1 Large cavity discharge signals at different time

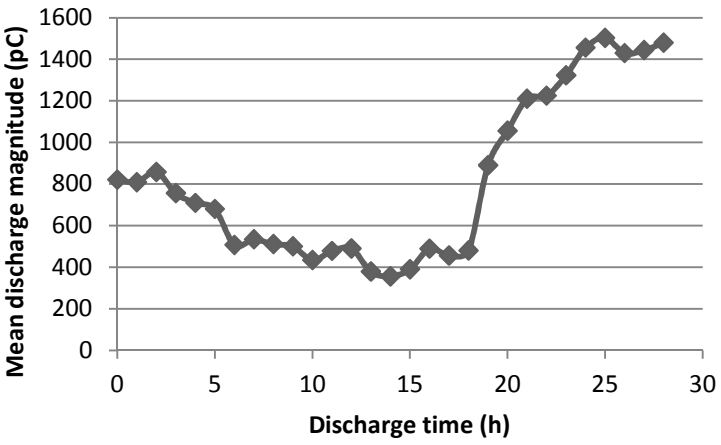
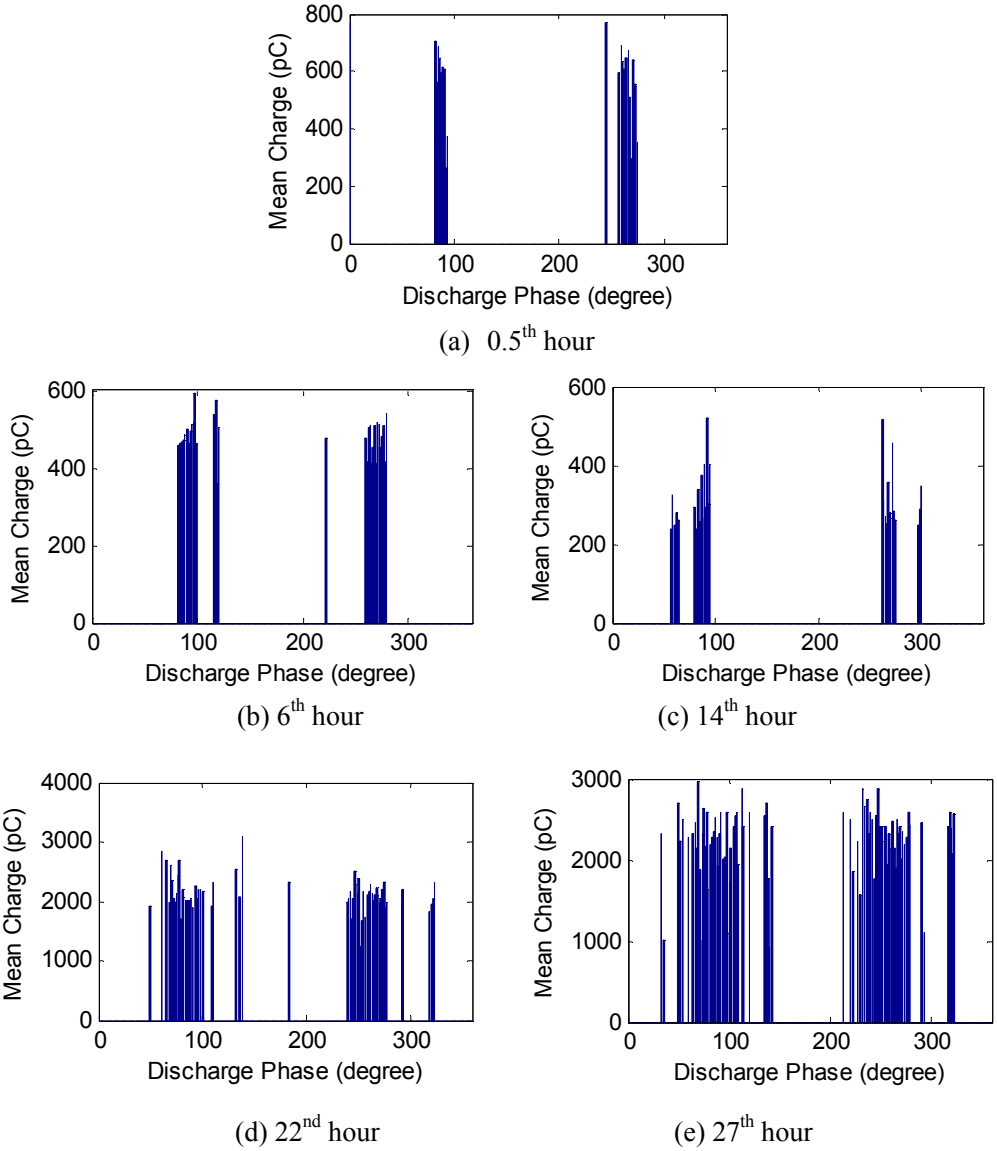


Figure 3. 2 Mean PD magnitude variation for large cavity

Figure 3. 2 indicates the variation of mean PD magnitude during whole process. In theory, damaged condition of the insulation can be identified through the magnitude variation. However in reality, this doesn't work. Firstly, different PD fault type or fault location in power transformer has diverse charge magnitude level. Compared to air-gap discharge, surface discharge as well as corona discharge possesses higher magnitude. PD happening in higher electric field area possesses higher magnitude. Secondly, relative costs of online real time PD monitoring systems compared to an asset value have resulted in their use being reserved for most critical apparatus [20]. Hence, only offline PD magnitude at a specific time can't figure out the discharge development stage. Besides, there may be more than one PD source within transformer. Only depending on magnitude variation is not able to separate PD source or identify PD type. Therefore, more PD feature information is needed. PRPD pattern, which takes PD magnitude, phase and number into account, is widely used in PD severity diagnosis. Four hundred discharge signals at one point were taken to build the four kinds of two-dimensional distribution images below, which are the maximum

charge magnitude distribution with respect to discharge phase $H_{q\max}(\varphi)$, mean charge magnitude distribution with respect to the phase $H_{qm}(\phi)$, distribution of pulse number versus the phase $H_n(\varphi)$ and distribution of mean charge magnitude versus related phase $H_n(q)$. $H_{qm}(\phi)$, $H_n(\varphi)$ and $H_n(q)$ of large cavity model at different time are shown as follows.



(a) 0.5th hour

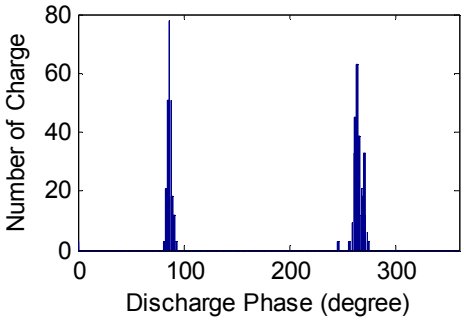
(b) 6th hour

(c) 14th hour

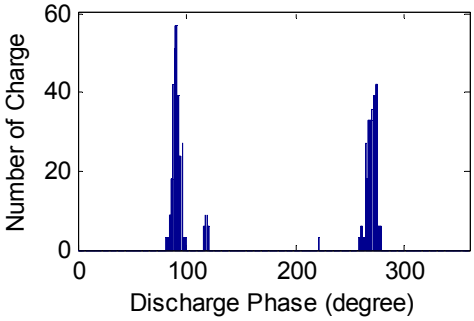
(d) 22nd hour

(e) 27th hour

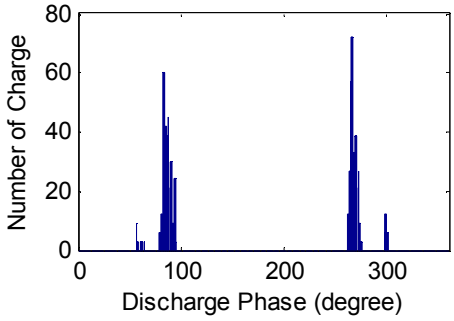
Figure 3. 3 $H_{qm}(\phi)$ of large cavity model



