Electric Power Systems Research 227 (2024) 109988

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Diagnostics analysis of partial discharge events of the power cables at various voltage levels using ramping behavior analysis method

Sambeet Mishra [a,*], Praveen Prakash Singh [b], Ivar Kiitam [b], Muhammad Shafiq [b,d], Ivo Palu [b], Chiara Bordin [c]

[a] University of South-Eastern Norway, Porsgrunn, Norway
[b] Tallinn University of Technology, Tallinn, Estonia
[c] UiT Arctic University of Norway, Tromsø, Norway
[d] Florida State University, FL, USA

ARTICLE INFO

Keywords:
Internal partial discharge
High voltage engineering
Power cable
Uncertainty quantification
Trends and patterns
Diagnostics methods

ABSTRACT

Partial discharge events can occur in high-voltage cables. It can be caused by defects in the cable insulation, contamination, or a combination of both. Partial discharge in cables can lead to insulation failure and cable failure. This investigation aims to identify the trends and patterns in the internal partial discharge (PD) occurrences in the power cables when exposed to different voltage levels - 6.4, 7.4, 9.4, and 11.3 kV. For pattern extraction, a well-established method, ramping behavior analysis, is implemented to extract and classify PD occurrences into sets of significant and stationary events. In this investigation, significant events correspond to an absolute peak (the discharge pulse) and subsequent oscillations from the measurement sensor. The stationary events represent a collection of noise that is recorded during the measurement. These noise signals are essentially small variations within a pre-determined threshold range. Furthermore, a comparative analysis is performed for each voltage level and for the voltage levels.

This investigation brings new knowledge on how internal partial discharge pulses occur at various voltage stress levels. Specifically, the emerging patterns and trends of internal partial discharge events. The results indicate that there is a positive correlation between the number of PD events and the increase in stress levels. Furthermore, negative PD peaks are more frequent at lower stress levels.

1. Background and objectives

Partial discharge (PD) is a phenomenon of substantial importance in high-voltage electrical equipment due to its relationship to insulation faults. Partial discharge is formally defined according to IEC 60270 as a localized electrical discharge that only partially bridges the insulation between conductors and which may or may not occur adjacent to a conductor. This implies the phenomenon is distinctly different from a full disruptive discharge, during which an electrical discharge occurs between phases or between phase and ground, i.e. a fault which precipitates short-circuit in the electrical system. One aspect of the importance of PD is that it can often precede this type of electrical fault in various types of electrical insulation, as it can be considered both a cause and a result of insulation aging.

PD occurs in areas where the local electric field strength is enhanced, and the dielectric strength of the insulation is decreased. This can occur, for example, inside gas-filled voids and cavities embedded into solid or liquid insulation, where the inequality of dielectric permittivities result in increased electric field strength inside the gaseous medium. Other locations prone to PD activity include interfaces between different dielectric media, areas surrounding sharp electrode edges in a gaseous environment or other similar circumstances. Three distinct types of PD are usually described based on where the discharge occurs relative to the electrodes and insulation. These PD types are internal discharge, surface discharge, and corona discharge.

When the discharge occurs adjacent to a solid or liquid dielectric material, it usually has a degrading effect on the insulation. Charged particles bombard the surface of the material, which undergoes chemical and physical changes, i.e. scission of intramolecular chemical bonds and chemical reactions, formation of cracks and chemical by-products, erosion, and carbonization. Materials can vary widely in their resilience to PD activity, but it is generally recognized that PD has a negative impact on high-voltage insulation and should be minimized whenever possible. This excludes corona discharge occurring on extra high voltage overhead lines, as air is a self-renewing insulation medium under normal circumstances, although steps are also taken to reduce corona-related power loss.

* Corresponding author.
E-mail addresses: sambeet.mishra@usn.no, to.sambeet@gmail.com (S. Mishra).

https://doi.org/10.1016/j.epsr.2023.109988
Received 6 July 2023; Received in revised form 7 October 2023; Accepted 5 November 2023
Available online 16 November 2023

Nomenclature

Δt	Duration of a ramp event
Δt_m	Duration of significant ramp event
Δw	Change Δ in magnitude (w) of voltage or change in the magnitude of partial discharge signal voltage
Δw_m	Magnitude of significant ramp event
Δw_s	Value difference between the predetermined threshold and the present value
Λ_m	Set of frequencies per featured event
σ_m	Mean of initial and final magnitudes of ramp event
τ	RBA_θ threshold
θ	Phase angle of PD occurrence
$\theta^{\Delta W_s}$	Angle between the horizontal and a line connecting the point at which the threshold is exceeded and the peak point
θ_m	Angle to peak of significant ramp event at the origin of the event
θ_p	Phase angle
$f(x)$	Curve fitting function
p_1	Polynomial coefficient
p_2	Polynomial coefficient
p_3	Polynomial y-axis intercept
t	Time
T_1	Start time of exceeding threshold
T_2	End time of exceeding threshold
w_1	Signal value upon start of exceeding threshold
w_2	Signal value upon end of exceeding threshold
$w_s(t)$	Peak value of a ramp event
Δt	Time difference
HFCT	High-frequency current transformer
PD	Partial discharge
RBA_θ	Ramping Behavior Analysis
RMSE	Root mean square error
XLPE	Cross-linked polyethylene

1.1. Background

Characteristics of PD activity depend on the type of PD source. Substantial research has been conducted to determine the pulse characteristics of different types of PD sources, i.e., corona, internal, and surface discharges, and also to identify particular types of equipment defects based on their PD signature [1–4]. Power cables are of particular interest concerning PD and it has been determined that the geometrical aspects of the void inside the insulation which acts as a PD source will have a profound impact on the PD activity [5]. Measuring PD is one of the few available approaches to collect information regarding the cable condition, which is an asset that is mostly difficult to examine due to its structure and reduced physical access. Therefore researchers have been focused on developing PD measurement systems for power cables and resolving the problems associated with their implementation [6–9]. PD measurement, applied both offline and online, is seen as favorable from the perspective of the electrical utility, by decreasing cable network maintenance costs over the long term [10]. In cable diagnostics, PD measurement is often combined with dissipation factor (tan δ) measurement, as both metrics respond to changes in cable condition [11].

PD has been identified as a factor that contributes to failure rates in both underground and submarine cables [12]. Cable faults are a problem in both the medium and high voltage grid, where a typical failure rate is around 1 fault per 100 km of cable per year, but this can vary substantially across regions [13]. Cable faults can also be difficult to locate and repair, which makes cable reliability a high priority for electrical utilities.

One of the difficulties in measuring PD in power cables is that PD pulses attenuate upon propagation from the discharge site to the point of measurement, which is usually situated at the terminations or joints of the cable system. Therefore the pulses can be too weak to be detected and require a different bandwidth of the measurement detector for optimal measurement sensitivity [14,15]. Also, locating the PD source along the cable can be challenging. Approaches such as time-domain reflectometry as well as methods based on the frequency content of measured pulses have been developed to determine the position of PD sources in cables [16–18]. The research focused on improving the accuracy of location algorithms is still ongoing [19].

When PD is detected, it is necessary to characterize it and determine the nature of the PD sources. There can sometimes be only one PD source, but sometimes several, which can be a combination of the various types of PD sources described previously. Recognition and classification of PD patterns include various different approaches, e.g., examining statistical parameters (mean, standard deviation, kurtosis, skewness, polarity and phase asymmetry, etc.) [20–24], principal component analysis [20,21,24,25], discriminant analysis [21,24], t-distributed stochastic neighbor embedding [20], density-based spatial clustering for applications with noise (DBSCAN) [20,25,26], inductive inference algorithms [22,27], rough set theory [22,28], fuzzy logic [22,23], self-organizing maps [22], support vector machines [20,22,24], neural networks and various other machine learning algorithms [20–24], including deep learning [24,29]. Using PD energy as an alternative to apparent charge values in the analysis of PD has also been proposed [30].

1.2. Motivation and reasoning

This paper presents the novel application of the Ramping behavior algorithm (RBA_θ) to PD analysis. The goal is to evaluate the effectiveness of this approach in characterizing the important aspects of PD activity and determine if this is feasible for application under various condition assessment scenarios. The principle of RBA_θ is determining and characterizing the peaks corresponding to significant changes in a process characterized by a time series. It has previously been applied in the analysis of wind farm output power data [31], which is an intrinsically stochastic weather-dependent process. The nature of PD activity also involves a significant degree of stochasticity, as it is sometimes observed that PD does not occur when the voltage exceeds the PD inception voltage, and when PD does occur, the phase angles at which PD pulses are identified and recorded are also variable. The variability is itself dependent on the type of PD source, e.g., corona discharge activity is more consistent compared to internal discharge activity across several AC voltage cycles.

