Feature extraction and ageing state recognition using partial discharges in cables under HVDC

Nathalie Morette, Thierry Ditchi, Yacine Oussar

Feature extraction and ageing state recognition using partial discharges in cables under HVDC

Nathalie Morette[a], Thierry Ditchi[a], Yacine Oussar[a,*]

[a] *Laboratoire de Physique et d'Étude des Matériaux (LPEM), ESPCI Paris, PSL Research University, CNRS, Sorbonne Université, 10 rue Vauquelin, 75005 Paris, France*

Abstract

PD detection is an effective way to evaluate the degradation state of cable insulation. The extraction and selection of relevant features from PD raw data have been mostly investigated to recognize the types of insulation defects in HV equipment. In this study, two different feature extraction methods combined with supervised classification techniques are implemented for ageing state recognition of a polyethylene-insulated cable under HVDC conditions. For this purpose, an original experimental setup is implemented. Experiments are performed on a long length 100 m coaxial cable subjected to high electric fields. PD events are detected by direct coupling and collected with a digitizing oscilloscope. Feature extraction based on PD pulse shape parameters represented in time domain as well as wavelet decomposition coefficients are used separately as input variables of Support Vector Machines classifiers (SVMs). A feature selection method is implemented to design optimized SVM classifiers that attribute an ageing state to the cable insulation. The classification performance achieved with both feature extraction methods are presented and compared. The results show satisfactory recognition rates of two ageing states of cable insulation, up to 100% with a small subset of variables, particularly when features are extracted from wavelet decomposition of PD experimental data.

Keywords: insulator, partial discharges, HVDC, polymeric-insulated cable, ageing state, feature extraction, feature selection, wavelet decomposition, supervised classification, support vector machines

1. Introduction

Partial Discharge (PD) phenomenon within the dielectric of High-Voltage (HV) power cables causes serious insulation damage and has a significant impact on the cable lifetime. Thus, it is of great interest to get an indication of the degradation state of the cable insulation before complete breakdown occurs. While PD measurements and analysis is a widely used method for commissioning tests and diagnosis of AC XLPE cable system, the DC voltage case is of increasing importance for electrical energy supply applications [1, 2]. In fact, in the perspective of a transition to renewable energies, the electrical energy will be transported in DC voltage. HVDC extruded cables are being installed in power grids and HVDC electrical systems require now efficient methods for PD detection and analysis at DC voltage [3]. However, since there exist few installed HVDC links, there is a limited experience regarding how cables age electrically under HVDC and a lack of data about their ageing and reliability [4]. In addition, PD measurements are not recommended in the latest related CIGRE Technical Brochure for DC Extruded Cable Systems [5] for commissioning tests. This is mostly because PD detection is complicated under DC stress [6]. As a consequence, there is little past knowledge on the severity of PD under DC stress and few HVDC links to build up a knowledge base. The relation between PD properties and the degradation of the cable insulation under DC is difficult to establish. Despite the CIGRE recommendations about PD under DC, few recent studies are running laboratory tests to build up a knowledge base for HVDC PD measurements [7, 8]. They focus on better understanding the PD process and the behavior of different defects under DC stress as well as their evolution during

*Principal corresponding author *
Email address: Yacine.Oussar@espci.fr (Yacine Oussar)

the service lifetime. For example, the PDIV (Partial Discharge Inception Voltage) and magnitude of PD under DC are evaluated and compared to PD under AC. However, these studies do not make use of PD data to provide an estimation of the cable degradation state. In fact, in the current state of knowledge, there is still no established methodologies to estimate the ageing state of HVDC cables insulation using PD measurements under DC voltage [6].

The studies conducted by CIGRE and IEEE Study Committees present phenomenological (or empirical) and physical models for life and reliability estimation of HVDC cables with XLPE insulation [5, 9]. The models give an estimation of the cable life (or "time to failure") as a function of the applied stress level. They are used as general models to predict the time to failure of any cable, knowing its design life, rated voltage and type (HVAC or HVDC extruded cable). The phenomenological life models proposed in these studies are based on Accelerated Life Tests (ALTs) results. Model parameters are estimated experimentally by ALTs in laboratory where cables are submitted to higher than design stresses during short duration (shorter than the cable design life). Parameters are obtained by applying best-fitting techniques to the life versus stress data [9]. Another approach proposed by IEEE committees are microscopic or physical life models that provide a description of the ageing mechanism at the microscopic level. These models describe the chemical and physical mechanisms that lead to localized degradation of cable insulation. One of the main purpose of life models proposed by [5, 9] is to estimate the cable life in service conditions. However, possible changes in ageing mechanisms during the service life make this goal difficult to achieve since a model valid for condition of ALTs might not be valid at service stresses [9]. In fact, phenomenological life models are capable of predicting life only within the limited stress range of the ALTs, making the extrapolation to service stresses more uncertain. For example, these life models assume that applied stresses are controlled and do not vary during the service lifetime of the cable, as it occurs during ALTs, whereas service stresses are never constant [9]. Moreover, phenomenological life models act as general models and are not specific to a particular cable given its own defects and the different stresses it has endured during its service lifetime. The ageing mechanisms that appear and evolve during service life are not taken into account. Regarding physical life models, they involve too much complexity because of their large range of parameters, which make them difficult to correctly fit a data set. These models are not practical for testing purpose but are used to provide information at the R&D stage to compare different compounds for improving the insulation. All these reasons make life models proposed by [5, 9] limited in predicting life of HVDC cables in service conditions.