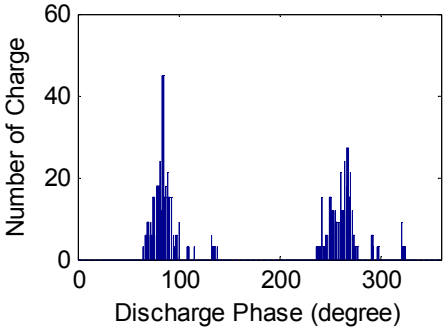
(a) 0.5th hour



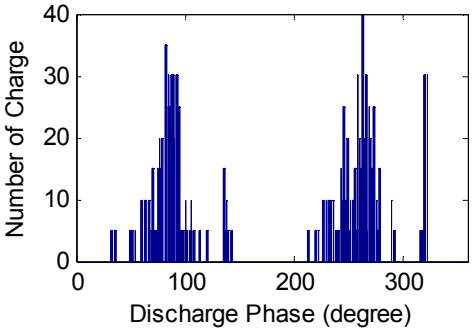
(b) 6th hour



(c) 14th hour

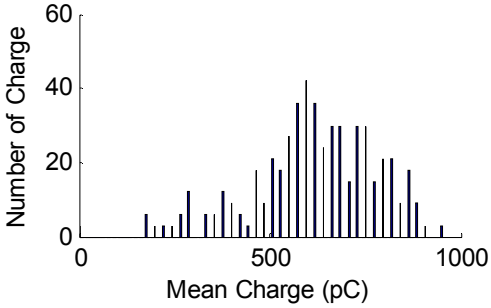


(d) 22nd hour

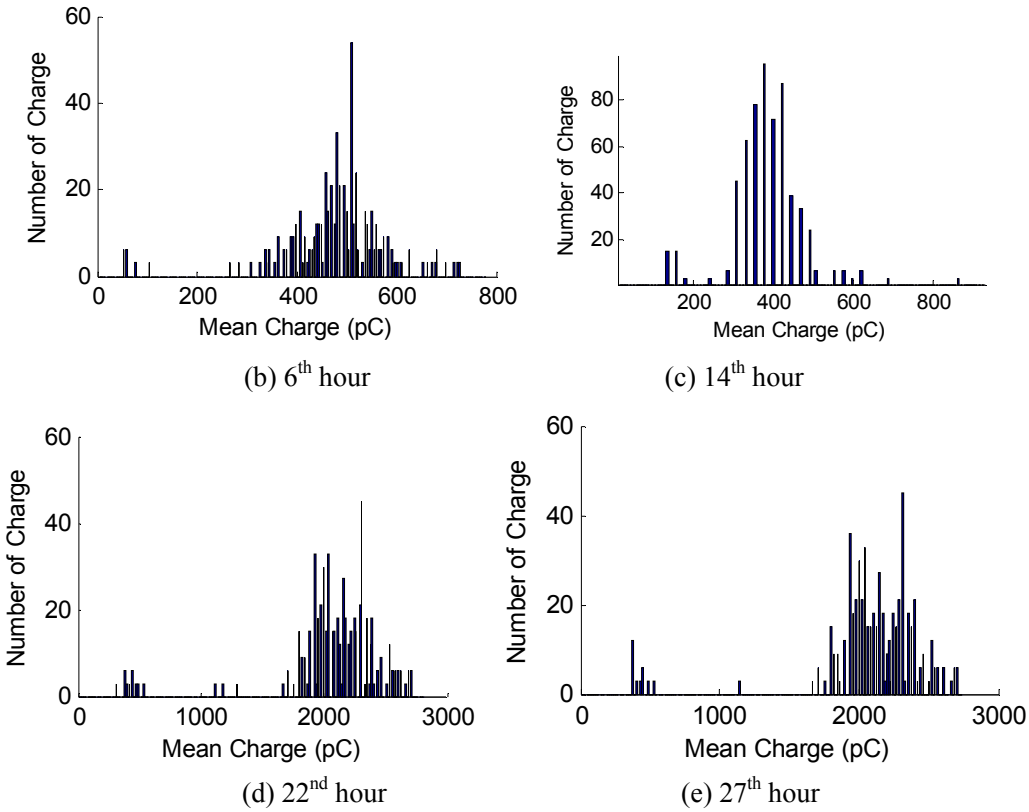


(e) 27th hour

Figure 3. 4 $H_n(\varphi)$ of large cavity model



(a) 0.5th hour



(b) 6th hour (c) 14th hour

(d) 22nd hour (e) 27th hour

Figure 3.5 $H_n(q)$ of large cavity model

At initial stage of PD, relatively high magnitude is due to enough gas molecules. The greater probability free electrons collide with gas molecules, the more electrons can be generated. According to Equation 2.5, more electrons lead to higher magnitude. As PD developed, O_2 and N_2 in the cavity were consumed, pressboards were corroded, HNO_3, H_2 and CH_4 were generated. The gas volume in cavity can decrease down to about eighty percent [21]. Meanwhile, masses of irregularly distributed cellulose filaments existing on the surface of insulation paper disappeared due to ablating and insulation surface became smooth. Three dimensional insulation surface roughness at different discharge time was calculated by atomic force microscope in [22]. It shows that the surface roughness in the medium term is the lowest. This makes electric field distortion due to surface roughness lower. Equation

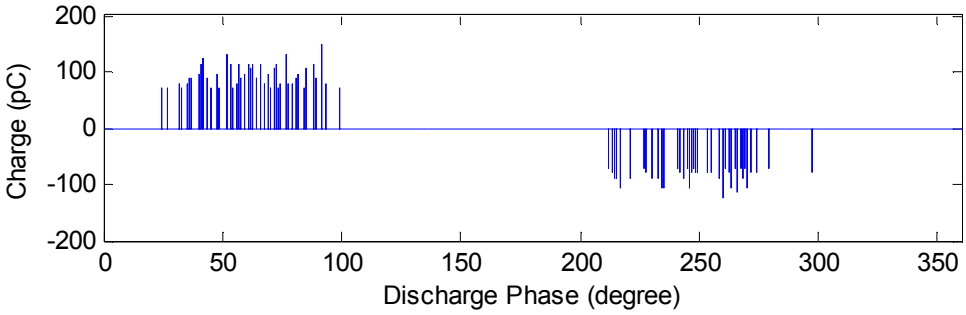
2.8 shows that lower field reduces the amount of free electrons due to surface emission. Therefore, both lower collision probability with gas and fewer free electrons result in lower PD magnitude in interim stage. As PD continuously developed, more oil was decomposed into H_2 and CH_4. Besides, pressboards were bombarded and corroded for a long enough time, so that large amount of CO as well as CO_2 were produced and surface roughness started increasing from the lowest point, becoming much higher [21, 22]. Those greatly increased the PD magnitude.

Throughout the whole power frequency, PD phase is symmetrical due to no corona effect. In addition, PD phase expanded along with increased time. That is because there are more electrons accumulated on the cavity wall and strong electronegative gas (oxygen), which can capture free electron and become anion, depleted. Both factors increase the probability of effective electron generation, leading to lower average statistical time leg and lower PD phase angle.

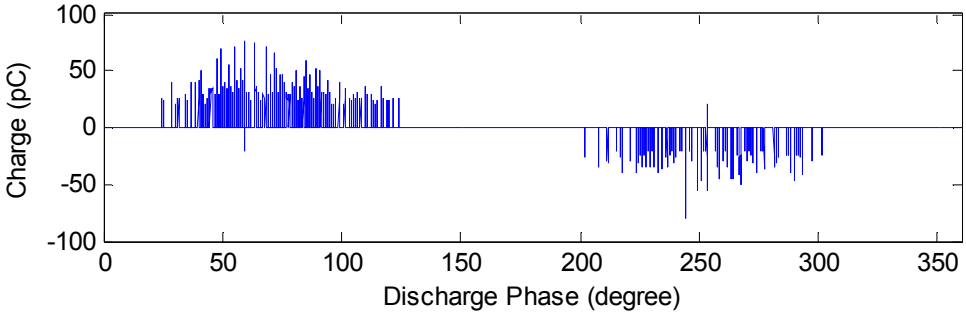
3.1.2 Small cavity model

Small cavity model was energized at $13kV$. It took a longer time, thirty-three hours to the final breakdown. Discharge signals at different time and mean PD magnitude variation are shown in Figure 3. 6 and Figure 3. 7 respectively. PD development characteristic of small cavity is similar to the large one's. PD magnitude variation is a concave curve and PD phase slowly expands. More statistical information are shown

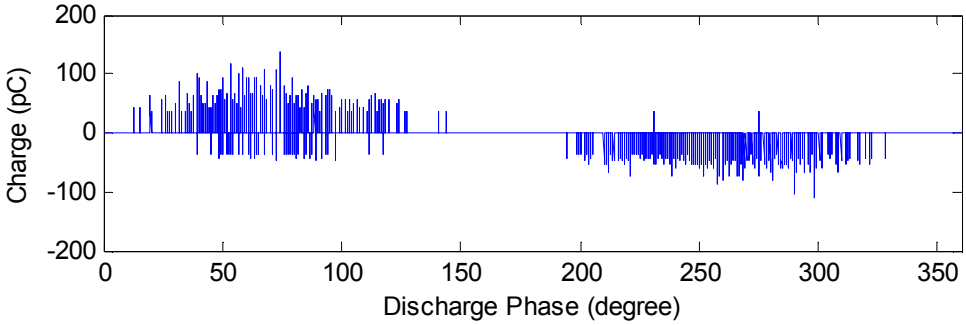
from Figure 3. 8 to Figure 3. 10.



