A Clustering Technique for Partial Discharge and Noise Sources Identification in Power Cables by Means of Waveform Parameters

F. Álvarez , J. Ortego , F. Garnacho and M.A. Sánchez-Urán

ABSTRACT

On-line partial discharge (PD) measurements have become a common technique for assessing the insulation condition of installed high voltage (HV) insulated cables. When on-line tests are performed in noisy environments, or when more than one source of pulse-shaped signals are present in a cable system, it is difficult to perform accurate diagnoses. In these cases, an adequate selection of the non-conventional measuring technique and the implementation of effective signal processing tools are essential for a correct evaluation of the insulation degradation. Once a specific noise rejection filter is applied, many signals can be identified as potential PD pulses, therefore, a classification tool to discriminate the PD sources involved is required. This paper proposes an efficient method for the classification of PD signals and pulse-type noise interferences measured in power cables with HFCT sensors. By using a signal feature generation algorithm, representative parameters associated to the waveform of each pulse acquired are calculated so that they can be separated in different clusters. The efficiency of the clustering technique proposed is demonstrated through an example with three different PD sources and several pulse-shaped interferences measured simultaneously in a cable system with a high frequency current transformer (HFCT).

1 INTRODUCTION

THE analysis of the degradation modes in different insulation materials of HV electrical systems has shown that the presence of PD is very common in all of them [1-3]. Once PD activity is detected it is important to identify the associated type of defect and its location to evaluate whether the discharges are harmful or not. The instant of time at which pulses are generated with respect to the reference signal of the test voltage, their amplitude and repetition rate are related with the nature of the type of defects or noise sources [4, 5]. Thus, by analyzing the phase-resolved partial discharge (PRPD) pattern obtained in a measurement, it is possible to identify the type of insulation defect involved in a PD degradation process [6, 7]. Similarly, different noise sources can be recognized by analyzing their characteristic pattern [8]. However, when on-line tests are performed, in many cases the interpretation of PRPD patterns proves to be very difficult or impossible and it is necessary to apply efficient techniques to filter the background noise and to classify the pulses corresponding to the sources involved [8-10].

Several approaches have been developed in order to discriminate different PD and noise sources, all of them are based on the extraction of characteristic parameters from individual registered pulses. In [11] a clustering technique based on the comparison of signal energies associated with particular wavelet-decomposition levels is presented. The same authors implement another source discrimination technique using a support vector machine and correlation analysis to group the discharges into different clusters [12].

Other investigations present separation techniques analyzing the signals in the time and frequency domain [10, 13, 14]. In [13, 14], two characteristic parameters are calculated and visualized in a two dimensional diagram in order to identify different clusters, while in [10] each pulse is characterized by a mathematical function that summarizes the time and frequency features of the signal and the classification method is based on the comparison of the resulted functions. Two different approaches have also been developed for the classification of the measured pulses analyzing their frequency spectra [15-17]. In [15] two power ratios in selected frequency bands are calculated for each pulse and represented in a diagram for clustering purposes, while in [16, 17] the spectral energy of the signals is evaluated simultaneously with three selected band-pass filters, drawing the corresponding charge values in a star diagram for the classification of pulses in different groups.

In this study a new pulse classification tool based on the waveform analysis of the recorded signals is presented. Three characteristic parameters (f, A_p and A_t) are calculated for each pulse; the first one, f, characterizes their frequency content while A_p and A_t describe the waveform of their normalized associated envelope. A graphical tool based on two and three dimensional representations of the characteristic parameters, makes it possible to identify in an effective way different types of defects and noise sources simultaneously present in a test object.

In sections 2 and 3 of this paper, the wave shape of different types of PD pulses is analyzed and mathematical models developed for their characterization and classification are presented. Section 4 describes the implementation of the clustering tool which has been developed. Finally experimental PD measurements are performed and the results of the clustering technique are analyzed in detail.
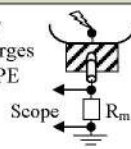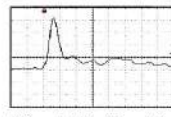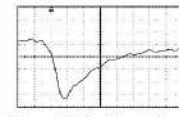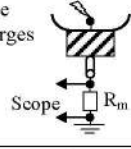
2 PD WAVEFORMS
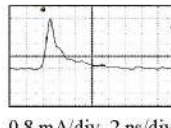
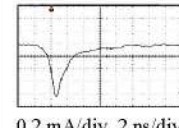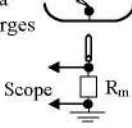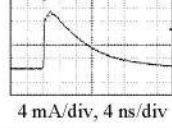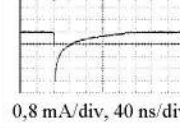The two main factors that differentiate captured PD signals arising in different sources are the physics of their origin (insulation medium, geometry, degradation level and local conditions) and their travelling paths until they reach the PD sensor. These two factors characterize the waveforms corresponding to each source, therefore it is interesting to analyze representative features related with the wave shapes of the measured pulses for classification purposes. Another important factor influencing the waveform of the captured pulses is the detection circuit implemented for their acquisition. Original pulses considered those measured close to the source site as shown in Table 1 [19] and pulses measured with RC wideband detection impedances in ideal conditions [20] present no oscillations. However, those measured by using RLC detection impedances such as the one described in

IEC 60270 [18] or by high frequency current transformers (HFCT) [11], present a damped oscillatory behavior.

Original PD signals are characterized by symmetrical or asymmetrical pulses with a duration width in the nanoseconds range (see Table 1). However, taking into account that the propagation phenomena cause signal attenuation and distortion, the waveforms change ostensibly from the PD source site to the sensor emplacement.

In general, when on-line PD measurements are performed on HV installations, HFCT sensors are placed in the grounding conductors. In power cables, the HFCTs are coupled in the cable earth connections and in cross-bonding configurations in the joints of the sectioned screens. Table 2 shows various PD pulses measured in two types of cable systems. Cable system 1 simulates one of the phases of a 15 kV distribution line. The PD tests were carried out in a laboratory with the cable configuration shown in Figure 1. PD pulses caused by different defects were measured with HFCT sensors coupled in the earth connections of the cable in the simulated metal enclosed switchgears. Cable system 2 consists of a three phase 45 kV distribution underground line, see Figure 2. An on-line PD test was performed measuring an internal defect with HFCT sensors installed in the cross-bonding of the joints. In both cases synchronized measurements were performed with the sensors placed in three different locations, in order to visualize the evolution with the distance of the waveforms along several hundred meters. The HFCT sensors have a bandwidth from 100 kHz to 20 MHz and the characteristics of the measuring instruments are 14 bits of vertical resolution, 100 MS/s of sampling rate and 50 MHz of bandwidth. The pulses measured present an oscillatory behavior with a duration in the microseconds range.

Table 1. Original PD pulse waveforms generated with a point to plane gap configuration and acquired with the measuring resistor R_m [19].

Type of defect	Positive current pulse	Negative current pulse
Cavity discharges in XLPE	0,8 mA/div, 2 ns/div	0,2 mA/div, 2 ns/div
Surface discharges	0,8 mA/div, 2 ns/div	0,2 mA/div, 2 ns/div
Corona discharges	4 mA/div, 4 ns/div	0,8 mA/div, 40 ns/div

It can be observed that most of the pulses generated present a progressively damped amplitude and a specific frequency

content. However, in some occasions, two consecutive oscillations can appear very close to each other, so that before the pulse goes down to zero a second oscillation appears. These waveforms may be due to
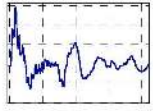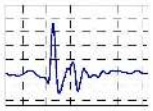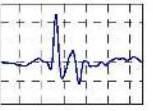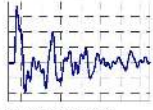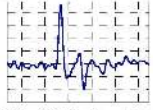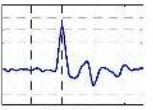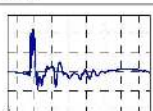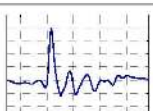
- the overlap of two consecutive different pulses (by random coincidence)
- the same original PD pulse that travels through an alternative path until the HFCT sensor
- reflections in the transmission path
- or by a distortion phenomena of the signal.