Most of the existing studies about PD measurements and ageing state recognition are conducted under AC [10, 11, 12]. In [10], the phase-angle and the magnitude of each discharge are recorded and are used to make statistical distributions, such as maximum discharge magnitude, average discharge magnitude, and number of discharges as a function of phase angle. The shapes of these distributions are characteristic for the ageing state of cable insulation. However, this statistical analysis requires the phase-resolved pattern of PD pulses (PRPD) that is not available in the DC case. In [13], the evolution of the PD mechanism inside voids during ageing is investigated using time-resolved discharge parameters. However, artificial cavities are created to simulate ageing. In general, database for PD analysis are built using PD measurements from artificial defects created in the cable insulation [3, 6, 8]. For example, in [8] PD under AC and DC voltages are performed and compared for artificial defects created in insulation based on phase-resolved and time-resolved PD patterns, respectively. Moreover, while damage during cable installation are considered as one of the main cause of cables failure, natural ageing of the cable under HVDC is still not well understood. For this reason, monitoring techniques enabling the collection of PD data over a long period of the cable service lifetime have the potential to play a useful role in the diagnostic of HVDC cable insulation [4].

One of the most challenging issues in classifying PD patterns according to the ageing state of the cable insulator is to extract informative features from PD measurements. Most of the existing researches on PD feature extraction is applied to PD pattern recognition for defect models classification in HV equipment [14, 15, 16, 17, 18, 19, 20, 21] or PD-noise discrimination [3]. Therefore, feature extraction and selection techniques are not sufficiently investigated for ageing state recognition. In [22], distributions and density functions of the amplitude and phase parameters are used to describe and discriminate PD patterns from various types of defects under AC. Under DC, the main observed parameters are repetition rates and magnitude of PD. In [4, 8, 21, 22], distributions of the maximum and mean of PD magnitude as a function of time interval between PD are used as well as density functions of discharge magnitude and time intervals between PD. Features are extracted from the shape of these distributions that are characteristic for the type of defect, which generates the discharges. However, the magnitude of PD pulses, which

is one of the two main parameters observed in PD analysis under DC, may be attenuated during the propagation of the PD pulse along the cable length. Thus, it would be preferable to consider a larger set of features to describe PD under DC. Standard for PD testing of HVDC equipment are still in development but the main aspect is the use of ramp type voltage or polarity reversal in order to emphasize PD activity under DC voltage [3, 4, 8].

Whereas PD occurring at polarity reversal have been studied extensively [8], little has been done to characterize PD occurring under constant polarity DC stress. In [8] for example, DC voltage steps are performed every 4 minutes. However, these techniques are only applicable in laboratory during off-line tests. The use of ramped voltage every 4 minutes differs from the normal operating conditions and PD that would not be active in service conditions could be activated during off-line ramped tests. Thus, the results of these studies have to be scrutinized.

In the present study, a methodology for ageing state recognition of PE-insulated cables under DC voltage is proposed, based on PD measurements performed at different moments of the cable's service life. For this purpose, a continuous monitoring system is developed in order to acquire typical PD signals occurring under DC voltage. No artificial defects are created in order to simulate ageing. Experimental conditions reproduce the natural ageing of a HVDC cable in service. No polarity reversal or ramping voltages are applied to enhance the PD activity. A low voltage cable is used to ensure the detection of PD in the insulation. In this way, only PD due to natural ageing or to already existing defects in cable insulation are collected. A set of PD measurements are acquired from "virgin" and "aged" states of the insulator and are used to assess the quality of the cable insulation. PD fault localization within the cable is performed based on the measured propagation times of PD pulses along the cable. It ensures the selection of internal PD that give more meaningful information about the cable insulation condition. This database is then used to implement classification techniques for ageing state recognition. As opposed to the general life models proposed in [5, 9], the estimation of the degradation state is specific to the cable under test, given its own defects and the different stresses and ageing mechanisms it has endured during its service lifetime. This methodology can be applied to any cable types. A key objective of this approach is the ability to correctly estimate the condition of the cable insulation in order to anticipate failure of the system or to avoid unnecessary intervention. In fact, in practice, one usually does not need to predict the lifetime very accurate far in the future. Often, the maintenance team only needs to know if the cable system will fail "soon".

Whereas most of the works regarding PD analysis under DC use the conventional statistical distributions of amplitude and time interval between PD for feature extraction, we propose to extract two different feature sets from the PD signals collected. Because the electrical ageing of the insulation induces modification of the PD physical process [23, 24, 25, 26], various PD pulse pattern characteristics evolve during ageing as well as the occurrence of PD events. Thus, a first feature set is built using physical variables such as pulse magnitude, rise-time, fall-time and time interval between PDs. Then, a second feature set is established using Discrete Wavelet Transform (DWT) of PD measurements. With regard to the PD pulse structure, non-periodic and fast transient features always exist in the PD signals detected, which cannot be revealed explicitly by conventional transform. For these reasons, wavelet analysis can reveal reliable information contained in PD signals that are useful for cable ageing state recognition. Thus, experimental PD signals are decomposed according to a wavelet transform and features are extracted from the first two moments, i.e. mean and standard deviation, of the wavelet coefficients distribution at each level of decomposition. Pulse shape features as well as features extracted from wavelet decomposition are then ranked using the Gram-Schmidt orthogonalization procedure [27] therefore selected according to a "wrapper approach" [28]. The recognition of the two ageing states of cable insulation is performed using Support Vector Machines (SVMs) [29]. The application of feature extraction, ranking, selection and classification techniques to ageing state recognition is a completely new approach.