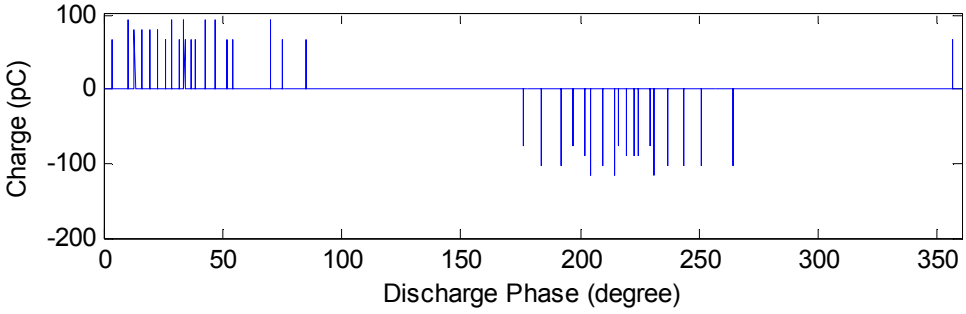
(a) 0.5th hour



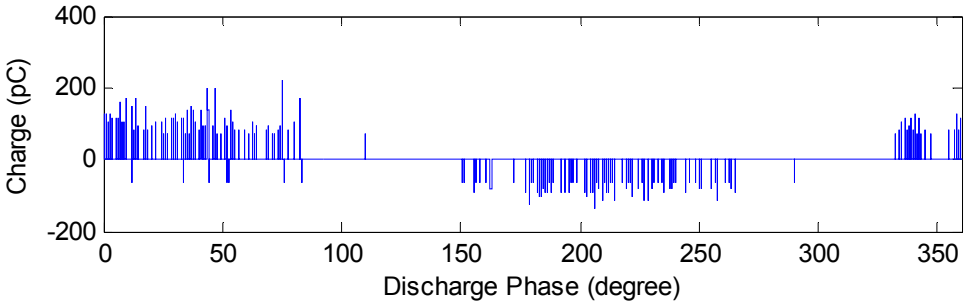
(b) 7th hour



(c) 15th hour



(d) 23rd hour



(e) 30th hour

Figure 3. 6 Small cavity discharge signals at different time

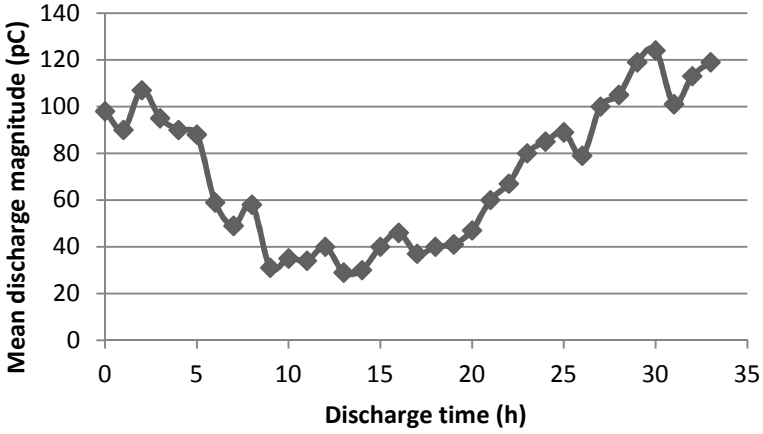
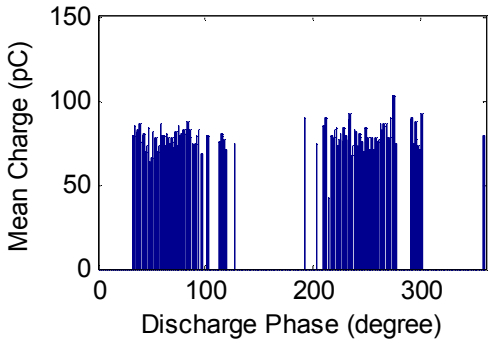
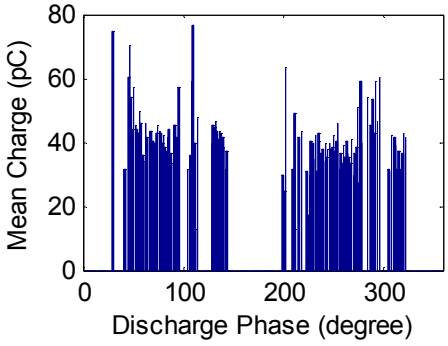


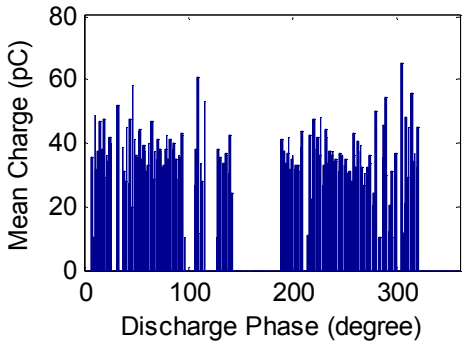
Figure 3. 7 Mean PD magnitude variation for small cavity



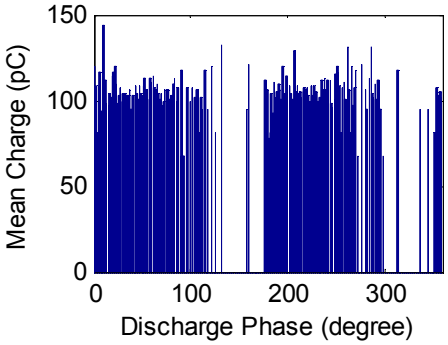
(a) 0.5th hour



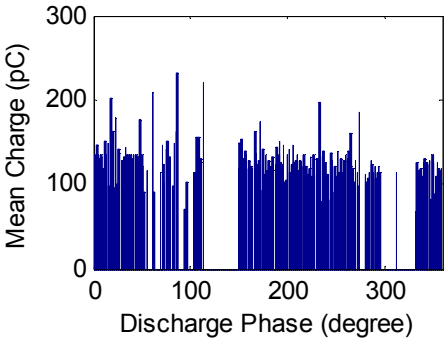
(b) 7th hour



(c) 15th hour

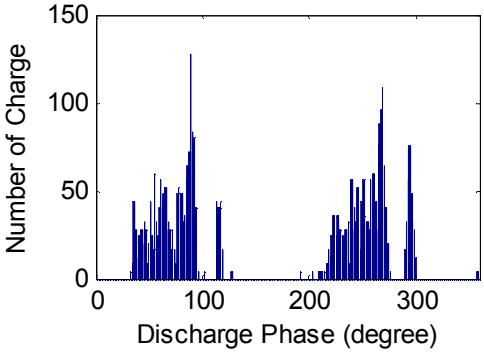


(d) 23rd hour

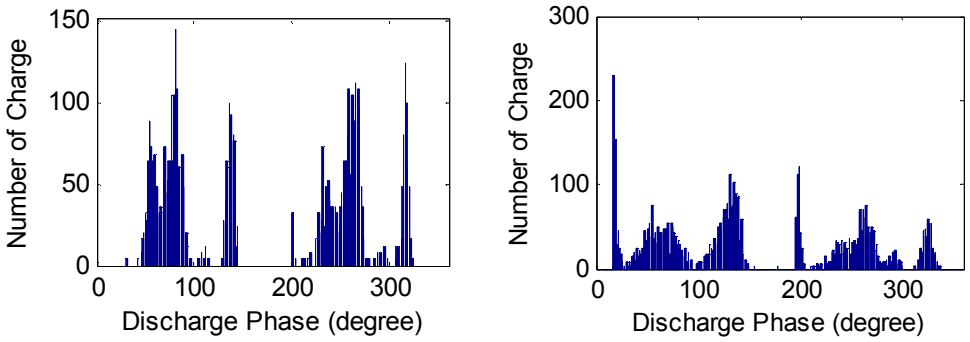


(e) 30th hour

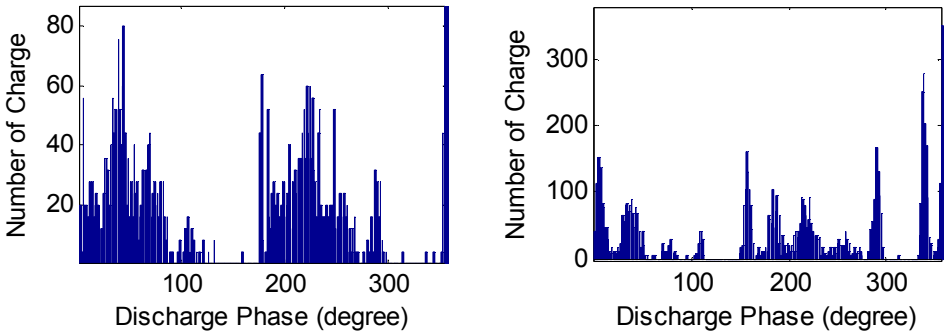
Figure 3. 8 $H_{qm}(\phi)$ of small cavity model



(a) 0.5th hour



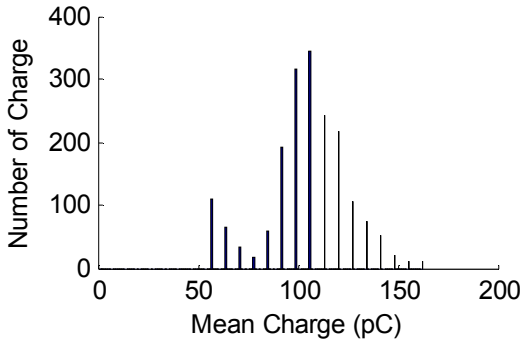
(b) 7th hour (c) 15th hour



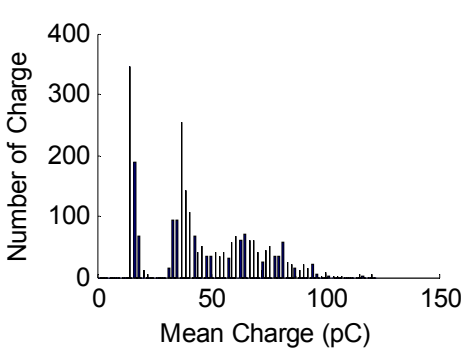
(d) 23rd hour (e) 30th hour

Figure 3. 9 $H_n(\varphi)$ of small cavity model

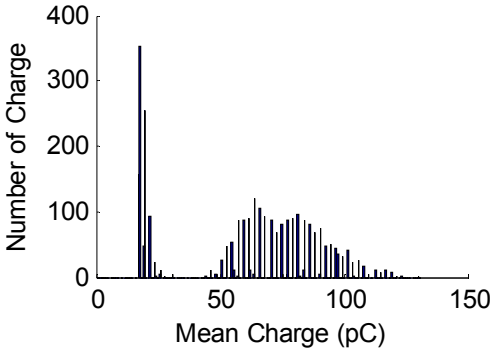
Due to the lower inception phase for small cavity interpreted in Chapter 2, $H_{qm}(\phi)$ and $H_n(\varphi)$ spanned much wider, making them look like less sharp. When PD came to later stage, the positive PD even expanded to the negative half cycle and vice versa. As shown in Figure 2. 2, opposite field, E_s will be established after PD occurs. For small cavity, more concentrated charge distribution and lower height result in the higher reverse field, which can make the resultant field in cavity, E_{cav} turn over before sinusoidal voltage drops to zero. Therefore, positive PD can occur at the end of negative half cycle.