Except in the case of reflections in the transmission path, in some occasions, the second oscillation could exhibit a greater amplitude (see in Table 2 the signal measured at 700 meters in cable system 2). In the case of consecutive pulses overlapped by random coincidence the repetition rate is negligible, however if the same waveform presents a significant repetition rate then the signals are caused by a particular PD source and can be identified applying clustering techniques.

It is also noted that when a PD pulse travels long distances, the high frequency components are filtered and the wave shape becomes more symmetric; an example of this behavior is shown in Table 2 for the signal measured at 1700 meters in the cable system 2.

Finally, different characteristic wave shapes are obtained for each defect. In order to classify these types of signals, in the next section their waveforms are characterized by a general mathematical model that allows the extraction of differentiable parameters.

Table 2. Waveforms of different PD pulses acquired by HFCT sensors placed in different positions of two cable systems.

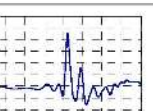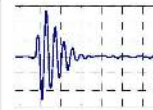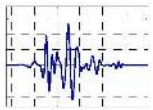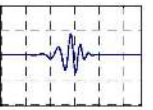| Cable system 1. Measurements in one phase of a 15 kV rated voltage line.			
PD source	Distance from PD source to sensor		
	2 m	400 m	800 m
Internal defect	20 mV/div, 500 ns/div	5 mV/div, 1 µs/div	5 mV/div, 1 µs/div
Surface defect	50 mV/div;500 ns/div	20 mV/div;1 µs/div	20 mV/div;1 µs/div
Corona in air	50 mV/div;500 ns/div	10 mV/div;1 µs/div	5mV/div;1 µs/div

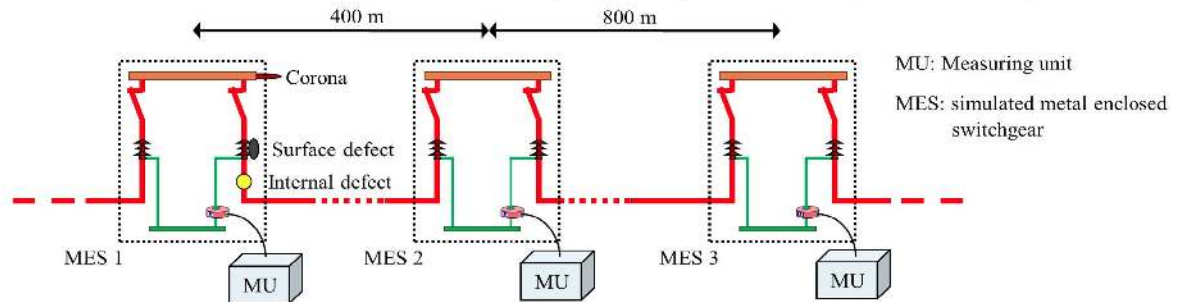| Cable system 2. Measurements in a three phase 45 kV distribution underground line.			
PD source	Distance from PD source to sensor		
	2 m	700 m	1700 m
Internal defect	10 mV/div;1 µs/div	2 mV/div;1 µs/div	0,5 mV/div;1 µs/div

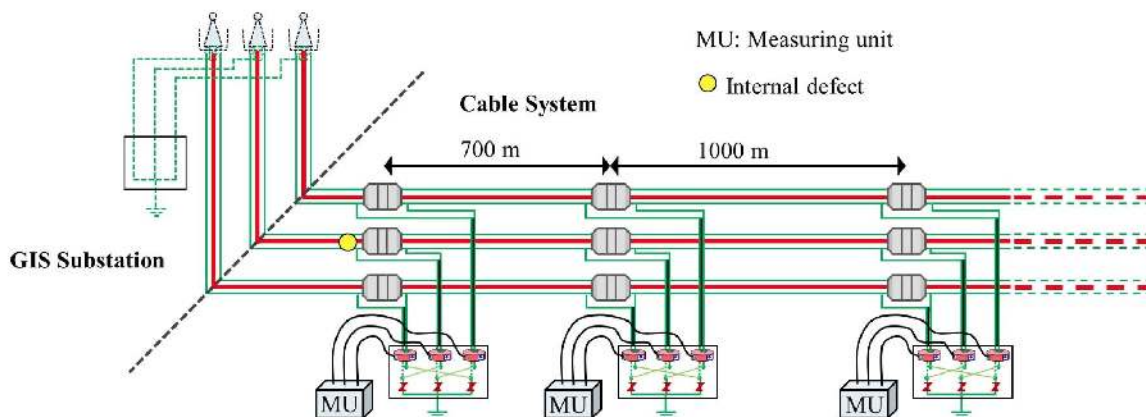**Figure 1.** Experimental setup for one phase of a 15 kV rated voltage line.

Figure 2. On-line PD test in a three phase 45 kV distribution underground line.

3 MATHEMATICAL MODELS

Mathematical models are used to characterize the pulse waveforms. Although this paper is focused on the wave shapes acquired by HFCT sensors, the original PD signals (see Table 1) are also analyzed in order to establish a correlation between them and the measured oscillating pulses. This analysis is useful to set up the basis of the mathematical models applied for damped oscillating signals.

3.1 ORIGINAL PD PULSES

3.1.1 SYMMETRIC PD PULSES

Symmetric PD pulses which originate in a defect can be modeled by the Gaussian function

$$i(t) = I_p \cdot e^{-\frac{t^2}{2 \cdot \sigma^2}} \tag{1}$$

where I_p is the maximum value of the current pulse and it is related with the involved charge Q by the following expression.

$$Q = I_p \cdot \sigma \cdot \sqrt{2 \cdot \pi} \tag{2}$$

The parameter σ is directly related with the pulse width at the half peak value $\Delta t_{50\%}$ (3).

$$\sigma = \frac{\Delta t_{50\%}}{2 \cdot \sqrt{2 \cdot \ln 2}} \tag{3}$$

As a result, based on the values of I_p and $\Delta t_{50\%}$ determined for a recorded pulse waveform, the Gaussian function $i(t)$ and the involved charge Q can be calculated with equations (2) and (3). A verification of the model with the PD pulse is convenient, checking the pulse width at 5% of the peak value ($\Delta t_{5\%}$), which can be determined by equation (4).

$$\Delta t_{5\%} = 2 \cdot \sigma \cdot \sqrt{2 \cdot \ln 20} \tag{4}$$

In a power cable the voltage pulse $V(t)$ that propagates on the first meters through the conductor can be expressed as

$$V(t) = V_p \cdot e^{-\frac{t^2}{2 \cdot \sigma^2}} \tag{5}$$

where:

$$V_p = I_p \cdot Z_c \tag{6}$$

being Z_c the characteristic impedance of the cable that can be considered approximately constant in a frequency range above 1 MHz.

This voltage source causes a current pulse through the insulation to the cable shield where the HFCT sensor is coupled, so the signal measured $i_e(t)$ is defined by the characteristics of this voltage pulse and by the electrical circuit presented on its way to the sensor.