This methodology can be implemented to the monitoring of DC cables systems including medium and High Voltage cables as well, either in service, during off-line or on-line PD measurements, or in laboratory ageing tests [6]. For on-line measurements, it requires the use of sensors installed at suitable, accessible locations on the HVDC network where PD activity is recorded from several points of attachment along the cable length. Some studies are working to improve on-line PD monitoring [4, 30] for HVDC systems with a particular emphasis on methods to overcome the interference problem. These include the use of a multi-sensor monitoring system, choice of detection frequency and signal discrimination techniques based on the wave shapes [30]. Once on-line PD monitoring technique for HVDC cables and accessories will be developed and various problems related to their measurement on field would be solved [6], our methodology for ageing state recognition using PD under DC has the potential to be applied to these new infrastructures.

This paper starts with the presentation of the experimental setup dedicated to the measurement of PD under DC voltage and the data acquisition procedure. The two feature extraction methods are then described in detail. The proposed classification methodology for ageing state recognition, based on SVM classifiers is presented as well as the feature ranking method used. Finally, the recognition rates obtained with the two different feature sets are compared and discussed.

2. Setup description and data acquisition

PD measurements under DC require long acquisition time due to the small repetition rate of PD [22]. Moreover, large bandwidth and high sampling rate are needed to record high frequency components of PD signals. Thus, a large memory of the detection system is necessary to perform long time measurement preserving high value of the sampling rate. Most of the existing studies regarding PD analysis under DC use classic PD detectors with detection bandwidth of 100 to 400kHz [8, 10] recommended by the standard IEC 60270 [31] that do not allow to record the real pulse shape of PD signals. Thus, information about PD process related to the ageing state of the insulation is not available using conventional detection methods. Therefore, PD are detected using non-conventional methods [32] that include modern digital PD measuring systems (Tektronix DPO3034) with 5 Mpts memory, 300 MHz bandwidth and 2.5Gs/s sample rate. This detection system allows to trigger the acquisition in a case of a transient event only and reduces the amount of memory needed. In this way, the measurement can be performed during a long time and the real pulse shape of PD signals are recorded. It allows to extract features from the real pulse shape of PD signals that are then used for classification purposes.

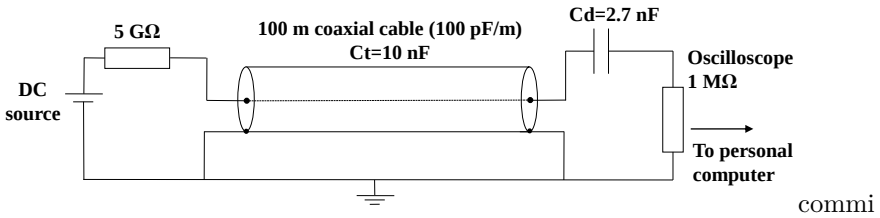

Figure 1: Experimental setup.

An overview of the experimental setup is shown in Fig. 1. A high voltage DC generator capable of providing up to 60 kV is used. The voltage applied to the cable is +15 kVdc. The high voltage is applied to the test object with global capacitance Ct through a 5 GΩ resistance. The test object is a solid polyethylene (PE) insulated coaxial cable which comprises a 0.9 mm diameter of solid soft annealed copper conductor. The thickness of PE insulation is 1.05 mm. The metallic shielding is a tinned copper braid and the overall jacket is PVC. The total length of the cable test object is 100 m and its linear capacitance is 100 pF/m, thus Ct=10 nF. Its characteristic impedance is 50 Ω. The cable ends are immersed in a cell test filled with transformer oil to avoid surface discharges at these locations. A High Voltage 2.7 nF decoupling capacitor Cd is placed between the test object and the measuring impedance of the oscilloscope (1 MΩ) that create a low-impedance path to detect PDs. The detected PD signals travel through a protection circuit before reaching the digitizing oscilloscope. The data are saved in a personal computer via Ethernet protocol. Matlab programs are developed for the automatic acquisition and analysis of PD experimental data. Time-domain reflectometry [33] is performed to localize PD faults within the cable. Only PDs occurring far away from the cable boundaries are selected to build the database. PD localization is an important aspect of the methodology because it ensures the selection of internal PD of the cable insulation. In fact, internal PD give more meaningful information about the insulation condition, as PD mechanism in internal cavities evolves during ageing due to the modification of the cavities properties [23, 24, 25, 26]. In this way, PD occurring at the cable ends are automatically eliminated because they do not reveal the intrinsic ageing of the cable insulation. When a PD occurs, the PD pulse propagates to both cable ends. At each of these points, total reflection occurs until the signal is totally attenuated in the cable. The PD event is localized by measuring the relative time of arrival of the