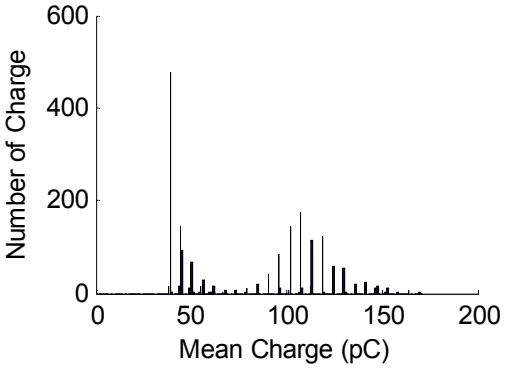
0.5th hour



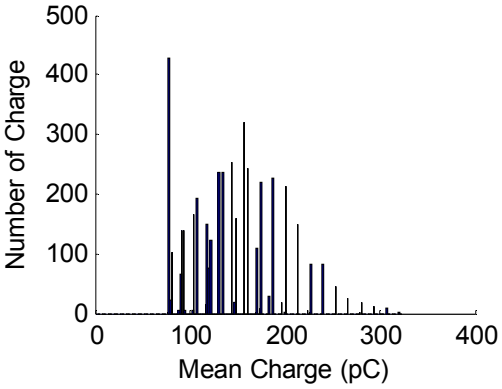
(b) 7th hour



(c) 15th hour



(d) 23rd hour



(e) 30th hour

Figure 3. 10 $H_n(q)$ of small cavity model

Compared with the $H_n(q)$ distribution for large cavity, number of low magnitude PD, for small one, accounts for much higher proportion.

3.2 Statistical parameter extraction of Phase Resolved Partial Discharge (PRPD) pattern

PD distribution images of large or small cavity at different time vary significantly, which can be used for diagnosis. However, only the operators with much knowledge and experience on PD can do the right diagnosis. Hence, quantification of the feature information is very necessary. Abstracting statistical parameters from the two dimensional distribution images talked about in session 3.1 is one of the most effective and sophisticated methods for quantitative analysis. It was successfully used to do PD diagnosis in [23-25]. The statistical parameters adopted in this paper are degree of skewness (Sk), degree of kurtosis (Ku), degree of asymmetry (Asy), number of local peak value $(peaks)$ and cross-correlation coefficient (Cc), which are detailed below.

Sk represents degree of skewness for a distribution image compared to Gaussian distribution,

$$Sk = \frac{\sum_{i=1}^{n} (y_i - \mu)^3 p_i}{\sigma^3} \tag{3.1}$$

where y_i is vertical coordinate value of distribution image, indicating PD magnitude or number. μ is the mean value, σ is the standard deviation and p_i is the probability of pulse occurring in *i*th phase window. In this paper, the power frequency cycle is divided into 256 phase windows, so n equals 128, each window

corresponding to approximate 1.4 degree.

$$\mu = \sum_{i=1}^{n} y_i p_i \tag{3.2}$$

$$\sigma = \sqrt{\sum_{i=1}^{n} (\mathrm{y}_i - \mu)^2 p_i} \tag{3.3}$$

If Sk equals to zero, it indicates that distribution image is symmetrical. If Sk is greater than zero, it indicates the image is left skewed to Gaussian distribution and vice versa.

Ku , which is defined in Equation 3.4, represents the degree of kurtosis compared to Gaussian distribution.

$$Ku = \frac{\sum_{i=1}^{n} (\mathrm{y}_i - \mu)^4 p_i}{\sigma^4} - 3 \tag{3.4}$$

The parameters within Equation 3.4 have the same definition with those in Equation 3.1. On condition that Ku is greater than zero, the distribution image should be sharper compared to Gaussian distribution and vice versa.

Asy represents asymmetry of PD magnitude or number in positive and negative cycle.

$$Asy = N_1 \sum_{i=1}^{N_2} y_i^- \Big/ N_2 \sum_{i=1}^{N_1} y_i^+ \tag{3.5}$$

where "y_i^+" and "y_i^-" are the vertical coordinate value for ith phase window in positive and negative cycle respectively, N_1 is the number of phase window, where

PD occurred, in positive cycle and N_2 stands for that in negative cycle. Higher Asy indicates more intense PD in negative cycle.

Cc is used to evaluate similarity of distribution image shape in positive and negative cycle. Cc equals one means that the shape is very similar and Cc equals zero signify that there exists significant difference.

$$Cc = \frac{\sum y_i^+ y_i^- - \sum y_i^+ \sum y_i^- / n}{\sqrt{[\sum (y_i^+)^2 - (\sum y_i^+)^2 / n][\sum (y_i^-)^2 - (\sum y_i^-)^2 / n]}} \tag{3.6}$$

According to the definitions above, 27 dimensional statistical parameters, which are shown in

Tabla 3. 1, can be abstracted from the four distribution images for a specific data collection time. 57 data sets were acquired for large cavity model and 67 data sets for small cavity model.

Tabla 3. 1 Statistical parameters

Distribution images	Parameters							
$H_{q\max}(\varphi)$	$sk+$	$sk-$	$ku+$	$ku-$	$peaks+$	$peaks-$	Asy	cc
$H_{qn}(\varphi)$	$sk+$	$sk-$	$ku+$	$ku-$	$peaks+$	$peaks-$	Asy	cc
$H_n(\varphi)$	$sk+$	$sk-$	$ku+$	$ku-$	$peaks+$	$peaks-$	Asy	cc
$H_n(q)$	sk		ku		$peaks$		—	—

3.3 Dimension reduction of statistical parameters based on Kernel based Principal Component Analysis (KPCA)

The 27 dimensional statistical parameters express sufficient information within the distribution images. Yet there may be some overlapping information. Taking all of them as input to the diagnosis classifier will definitely make the classifier more complicated and harder to train. Therefore, kernel principal component analysis (KPCA), which can achieve parameter dimension reduction and keep original information, was applied.

3.3.1 Theory of KPCA

KPCA is an extension of principal component analysis (PCA) based on kernel trick. Theory of PCA is described as follows.

①PCA

Given a dataset $X=[\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3\cdots\mathrm{X}_n]$ where X_i is a m dimensional column vector and m is the number of samples. PCA is an orthogonal linear transformation that transforms X into a new coordinate system such that the greatest variance lies on the first coordinate (called the first principal component), the second greatest variance on the second coordinate and so on. X can be transformed into a p dimensional row vector, Y through multiplying a transformation matrix, T with n rows and p columns. How to get value of p and the transformation matrix are

detailed below.

$$Y_{1\times p} = X_{1\times n} T_{n\times p} \tag{3.7}$$

The n by n covariance matrix, $\sum$ of data set X is calculated by Equition 3.8.

$$C_x = \frac{1}{m}[(X - E[X])(\mathrm{X} - \mathrm{E[X]})^T] = \frac{1}{m}\sum_{i=0,j=0}^{i=n,j=n}[(X_i - \mu_i)(X_j - \mu_j)^T] \tag{3.8}$$

$$C_x U_i = \lambda U_i, \quad i = 1, 2\cdots n \tag{3.9}$$

Then eigenvalue, λ_i and eigenvectors, U_i of C_x are calculated and λ_i are sorted from high to low. Accumulating contribution rate is defined in Equation 3.10.

$$\eta = \frac{\sum_{i=1}^{p}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \tag{3.10}$$

When η is high enough, typically higher than 90 percent, corresponding p is set and corresponding top- p eigenvectors generate the n by p transformation matrix.

②KPCA

Using a kernel, the originally linear operations of PCA are done in a reproducing kernel Hilbert space with a non-linear mapping, ϕ. Hence, the components obtained by KPCA can be taken as the non-linear principle components in original space. First, original dataset X is mapped into high dimensional feature space through ϕ function. Then covariance matrix is calculated.

$$C_{\phi(x)} = \frac{1}{m}\sum_{i=0,j=0}^{i=m,j=m}[\phi(X_i - \mu_i)][\phi(X_j - \mu_j)^T] \tag{3.11}$$

However, we don't know where the feature space is and ϕ function is never calculated explicitly. Kernel method, which can transform the inner product operation

in feature space into kernel function calculation in original space, is introduced to solve the problem.

$$K(X_i - \mu_i, X_j - \mu_j) = [\phi(X_i - \mu_i)][\phi(X_j - \mu_j)^T] \tag{3.12}$$

Polynomial kernel function, Equation 3.13 and Gaussian kernel function, Equation 3.14 are widely used.

$$K(x, y) = (x^T y + c)^2 \tag{3.13}$$

$$K(x, y) = \exp\left[-\frac{\|x - y\|^2}{\sigma^2} \right] \tag{3.14}$$

After obtaining covariance matrix, likewise, the transformation matrix is computed as the procedures discussed about in PCA.

$$(v^k \cdot \phi(\mathrm{x})) = \sum_{i=1}^{m} (\alpha_i)^k K(x_i, x) \tag{3.15}$$

3.3.2 Dimension reduction result of statistical parameters

The 27 dimensional statistical parameters from large cavity model and small one are reduced dimensions together by using PCA and KPCA. Accumulating contribution rate through PCA and KPCA are shown in Figure 3. 11 and Figure 3. 12 respectively. The contribution rate of the first six components coming from PCA is 86.11%, which means the six components preserve 86.11 percent information of the original 27 dimensional parameters. Compared with PCA, KPCA is more effective. The first six components preserve 96.1% of all the original information. So the 27 dimensional statistical parameters are reduced to six dimensions with the help of KPCA. And the 3-d visualization of KPCA results is shown in Figure 3. 13.