3.1.2 ASYMMETRIC PD PULSES

Asymmetric PD pulses can also be modeled as the difference of two exponential functions

$$i(t) = I_0 \cdot \left(e^{-\frac{t}{\tau_a}} - e^{-\frac{t}{\tau_b}} \right) \tag{7}$$

being τ_a and τ_b the time constants of the exponential terms. I_0 is associated to the charge Q and to the peak of the current I_p through equations (8) and (9).

$$Q = I_0 \cdot (\tau_a - \tau_b) \tag{8}$$

$$I_0 = I_p \cdot \frac{\tau_a}{\tau_a - \tau_b} \cdot \left(\frac{\tau_a}{\tau_b} \right)^{\frac{\tau_b}{\tau_a - \tau_b}} \tag{9}$$

The time parameters τ_a and τ_b are related with the time to the peak value t_p of the pulse and to the pulse width at the half peak value $\Delta t_{50\%}$.

$$t_p = \frac{\tau_a \cdot \tau_b}{\tau_a - \tau_b} \cdot \ln \left(\frac{\tau_a}{\tau_b} \right) \tag{10}$$

If the condition $\tau_b < 0.2 \cdot \tau_a$ is satisfied, the following simple analytic expression can be established

$$\Delta t_{50\%} \approx t_p + \tau_a \cdot \ln \left(\frac{2 \cdot \tau_a}{\tau_a - \tau_b} \right) \tag{11}$$

The values of I_p, t_p and $\Delta t_{50\%}$ can be determined analyzing the waveform of the recorded pulse. Once these values are obtained the parameters τ_a and τ_b are calculated solving the system of equations formed by (10) and (11). Finally the value of I_0, the modeled pulse $i(t)$ and its charge Q can be determined with equations (9), (7) and (8) respectively.

A verification of the model with the PD pulse can be performed, checking the pulse width at 5% of the peak value ($\Delta t_{5\%}$), which can be determined by equation (12).

$$\Delta t_{5\%} \approx t_p + \tau_a \cdot \ln \left(\frac{20 \cdot \tau_a}{\tau_a - \tau_b} \right) \tag{12}$$

Some figures for I_p, V_p and $\Delta t_{5\%}$, are given in Table 3 for different symmetrical and asymmetrical PD pulse waveforms, considering a reference fix value of charge of 10 pC and assuming a cable characteristic impedance Z_c of 25 Ω for frequencies higher than 1 MHz.

Table 3. Values of I_p, V_p and $\Delta t_{5\%}$ for different waveforms defined by means of t_p and $\Delta t_{50\%}$.

Q=10 pC	t_p (ns)	$\Delta t_{50\%}$ (ns)	I_p (mA)	V_p (mV)	$\Delta t_{5\%}$ (ns)
Symmetrical	0	1	9,4	235	2,1
	0	1,5	6,3	157	3,1
Asymmetrical	0,2	1	7,9	197	3,4
	1	5	1,6	40	17

It is important to note that for a fix charge value the peak values for the current I_p and voltage V_p decrease when the pulse width increases or when it is asymmetrical. Given that the sensitivity of the measurement improves when the signal V_p propagating through the cable is higher, more sensitive measures can be achieved when narrow and symmetrical pulses are generated.

3.2 PD PULSES ACQUIRED BY HFCT SENSORS

In this section mathematical expressions are proposed to model the oscillating transient waveforms measured in cable systems with HFCT sensors. The study of two preliminary mathematical models leads to a generic model that allows the extraction of representative parameters to efficiently classify different types of pulses.

3.2.1 SINUSOIDAL MODEL DAMPED BY A SIMPLE EXPONENTIAL FUNCTION

The expression that describes the behavior of a simple RLC circuit allows damped sinusoidal waveforms to be obtained for the earth current measured by a HFCT (see Figure 3). This simple RLC model obtains an initial reasonable approximation to PD pulses when the measurements are performed in single cables and the PD sources are close to the HFCT sensor. A characteristic PD pulse of this type corresponding to an internal defect in a power cable is shown in Figure 4. The pulse was measured in the cable earth connection with a HFCT sensor and the PD source was close to the sensor.

The circuit elements that justify the existence of the RLC circuit are the cable capacitance C and the equivalent resistance R of the two semiconducting layers seen by the traveling PD pulse during its propagation path, in series with the equivalent inductance L of the cable earth connection where the HFCT sensor is installed. The effect of the HFCT on the cable earth connection must be considered in the equivalent inductance L. The earth current $i_e(t)$ to be measured by the sensor can be modeled by the following mathematical function, where a sinusoidal wave is damped by the exponential function $h_1(t)$.

$$i_e(t) = h_1(t) \cdot sin(\omega \cdot t) \qquad (13)$$

$$h_1(t) = I_{e0} \cdot e^{-a \cdot t} \qquad (14)$$

being ω the angular frequency of the sinusoidal signal, a the damping coefficient of $h_1(t)$ and I_{eo} the value of $h(t)$ for $t=0$.

In an RLC circuit for a given pulse the parameters a and ω are defined by

$$a = \frac{R}{2 \cdot L} \qquad (15)$$

$$\omega = \sqrt{\frac{1}{L \cdot C} - a^2} \qquad (16)$$

with the condition $\frac{1}{L \cdot C} > a^2$

In practice, to determine the modeled function $i_e(t)$ is not necessary to calculate the values of R, L and C. The values of the parameters I_{e0}, a and ω that define $i_e(t)$ can be adjusted in order to obtain the best fitting of this function to the measured pulses. The waveform of the traveling pulse shown in Figure 4 was modeled by the damped sinusoidal waveform $i_e(t)$ fitting the parameters I_{e0}, a and ω. It can be observed that the exponential function $h_1(t)$ characterizes the wave shape of the pulse tail. The dashed line represents the measured pulse in absolute value.

In reality, only this type of pulses representing a short time to peak in the first period and asymmetric forms skewed to the left can be modeled with equation (13). A more generic expression is required in order to characterize the rest of the pulse types shown in Table 2.
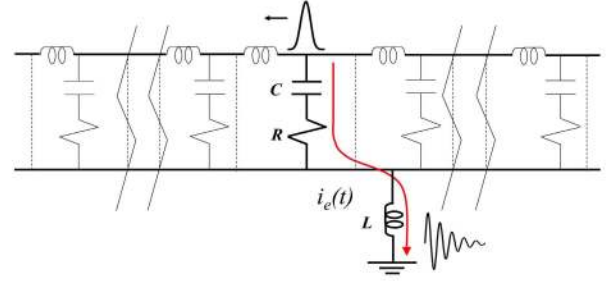

Figure 3. Equivalent RLC circuit to simulate damped sinusoidal waveforms.
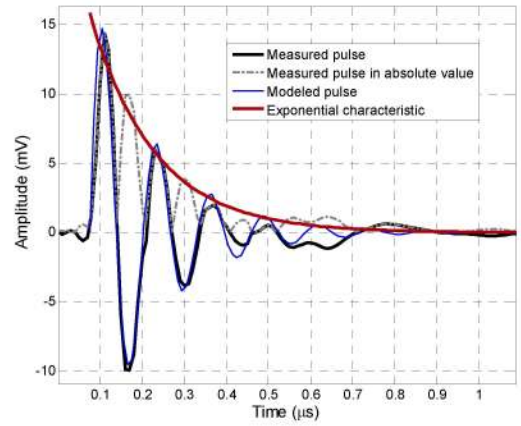
Figure 4. Modeled pulse and exponential characteristic for a PD measured pulse.

3.2.2 SINUSOIDAL MODEL DAMPED BY THE DIFFERENCE OF TWO EXPONENTIAL FUNCTIONS

In cross-bonding sheath cable configurations two RLC transient currents with different time attenuation constants traveling in opposite direction through the HFCT sensor, can justify the following sinusoidal waveform modulated by the difference of two exponential functions.

$$i_e(t) = h_2(t) \cdot sin(\omega \cdot t - \varphi) \qquad (17)$$

where

$$h_2(t) = I_0 \cdot \left(e^{-a \cdot t} - e^{-b \cdot t} \right) \qquad (18)$$

being φ the phase displacement in the instant $t=0$.

The parameter I_0 is related with the peak value I_p of $h_2(t)$ by the following expression.

$$I_0 = I_p \cdot \left(\frac{b}{(b-a)} \cdot \left(\frac{b}{a} \right)^{\frac{a}{(b-a)}} \right) \qquad (19)$$

In this case, a good model for oscillating pulses that do not reach the maximum value in the first peak can be achieved. Furthermore, $h_2(t)$ represents an enveloping function that characterizes the wave shape of the pulses.