When determining the nature of PD sources from measured data, there is a limited number of aspects based on which the PD activity can be characterized and assessed. When measuring PD using a suitable sensor, e.g., a high-frequency current transformer (HFCT), the pulsed PD activity appears as a series of distinct peaks, provided that the level of background noise is sufficiently low so as to not obscure the PD. As mentioned previously, these pulses have distinct characteristics like peak value, rise-time, fall-time, frequency, time interval, etc., due to which these very suitable targets for investigation using RBA_θ. Another advantage of RBA_θ is that the algorithm contains an in-built threshold, which filters out changes with a small amplitude, which represents measurement noise in the context of PD. Due to the high degree of the intrinsic variability of internal discharges and the high relevance to power equipment reliability of this phenomenon, we have analyzed the usefulness of RBA_θ in characterizing this type of PD.

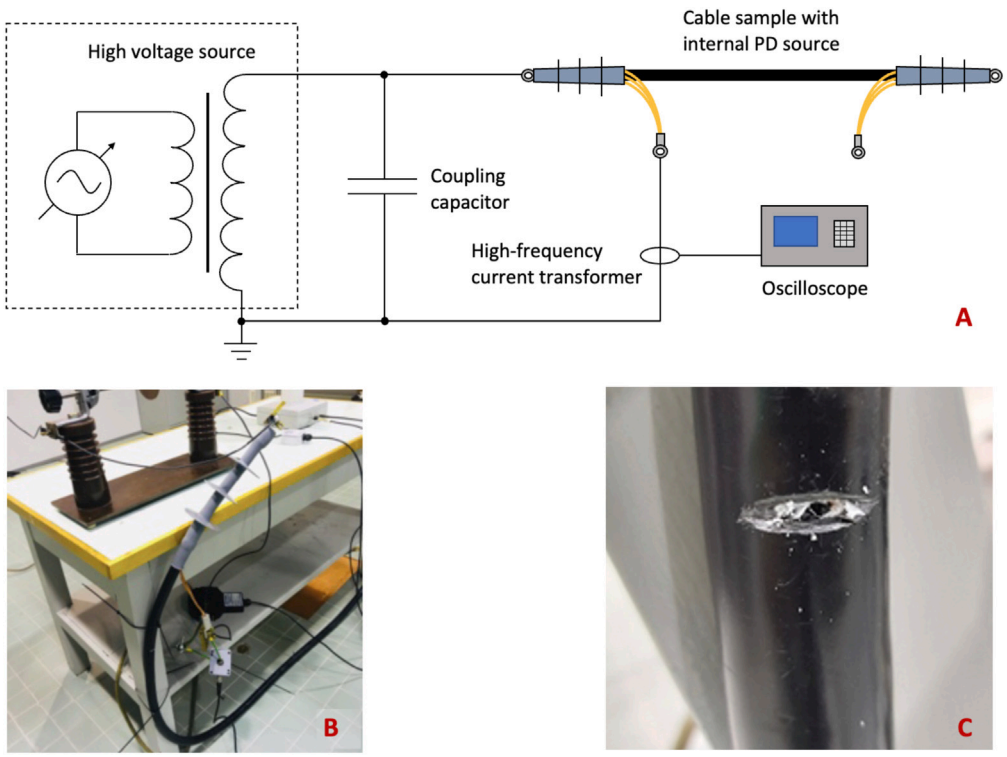

Fig. 1. Test setup for measuring internal partial discharge activity in medium voltage power cable sample. (A) Test circuit schematic; (B) Photo of test setup; (C) Internal PD source on cables.

1.3. Key contributions

The paper is organized into four sections starting with an introduction covering the background of the investigation (Section 1), reasoning, and key contributions. Section 2 describes the experimental method and its application. Following that, Section 3 elaborates on the results and discusses the findings of this work. Finally Section 4 concludes by summarizing the results and the main findings with remarks for future research.

The primary contributions of this work are related to the comprehensive characterization of internal discharge activity by means of using the RBA_θ, which is a novel way to approach PD analysis. This adds to the toolkit available to accomplish PD source recognition and discrimination, either as a standalone method or in combination with other approaches described in the literature. This can aid in creating PD analysis algorithms, which can assist human experts in evaluating PD measurement results. PD measurements can be difficult to interpret, depending on various measurement conditions. Some of the strengths of the RBA_θ algorithm, which make it a favorable candidate for PD analysis, include that it characterizes the rise and fall times of pulses in addition to the peak values and also includes a thresholding function to dismiss measurement noise, which might be misinterpreted as PD. The basic components of RBA_θ involve segmentation, identification, characterization, quantification, and comparison of analyzed data. As a result, using the RBA_θ on test data gathered from an internal discharge source at various test voltages provides an expanded understanding of its behavioral characteristics.

2. Method: Analysis of internal discharge properties

This section describes the measurement of PD in an artificially created source intended to produce internal discharge. The measurements were performed in a laboratory environment and the results were processed using the RBA_θ algorithm.

2.1. Experimental setup

The PD data were gathered on a test sample with an internal discharge source using a high-frequency current transformer (HFCT) to measure the PD activity (Fig. 1). The test object was a sample of a 20 kV medium voltage cable, which has a cut on the outer jacket of the cable, which reached into the main insulation of the cable. The HFCT is connected to the screen of the cable, and the setup also includes a coupling capacitor, which is added to stabilize voltage and supply the charge during PD events. The HFCT has a bandwidth of 0.5–80 MHz (−3 dB) and a transfer ratio of 1:10. The output of the HFCT is measured and data are recorded using a digital storage oscilloscope. Various voltage levels are applied to the test object and the PD activity is recorded and analyzed. An example of a PD measurement obtained using this setup is depicted in Fig. 2, where the PD pulses appear as sharp peaks and are plotted in relation to the AC voltage used to energize the test sample to obtain information regarding the phase angle at which the PD occurs.

The measurements were performed at various voltages at and beyond the level at which PD begins to occur (i.e., PD inception voltage). The data were gathered at 6.4, 7.4, 9.4, and 11.3 kV and the PD activity was recorded for a total of 32 cycles of AC voltage at each voltage level. The sampling rate used in the experiments was 156 MS/s. Examples describing the typical PD measurement results at different voltage levels are depicted in Fig. 3. In the figure, the test voltage waveform is provided purely for visualization of phase reference. Upon increasing the test voltage, the PD activity undergoes some noticeable changes:

- The magnitude of the largest PD pulses tends to increase
- The number of PD pulses over one AC voltage cycle tends to increase
- The range of phase angles at which PD activity occurs tends to expand

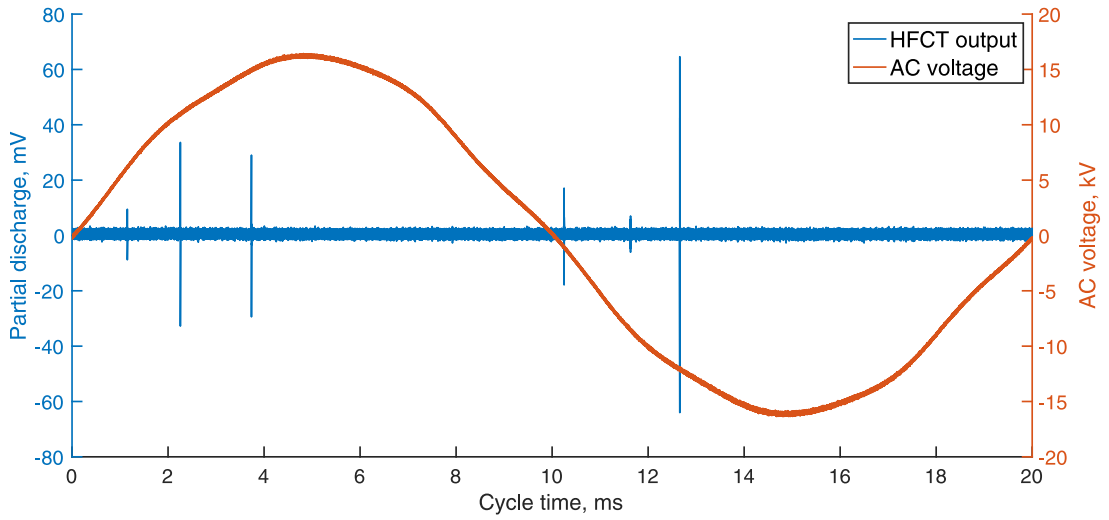

Fig. 2. Measurement data set recorded at 11.3 kV. Partial discharge activity is plotted in relation to the 50 Hz energizing voltage.

The figure also demonstrates the stochastic nature of PD activity. The phase angles at which PD pulses are observed tend to change as the voltage increases. If individual AC voltage cycles recorded at any voltage level beyond the PD inception voltage are compared, then the magnitude, occurrence time, and the number of pulses exhibit noticeable variation. As the PD measurement was accomplished using an HFCT, which does not measure the apparent charge of PD pulses, no charge magnitude calibration was performed. The PD pulses are quantified based on the peak response of the HFCT output in mV.