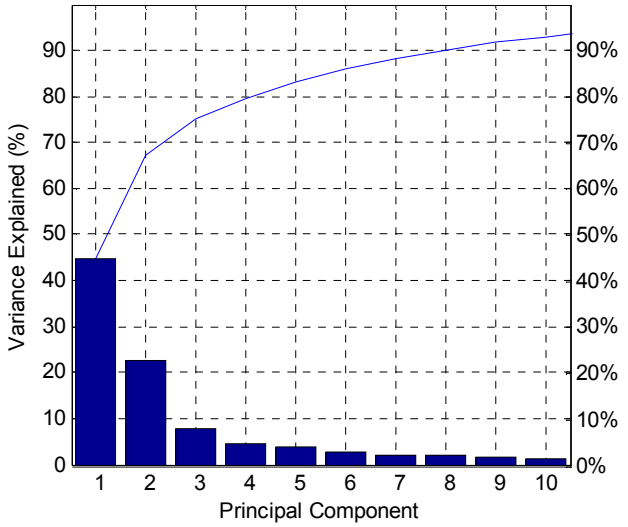


Figure 3. 11 Accumulating contribution rate through PCA

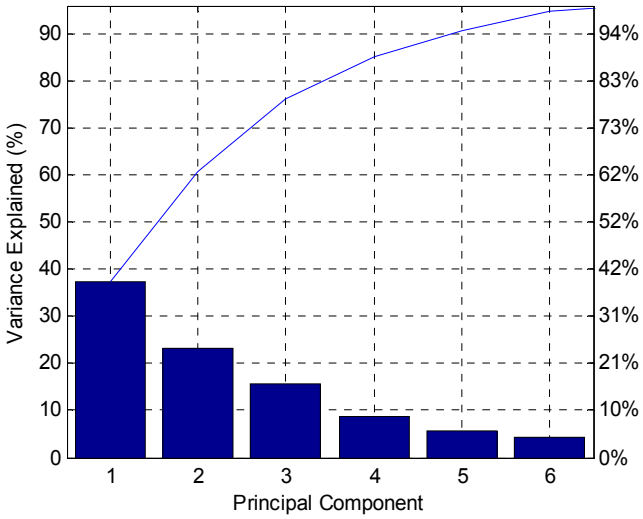


Figure 3. 12 Accumulating contribution rate through KPCA

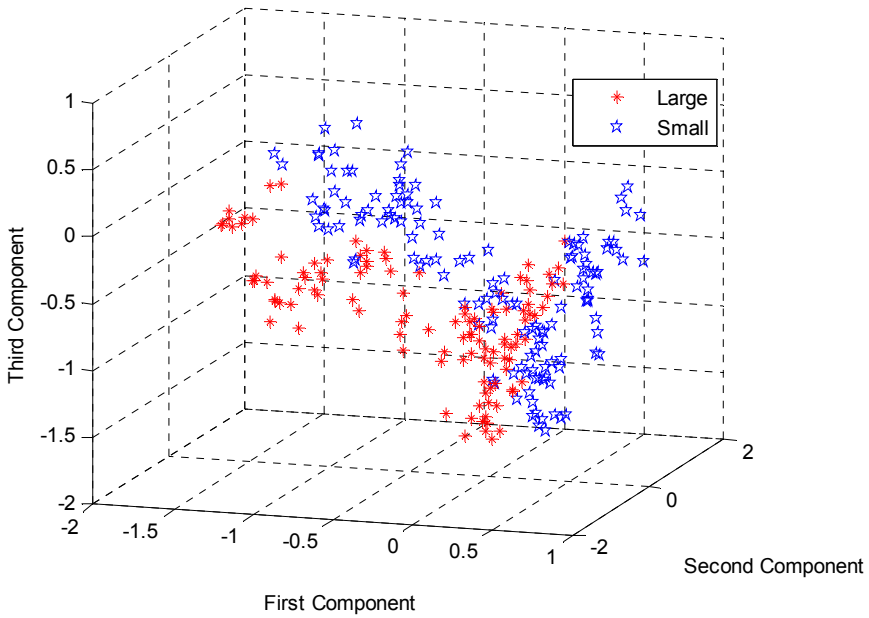


Figure 3. 13 3-d visualization of KPCA results

3.4 Summary

The development characteristics of large cavity and small cavity PD are studied in this chapter. Furthermore, statistical parameters are extracted from PRPD pattern for quantitative analysis.

1) During Air-gap PD development process, charge magnitude variation of large and small cavity model both presents concave curve shape with respect to time. Besides, discharge phase continuously expends and the phase span of large cavity model is much narrower than that of small cavity model. When air-gap PD comes to the last stage, for small cavity model, the positive PD can even expand to the negative half cycle and vice versa.
2) 27 dimensional statistical parameters from PRPD pattern are carried out dimension reduction through PCA and KPCA. KPCA shows a better performance,

keeping 96.1% of original feature information in the first six principle components.

Chapter 4 Air-gap PD development stage unsupervised classification and identification considering cavity size

Discharge process was accelerated due to higher field in lab experiment and PD development is a relatively random event. Besides in site operation, it is highly costly and electromagnetic interference sensitive to implement on-line real-time PD monitoring for power transformers. What operators can obtain is the PD signal at a specific time when dissolved gas analysis (DGA) shows an abnormal gas level. Therefore, it's not practical to use the lab experiment data of PD development process and some site operation data to predict a precise time when a power transformer will break down.

In order to diagnose the insulation condition, what is feasible is to divide the PD development into several stages according to the whole discharge experiment data, so unsupervised classification method, clustering analysis is put to use in this chapter to do PD development stage classification. Then the clustering results can be used to train a supervised classifier, which can identify the development stage that an unknown PD sample is at.

4.1 Air-gap PD development stage classification based on clustering analysis

Clustering analysis is an important cognition technique for mankind to discover the unknown world. It is the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. In this section, hierarchical clustering and K means clustering are used to divide air-gap PD development stage.

4.1.1 Hierarchical clustering

Hierarchical clustering is a method of cluster analysis that seeks to build a hierarchy of clusters. Agglomerative strategy is used for hierarchical clustering. This is a "bottom up" approach: each observation starts in its own cluster and pairs of clusters are merged as one moves up the hierarchy. The results of hierarchical clustering are usually presented in a dendrogram. In order to decide which clusters should be combined, a measure of similarity between is required. This is achieved by use of an appropriate metric, a measure of distance between pairs of observations, and a linkage criterion which specifies the similarity of sets as a function of the pairwise distances of observations in the sets.

Some commonly used metrics for hierarchical clustering are shown below. X and Y are two observations.

1) Euclidean distance

$$D(X,Y)=\sqrt{\sum_{i=0}^{n}(\mathrm{X}_i-Y_i)^2} \tag{4.1}$$

2) Manhattan distance

$$D(X,Y)=\sum_{i=0}^{n}\left|X_i-Y_i\right| \tag{4.2}$$

3) Mahalanobis distance

$$D(X,Y)=(X-Y)^T\times\sum{}^{-1}\times(X-Y) \tag{4.3}$$

where $\sum$ is covariance matrix of samples.

Then common linkage criterions that determine the distance between two clusters, G_1 and G_2, are introduced.

1) Minimum linkage clustering

Distance between the two nearest points in two clusters determines the similarity of the two clusters.

$$D(G_1,G_2)=\min\{D(X,Y)\mid X\in G_1,Y\in G_2\} \tag{4.4}$$

2) Maximum linkage clustering

Distance between the two farthest points in two clusters determines the similarity of the two clusters.

$$D(G_1,G_2)=\max\{D(X,Y)\mid X\in G_1,Y\in G_2\} \tag{4.5}$$

3) Mean linkage clustering

Mean distance between two points in two clusters determines the similarity.

$$D(G_1, G_2) = \frac{1}{n_1 \times n_2} \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} D(X_i, Y_j) \tag{4.6}$$

4) Centroid linkage clustering

Distance between the centroids in two clusters determines the similarity.

$$D(G_1, G_2) = D(\overline{X}, \overline{Y}) \tag{4.7}$$

Hierarchical clustering is conducted as following steps.

① Construct an initial distance matrix, where the number in the i-th row j-th column is the distance between the i-th and j-th observations. In this step, every observation is a cluster.

② Find the minimum mean distance between two clusters in distance matrix, merge one cluster into the other becoming a new cluster. Total cluster number reduces one.

③ Recalculate the distance matrix.

④ Repeat ② and ③ until total cluster number drops to one.

⑤ Draw the dendrogram.

⑥ Stop clustering when the clusters are too far apart to be merged or when there is a sufficiently small number of clusters.

4.1.2 K means clustering

In hierarchical clustering, once an observation is partitioned to a specific cluster, it can't be assigned to other later. This requires a very explicit clustering at first. Besides, it is not economical to calculate the distance matrix when data is very big. K means

clustering is an iterative algorithm, originally from signal processing, which can dynamically partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Steps of K means clustering are introduced as follows.

① Select k observations as initial clustering centers.

② Assign each observation into the optimal cluster center which yields minimum squared Euclidean distance from the observation.

③ Calculate the new means to be the centers of the observations for next clustering.

④ Stop clustering until the centers stay unchanged.

4.1.3 Clustering results

Hierarchical clustering and K means clustering are implemented with the help of software PASW Statistics 18. Hierarchical clustering results of six dimension parameters throughout whole span of experiments for large cavity model and small cavity model are shown in Figure 4. 1 and Figure 4. 2 respectively. Euclidean distance and mean linkage clustering are employed. For large cavity model, when rescaled distance is 13.3, 57 samples, which are marked as t1, t2, t3…t57, are classified into three clusters. The first cluster is from t1 to t14, the second one is from t15 to t38 and the third one is from t39 to t57. Similar clustering results for 67 samples from small cavity model are obtained, t1 to t16, t17 to t44, t45 to t67 except t56 that is classified into the second cluster.