Figure 5 shows an example of two pulses modeled by equation (17) that were measured with HFCT sensors. The first pulse (Figure 5a) corresponds to the one shown in Table 2 measured in cable system 2 close to the source for the internal defect. The second pulse (Figure 5b) was measured in the experimental setup presented in subsection 5.2 for the surface defect described in subsection 5.1.

Although a good fitting is achieved for pulses presenting one or more oscillations before the peak of the signal (see Figure 5), when the pulses become more symmetric or even asymmetric skewed to the right, it is not possible to obtain a suitable envelope applying the mathematical function $h_2(t)$ and a more general mathematical model is required. An example of these last types of pulse is shown in Table 2 for the cable system 2; the signal measured at 700 m presents an asymmetric skewed to the right waveform while the signal measured at 1700 m is more symmetric.
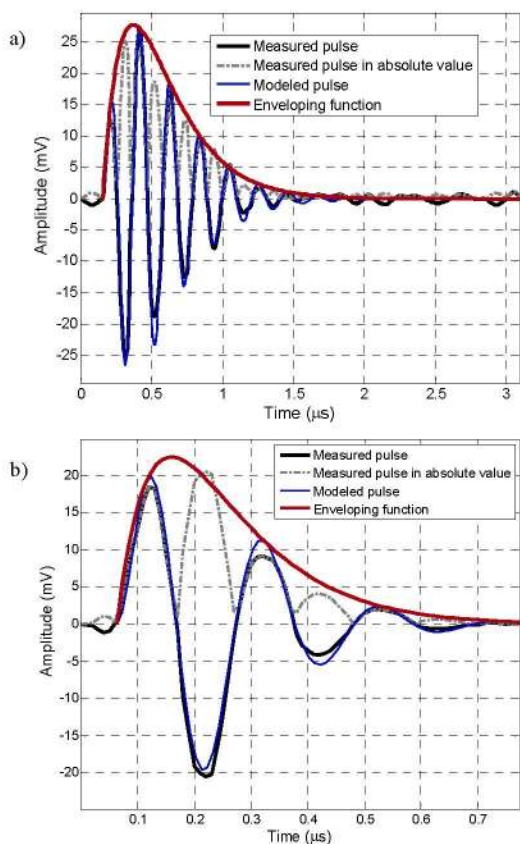

Figure 5. PD pulses modeled on the basis of a sinusoidal waveform damped by the difference of two exponential functions. a) Oscillating pulse with the maximum value in the second oscillation. b) Oscillating pulse with the maximum value in the second peak.

3.2.3 GENERALIZATION

A generic model can be achieved with a mathematical function $i_e(t)$ composed of a sinusoidal function $g(t)$ modulated with the enveloping function $h_3(t)$.

$$i_e(t) = g(t) \cdot h_3(t) \qquad (20)$$

$$g(t) = sin(\omega \cdot t - \varphi) \qquad (21)$$

$$h_3(t) = \frac{I_0}{e^{\alpha \cdot (t - t_0)} + e^{-\beta \cdot (t - t_0)}} \qquad (22)$$

The sinusoidal function $g(t)$ associated with each measured pulse is defined by the frequency $(f = \omega/2 \cdot \pi)$ and its phase displacement φ in the instant $t=0$.

The enveloping function $h_3(t)$ is named asymmetric hyperbolic secant (AHS), due to its resemblance with the hyperbolic secant when the damping coefficients α and β of the exponential terms are equal. The parameter t_0 represents the temporal displacement of the AHS at the instant $t=0$ and I_0 characterizes its amplitude. The value of I_0 is related with the peak value I_p of the AHS by the equation (23). By adjusting the values of α, β, t_0 and I_0 it is possible to obtain a suitable envelope to modulate the sinusoidal waveform; the setting procedure of these parameters is described in next section.

$$I_p = \frac{I_0 \cdot \beta}{(\alpha + \beta)} \cdot \left(\frac{\alpha}{\beta}\right)^{\frac{\alpha}{(\alpha + \beta)}} \qquad (23)$$

The time to peak value is defined by the following expression.

$$t_p = \frac{ln(\beta / \alpha)}{(\alpha + \beta)} \qquad (24)$$

As shown in Figure 6, by applying the AHS function more flexibility is achieved for obtaining both symmetrical envelopes ($\alpha = \beta$), or asymmetrical skewed to the left or to the right ($\alpha \neq \beta$) envelopes. Therefore, it is possible to characterize all types of pulses according to their waveform, evaluating the values obtained for α and β.
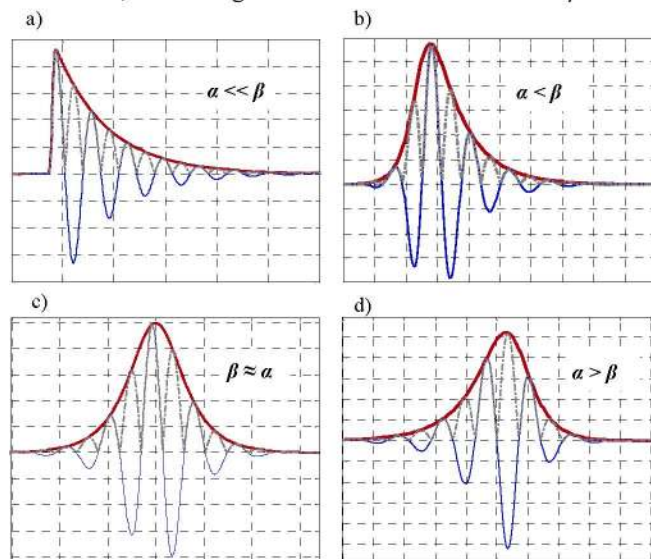
Figure 6. Enveloping functions for different types of pulses. a) Asymmetric pulse skewed to the left with the maximum value in the first peak. b) Asymmetric pulse with an oscillation before the maximum value. c) Symmetric pulse. d) Asymmetric pulse skewed to the right.

4 PD CLUSTERING IMPLEMENTATION

Despite the stochastic behavior of PD phenomena and noise signals, by analyzing the waveforms of the pulses corresponding to different defects and noise sources, certain similarities in the frequency spectra and in the nature of the envelope for those signals belonging to the same source have been detected. In the processing method implemented, the

damped oscillating transient wave $i_e(t)$ defined by equation (20) is associated with each generic recorded pulse.

The frequency ($f=\omega/2\cdot\pi$) and the phase displacement φ parameters that define the sine function $g(t)$ in equation (21), are selected to obtain a good fitting in the zero crossings of the captured signals. The frequency is determined by analyzing the frequency spectrum of the sampled signals applying the Fourier transform $F(f)$ and by means of equation (25). For each pulse, f corresponds to the arithmetic mean of the values in the interval where the energy of the spectrum is more significant. This interval is defined by the frequency width at the 70% of the maximum value in amplitude of the spectrum, see Figure 7.

$$f = \frac{\int_{f_1}^{f_2} f \cdot F(f) \cdot df}{\int_{f_1}^{f_2} F(f) \cdot df} \tag{25}$$

The enveloping function $h_3(t)$ of equation 22 is fitted to the absolute values of the local maxima of each pulse, characterizing its wave shape. The values for the parameters I_0, α, β and t_0 that enable the best fitting are determined by applying the least squares approach.