The measurements obtained using the experimental setup capture the internal partial discharge in the cable using a data acquisition system. However, the recorded PD pulses are accompanied by oscillations along with superimposed noise. Therefore, to separate these unwanted signals and extract only PD peaks occurring at several instances without losing significant information, a ramping behavior algorithm (RBA_θ) is applied. The RBA_θ is only applied to the time-series PD data set. The RBA_θ algorithm extracts two sets of events, which are termed stationary and significant. The diversification of events into segments facilitates circumventing the noise from the PD measurements, which can be further investigated with ease. This algorithm is applied for different sets of measurements taken at each stress level. The following subsection will provide a high-level overview and the steps of the RBA_θ method. Subsequently, how the RBA_θ is applied to the internal PD measurements is explained.

2.2. Working principle of ramping behavior analysis (RBA_θ) method

The RBA_θ methodology, which has been proposed in [32,33], and applied in [31], was originally utilized in the process of wind ramp event extraction. A swing across the reference axis is one way to describe what is known as a ramp event. RBA_θ categorizes the ramp events as either significant or stationary depending on their duration. Significant events are those that exhibit abrupt and significant fluctuations, whereas stationary events have swings of a relatively smaller magnitude and are consistent across time. The original application of the RBA_θ methodology focuses on the significant events that can be found in time series data pertaining to wind power production. Fig. 4 is a graphic representation of the characteristics that define an event that is deemed to be of major importance, therefore significant. Five distinguishing characteristics are used to summarize each significant event. The peak value of a ramp event, which is the peak value of the capacity factor at the time of vertex, is denoted by the $w_s(t)$ variable. The value of t indicates the time at which the ramp event took place. The size of a ramp event is measured by the Δw_s value, which is the value change that occurs between the predetermined threshold and the present value of the wind capacity factor. The duration of time that a ramp event has been present is denoted using the Δt variable. Finally $\theta^{\Delta W_s}$ represents the angle that must be covered in order to get from the initial departure point that intersects the threshold to the peak point $w_s(t)$.

The RBA_θ methodology introduced above was originally developed for wind-related applications, with a particular focus on extracting significant events from wind power production curves. Given the inherent uncertainty of wind power production, as well as the promising results obtained by using RBA_θ for wind power feature extractions, it is reasonable to assume that the same methodology could have a broader range of applications where stochasticity is prominent and the extraction of significant events is of interest. The work proposed in [34] provides insights on the value of extracting events from time-series data to perform predictions that are less computationally intensive both from the computational time aspects and from the energy consumption aspects. Such work is already a demonstration that the RBA_θ key idea of events extraction has great potential for a wide range of applications that go beyond the ramping behavior analyses limited to wind power production. In particular, the insights gathered from RBA_θ can be utilized by decision support system tools based on mathematical optimization for power and energy networks. The subject of energy and power systems modeling [35] as a subfield of the broad domain of energy informatics [36], would greatly benefit from data manipulation methodologies aimed at identifying significant events in the input time series data. Significant events could be utilized in place of the more traditional typical periods that are usually adopted in literature to reduce the computational time of large-scale models. An example is provided in [37] aimed at identifying needs for network reinforcement, restructuring, and reconfiguration. Here events' identification could improve the dataset utilized as model input, and have a positive impact on the overall computational time. Furthermore, decision support system models dealing with renewable integration, novel technologies, and non-linearities such as [38] can greatly benefit from the RBA_θ key concept of events extraction.

From this perspective, the identification of trends and patterns in the internal partial discharge occurrences in the power cables represents another promising area of application, and is the main focus of this article, as further discussed in the following paragraphs.

2.3. Application of ramping behavior analysis method

The data set obtained from the measurements consists of a large number of samples, time (in ms), and PD activity measured using

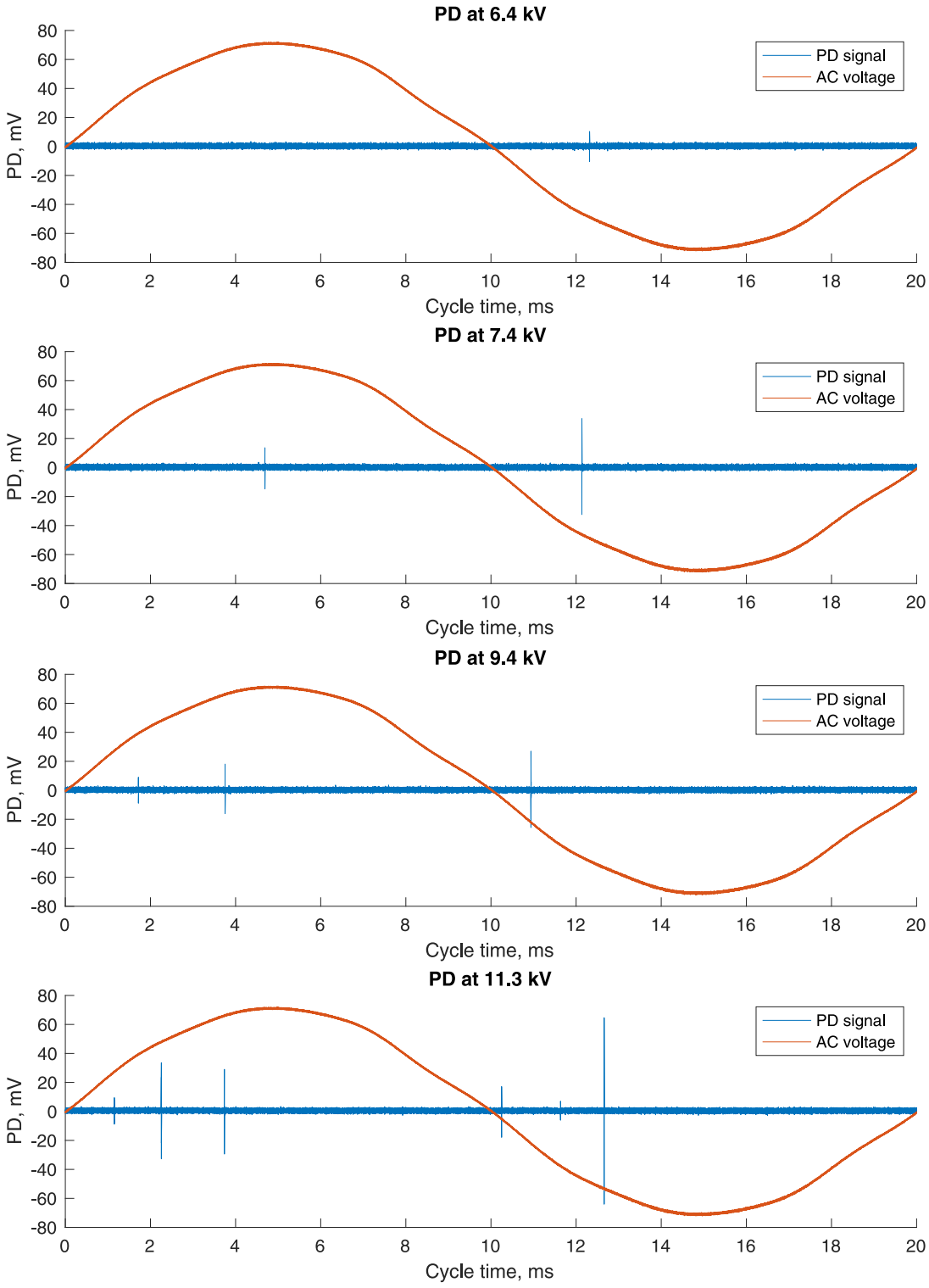

Fig. 3. Examples of internal partial discharge activity at various voltage levels.

an HFCT superimposed with noise (in mV). The used sampling rate generated a total number of data points of around 3.12 million per AC cycle and the size of the file was 107 MB for each data set at different voltage stress levels. The output data from the obtained measurements was mapped as RBA_θ input given that it is based on the number of samples and the PD magnitude.