The three clusters for PD development stage show good agreement with rough stage classification by the mean PD magnitude variation curve in Figure 3. 2 and Figure 3. 7, by gas volume variation in cavity[21] and by three-dimensional insulation surface roughness variation[22]. The magnitude in first stage, 0^{th} hour to 6.5^{th} hour for large cavity model and 0^{th} hour to 7.5^{th} hour for small cavity model, is relatively high, but the high magnitude level does not last very long. This stage is called initial discharge stage. Then PD development turns into another stage, magnitude becoming lower and PD phase extending more widely. PD is not stable, since during sometime in 10^{th} hour to 15^{th} hour, oscilloscope barely can detect effective PD pulse. The second stage, 7^{th} hour to 18.5^{th} hour for large cavity model and 8^{th} hour to 21.5^{th} hour for small cavity model, is called weak discharge stage. Afterward the magnitude increases sharply and PD phase continues expanding. PD becomes fierce, because pressboards are long lastingly bombarded by charged particles and corroded by discharge byproducts. “Chi Chi” discharge sound can be heard. The last stage is called pre-breakdown stage.

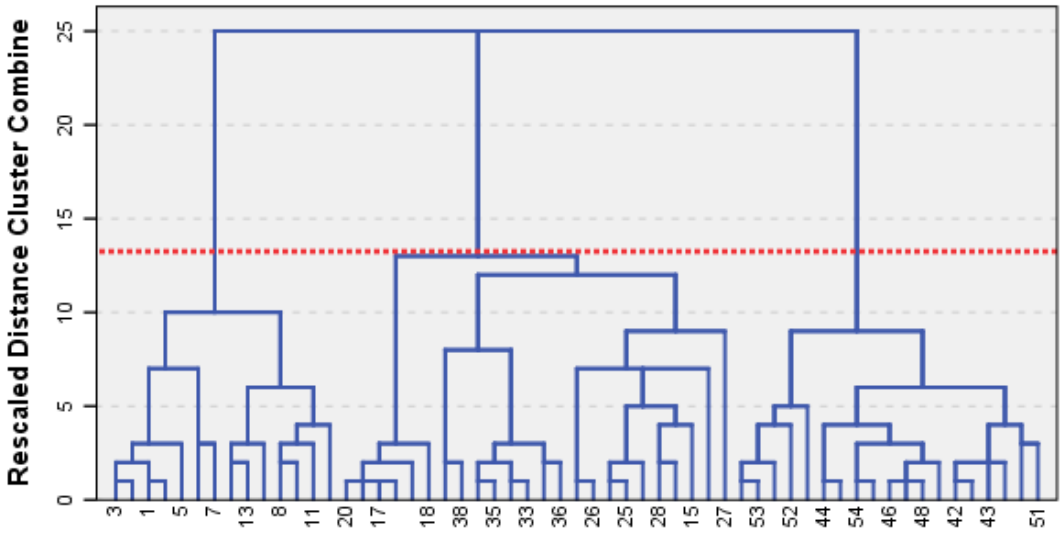


Figure 4. 1 Hierarchical clustering results of large cavity model

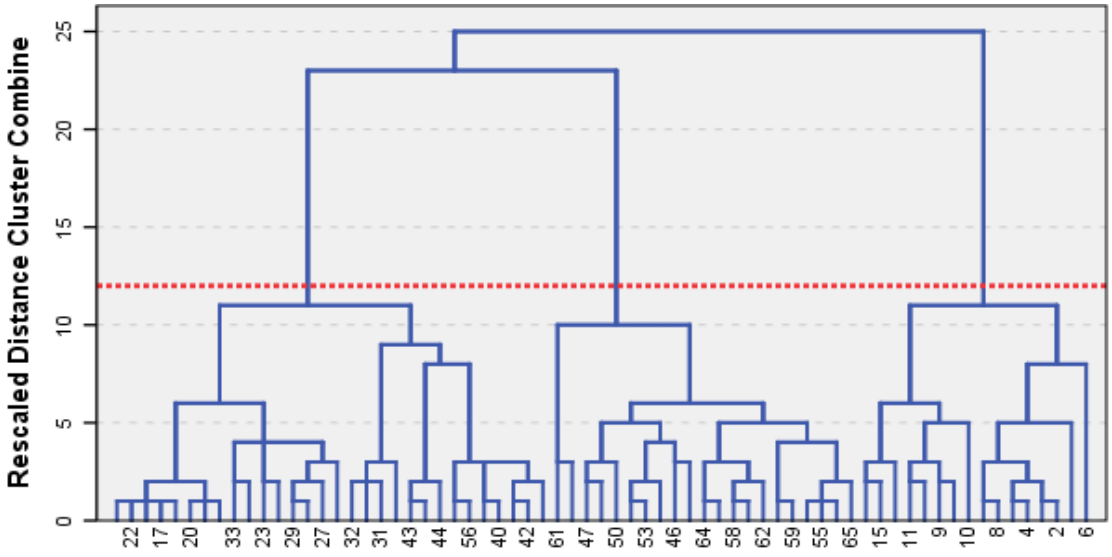


Figure 4. 2 Hierarchical clustering results of small cavity model

Same clustering results are obtained for PD development stage classification by using K means clustering. Three initial cluster centers are set by software, and after four times iteration for large cavity data and six times for small cavity data, cluster centers stay unchanged. Initial and final cluster centers for large cavity data as well as small cavity data are shown in Table 4. 1 and Table 4. 3. Class members and distances to

cluster centers for the two cavity configuration data are shown in Table 4. 2 and Table 4. 4 respectively. And 3-d visualization of clustering results for large and small cavity PD development stage is shown in Figure 4. 3

Table 4. 1 Initial and final cluster centers for large cavity data

Parameters	Initial cluster centers			Final cluster centers		
	1	2	3	1	2	3
X1	-1.6001	0.1899	0.1673	-1.2206	0.1431	-0.0266
X2	-0.1147	1.6325	-1.1493	0.0425	0.8647	-0.4799
X3	0.3170	-0.6432	-1.3449	-0.4968	-0.9915	-0.9623
X4	0.4080	-0.3450	-0.1469	-0.0528	-0.4626	-0.2195
X5	-0.2304	0.2310	0.2662	-0.0148	0.0544	-0.3948
X6	-0.1416	0.1831	-0.4248	0.0203	0.1118	-0.3333

Table 4. 2 Class membership and distance to cluster center for large cavity data

Samples	Clusters	Distance	Samples	Clusters	Distance	Samples	Clusters	Distance
1	1	0.4546	20	2	0.8046	39	3	0.7193
2	1	0.5638	21	2	0.9337	40	3	0.8081
3	1	0.4914	22	2	0.5971	41	3	0.6629
4	1	0.4511	23	2	0.2771	42	3	0.5719
5	1	0.5826	24	2	0.8944	43	3	0.6683
6	1	0.9831	25	2	0.4827	44	3	0.6803
7	1	0.5495	26	2	0.9874	45	3	0.2299
8	1	0.3381	27	2	1.1149	46	3	0.5283
9	1	0.5288	28	2	0.8622	47	3	0.6471
10	1	0.6402	29	2	0.6523	48	3	0.6841
11	1	0.7449	30	2	0.3846	49	3	0.5201

12	1	0.7778	31	2	0.5753	50	3	0.5933
13	1	0.6898	32	2	0.6239	51	3	0.6816
14	1	0.6014	33	2	0.5116	52	3	0.8342
15	2	0.6772	34	2	0.8618	52	3	0.5768
16	2	1.0815	35	2	0.7525	54	3	0.4191
17	2	0.8033	36	2	0.7217	55	3	0.6745
18	2	1.0274	37	2	0.8896	56	3	0.3686
19	2	0.8716	38	2	0.8991	57	3	0.7007

Table 4. 3 Initial and final cluster centers for small cavity data

Parameters	Initial cluster centers			Final cluster centers		
	1	2	3	1	2	3
X1	-0.9431	0.3651	0.7854	-0.6121	0.1431	0.7716
X2	0.2963	0.5312	-0.9124	0.5574	0.7647	-0.7062
X3	0.2490	-0.2963	-0.8143	-0.1336	-0.2915	-0.8092
X4	0.7879	-0. 6256	0.3055	0.8326	-0.4626	-0.0729
X5	-0.5907	0.3935	0.0589	-0.4139	0.2544	-0.1354
X6	-0.3168	0.3102	-0.7292	0.0203	0.2151	-0.6618

Table 4. 4 Class membership and distance to cluster center for small cavity data

Samples	Clusters	Distance	Samples	Clusters	Distance	Samples	Clusters	Distance
1	1	0.4701	24	2	0.4088	47	3	0.5589
2	1	0.5673	25	2	0.8916	48	3	0.5720
3	1	0.4913	26	2	0.4790	49	3	0.1786
4	1	0.4492	27	2	0.5555	50	3	0.4494
5	1	0.8329	28	2	0.7173	51	3	0.5534
6	1	0.9677	29	2	0.4757	52	3	0.6577
7	1	0.4830	30	2	0.9547	53	3	0.2652

8	1	0.5514	31	2	0.7361	54	3	0.4781
9	1	0.3671	32	2	0.6682	55	3	0.6447
10	1	0.5352	33	2	0.5307	56	2	0.9942
11	1	0.7001	34	2	0.7582	57	3	0.6916
12	1	0.6614	35	2	0.3325	58	3	0.4936
13	1	0.7452	36	2	0.4089	59	3	0.6684
14	1	0.7680	37	2	0.5043	60	3	0.7365
15	1	0.6702	38	3	0.6553	61	3	0.8197
16	1	0.8134	39	2	0.5952	62	3	0.3151
17	2	0.7532	40	2	0.7610	63	3	0.4631
18	2	0.6810	41	2	0.7128	64	3	0.4290
19	2	0.5208	42	2	0.6934	65	3	0.6831
20	2	0.6158	43	3	0.8871	66	3	0.5169
21	2	0.6247	44	2	0.9230	67	3	0.7323
22	2	0.6100	45	3	0.9924			
23	2	0.5886	46	3	0.6560			