In Figures 8a, 8b and 8c the enveloping functions $h_3(t)$ calculated for measured pulses showing different wave shapes are presented. In Figure 8a an asymmetrical skewed to the left pulse is shown, this pulse was measured in the experimental setup presented in subsection 5.2 for the internal defect described in subsection 5.1. The wave shape is characteristic of defects measured in a single phase test

of a cable system; see the signals obtained for cable system 1 in Table 2. In Figure 8b an asymmetrical skewed to the left pulse is shown. This pulse was measured at 700 m from the PD source, in the on-line test performed for the three phase 45 kV distribution underground line of section 2; the pulse corresponds to the one shown in Table 2 (cable system 2, distance PD source - sensor 700 m). Finally, in Figure 8c a pulse with a nearly symmetric waveform is characterized. This pulse corresponds to the one shown in Table 2 measured at 1700 meters from the PD source in cable system 2. In the three cases, the application of the AHS function has resulted effective to characterize the wave shapes of the measured pulses.
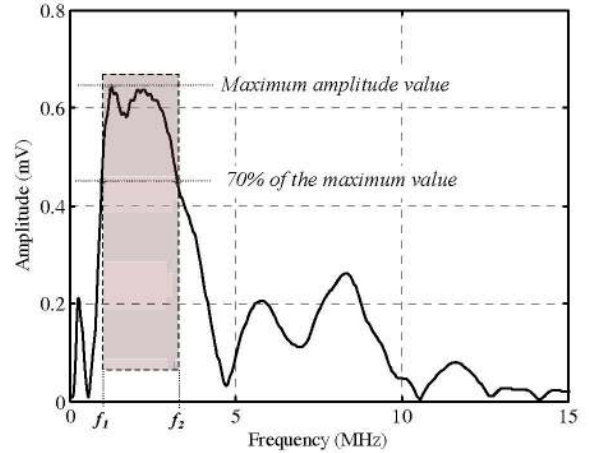
Figure 7. The frequency spectrum of a PD pulse and the frequency interval considered for the estimation of the parameter f.
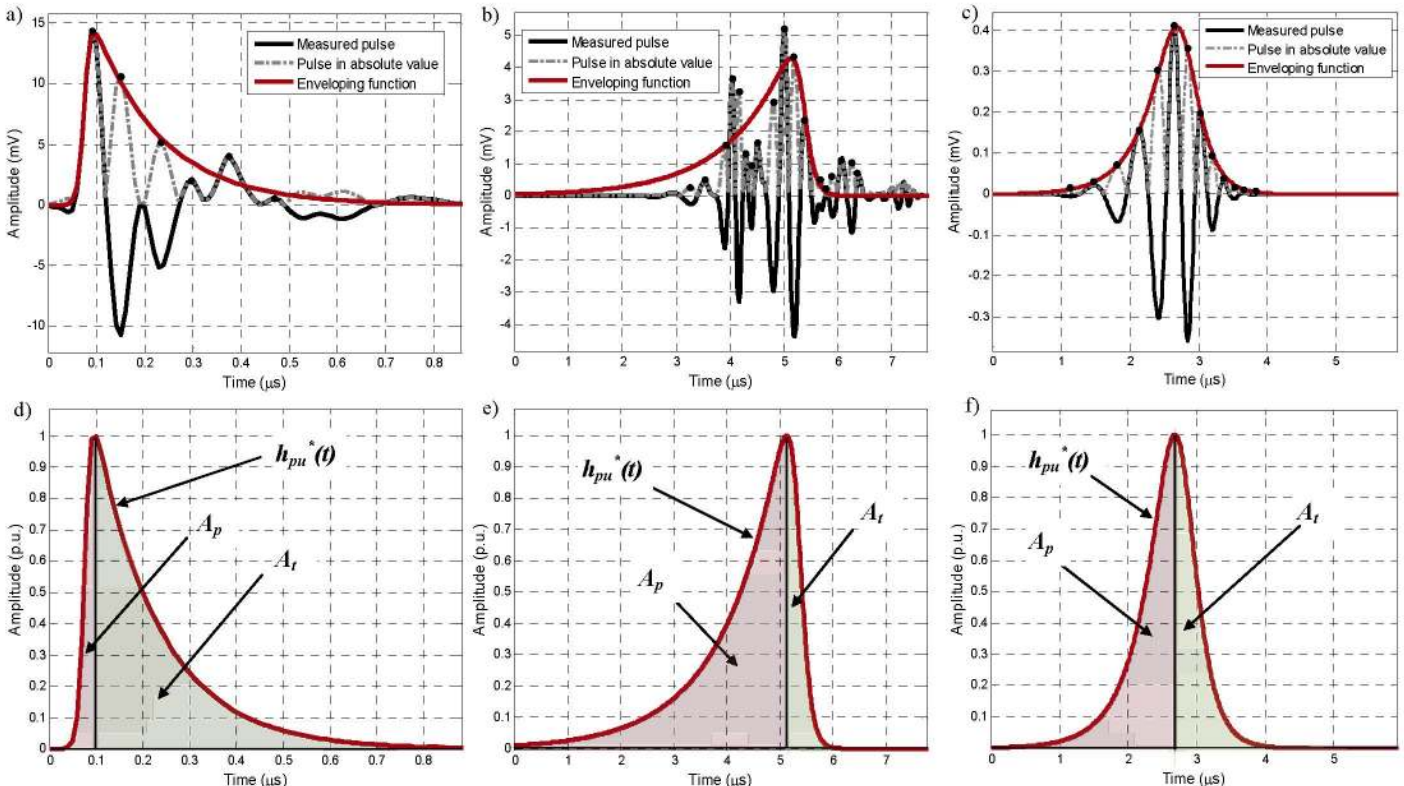
Figure 8. Enveloping functions calculated for different shapes of measured PD pulses. a) Pulse type 1: asymmetric skewed to the left. b) Pulse type 2: asymmetric skewed to the right. c) Pulse type 3: symmetric. d) Normalized envelope for pulse type 1. e) Normalized envelope for pulse type 2. f) Normalized envelope for pulse type 3.

The shape of the enveloping function is mainly described by the parameters β and α, which are related with the steepness in the increasing and decreasing intervals, respectively. Therefore, the principal parameters of the modeled pulses that characterize their waveform are f, α and β. These parameters are calculated automatically for each pulse-shaped signal measured; i.e. no previous settings in the measuring instrument are required to obtain them.

For asymmetrical PD signals skewed to the left, the parameter α defines the wave tail and both parameters α and β define the front of the pulse. Taking into account that the first samples acquired by the digital recorder, corresponding to the pulse front, can be very changing due to the frequency content dependency and to the sampling rate of the digitizer, in some occasions the parameter β can present a large dispersion especially for pulses with fast rise time. Although α and β provide useful information for PD clustering, by analyzing two new parameters related with the front and tail areas of the normalized enveloping function $h_{pu}^{*}(t)$ (26), lower dispersions are obtained and an improvement in the classification process is achieved.

$$h_{pu}^{*}(t) = \frac{h_3(t)}{I_p} \qquad (26)$$

These two new parameters, called peak area A_p and tail area A_t, correspond to the areas enclosed by the function $h_{pu}^{*}(t)$ from its beginning to the maximum amplitude value and from this point to its end, respectively. Both area parameters are independent from each other and less dispersion is obtained in their values in comparison to the dispersion of parameters α and β.

The normalized functions $h_{pu}^{*}(t)$ corresponding to the pulses of Figures 8a, 8b and 8c are shown in Figures 8d, 8e and 8f, the peak and tail areas corresponding to A_p and A_t parameters respectively are indicated by the shaded areas.

Different clusters associated to the pulse sources present in a PD test can be distinguished through the representation of the waveform parameters f, A_p and A_t in two dimensional (2D) and three dimensional (3D) diagrams. The PD sources can be identified by employing an interactive procedure, selecting each cluster and analyzing the associated PRPD pattern.

5 LABORATORY EXPERIENCE

In order to analyze the sensitivity of the representative parameters (f, A_p and A_t) to discriminate different sources of pulses, three types of PD sources were generated (a cavity in XLPE, an external surface insulation defect and corona in air) in the presence of pulse-shaped noise interferences.

Internal discharges (voids inside solid dielectrics) are considered the most harmful for the dielectric elements and lead to insulation failures. Surface discharges in air, although considered less dangerous (the inner part of the insulation does not degrade), can also lead to failure. Corona discharges pose no risk if they are generated in air, however, if they are produced close to silicone insulators or in SF_6, the consequences can also lead to insulation faults.