The input data obtained from the measurement system is normalized between 0 to 1 by the RBA_θ for computation (in this case 1 is 30 mV). The user-defined threshold (τ) based on the nominal value segregates the time series data set into significant and stationary events. The significant accounts for the PD peak formation events and the stationary events account for the noise interference. This process of filtering is based on the selection of optimal threshold (τ) and is achieved iteratively by analyzing the peak noise amplitude manually. Furthermore, the determination of features based on the amplitude of the data signal in the significant event can further be classified as a significant up or significant down event depending on the direction of variation. Due to the characteristics of the recorded measurements, no

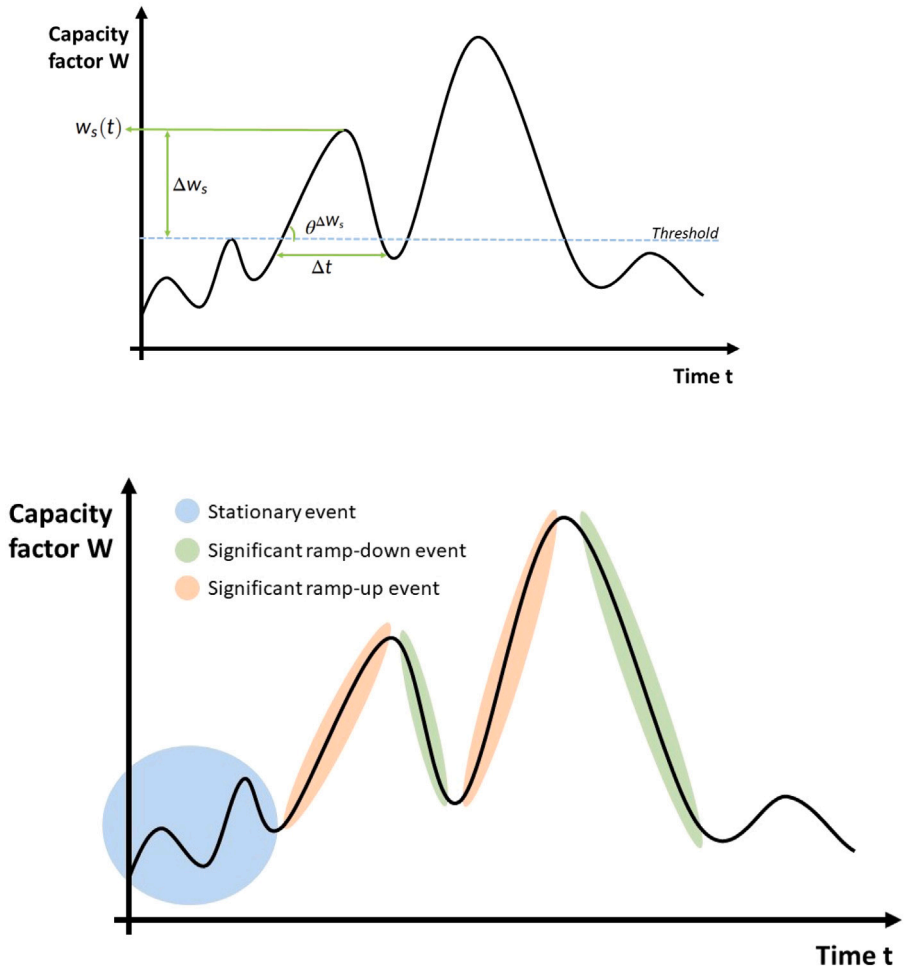

Fig. 4. Main features of ramp events, and examples of stationary and significant events, for the original wind application of RBA.

further conditioning was necessary prior to RBA. However, depending on background noise levels, denoising of the recorded PD data might be necessary under some circumstances. A number of methods have been investigated to accomplish this more commonly, e.g. fast Fourier transform (FFT), short-time Fourier transform (STFT), low-pass and notch filtering, as well as wavelet-based filtering methods [39].

Algorithm 1 Event extraction using Ramping Behavior Algorithm

Require: $N_l > 0$
Ensure: $N = 0 \dots 1$
Ensure: *Set* τ
  while $N_l \geq 0$ **do**
    if $N_i \geq \tau$ **then**
      $E^s \leftarrow N_i$
      $E^s \leftarrow \Delta t, \Delta t_m, \Delta w, \Delta w_m, \theta, t, T_1, T_2$
    else if $N_i \leq \tau$ **then**
      $E_s \leftarrow N_i$
      $E_s \leftarrow \Delta w_s, \theta^{\Delta w_s}, w_s, T_1, T_2, \sigma_s$
    end if
  end while

Implementation of RBA$_\theta$ algorithm on partial discharge measurement dataset is demonstrated as pseudocode in Algorithm 1. The measurement data from four stress levels are sequentially fed into the algorithm, and extracted events are stored. In each iteration, the length of data (N_l) is required to be positive and the dataset (N) is normalized between 0 and 1. Afterward, a threshold τ is set that will distinguish between significant (E^s) and stationary (E_s) events. For this analysis, the significance corresponds to partial discharge events and stationary noise in the measurement. Each significant and stationary event is made up of several features that describe an individual event of either type as listed in the algorithm above. A complete description of the RBA$_\theta$ algorithm, features and the working principle is covered in [31].

The working procedure of the RBA$_\theta$, starting from input data series, PD peak events, to extraction of significant events with the feature description is depicted in Fig. 5. When the variation between consecutive data set points exceeds the threshold value, the instant at which this change occurs is recorded as the start time (T_1) of the significant event and the end of the event as (T_2). The time interval (Δt) is used to calculate the total number of samples recorded when the PD peak changes from its initial value to a maximum value of the PD pulse given by the difference between (T_1) and (T_2). The initial magnitude of the PD pulse is denoted by w_1, captured at instant T_1, and w_2 corresponds to the PD peak magnitude observed at T_2. Furthermore, the difference between the initial and final PD magnitudes Δw is recorded, also the mean value of these σ_m. The slope formed between w_1 and w_2 is calculated in relation to the horizontal from the data point where T_1 is marked and the angle denoted is by θ_m. The set of the frequencies per featured event denoted by Λ_m for Δt_m, Δw_m, θ_m, and σ_m are also calculated for the given PD data set.

The previously described mapping process results in two output data sets, i.e., the significant events which capture the occurrence of PD peaks and stationary events which account for noise. Hence for further evaluation stationary events are not considered. It should also be noted that the PD pulses are also subjected to noise interference (although with different frequencies) which affects the peak value of the PD pulses along with their waveform. Table 1 represents the output of the

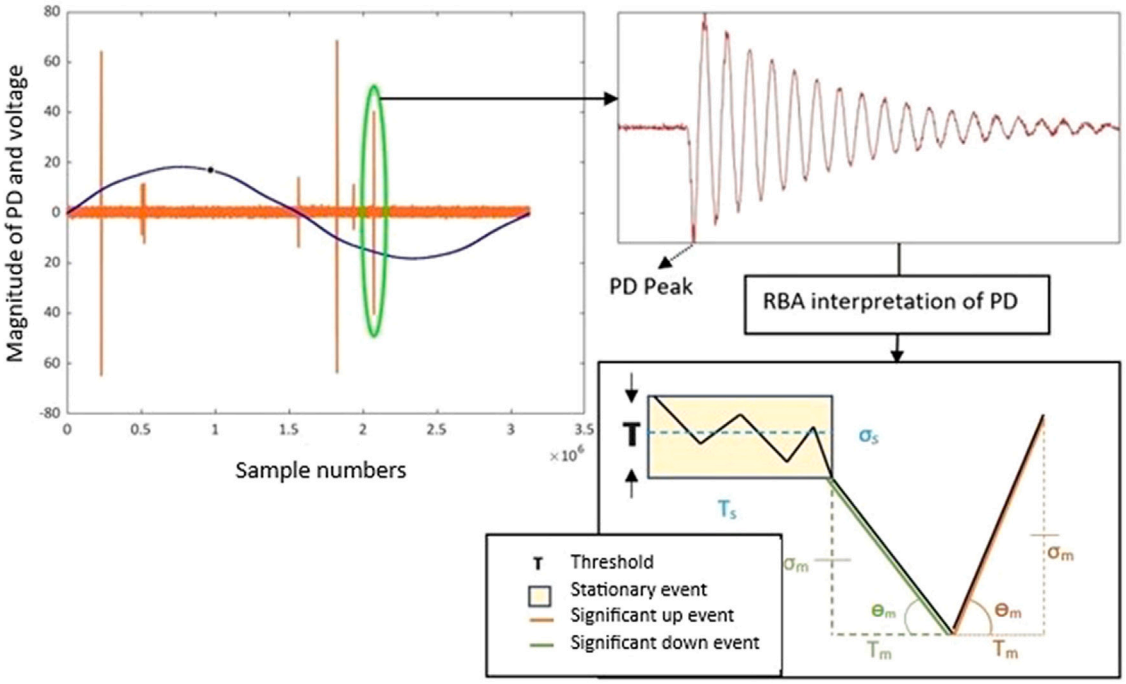

Fig. 5. Mapping of partial discharge (PD) measurements with RBA_θ algorithm.