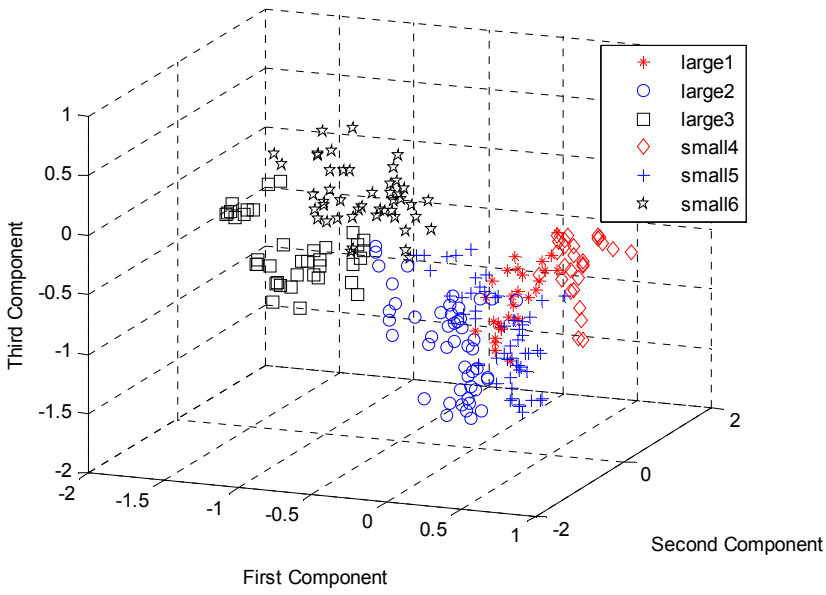


Figure 4. 3 3-D visualization of clustering results

4.2 Air-gap PD development stage identification considering cavity size

Air-gap PD development stage identification is a typical pattern recognition problem. In this section, conventional Radial Basis Function (RBF) neural network and Random Forests which has not been introduced to deal with PD pattern recognition problem are used to identify different cavity size's PD development stage.

4.2.1 RBF Neural network

Artificial neural network (ANN) is computational models inspired by animals' central nervous systems that are capable of machine learning and pattern recognition. Since ANN can achieve complex nonlinear mapping by interconnected group of artificial neurons, it is widely used in classification problem. An artificial neuron, which is a mathematical function characterizing biological neuron, is the constitutive unit in an artificial neural network. The artificial neuron receives one or more inputs (representing one or more dendrites) and sums them to produce an output (representing a biological neuron's axon). Usually the sums of each node are weighted, and the sum is passed through a non-linear function known as an activation function. The artificial neuron model is shown as follows.

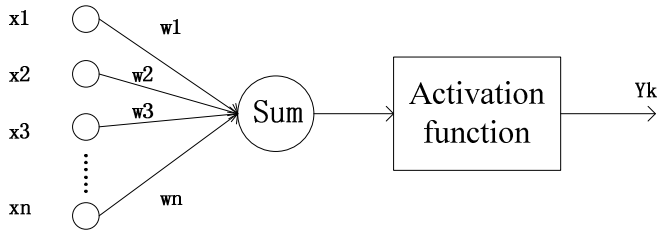


Figure 4. 4 Artificial neuron model

$$y_k = \varphi(\sum_{j=0}^{m} \omega_{kj} x_j) \tag{4.8}$$

Multilayer feedforward network that consists of multiple layers of computational units is widely used. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a sigmoid function as an activation function. Back propagation algorithm is a common method of training artificial neural networks. Since during training, weights in every layer need to be recalculated each time, multilayer feedforward network is hard to train and easy to trap in local minimum [26]. Hence, a lot of improvement has been done on the multilayer network and RBF network is one of the most sophisticated improved networks.

RBF neural network, shown in Figure 4. 5, is a three layers' network that uses radial basis functions as activation functions in the hidden layer and linear activation functions in the output layer. Gaussian function is commonly taken as the radial basis activation function, shown in 4.9. Compared to sigmoid function, shown in 4.10, the output of Gaussian function only depend on the distance from a center vector, when

the input is far from the center, very small output would be generated. This makes RBF network faster to train than multilayer feedforward network using sigmoid activation function.

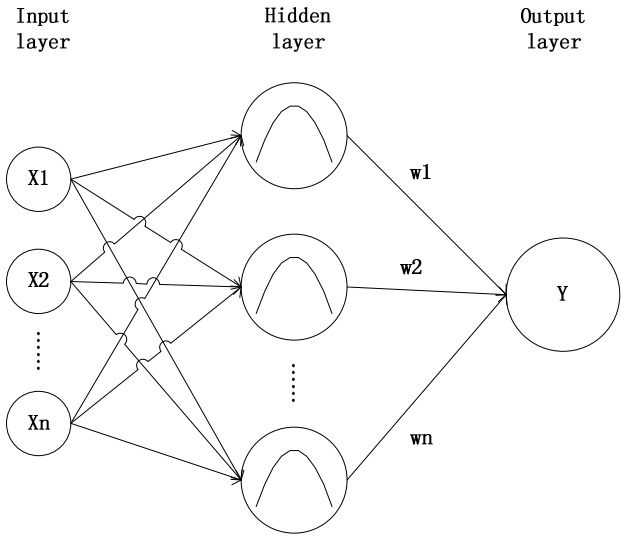


Figure 4. 5 structure of RBF neural network

$$\varphi(x) = \exp[-(x - c_i)^2 / \sigma_i^2] \tag{4.9}$$

$$\varphi(x) = 1 / 1 + e^{-x} \tag{4.10}$$

During the training, the center vector c_i, width of the radial function σ_i in the hidden layer and the weight ω_i connecting hidden layer and output layer need to be decided. RBF networks are typically trained by a two-step algorithm. In the first step, the center vectors can be estimated by k-means clustering method that is detailed in Section 4.1.2. Then width of the radial function is decided by the distance between center vectors. The second step determines the weights typically using Levenberg–Marquardt algorithm that combines the advantage of Gauss–Newton algorithm and the method of gradient descent[27].

4.2.2 Random forests

It is natural and intuitive to classify a pattern through a sequence of questions, in which the next question asked depends on the answer to the current question. Decision tree is another kind of classifier which comes from this theory. It is a flowchart-like structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label. A path from root to leaf represents classification rules. Decision tree is simple to understand and interpret and can be trained very fast.

To maintain advantages of decision tree while increasing accuracy, random forests was proposed. Random forests is an ensemble learning method for classification (and regression) that operates by constructing a multitude of decision trees at training time and outputting the most popular class that each tree casts a unit vote for. The main principle behind this ensemble method is that a group of “weak learners” can come together to form a “strong learner”. Leo mentioned the advantages of random forests in [14]. ① It is unexcelled in accuracy among current algorithms. And the Accuracy is as good as Adaboost and sometimes better ② It runs efficiently on large data bases. It can be trained very fast and easily parallelized. ③ It can handle thousands of input variables without variable deletion. ④ It is resistant to over training. ⑤ It is relatively robust to the noise.

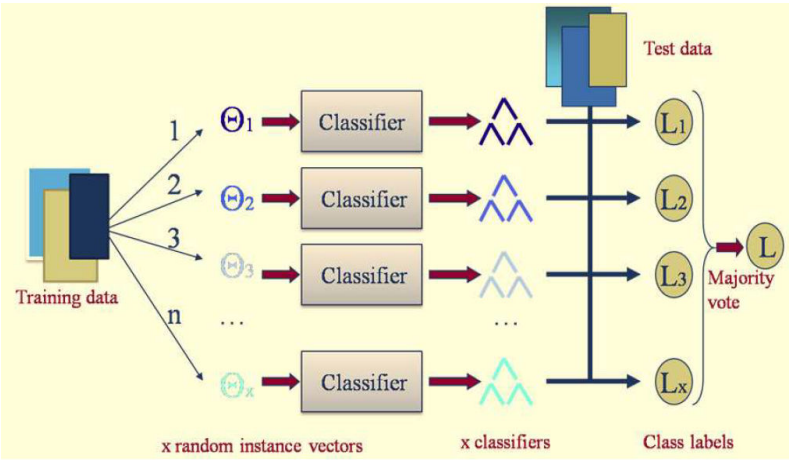


Figure 4. 6 Random forests

In order to get strong learner, random forests is implemented as follows.

① The training set for each tree is selected by Bootstrap aggregating method, also called bagging. The training subset is sampled from whole training dataset uniformly and with replacement. Typically, it accounts for 63.2% of the whole dataset.

② Each tree is grown to as large size as possible without pruning based on CART algorithm. The optimal feature chosen to split a node is the one that can generate a child node with lowest Gini impurity, $I(\mathrm{n})$.

$$I(n) = 1 - \sum_{j=0}^{n} p(j/t)^2 \tag{4.11}$$

where $p(j/t)$ is the probability of a specific class within the child node. Gini impurity should be zero if the node is all one class. In normal decision tree, the optimal feature to split a node is selected from all features. However, for random forests, the optimal feature is selected from a random subset of features. Typically, for a dataset with p features, $\sqrt{p}$ features are used in each split.

③ The ultimate output of random forests is obtained by the majority vote of the each decision tree.