The presence of electric noise disturbances when PD tests are performed represents a drawback due to the loss of sensitivity, especially when PD pulses with low energy and amplitude are measured [8, 10]. The electric noise consists of a set of unwanted signals generated by electric and electronic devices and by radiofrequency emissions that are overlaid with the signals intended to process [21, 22]. In some cases some noise signals may have an impulsional behavior similar to that of PD pulses, common sources of these pulse-type signals are power electronic devices [8]. The set of the rest noise signals presents a random behavior and constitutes a continuous wave disturbance with a characteristic frequency content.

The measuring device used in the PD tests includes a noise filtering tool based on the wavelet transform. This filter allows pulse-shaped signals to be discriminated from continuous background noise even in high disturbance environments [20, 23, 24].

5.1 GENERATION OF PULSE-SHAPED SIGNALS

Internal discharges. A cavity type defect was caused in a cable termination by a transversal cut in the main insulation, at the end of the cable semiconducting external layer, as shown in Figure 9a. This type of defect can be caused due to an incorrect handling of the splicing tools at the edges of semiconducting shield cut-backs, in accessories assembling processes [25].
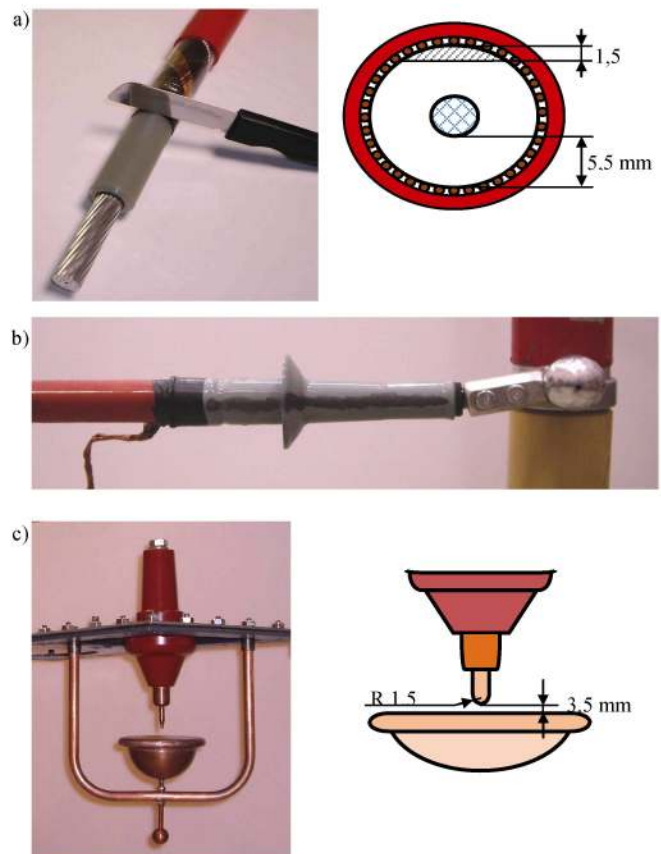

Figure 9. PD sources. a) Cavity in a cable termination, b) surface insulation defect and c) point-plane corona effect.

Surface discharges. External surface discharges can appear in outdoor insulation elements when they are dirty and/or wet. This type of defect was produced by the artificial contamination created with carbonization traces on the external surface of a cable termination, as shown in Figure 9b.

Corona in air. Corona discharges were caused by means of a point-plane configuration assembled in a holding device that allows the adjustment of the gap between the point and the plane electrodes (see Figure 9c).

Pulse-shaped noise interferences. Three different types of impulsional noise were present during the test. A noise source was generated by the IGBTs of the resonant system used to apply high voltage; these interferences are synchronized with the test voltage signal. Another electronic pulse-type noise with repetitive behavior was produced by the switching of the transistors of an electronic power device connected to the mains of the laboratory. Finally, a significant amount of random pulse-shaped signals conducted by the earth cable network were measured.

5.2 EXPERIMENTAL SETUP

The experimental arrangement implemented in a high voltage laboratory consists on a 12/20 kV XLPE cable system of 1500 meters length and a 240 mm² aluminum conductor. The three PD sources were placed at the beginning of the cable as shown in Figure 10.

A resonant generator of 36 kV with a 22.9 H reactor was used to apply a test voltage of 50 Hz and 15 kV. This voltage level is above the inception voltage of each defect and provides stable measurements. The synchronization reference signal was acquired by a capacitive divider.

The PD measuring device is equipped with an acquisition board of 14 bits of vertical resolution, a bandwidth of 50 MHz and a sampling frequency of 100 MS/s. A clamp-type HFCT sensor with a bandwidth from 100 kHz to 20 MHz was coupled in the earth connection of the cable end connected to the high voltage generator. In this frequency range (0,1-20 MHz), the classification procedure presented is able to discriminate successfully pulses from different sources measured in the cable sheath of power cables.
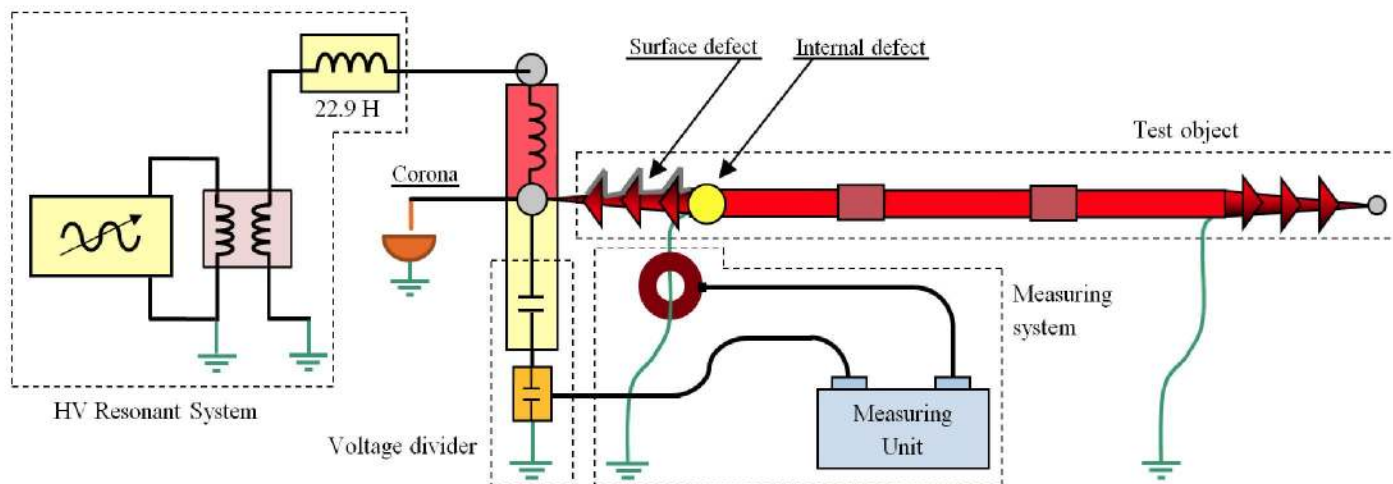

Figure 10. Experimental setup. High voltage resonant generator, cable system under test with three insulation defects and measuring system composed of a HFCT sensor and a measuring instrument with signal processing tools.

5.3 HIGH VOLTAGE PD MEASUREMENTS

The insulation defects described in section 5.1 were measured in the presence of the noise interferences with the aim to characterize the clustering tool. A significant amount of pulses were processed in order to obtain representative clusters and PRPD patterns for the individual sources.

The PRPD pattern for the complete acquisition using a conventional filtering by means of a band-pass filter plus a threshold level setting of 1 mV is shown in Figure 11a. The same data processed using the wavelet filter are shown in Figure 11b, in this case only the pulse-shaped signals are represented as they are filtered from continuous background noise. In the presence of high levels of background noise, without the wavelet filter any PRPD pattern may be hidden. The area below the noise level in Figure 11a becomes a dead zone, because the large number

of accumulated PD and mainly random noise signals makes it difficult the identification of patterns. In many occasions, when the classification tool is applied selecting the signals with a low threshold level, the large number of accumulated random noise signals in comparison with the least number of PD pulses does not allow a proper identification of clusters. Improved results can be obtained by setting the threshold in a higher level, at the expense of a loss of sensitivity in the detection of PD pulses with amplitude values below the adjusted level.