Table 1
Example of significant events obtained from RBA_θ output.

t_1	t_2	Δt_m	$w_m(t_1)$	$w_m(t_2)$	Δw_m	σ_m	θ_m (°)
158 305	158 316	11	0.06842	0.96102	0.89259	0.51472	82.97
318 594	318 601	7	0.10773	1.11090	1.00317	0.60931	86.00
418 648	418 652	4	0.02606	−0.1951	−0.2211	−0.0845	−79.74
666 536	666 546	10	0.07188	0.76551	0.6936	0.4187	81.79
1 544 996	1 545 010	14	0.02936	−0.4496	−0.4789	−0.2101	−73.70
1 750 350	1 750 362	12	0.00325	−0.4299	−0.4331	−0.2133	−74.51
1 874 555	1 874 563	8	0.03584	−1.0946	−1.1304	−0.5293	−85.95
2 090 801	2 090 809	8	−0.0553	−0.876	−0.8209	−0.4658	−84.43
2 117 052	2 117 054	2	−0.0584	0.21464	0.27309	0.0781	85.811
2 161 524	2 161 532	8	−0.0358	−0.4073	−0.3714	−0.2215	−77.84

significant event captured by mapping it from the measurement file at an 11.3 kV voltage stress level. Not only are the PD peaks captured, but the oscillations accompanying the PD are also shown in the table. The occurrence of oscillations is a by-product of the measurement process and attention should be primarily directed towards the initial PD pulse and its waveform. The first peak that occurred during the event is the actual partial discharge. This peak formation may occur in one or more steps depending on the noise interference. The point at which the PD peaks magnitude changes its polarity is starting point of the oscillation of any given PD event. This sign reversal as shown in Table 1 (starting from column 3) is the significant event which is accounting for PD oscillations.

Fig. 6 illustrates the total number of significant events (secondary y-axis) and PD events (primary y-axis) extracted from the measurements in each data set at different voltage stress levels. The variation of the total number of events observed ranges between 5 to 22. The PD peaks observed at 6.4 kV voltage stress level consist of only negative polarity and each observation has only one PD peak. At a voltage stress level of 7.4 kV, the variation of the total number of events ranges between 7 to 54. The occurrence of PD peaks with negative polarity is prominent in each evaluated data set and the positive polarity peaks occur in only a few data sets. The total PD peaks range between 1 and 2. After further increment in voltage stress level (i.e., at 9.4 kV), the occurrence of PD peaks having positive polarity occurrence becomes more pronounced and is observed in each data set. However, the negative polarity PD peak occurrence frequency is still higher. At 11.3 kV, both positive and negative polarity PD peaks increase and range between 1 to 4 and 2 to 4 respectively.

The frequency of occurrence of PD peaks at different phase angles and their respective peak magnitude is shown in Fig. 7 for different voltage stress levels. The phase angle is computed by equating the total number of samples to 2π. The polarity of the voltage cycle has an effect on the PD peak polarity. In the positive cycle, the occurrence of positive PD peaks are observed, and vice-versa in the negative voltage cycle, as expected. Since at low voltage stress levels, positive peaks are absent, further analysis is carried out for the negative voltage cycle only.

At 11.3 kV voltage stress level, the frequency of occurrence of PD magnitudes (normalized) between 0 to −1 and −1 to −2 are almost equal (i.e. 9 and 10 respectively) which occurs in the voltage phase angle ranging from 150° to 200° and 200° to 250° respectively. The highest PD peak magnitudes having frequency of occurrence 2, are also realized in the latter range. The frequency of occurrence of PD magnitudes is scattered more at the 9.4 kV stress level. The highest occurring frequency (i.e. 7) of the normalized PD magnitudes ranges from 0 to 1 which occurs from 150° to 200° of voltage phase angle. High magnitude PD peaks arise in all the voltage phase angles as shown in Fig. 7 with almost similar frequency of occurrence. However, the voltage phase angle range from 200° to 250° has slightly higher occurrence.

For 7.4 kV stress level, the range from 180° to 210° has the highest frequency of occurrence of PD pulses with magnitudes ranging from 0 to −1. The normalized PD magnitudes between −1 to −2 occur mostly between 220° to 240° while one PD peak is also identified in the range of 180° to 200°. The PD peaks ranging from 0 to −0.5 occur mostly within the voltage phase angle range of 220° to 240°. The highest PD peaks having normalized magnitude of −1 occur dominantly in the range of 240° to 260°.

Fig. 8 represents the variation of the time difference (or sample number difference) for each PD pulse, measured at different voltage

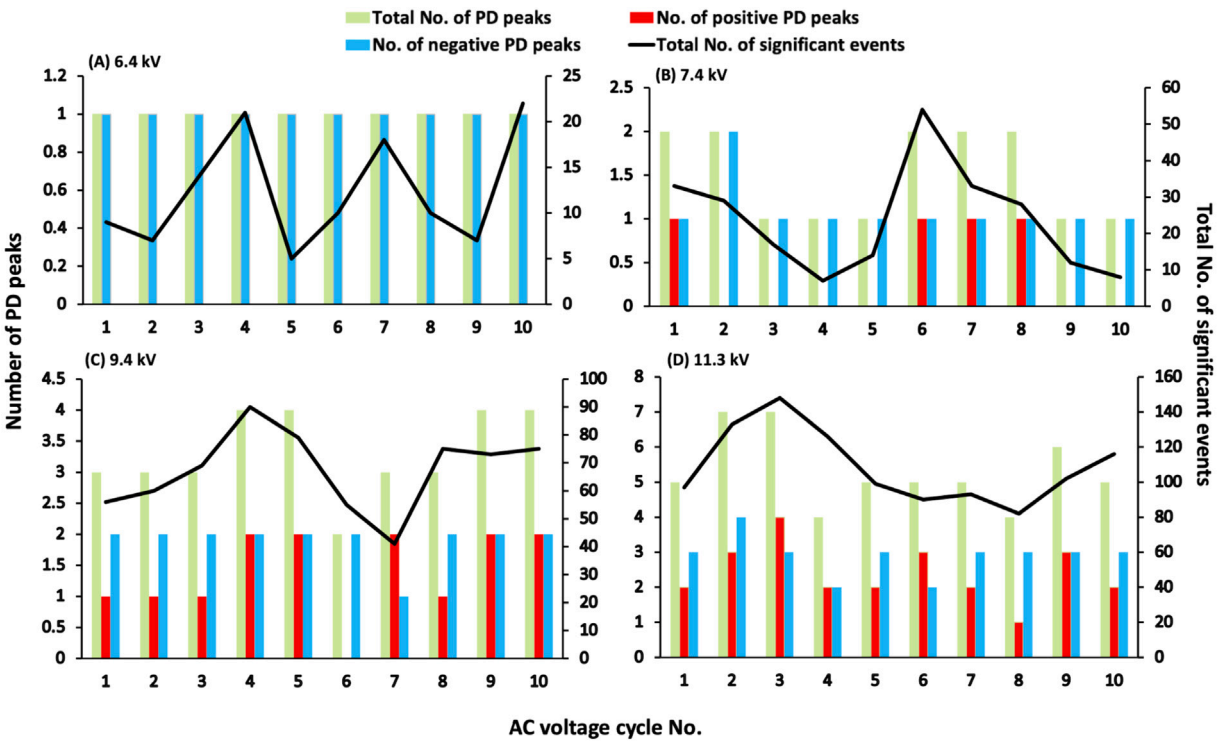

Fig. 6. Occurrence of positive and negative partial discharge peaks and the total number of identified significant events for each voltage stress level across 10 AC voltage cycles.

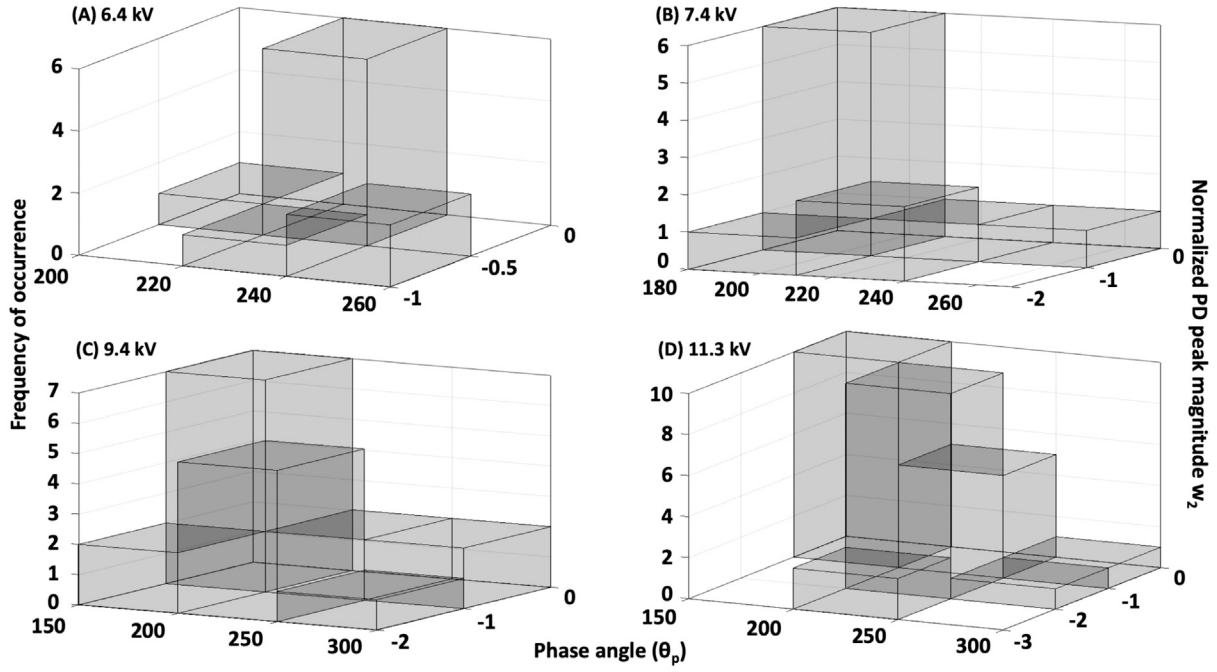

Fig. 7. Frequency of occurrence of PD peak magnitudes at different voltage phase angle (θ_p) for each voltage stress level.