multiple reflections. Fig. 2 shows a PD signal that is intrinsic to the cable insulation. The PD signals were filtered using wavelet transform and symlets as the mother wavelet with a threshold level equal to 8 [34]. The MATLAB routine 'wden' was used to implement the wavelet transform. The second pulse (Fig. 2) is the pulse reflected back from the far end after having additionally traveled twice the distance between the PD fault and the cable far end. Thus, the distance of the PD fault from the far end is deduced from the time difference $t2 - t1 = 700$ ns and the signal speed v of 2.10^8 ms^{-1} for polyethylene-insulated cable. It can be calculated as follow: $d_{far_{end}} = v\frac{t2-t1}{2} = 70$ m. Consequently, the distance to the cable near end is $d_{near_{end}} = 100 - 70 = 30$ m (Fig. 2). The time difference between first and third pulse, $t3 - t1 = 1\ \mu s$ (Fig. 2) corresponds to the propagation along twice the total length of the cable (l=100 m): $t3 - t1 = \frac{2l}{v}$. The fast response of the unconventional detection system used in this work (with 300 MHz bandwidth) allows performing PD localization using the Time Domain Reflectometry method without superposition of pulses. The time resolution of the detection system is of the order of few nanoseconds whereas the time resolution of a conventional detection system is less than 2 μs [31].

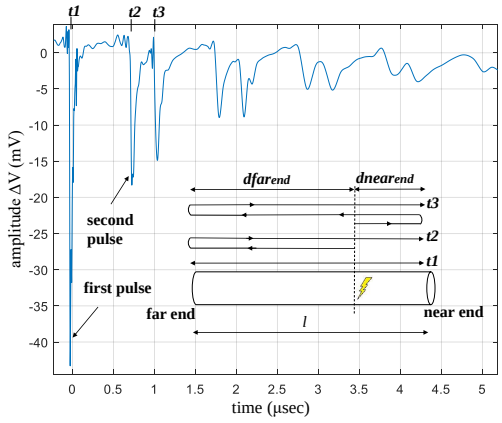

Figure 2: PD signal, propagation and attenuation. PD occurring at 70 m from the cable far end.

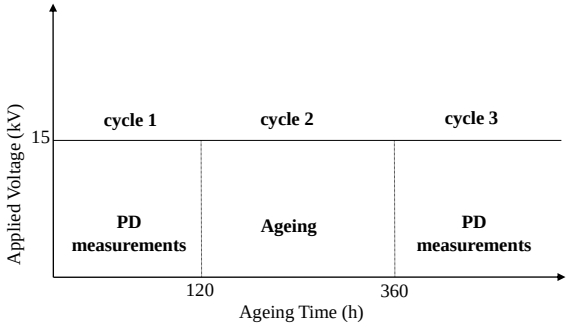

Figure 3: Ageing test cycles.

The ageing and data acquisition protocol is divided into three cycles. The ageing test voltage is constant and held to 15 kV during a total time of 1800 hours to ensure PD ignition and ageing of cable insulation. The cable is submitted to high electrical stresses during short duration. In this way, an ageing cycle of few days performed in laboratory is equivalent to the natural ageing of the cable during several years of service [9]. During the first 120 hours, PD measurement is performed (cycle 1 on Fig. 3). A total of 59 PD events are collected and correspond to the state "virgin" of the cable insulation. The second cycle (cycle 2 on Fig. 3) corresponds to an ageing phase during 240 hours (10 days). No PD measurement is performed. After this cycle, the insulation is considered to be "aged". PD events are collected again after a total of 360 hours of polarization (cycle 3 on Fig. 3). The PD

repetition rate is much higher, but only the first 59 measurements are selected to build the database and correspond to the state "aged" of the cable insulation. Hence, a total of 59 measurements are collected for each of the two ageing states.

3. Feature extraction from PD measurements

Feature extraction using physical parameters of PD signals

From experimental PD data, several features can be extracted. First, a set of ten variables is built and formed by the following physical features: magnitude, rise-time, fall-time, area and width at half maximum of the first and second pulses, distance of the PD event to the cable far end (Fig. 4), and time interval between PDs.

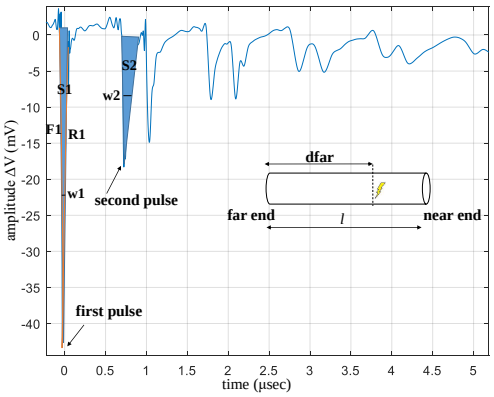

Figure 4: Physical features extraction.

Each of those features is calculated for the 118 PD measurements. The set of data is formed by 59 samples with label "virgin" and 59 samples with label "aged". The ten extracted features are summarized in Table 1.