In the case studied, the background noise could be filtered adjusting the threshold level slightly above the noise level. In this case, only the pulses overlaid in the continuous noise signal with amplitude values exceeding the adjusted threshold level are recorded for post-processing, the rest are not detected, being necessary a

longer measuring time to obtain clusters with a significant density. The proposed clustering method is able to provide an effective separation for all the pulses detected removing the background noise with the setting of the threshold level slightly above the background noise level. However, it is recommended to apply the clustering method to the pulses detected with the wavelet filter, thus the processed pulses are not influenced by the background noise and in general a lower number of periods of the voltage signal are required to obtain representative results. Furthermore, with the wavelet filter the pulses are selected without the requirement to adjust any threshold level.

Although a better pattern recognition can be performed when the wavelet filter is applied (see figure 11b), it is still not possible to differentiate the sources involved due to the superposition of the pulses generated.
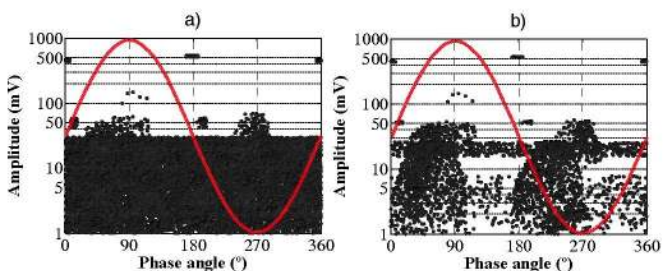

Figure 11. a) PRPD pattern using a conventional filter. b) PRPD pattern applying the wavelet filter.

For further analysis, the parameters of the pulse classification procedure described in section 4 f, α, β, A_p and A_t are calculated for the pulses detected with the wavelet filter. These pulses are represented in the 2D diagrams (β-α and A_p-A_t) and in the 3D diagrams (f-β-α and f-A_p-A_t) shown in Figure 12. The 3D diagrams can be rotated manually in the software application in order to discern visually the best separation between clusters.
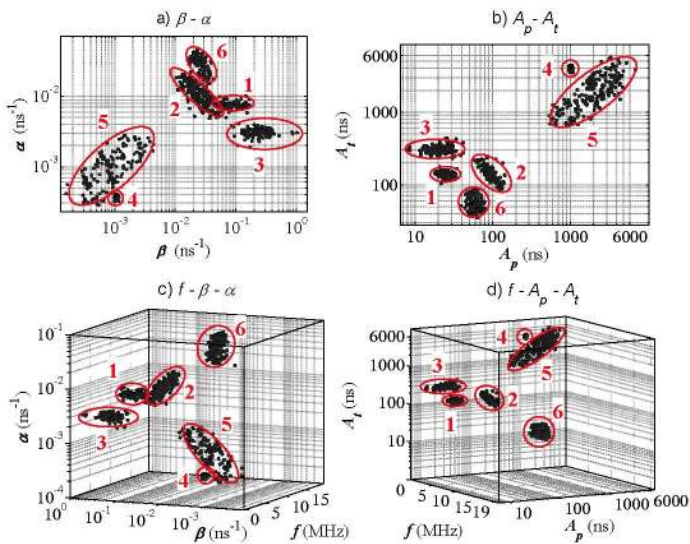

Figure 12. Comparison of the parameters α-β with A_p-A_t. a) 2D diagram for β and α. b) 2D diagram for A_p and A_t. c) 3D diagram for f, β and α. d) 3D diagram for f, A_p and A_t.

The values of the normalized standard deviation of the parameters α, β, A_p and A_t calculated for each cluster are shown in Table 4. The parameters A_p and A_t present values of normalized standard deviation lower than the parameters β and α respectively, $\sigma_n(A_p) < \sigma_n(\beta)$ and $\sigma_n(A_t) < \sigma_n(\alpha)$. Using the parameters A_p and A_t more concentrated and separated clusters are obtained (see Figures 12a and 12b) and the clustering method results more effective.

Table 4. Normalized standard deviation calculated for the classification parameters A_p, β, A_t and α.

Pulse source	Normalized standard deviation			
	$\sigma_n(A_p)$	$\sigma_n(\beta)$	$\sigma_n(A_t)$	$\sigma_n(\alpha)$
1 Internal defect	0,1614	0,3502	0,0683	0,0788
2 Surface defect	0,1782	0,4014	0,2264	0,3363
3 Corona effect	0,3070	0,5267	0,1145	0,1163
4 IGBT pulses	0,0320	0,0357	0,0422	0,0461
5 Random pulse-shaped noise signals	0,5247	0,5829	0,4028	0,4861
6 Pulses from power electronic	0,1149	0,2104	0,2272	0,2778

The 3D diagram (f-A_p-A_t) and the 2D diagrams (A_p-A_t, f-A_p and f-A_t) shown in Figure 13, make it possible to clearly visualize the clusters formed. Figures 13d and 13f show a high resolution enlargement of the area marked in red in Figures 13c and 13e in order to better visualize the clusters which have been formed.
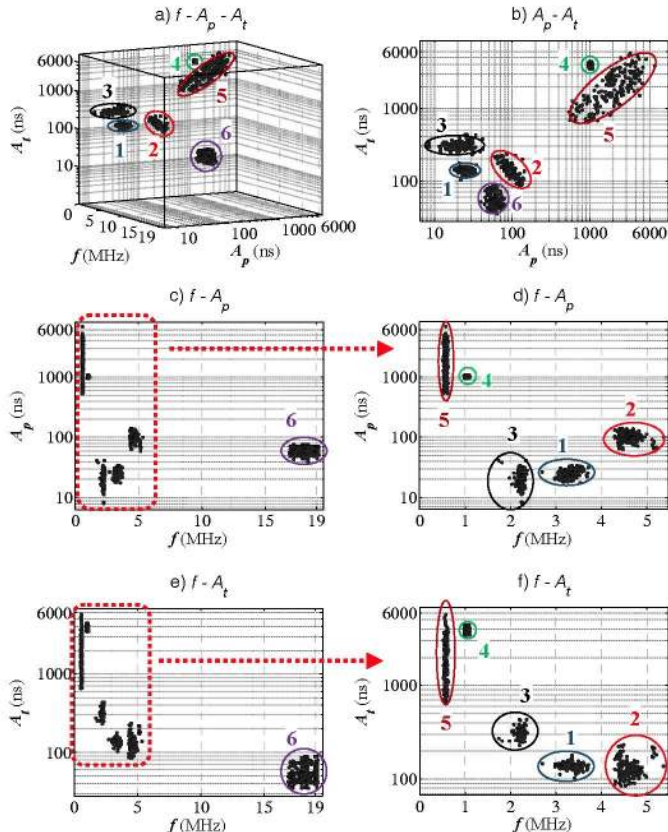

Figure 13. 3D and 2D diagrams for the parameters f, A_p and A_t. a) 3D (f-A_p-A_t), b) 2D (A_p-A_t), c) 2D (f-A_p), d) detailed view of the diagram f-A_p, e) 2D (f-A_t) and f) detailed view of the diagram f-A_t.