stress levels, to obtain its peak amplitude from its initial value. At 11.3 kV, the variation of Δt ranges from 3 to 19, however, for most of the PD pulses, Δt ranges between 8 to 14. The variation of Δt at 9.4 kV stress level has almost a similar range (i.e. between 3 to 18). However, for most of the PD peaks Δt ranges between 8 to 12. The magnitude of Δt reduces at 7.4 kV stress level, having the highest value of 13. However, at 6.4 kV stress level, the Δt highest value is 17. Most of the PD peak formation for this stress level occurs at 12.

3. Results: Patterns and trends of internal partial discharge

The aggregate percentage of the positive and negative PD peaks with respect to the total number of peaks and the total number of significant events captured in each data set is presented in Table 2. It can be inferred that the occurrence of negative PD peaks is predominant at each voltage stress level. Also, at lower stress levels, positive PD peaks are not observed with increasing stress levels. Hence only the negative voltage half-cycle is analyzed. Additionally, the total number

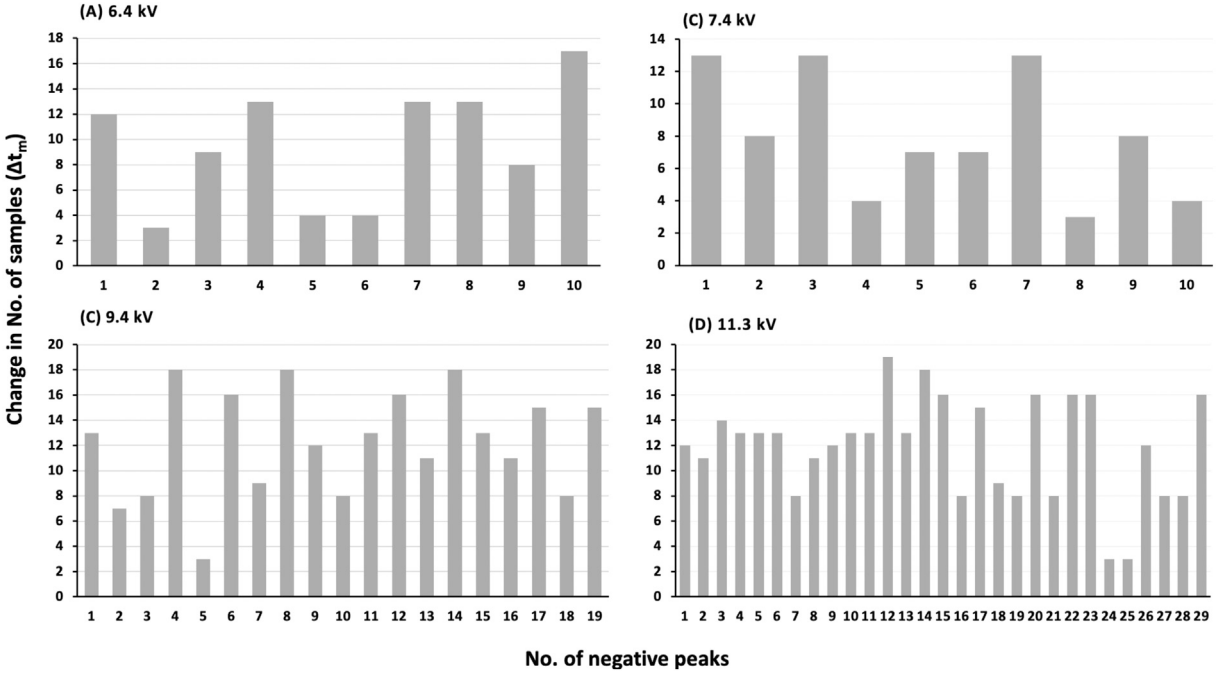

Fig. 8. Negative PD pulse rise-times across 10 AC voltage cycles as determined by RBA_{θ} algorithm.

Table 2
Aggregated percentage change of partial discharge peaks as significant events at different voltage levels.

Voltage (kV)	6.4	7.4	9.4	11.3
Negative PD peaks (%)	100	80	60	55.5
Significant events corresponding to negative PD peaks (%)	11.3	6.6	2.8	2.7
Positive PD peaks (%)	0	20	40	44.5
Significant events corresponding to positive PD peaks (%)	0	1.1	2.1	2.2

Table 3
Comparison of aggregate percentage change in PD peaks and significant events.

Voltage (kV)	6.4	7.4	9.4	11.3
Phase angle (%)	99.82	99.89	103.71	103.06
Δt_m (%)	129.92	122.21	117.45	104.85
$w_m(t_2)$ (%)	116.63	122.96	136.78	153.02
θ_m (%)	100.63	99.43	94.77	100.45

of significant events grows with increasing stress levels which correspond to the lower percentage change in negative PD peaks with increasing stress levels.

Table 3 represents the rolling average percentage of the RBA_{θ} algorithm output parameters for different voltage stress levels. The percentage of voltage phase angle (corresponding to t_2) decreases from 103.06% at 11.3 kV to 99.82% at 6.4 kV stress level. However, the percentage in Δt increases with decreasing stress levels. The rolling average percentage of Δw (corresponding to a change in PD peak from its initial value to final value) has similar behavior to change in voltage phase angle and inverse relation with change in Δt. At 11.3 kV, the aggregate rolling percentage change is 153.02% which reduces with decreasing stress level and reaches 116.63% at 6.4 kV voltage stress level. The percentage of θ_m (the angle at which the PD peak is generated) has a value of 100.45% at 11.32 kV which reduces to 99.43% at 7.4 kV. However, the aggregate percentage change again increases to 100.63% at 6.4 kV. Successive voltage cycles and the occurrence of the PD peaks at each voltage stress level explain the change in the occurrence of the PD peaks. Then, a rolling average method is used where an average value is calculated over a time horizon.

To determine a trend in the variation of PD peak occurrence at different voltage stress levels, linear polynomial fitting is applied to variables Δt and Δw. The mathematical equation for the curve is given as presented in (1).

$$\Delta w(\Delta t) = p_1(\Delta t)^2 + p_2\Delta t + p_3 \tag{1}$$

Fig. 9 shows the variation of the curve for different voltage stress levels. This variation can be explained in two parts — first, the number of PD peaks varies at each stress level, and second, the variations of Δt which are affected by the noise (higher voltage stress level has more effect) and the Δw (normalized PD peaks) have inverse relation in their rolling aggregate percentage as depicted in 4. These factors contribute to different curve formations after applying the polynomial fitting. The parameters of the equation along with their R^2 and RMSE (root mean squared error) values are presented in Table 4. It can be observed that the RMSE value increases with voltage and is highest (0.415) at the 11.3 kV stress level. Concurrently, the R^2 value decreases with voltage. Coefficient p_1, which is associated with the second degree and the constant value p_3 have negative values while p_2 is positive. With decreasing voltage stress level, the RMSE value decreases and the R^2 score improves by reaching values 0.13 and 0.82, respectively. At 9.4 kV, the signs of p_1 and p_3 change from positive to negative, and p_2 becomes negative instead of positive. At 7.4 kV, the sign of all the constants becomes negative. Finally, at the 6.4 kV stress level, the sign convention becomes similar to 11.3 kV, although having different values as presented in Table 4.