Variables	physical description
a1	magnitude of the first pulse
a2	magnitude of the second pulse
dfar	distance of the PD event to the cable far end
R1	rise-time of the first pulse
F1	fall-time of the first pulse
S1	area of the first pulse
S2	area of the second pulse
w1	width at half maximum of the first pulse
w2	width at half maximum of the second pulse
δt	time interval between PDs

Table 1: Physical features extracted from PD measurements.

Feature extraction using Discrete Wavelet Transform of PD signals

Traditionally, the techniques from the signal processing world act in either the time or frequency domain to analyze and extract relevant features from data. For instance, the Fourier Transform decomposes the PD signal into its frequency components; however, the time localization of the signal components is not available. One solution is to adopt Short-Time Fourier Transform (STFT) that gets frequency components of local time intervals of fixed duration. The problem is that most PD signals have high frequency content for short duration and fast

transient components tend to be ignored by this kind of transforms. Unlike the Fourier Transform or the Short-Time Fourier Transform, the wavelet transform analyzes the signal at different frequencies with different resolutions [35]. Therefore, wavelet transform is more advantageous for the analysis of transient signals. Wavelet Transform (WT) has been widely adopted as a signal de-noising tool in PD measurements [36, 37, 38]. It has also been applied for the extraction of representative features of different PD patterns for insulation defects classification in HV equipment [39]. In this work, Discrete Wavelet Transform (DWT) is used for feature extraction of PD signals in order to discriminate two ageing states of cable insulation. There are different types of wavelets, which are available for such implementation, such as Daubechies, Symlets, Coiflets, Gaussian and Shannon. Among the wavelets available, the Daubechies wavelet is the most used in partial discharge studies [40] because it has almost all of the required properties such as compactness, limited duration, orthogonality and asymmetry for analysis of fast transient, non-periodic pulses [41]. Fig. 5 shows the block diagram of wavelet decomposition. The original signal is decomposed into approximation and detail coefficients, which represent respectively low frequency and high frequency content of the signal. The decomposition is repeated to further increase the frequency resolution until the desired decomposition level is achieved.

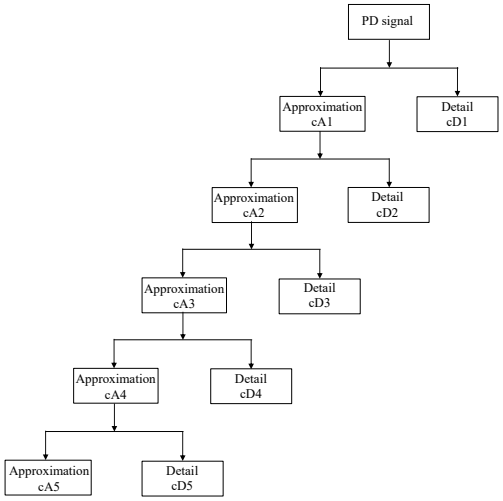

Figure 5: Block diagram of wavelet decomposition.

In this paper, each of the 118 PD signals collected are decomposed using the 'db10' version of daubechies wavelet. It decomposes the signal into multi-level details according to their respective frequencies. The characteristic times of the PD pulse (rise-time, fall-time, width at half maximum) vary according to the degradation state of the insulator and are of the order of nanoseconds [24]. Hence, PD signals are decomposed up to five levels to recover this range of frequency. From these decomposed signals, the five detail coefficients cD1, cD2, cD3, cD4 and cD5 are used to calculate the corresponding features for ageing state classification. In order to reduce the dimensionality of wavelet decomposition, this paper computes only the first two moments including mean and standard deviation for each of the five distributions composed by the detail coefficients cD1, cD2, cD3, cD4 and cD5. Each of the five decomposition levels has two descriptors defined by:

$$\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x(n)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^{N} [x(n) - \bar{x}]^2}$$

where $x(n)$ is the wavelet coefficient at location n and N is the total number of wavelet coefficients at each level. Hence, a total of 10 features are calculated for each PD signals.

4. Supervised classification

Statistical classification techniques are usually implemented to separate objects into different categories called *classes*. An object i is defined by a vector of variables x_i. In this study, each of the ageing state "virgin" and "aged" corresponds to a class [42]. Two different sets of variables are extracted for each object. Classifiers are mathematical functions which assign a predefined class to objects using relevant features. In this study, the classifier is built from a set of measurements containing several objects whose class ("virgin" or "aged") is known. This type of classification is called supervised classification because the whole available samples are labeled. Designing a classifier consists in defining a function $f(x_i, \theta)$ and a set of parameters θ, which give an optimal separation between classes. A classifier is adjusted during the training phase, using the samples that form the training set and an efficient training algorithm. Afterwards, its performance is evaluated with the samples that form the validation set [42]. The validation rate is defined as the percentage of well classified samples when used for validation. It provides an estimate of the performance of the classifier on examples that were not used for the training. If the training algorithm is sensitive to the initial values of the parameters to be adjusted, this process is repeated several times for various values of θ. Finally, the winner is the classifier with vector θ that gives the best validation score. In this study, we use the cross-validation technique [43], so that each sample of the database is used for both training and validation. The database is divided into L sets called folds. Only one fold is used for the validation and the classifier is trained on the $L-1$ remaining folds. The training and validation phases are repeated L times. The validation fold changes at each training and the global validation rate is the average score of the validation over the whole database.