In the four diagrams, the six groups of pulses indicated with the oval shapes have been differentiated. With the software implemented any arbitrary volume or area of pulses can be selected and its corresponding PRPD pattern is visualized. For each cluster, its individual PRDP pattern has been obtained (see Figure 15). The six differentiated PRPD patterns enable the identification of the six pulse-shaped signal sources involved in the test. Figure 14 shows a representative pulse of each PD and noise source together with the calculated enveloping function. The values of the classification parameters obtained for these pulses are indicated in Table 5. An accurate fitting is achieved when the pulses present a predominant oscillation and attenuation (see the enveloping functions of Figures 14a, b, c, e and f). However, in some cases it results difficult to get an accurate fitting, see the enveloping functions of Figures 8b and 14d. Despite this, as this type of pulses are originated in the same source their waveform is very similar, and consequently, the enveloping functions result very repetitive; this enables to obtain a differentiated cluster when the pulse clustering process is applied.
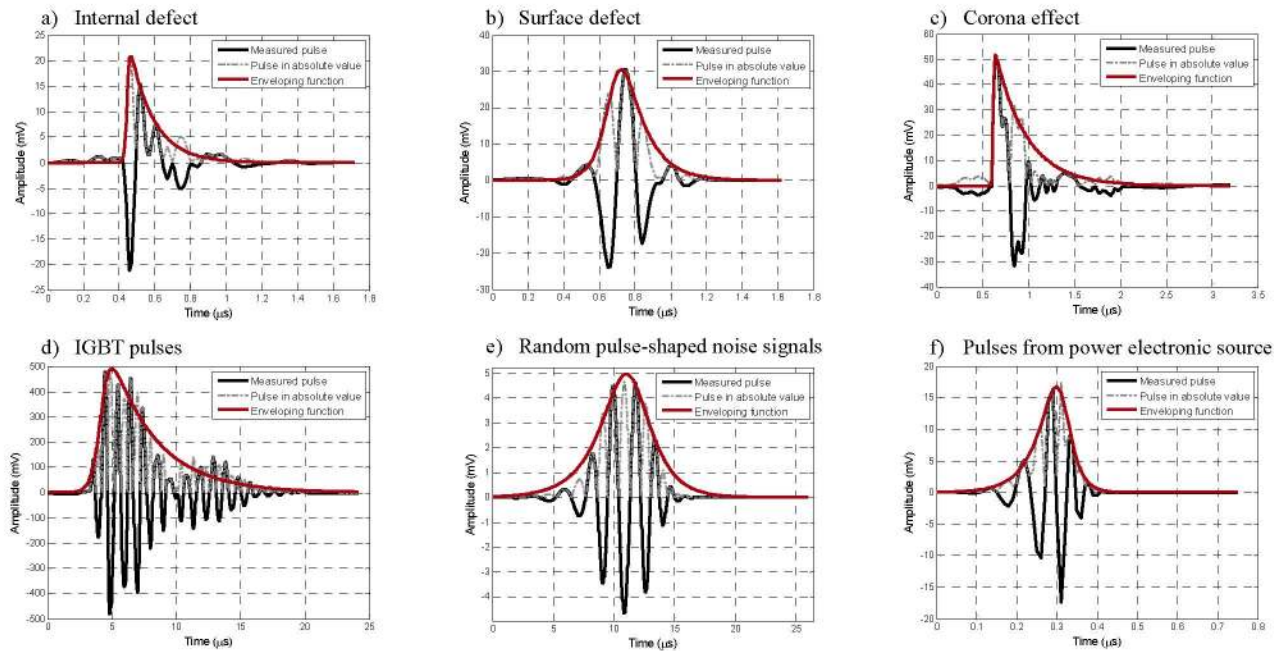

Figure 14. Waveform of representative pulses for each PD and noise source and enveloping functions.

Table 5. Value of the classification parameters f, α, β, A_p and A_t for each characteristic pulse.

Pulse source	f (MHz)	α (ns^{-1})	β (ns^{-1})	A_p (ns)	A_t (ns)
1 Internal defect	3,27	7,390E-03	1,265E-01	28,34	138,31
2 Surface defect	4,41	8,181E-03	2,428E-02	108,4	147,21
3 Corona effect	2,15	2,912E-03	1,976E-01	22,4	340,09
4 IGBT pulses	1,03	3,415E-04	1,057E-03	1087	3959
5 Random pulse-shaped noise signals	0,56	6,218E-04	5,010E-04	2959,2	2650,6
6 Pulses from power electronic	17,7	4,342E-02	2,107E-02	65,01	45,07

By analyzing in detail the characteristics of each PRPD pattern (see Figure 15) it can be stated that:

a) A characteristic pattern of internal discharges is obtained for cluster 1. PD occur slightly before the zero crossings and in the increasing intervals of the test voltage. There is a certain symmetry and a similar repetition rate when comparing the patterns of both half-periods.

b) The pattern displayed for cluster 2 corresponds with a surface defect. PD occur in the increasing intervals and in the peak level of the test voltage. There is an asymmetry when comparing the patterns of both half-cycles.

c) A pattern of corona discharges (point-plane configuration) is obtained for cluster 3. PD appear in the phase position where the test voltage is maximum. There are stable amplitude values within a certain range in both half-periods. There are less pulses and with a higher amplitude in the positive half-cycle.

d) The pattern corresponding to cluster 4 is characteristic of the pulses generated by the IGBTs of the resonant system. The pulses are synchronized with the test voltage and positioned close to the zero crossings.

e) The pattern obtained for cluster 5 corresponds with the random pulse-shaped signals conducted by the earth cable network. Pulses with different amplitude values that are not synchronized with the test voltage. This type of pulses usually forms a dispersed cluster in the 2D and 3D diagrams due to their random behavior.

f) Finally, the pattern of cluster 6 can be associated with the noise source generated by the power electronic device.

Pulses of similar amplitude values within a defined range and not synchronized with the test voltage.
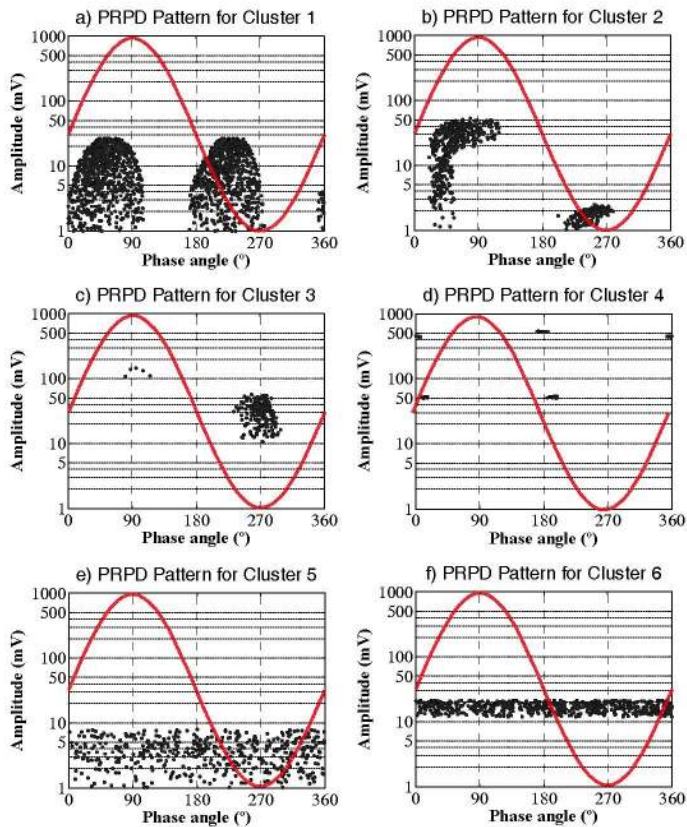

Figure 15. PRPD patterns which correspond with the clusters obtained by means of the classification tool.

With the pulse classification method proposed, the insulation defects and the noise sources have been successfully identified. Furthermore, the analysis of the individual patterns has enabled making an accurate diagnosis of the cable system.

6 CONCLUSIONS

For on-line PD measurements performed in the simultaneous presence of PD and noise sources, an effective pulse classification tool is essential for achieving a proper sensitivity and a correct diagnosis.

In order to identify different PD and noise sources in cable systems, a clustering tool based on the waveform analysis of the captured signals has been developed by implementing a mathematical model. The procedure for obtaining the classification parameters is fully automatic, i.e. no previous settings in the measuring instrument are required.

Pulse source identification has been verified with successful results measuring common PD sources (corona, surface and internal discharges) in the presence of several noise disturbances.

With the classification of the measured signals, misjudgements in fault detection and in the identification of PD sources can be avoided and a more reliable diagnosis of

the failure mechanisms involved in HV cable systems can be achieved.