For comparison, the smoothing spline fitting was also applied. In Fig. 9, it can be observed that as the voltage increases, so does the disparity between the two fitted curves. It is evident that the degree of stochasticity in the PD activity increases substantially with the increase in voltage, particularly due to the increased variability of Δw. The higher applied voltage facilitates larger PD peak magnitudes, while the possibility for smaller PD pulses to occur is also present. This results in a higher degree of variability in the dataset, which will necessarily have a rather poor fit to a polynomial function. From Table 4, it is also evident that the smoothing spline yields a better fit compared to the second degree polynomial, with higher R^2 values and lower RMSE values at all voltage levels. The spline fitting also lacks the monotonic nature of the polynomial observed for these data.

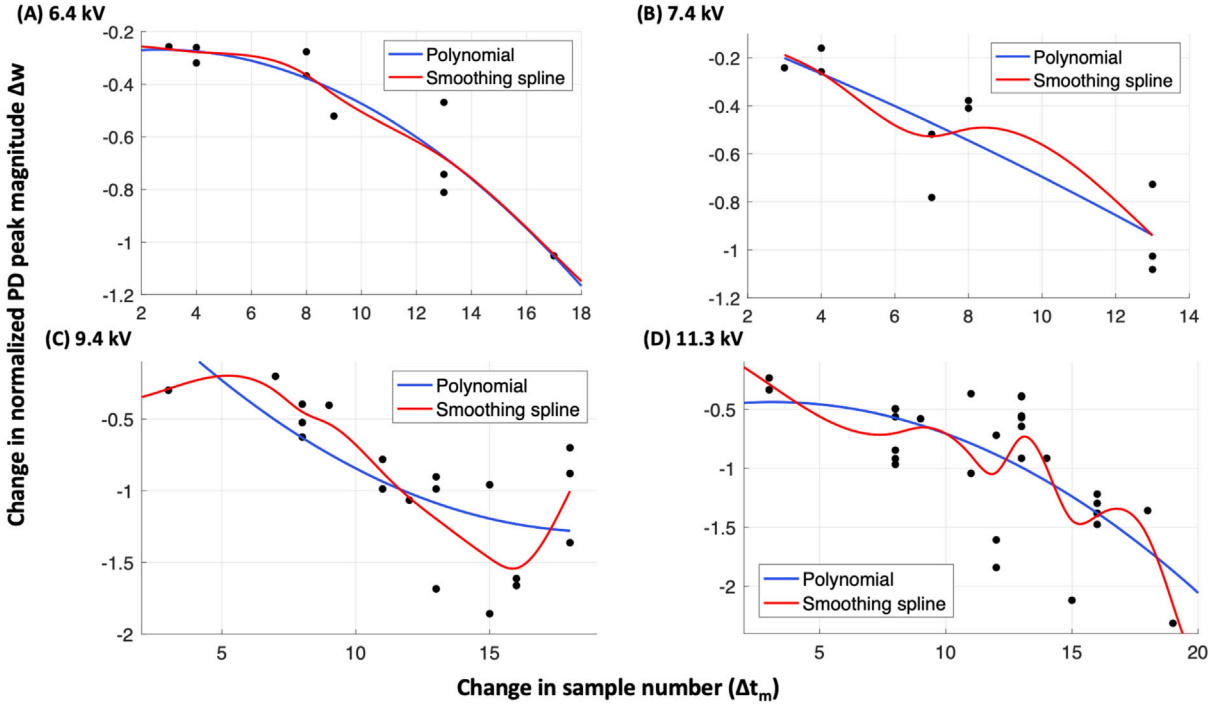

Fig. 9. 2nd degree polynomial and smoothing spline fitting applied to partial discharge RBA algorithm results. (REVISED CAPTION).

Table 4
Polynomial fitting equation coefficients, R^2 score and RMSE for polynomial and smoothing spline (S) fitting at different voltage stress levels.

Voltage (kV)	p_1	p_2	p_3	R^2	RMSE	R^2 (S)	RMSE (S)
6.4	−0.003	0.025	−0.312	0.821	0.130	0.877	0.122
7.4	−0.001	−0.057	−0.019	0.764	0.181	0.823	0.166
9.4	0.005	−0.201	0.645	0.512	0.368	0.748	0.305
11.3	−0.005	0.035	−0.495	0.469	0.415	0.698	0.344

4. Concluding remarks

The objective of this work was to quantify the patterns and trends in the occurrences of internal partial discharge. Internal partial discharge tests were conducted at four stress levels: 6.4, 7.4, 9.4, and 11.3 kV to collect the data set. The data set contains actual partial discharge events and noise from the measurement.

The ramping behavior analysis (RBA_θ) method was developed to quantify variations and characterize the data by segregating them into significant and stationary events. The RBA_θ method has been well established in wind power production data sets, which are inherently stochastic. Thus, such a method aligns with the objective of this work in identifying internal partial discharge events. RBA_θ method is implemented in parallel and recursive programming form. This results in faster computational processing. Beyond that, the key advantage of RBA_θ method is to provide insight from data with easier recognizable results. That is to imply significant and stationary events are much easier to recognize and therefore plan for. In addition, the number of events is much smaller than the number of data points, thereby providing a holistic insight into the phenomenon with smaller storage size requirements.

The following steps are involved in the RBA_θ method:

- Segmentation: differentiating between the noise and partial discharge events by identifying and setting the threshold
- Identification: identifying PD peak events with oscillations by capturing the whole cycle from the origin until the end of oscillations.
- Characterization: Characterizing PD peak formation which takes place in several steps due to noise interference and removing the oscillations.
- Quantification: Quantifying the events into significant and stationary events.
- Comparison: Comparative analysis is performed using the parameters phase angle (time or observations), magnitude, and angle variations associated with PD peaks for each stress level to quantify the changes.

The following are the main findings of the analysis.

- The occurrence of PD events is voltage-polarity dependent. Negative PD peaks occur during the negative voltage half-cycle, as expected. The number of PD peaks increases with increasing stress levels. At lower stress levels, PD peaks occur mostly during the negative polarity. With increasing voltage, the occurrence of positive PD peaks becomes more frequent.
- The number of significant events is associated with the occurrence of PD according to the RBA_θ setting. The number of significant events is higher than the PD peak number, as it accounts for the oscillations accompanying PD peaks.
- The stationary events account for the noise. The range of phase angles at which PD occurs is mostly between 200°–250° at 11.32 kV and 9.4 kV. For 7.4 kV and 6.4 kV stress levels, the PD occurrence frequency is highest between the phase angle range from 220°–240°. The time duration from an initial value of the PD peak to the final value is higher at high-stress levels.
- The aggregate percentage change of negative PD peaks with respect to significant events increases with decreasing stress levels as positive PD peaks increase (from Table 2).
- The percentage change in phase angle increases with decreasing stress levels. Subsequently, the change in PD peak magnitude is in descending order in correspondence with the decreasing stress levels (from Table 3).
- The finding from curve fitting is that the RMSE of curve fitting increases with increasing stress level (i.e., stochasticity increases with increasing voltage). The polynomial fitting curve constants

also change signs, resulting in different types of curves when fitted.

The scope of this study is to diagnose the partial discharge events from the measurements collected by the sensors. Then, the data source is limited by the observations from the sensor. The noise data recorded by the sensor is characterized as stationary events and not included in this study. In future research work, more continuous observations could be collected and stationary events could be utilized in investigating how a partial discharge pulse builds on over time.

This work demonstrates the application of RBA_θ to the analysis of a specific type of internal PD defect. Further work on this subject can include studying the characteristics of different types of PD defects as determined using RBA_θ, both of internal and external nature and also corona discharge, to determine how these sources can be differentiated during PD signal interpretation. The impact of various types of noise on the results of the RBA_θ analysis, which can be present during any PD measurement both in a laboratory or on-site setting, should also be investigated.

CRediT authorship contribution statement

Sambeet Mishra: Conceptualization, Methodology, Supervision, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Praveen Prakash Singh:** Data curation, Writing – original draft, Investigation, Formal analysis. **Ivar Kiitam:** Writing – original draft, Writing – review & editing, Investigation, Formal analysis. **Muhammad Shafiq:** Conceptualization, Data curation, Funding acquisition. **Ivo Palu:** Funding acquisition, Resources. **Chiara Bordin:** Methodology, Investigation, Formal analysis, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The source of funding received for partial fulfillment of this work is duly cited in the acknowledgment section.

Data availability

Data will be made available on request.

Acknowledgments

This work was partially supported by the Estonian Research Councilgrant PSG632.

References

[1] H. Illias, T.S. Yuan, H. Mokhlis, G. Chen, P.L. Lewin, et al., Partial discharge patterns in high voltage insulation, in: 2012 IEEE International Conference on Power and Energy (PECon), IEEE, 2012, pp. 750–755.

[2] I. Kiitam, M. Shafiq, P. Taklaja, M. Parker, I. Palu, L. Kütt, Characteristic pulse pattern features of different types of partial discharge sources in power cables, in: 2021 IEEE PES/IAS PowerAfrica, IEEE, 2021, pp. 1–5.