Ranking and selecting variables

In order to determine among all the extracted PD features which are the most relevant for building an efficient classifier, we start by ranking the variables using the Gram-Schmidt orthogonalization procedure [27]. This method proceeds by selecting the variables which are the most correlated with the vector of labels (the output). The most relevant variable is the one which maximizes the following quantity:

$$\cos^2(\mathbf{x}^k,\mathbf{y}) = \frac{({}^T\mathbf{x}^k\mathbf{y})^2}{({}^T\mathbf{x}^k\mathbf{x}^k)(\mathbf{y}\mathbf{y})}$$

where $\mathbf{x}^k$ is the vector containing the P measurements of the k-th variable and $\mathbf{y}$ is the vector of labels.

Then, the output and all other variables are projected on the subspace orthogonal to the selected variable. The same calculation is repeated from this subspace to determine among the projected features, which is the most correlated with the projection of the output. This ranking method possesses the advantage to avoid the use of redundant features. Once all the variables are ranked according to their level of relevance, separation methods are implemented using an increasing number of variables starting with the most relevant, and after with the two most relevant, and so on. There are as many subsets of variables to consider as the number of variables themselves. This variable selection method is called the wrapper approach [28].

Support Vector Machines (SVM)

Since classifiers are built using real world measurements, data probably includes noise. Therefore, the implementation of machine learning techniques that include a regularization process is of great benefit. A Support Vector Machines (SVM) training algorithm has a built-in regularization mechanism that permits to avoid overfitting and then to maximize the generalization capabilities [42]. For two-class classification problems, the Support Vector Machines algorithm is used to find an optimal separation between the two classes : the maximum margin hyperplane [29]. The set of examples that are sufficient to determine the maximum margin hyperplane are called the support vectors. If the data are linearly separable, a linear SVM classifier is sufficient. Otherwise, if the data are not linearly separable, SVM classification proceeds by projecting the input vectors in a high-dimensional space called the feature space where a linear separation is possible. In practice, this data conversion leads to the use of a kernel function. To be a SVM kernel, a function has to verify a set of conditions listed in [29]. A typical kernel function used in this work is the Gaussian kernel. It introduces an additional parameter to estimate: the standard deviation of the Gaussian function. In this study, the two ageing states are separated according to the sign of the SVM discriminant function:

$$f(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i y_i k(\mathbf{x}, \mathbf{x}_i) + b$$

where: k is the kernel function, $\mathbf{x}_i$ are the support vectors, y_i are the corresponding class labels (± 1) and M is the number of support vectors. Note that α_i and b are the parameters of the classifier adjusted during the training process that leads to maximizing:

$$L(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j)$$

subject to

$$\sum_{i=1}^{M} \alpha_i y_i = 0$$

and

$$0 \leq \alpha_i \leq C, \ for \ 1 \leq i \leq M$$

The regularization parameter C is called a hyperparameter. It controls the trade-off between classification errors on training data and margin maximization, thus regularization. Another hyperparameter is the standard deviation of the Gaussian kernel when using nonlinear SVM. Both of them are tuned using a grid search optimization with the cross-validation score as a criterion to be minimized. Thus, the training of SVM leads to a kind of bi-level optimization where the hyperparameters are the upper-level variables to adjust and the α_i and b the lower-level ones [37].

Application to experimental PD data and results

Classification results using physical features of PD signals: The feature ranking method implemented is the Gram-Schmidt orthogonalization procedure, as previously described. The ranking obtained on the ten physical features is shown in Table 2. It can be observed that δt is the most relevant feature, followed by w2, w1, dfar and so on.

Feature	Rank
a1	7
a2	9
dfar	4
R1	8
F1	5
S1	6
S2	10
w1	3
w2	2
δt	1

Table 2: Ranking of the ten physical features according to their level of relevance.

Following the feature ranking, linear SVM and nonlinear SVM classification methods are implemented using an increasing number of variables starting with the most relevant (δt), and after with the two most relevant (δt, w2), and so on. Ten classifiers are then built and their parameters estimated. The cross-validation technique described above is implemented with $L = 5$ folds in order to evaluate the generalization capabilities of the classifiers [42]. 100 random partitionings are drawn to make the cross-validation score independent from the data partitioning in the folds. The cross-validation score is computed as many times. Table 3 shows the global cross-validation scores that are the averages over the 100 values obtained. The results of Table 3 show that classifiers that implement nonlinear

SVM allow to achieve better recognition rates compared to those with linear SVM whatever the selected features. With linear SVM, the best score (97.63%) is obtained when the six most relevant features according to the Gram-Schmidt ranking are used as input variables that are (δt, w2, w1, dfar, F1, S1). Thus, the feature selection permits to eliminate four variables since they do not improve the cross-validation score. With nonlinear SVM, the best performance (99.63%) is obtained with the seven most relevant variables according to the Gram-Schmidt ranking that are (δt, w2, w1, dfar, F1, S1, a1). In this case, the feature selection permits to eliminate three variables. For a comparison, all the possible classifiers having one input variable are built in order to determine which variable is the most informative. Since ten variables are available, the number of classifiers is 10. This method selects the variable among 10 that gives the best classification performance. To make the procedure independent from data partitioning in the five folds, 100 random partitionings are performed over the five folds. The results are shown in 3 (third column). For linear SVM as well as nonlinear SVM, the feature that gives the higher score (94.92% and 98.02% respectively) is δt which is the most relevant variable according to the Gram-Schmidt ranking (Table 2). A recognition rate up to 98% is achieved when only δt is used as input of a nonlinear SVM classifier. This result confirms the feature ranking and selection method following which δt is the most informative variable. The classification score is slightly enhanced (99.63%) when δt is combined with the six other most relevant variables according to the Gram-Schmidt ranking.