4.2.3 Identification results

In order to identify different cavity size's PD development stage, the WEKA implementation of RBF neural network and random forests are used in this work. The training inputs of the classifiers are the six dimensional statistical parameters that are derived from KPCA and the training outputs are PD development stage clustering results, initial discharge stage, weak discharge stage and pre-breakdown discharge stage, of large and small cavity which are labeled as L1, L2, L3, S1, S2 and S3. For either discharge model, two-times discharge development process data, which is shown in Table 4. 5, is used to train the two classifiers. And 10-fold Cross-validation method is employed to evaluate the classifier identification performance.

Table 4. 5 Training dataset

Development stage and cavity size	L1	L2	L3	S1	S2	S3
Sample quantity	29	50	39	32	54	44
Total quantity	248					

RBF neural network has six input neurons, one output neuron and adjustable hidden neurons. After parameter modulation, the optimal hidden neuron number is six or nine that gives total identification accuracy of 87.90%. For random forests, decision tree number and feature number used in each split are adjustable. The identification results are detailed in Table 4. 7. Random forests classifier shows a better identification performance. When decision tree number is 250 and feature number is 3, random

forest got the best result that total accuracy is 93.15%. And the confusion table is shown in Table 4. 7.

Table 4. 6 Identification results

RBF neural network		Total accuracy	Random forests		Total accuracy
Hidden neurons			features	trees	
3		84.68%	2	50	91.53%
4		85.08%	2	150	92.34%
5		80.65%	2	250	92.74%
6		87.90%	3	50	91.94%
7		85.89%	3	150	92.74%
8		86.69%	3	250	93.15%
9		87.90%	4	50	91.13%
10		85.08%	4	150	91.94%
11		83.47%	4	250	91.34%

Table 4. 7 Confusion table for 3 features and 250 trees

Samples	Confusion table						Accuracy	Total accuracy
	L1	L2	L3	S1	S2	S3		
L1	27	1	1	0	0	0	93.10%	93.15%
L2	1	47	1	0	1	0	94%	
L3	0	3	35	1	0	0	89.74%	
S1	0	0	2	30	0	0	93.75%	
S2	0	0	0	1	53	0	98.11%	
S3	0	0	0	2	1	41	93.18%	

4.3 Summary

This chapter presents the clustering results of large and small cavity PD development stage. In addition, RBF neural network and random forests are used to identify the air-gap PD development stage.

1) Same PD development clustering results are obtained by Hierarchical clustering and K means clustering method. The PD development stage for large and small cavity model are both divided into three stages, initial discharge stage, weak discharge stage and pre-breakdown stage.
2) For development stage identification of unknown air-gap PD samples, random forests classifier shows a better performance than RBF neural network.

Chapter 5 Conclusions

In order to evaluate air-gap PD severity more comprehensively and accurately, the effect of cavity size on PD activity and PD development characteristics are studied in this work through experiments. Then statistical parameters are extracted from PRPD pattern, being carried out dimensions reduction with the help of KPCA. At last, software PASW Statistics and WEKA are used to implement the PD development stage division and identification. The main conclusions are drawn as follows.

1) Large cavity PD possesses lower inception field. Higher cavity height contributes more to the lower inception field than larger cavity diameter. Large cavity PD possesses higher charge magnitude. Charge magnitude mainly depends on the cavity height. Large cavity PD possesses higher inception phase. Both cavity height and diameter have impact on the inception phase. PD happening large cavity is more harmful than that happening in small cavity.
2) During Air-gap PD development process, charge magnitude variation of large and small cavity model both presents concave curve shape with respect to time. Besides, discharge phase continuously expends and the phase span of large cavity model is much narrower than that of small cavity model. When air-gap PD comes to the last stage, for small cavity model, the positive PD even can expand to the negative half cycle and vice versa. Cavity size should be distinguished in Air-gap PD severity diagnosis.

3) 27 dimensional statistical parameters from PRPD pattern are carried out dimension reduction through PCA and KPCA. KPCA keeps more feature information within six dimensional parameters than PCA.
4) Same PD development clustering results are obtained by Hierarchical clustering and K means clustering method. The PD development stage for large and small cavity model are both divided into three stages, initial discharge stage, weak discharge stage and pre-breakdown stage.
5) For development stage identification of different-size air-gap PD samples, total identification accuracy of random forests classifier is 93.15%, showing a better performance than RBF neural network does.

References

[1] B. Richardson, "Transformer life management bushings and tapchangers," in *Transformer Life Management (Ref. No. 1998/510), IEE Colloquium on*, 1998, pp. 8/1-8/4.

[2] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *Electrical Insulation Magazine, IEEE,* vol. 18, pp. 12-25, 2002.

[3] M. Arshad, S. M. Islam, and A. Khaliq, "Power transformer aging and life extension," in *Probabilistic Methods Applied to Power Systems, 2004 International Conference on*, 2004, pp. 498-501.

[4] T. Zhang and J.-d. Cai, "Study on Life Prediction Method of Oil Immersed Power Transformer," in *Electrical Insulation and Dielectric Phenomena, 2008. CEIDP 2008. Annual Report Conference on*, 2008, pp. 486-489.

[5] T. Tanaka, "Internal Partial Discharge and Material Degradation," *Electrical Insulation, IEEE Transactions on,* vol. EI-21, pp. 899-905, 1986.

[6] G. C. Montanari and A. Cavallini, "Partial discharge diagnostics: from apparatus monitoring to smart grid assessment," *Electrical Insulation Magazine, IEEE,* vol. 29, pp. 8-17, 2013.

[7] R. Sarathi, I. P. Merin Sheema, and J. Sundara Rajan, "Understanding surface discharge activity in copper sulphide diffused oil impregnated pressboard under AC voltages," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 21, pp. 674-682, 2014.

[8] M. Tozzi, A. Cavallini, G. C. Montanari, V. Iuliani, and C. Serafino, "Global monitoring approach in a large autotransformer through PD, DDF and DGA analysis: PD source location and maintenance action planning," in *Electrical Insulation Conference (EIC), 2011*, 2011, pp. 206-210.

[9] M. Di Lorenzo del Casale, R. Schifani, and J. T. Holboll, "Partial discharge tests using CIGRE method II," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 7, pp. 133-140, 2000.

[10] L. Niemeyer, "A generalized approach to partial discharge modeling," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 2, pp. 510-528, 1995.

[11] F. Gutfleisch and L. Niemeyer, "Measurement and simulation of PD in epoxy voids," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 2, pp. 729-743, 1995.

[12] R. Schifani, R. Candela, and P. Romano, "On PD mechanisms at high temperature in voids included in an epoxy resin," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 8, pp. 589-597, 2001.

[13] H. A. Illias, G. Chen, and P. L. Lewin, "Partial discharge within a spherical cavity in a dielectric material as a function of cavity size and material temperature," *Science, Measurement & Technology, IET,* vol. 6, pp. 52-62, 2012.

[14] L. Breiman, "Random forests," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[15] F. Puletti, M. Olivieri, A. Cavallini, and G. C. Montanari, "Risk Management of HV Polymeric Cables Based on Partial Discharge Assessment," in *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, 2006, pp. 626-633.

[16] M. Runde, R. Hegerberg, N. Magnusson, E. Ildstad, and T. Ytrehus, "Cavity formation in mass-impregnated HVDC subsea cables-mechanisms and critical parameters," *Electrical Insulation Magazine, IEEE,* vol. 30, pp. 22-33, 2014.

[17] "Test Method for Dielectric Breakdown Voltage and Dielectric Strength of Solid Electrical Insulating Materials at Commercial Power Frequencies," *ASTM Standard D149-09,* 2009.

[18] W. Kai, T. Ijichi, T. Kato, Y. Suzuoki, F. Komori, and T. Okamoto, "Contribution of surface conductivity to the current forms of partial discharges in voids," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 12, pp. 1116-1124, 2005.

[19] "IEEE Guide for the Statistical Analysis of Electrical Insulation Voltage Endurance Data," *ANSI/IEEE Std 930-1987,* p. 0_1, 1987.

[20] P. C. Baker, B. Stephen, and M. D. Judd, "Compositional Modeling of Partial Discharge Pulse Spectral Characteristics," *Instrumentation and Measurement, IEEE Transactions on,* vol. 62, pp. 1909-1916, 2013.

[21] C. S. Kim, T. Kondo, and T. Mizutani, "Change in PD pattern with aging," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 11, pp. 13-18,

2004.

[22] R. Liao, J. Yan, L. Yang, M. Zhu, and B. Liu, "Study on the relationship between damage of oil-impregnated insulation paper and evolution of phase-resolved partial discharge patterns," *European Transactions on Electrical Power,* vol. 21, pp. 2112-2124, 2011.

[23] B. Fruth and L. Niemeyer, "The importance of statistical characteristics of partial discharge data," *Electrical Insulation, IEEE Transactions on,* vol. 27, pp. 60-69, 1992.

[24] B. Yazici, "Statistical pattern analysis of partial discharge measurements for quality assessment of insulation systems in high-voltage electrical machinery," *Industry Applications, IEEE Transactions on,* vol. 40, pp. 1579-1594, 2004.

[25] E. Gulski and F. H. Kreuger, "Computer-aided recognition of discharge sources," *Electrical Insulation, IEEE Transactions on,* vol. 27, pp. 82-92, 1992.

[26] F. Lampariello and M. Sciandrone, "Efficient training of RBF neural networks for pattern recognition," *Neural Networks, IEEE Transactions on,* vol. 12, pp. 1235-1242, 2001.

[27] Y. Fan and M. Paindavoine, "Implementation of an RBF neural network on embedded systems: real-time face tracking and identity verification," *Neural Networks, IEEE Transactions on,* vol. 14, pp. 1162-1175, 2003.