[3] D. Fynes-Clinton, C. Nyamupangedengu, Partial discharge characterization of cross-linked polyethylene medium voltage power cable termination defects at very low frequency (0.1 hz) and power frequency test voltages, IEEE Electr. Insul. Mag. 32 (4) (2016) 15–23.

[4] R. Gillie, A. Nesbitt, R. Ramirez-Iniguez, B. Stewart, G. Kerr, Analysis of HV cable faults based on correlated HFCT and IEC60270 measurements, in: 2014 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), IEEE, 2014, pp. 168–171.

[5] E. Shaalan, S. Ward, A. Youssef, Effect of cavity position, size and geometry on partial discharge behaviour inside 18/30 kv XLPE cables, Int. J. Electr. Eng. Inform. 13 (2021) 495–507.

[6] F. Steennis, P. Wagenaars, P. van der Wielen, P. Wouters, Y. Li, T. Broersma, D. Harmsen, P. Bleeker, Guarding MV cables on-line: With travelling wave based temperature monitoring, fault location, PD location and PD related remaining life aspects, IEEE Trans. Dielectr. Electr. Insul. 23 (3) (2016) 1562–1569.

[7] M. Shafiq, I. Kiitam, P. Taklaja, L. Kütt, K. Kauhaniemi, I. Palu, Identification and location of PD defects in medium voltage underground power cables using high frequency current transformer, IEEE Access 7 (2019) 103608–103618.

[8] P.C. Van Der Wielen, J. Veen, P.A. Wouters, E.F. Steennis, On-line partial discharge detection of MV cables with defect localisation (PDOL) based on two time synchronised sensors, in: CIRED 2005-18th International Conference and Exhibition on Electricity Distribution, IET, 2005, pp. 1–5.

[9] A. Milioudis, G. Andreou, D. Labridis, On-line partial discharge monitoring system for underground MV cables–part II: Detection and location, Int. J. Electr. Power Energy Syst. 109 (2019) 395–402.

[10] O. Siirto, J. Vepsäläinen, A. Hämäläinen, M. Loukkalahti, Improving reliability by focusing on the quality and condition of medium-voltage cables and cable accessories, CIRED-Open Access Proc. J. 2017 (1) (2017) 229–232.

[11] O. Arikan, C.C. Uydur, C.F. Kumru, Insulation evaluation of MV underground cable with partial discharge and dielectric dissipation factor measurements, Electr. Power Syst. Res. 220 (2023) 109338, http://dx.doi.org/10.1016/j.epsr.2023.109338.

[12] W. Wang, X. Yan, S. Li, L. Zhang, J. Ouyang, X. Ni, Failure of submarine cables used in high-voltage power transmission: Characteristics, mechanisms, key issues and prospects, IET Gener. Transm. Distrib. 15 (9) (2021) 1387–1402.

[13] R.G. Nordic, 2021 Nordic and Baltic Grid Disturbance Statistics. European Network of Transmission System Operators for Electricity, 2022, 2022, Accessed: April 2023, https://eepublicdownloads.entsoe.eu/clean-documents/SOC20documents/Nordic/2022/2021_Nordic_and_Baltic_Grid_Disturbance_Statistics_FOR_PUBLISHING.pdf.

[14] S. Chandrasekar, A. Cavallini, G.C. Montanari, F. Puletti, Bandwidth and sensitivity issues in PD detection in power cables, IEEE Trans. Dielectr. Electr. Insul. 14 (3) (2007) 735–743.

[15] M. Tozzi, Partial discharge in power distribution electrical systems: Pulse propagation models and detection optimization, 2010.

[16] G. Robles, M. Shafiq, J.M. Martínez-Tarifa, Multiple partial discharge source localization in power cables through power spectral separation and time-domain reflectometry, IEEE Trans. Instrum. Meas. 68 (12) (2019) 4703–4711.

[17] A. Ragusa, H.G. Sasse, A. Duffy, On-line partial discharge localization in power cables based on electromagnetic time reversal theory-numerical validation, IEEE Trans. Power Deliv. 37 (4) (2021) 2911–2920.

[18] A.R. Mor, P. Morshuis, P. Llovera, V. Fuster, A. Quijano, Localization techniques of partial discharges at cable ends in off-line single-sided partial discharge cable measurements, IEEE Trans. Dielectr. Electr. Insul. 23 (1) (2016) 428–434.

[19] R. Liang, Z. Zhang, H. Li, P. Chi, G. Li, Y. Tao, Partial discharge location of power cables based on an improved single-terminal method, Electr. Power Syst. Res. 193 (2021) 107013, http://dx.doi.org/10.1016/j.epsr.2020.107013.

[20] M. Wu, H. Cao, J. Cao, H.-L. Nguyen, J.B. Gomes, S.P. Krishnaswamy, An overview of state-of-the-art partial discharge analysis techniques for condition monitoring, IEEE Electr. Insul. Mag. 31 (6) (2015) 22–35.

[21] A. Krivda, Automated recognition of partial discharges, IEEE Trans. Dielectr. Electr. Insul. 2 (5) (1995) 796–821.

[22] W.J.K. Raymond, H.A. Illias, H. Mokhlis, et al., Partial discharge classifications: Review of recent progress, Measurement 68 (2015) 164–181.

[23] N. Sahoo, M. Salama, R. Bartnikas, Trends in partial discharge pattern classification: a survey, IEEE Trans. Dielectr. Electr. Insul. 12 (2) (2005) 248–264.

[24] S. Govindarajan, A. Morales, J.A. Ardila-Rey, N. Purushothaman, A review on partial discharge diagnosis in cables: Theory, techniques, and trends, Measurement (ISSN: 0263-2241) 216 (2023) 112882.

[25] L. Hao, A. Contin, J. Hunter, D. Swaffield, P. Lewin, C. Walton, M. Michel, A new method for automatic multiple partial discharge classification, 2011.

[26] V. Javandel, M. Vakilian, K. Firuzi, Multiple partial discharge sources separation using a method based on laplacian score and correlation coefficient techniques, Electr. Power Syst. Res. 210 (2022) 108070, http://dx.doi.org/10.1016/j.epsr.2022.108070.

[27] T. Abdel-Galil, R. Sharkawy, M. Salama, R. Bartnikas, Partial discharge pulse pattern recognition using an inductive inference algorithm, IEEE Trans. Dielectr. Electr. Insul. 12 (2) (2005) 320–327.

[28] S. Biswas, D. Dey, B. Chatterjee, S. Chakravorti, An approach based on rough set theory for identification of single and multiple partial discharge source, Int. J. Electr. Power Energy Syst. 46 (2013) 163–174.

[29] S. Mantach, A. Lutfi, H. Moradi Tavasani, A. Ashraf, A. El-Hag, B. Kordi, Deep learning in high voltage engineering: A literature review, Energies 15 (14) (2022) 5005.

[30] A. Deshpande, H. Mangalvedekar, A. Cheeran, Partial discharge analysis using energy patterns, Int. J. Electr. Power Energy Syst. 53 (2013) 184–195.

[31] S. Mishra, E. Ören, C. Bordin, F. Wen, I. Palu, Features extraction of wind ramp events from a virtual wind park, Energy Rep. 6 (2020) 237–249.

[32] S. Mishra, M. Leinakse, I. Palu, Wind power variation identification using ramping behavior analysis, Energy Procedia 141 (2017) 565–571.

[33] S. Mishra, M. Leinakse, I. Palu, J. Kilter, Ramping behaviour analysis of wind farms, in: 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), IEEE, 2018, pp. 1–5.

[34] S. Mishra, C. Bordin, K. Taharaguchi, A. Purkayastha, Predictive analytics beyond time series: Predicting series of events extracted from time series data, Wind Energy 25 (9) (2022) 1596–1609.
[35] C. Bordin, Mathematical optimization applied to thermal and electrical energy systems, 2015.
[36] C. Bordin, A. Håkansson, S. Mishra, Smart energy and power systems modelling: an iot and cyber-physical systems perspective, in the context of energy informatics, Procedia Comput. Sci. 176 (2020) 2254–2263.
[37] C. Bordin, S. Mishra, I. Palu, A multihorizon approach for the reliability oriented network restructuring problem, considering learning effects, construction time, and cables maintenance costs, Renew. Energy 168 (2021) 878–895.
[38] A.A. Ibrahim, B. Kazemtabrizi, C. Bordin, C.J. Dent, J.D. McTigue, A.J. White, Pumped thermal electricity storage for active distribution network applications, in: 2017 IEEE Manchester PowerTech, IEEE, 2017, pp. 1–6.
[39] S. Sriram, S. Nitin, K. Prabhu, M. Bastiaans, Signal denoising techniques for partial discharge measurements, IEEE Trans. Dielectr. Electr. Insul. 12 (6) (2005) 1182–1191.