Classifier	The N most relevant features according to the Gram-Schmidt ranking	The most relevant feature
Linear SVM	(97.63 ± 0.78) % N=6	(94.92 ± 0.00) %
Nonlinear SVM	(99.63 ± 0.60) % N=7	(98.02 ± 0.54) %

Table 3: Classification results.

Thus, the best performing combination consists of nonlinear SVM classifiers using (δt, w2, w1, dfar, F1, S1, a1) as input variables. When combined together, these seven variables appear to be the most relevant physical features to best recognize two degradation states of the cable insulation with a performance of 99.63%.

The classification errors obtained can be explained by various hypothesis such as a lack of examples, especially in the region of the maximum margin hyperplane, the presence of measurement noise in the data or a sub-optimal selection procedure for the hyperparameters C and standard deviation of the Gaussian kernel. For nonlinear SVM, the nature of the kernel function also has an influence on the classification scores obtained. In fact, the separation boundary is not exactly the same for a Gaussian kernel function compared to a polynomial one. In addition, the number of folds as well as the number of random partitioning of the data into the folds influence the validation scores. Moreover, the selected features may not contain all the necessary information to separate data without errors. All these reasons could explain why a recognition rate of 100% is not achieved for both linear and nonlinear SVM.

Classification results using features from wavelet decomposition of PD signals: The ranking obtained on the ten extracted features from wavelet decomposition of PD pulses is shown in Table 4. It can be observed that std(cD5) is the most relevant feature, followed by std(cD4), mean(cD4), std(cD3) and so on.

Feature	Rank
mean(cD1)	10
std(cD1)	5
mean(cD2)	7
std(cD2)	8
mean(cD3)	9
std(cD3)	4
mean(cD4)	3
std(cD4)	2
mean(cD5)	6
std(cD5)	1

Table 4: Ranking of the ten wavelet features according to their level of relevance.

The classification methods remain the same as described above. The results of Table 5 show that classifiers that implement nonlinear SVM allow to reach slightly better recognition rates compared to those with linear SVM. With linear SVM, the best score (99.51%) is obtained when only the first feature according to the Gram-Schmidt ranking is used as input, i.e. std(cD5). The feature selection permits to eliminate nine variables since they do not improve the cross-validation score. With nonlinear SVM, the best performance (100%) is also obtained with std(cD5). To validate these results, we propose to build all the possible classifiers having one input variable in order to know if there exists another variable, not ranked first by the Gram-Schmidt orthogonalization procedure, that gives the same higher score. To make the procedure independent from data partitioning in the five folds, 100 random partitionings are performed over the five folds. The results are exactly the same as the previous approach. For linear SVM as well as nonlinear SVM, the best recognition rates are 99.51% and 100% respectively, using std(cD5) as input variable which confirms the Gram-Schmidt ranking and feature selection results.

Classifier	The N most relevant features according to the Gram-Schmidt ranking
Linear SVM	(99.51 ± 0.42) % N=1
Nonlinear SVM	(100 ± 0.00) % N=1

Table 5: Classification results.

The best performing combination consists of nonlinear SVM classifier using only one variable as input that is the most informative according to the Gram-Schmidt procedure. The best achieved recognition rate is 100%. Thus, the standard deviation of the detail coefficients distribution at level 5 is the most relevant feature to best recognize two degradation states of the cable insulation. As the detail coefficients of each decomposition level represent a specific frequency band of the original signal, it can be concluded that the frequency band corresponding to the decomposition level 5 of PD pulses is discriminant for the two different ageing states.

5. Conclusion

In this work, feature extraction from real PD data and supervised classification methods are applied for ageing state recognition of cable insulation submitted to DC voltage. The use of unconventional measurement system allows to record the real pulse shape of PD pulses and to extract variables that contain relevant information about ageing state of cable insulation. First, pulse shape parameters as well as time interval between PDs are extracted from experimental PD data to build a first set of physical features formed by 10 variables. Then, PD data are transformed using DWT and decomposed up to five levels. A second set of numerical features formed by the mean and variance of the wavelet transform coefficients distribution at each level of decomposition is used for ageing state recognition. With pulse shape features, the best performing combination consists of a nonlinear SVM classifier using the seven most relevant features according to the Gram-Schmidt orthogonalization procedure as input variables (δt, w2, w1, dfar, F1, S1, a1). The best achieved recognition rate exceeds 99%. However, one can notice the efficiency of the most relevant variable δt that is sufficiently informative to achieve a recognition rate of 98% by itself when used as input of a nonlinear SVM. The classification performance is slightly enhanced when features are extracted from the wavelet decomposition of PD signals. A recognition rate of 100% is achieved using nonlinear SVM classifier with only one feature as input. The standard deviation of the decomposition coefficients distribution at level 5 is the most informative to best recognize two degradation states of cable insulation. This feature is ranked as the most relevant variable according to the ranking performed using an orthogonalization procedure.

This study suggests two different feature extraction methods and a feature ranking technique, which when combined with SVM classifiers, reduce the dimensionality of the feature space and leads to successful separation of two degradation states of cable insulation with high accuracy. Whereas PD measurements were not recommended by the CIGRE guide for DC cable systems, the classification results prove the efficiency of the methodology that makes use of PD signals for ageing state recognition of DC cable systems, especially when features are extracted from the DWT of PD measurements.

A typical application of this study would be a situation where PD are collected punctually or at different moments of a DC cable lifetime, either in service or in laboratory tests. Once PD are collected, the data are processed using the feature extraction, selection and classification methods presented to estimate the ageing state of the insulation at a particular moment of the cable lifetime.

Our methodology for ageing state recognition using PD under DC recorded at different moments of the cable lifetime has the potential to be applied in existing and new settings regarding HVDC cable monitoring systems. First, our methodology could be performed during commissioning test to verify the degradation state of the cable insulation in order to rectify manufacturing or assembly defects at commissioning, before they lead to complete failure or secondary damage once in service. Moreover, the use of a non conventional detection system opens the possibility to perform PD measurements during long acquisition time and store PD patterns representative of different ageing states of the cable insulation in an analyzer. In this way, cables can be monitored at various phases of their life. Moreover, since high frequency components of PD signals must be recorded and long acquisition time are needed, both high sampling rate and high memory of the detection system are required that could lead to huge equipment costs [4].

However, there are some limitations to use the proposed method in on-line PD measurement settings, first because on-line PD measurements to installed HVDC equipment is rare [30]. In fact, HVDC stations are a particularly challenging environment for the performance of on-line PD measurement due to high levels of electromagnetic interference and multiple PD sources occurring simultaneously [44]. Furthermore, cables are usually connected to overhead transmission lines, which act as a huge receiving antenna, picking up a wide range of signals such as radio, television and mobile phone transmissions, especially during on-line measurements. Moreover, because our methodology is based on internal PD signals, noise data and parasite discharges should be automatically eliminated in order to avoid misguiding in the diagnostic of the cable insulation [30]. In addition, PD under DC occur less frequently than under AC. Thus, long acquisition time are required in order to acquire enough PD data for the diagnosis of cable insulation. Therefore, the risk of triggering the acquisition on a noise signal instead of a PD is much more important and errors in the interpretation of PD measurements are more likely to happen under DC voltage that may lead to false conclusions in the diagnostics of the cable insulation (e.g., unnecessary disconnections of the equipment or unexpected failures) [4]. Thus, in parallel of our methodology, techniques should be implemented for the automatic discrimination of PD pulses from noise signals. Moreover, on-line PD measurement devices should be developed in order to perform PD measurements along the cable length. Despite online detection of PD over long cables is a big challenge [45], it has proved efficiency on cables terminations and joints where sensor monitoring systems are integrated [30]. In this case, the methodology developed in this study can be extended to the monitoring of DC cables systems including cables and their accessories such as joints and terminations that are more likely locations for insulation defects [4] and occurrence of PD.

Finally, the perspective of transferring a cable diagnostic model from one environment (e.g, one cable type and a particular PD detection system) to another would be of great interest for both economic and technical reasons [46]. Thus, the generalization capabilities of the methodology presented should be evaluated in order to predict ageing state of different cables types using PD data acquired under different experimental conditions. For this purpose, a work on domain adaptation techniques is currently ongoing.

References

[1] Eigner A, Rethmeier K 2016 An overview on the current status of partial discharge measurements on AC high voltage cable accessories IEEE El. Ins. Mag. 32 48-55, doi: 10.1109/MEI.2016.7414231

[2] Sabat A, Karmakar S 2011 Simulation of partial discharge in high voltage power equipment IJEEI 2 234-247

[3] Niu H. Q, Cavallini A, Montanari G. C 2008 Identification of partial discharge phenomena in HVDC apparatus IEEE Int. Sympos. Electr. Insul. 373-376, doi:10.1109/ELINSL.2008.4570352

[4] Judd M, Siew W, Hu X, Corr E, Zhu M, Reid A, Mountassir O, Cristobal M, Giussani R and Seltzer-Grant M (2015), "Partial Discharges under HVDC Conditions", 12, 2015

[5] CIGRE WG B1.32 2012 Recommendations for Testing DC Extruded Cable Systems for Power Transmission at a Rated Voltage up to 500 kV, CIGRE Technical Brochure No. 496

[6] Takahashi T, Wibowo AS, Cavallini A, Montanari GC, Boyer L, Luton M-H and Mirebeau P 2016 AC and DC partial discharge measurements on defective cables, IEEE Electrical Insulation Conference (EIC), doi: 10.1109/EIC.2016.7548617

[7] Morshuis P, Cavallini A, Fabiani D, Montanari G and Azcarraga C 2015 Stress conditions in HVDC equipment and routes to in service failure, IEEE Transactions on Dielectrics and Electrical Insulation., 22(1) 81-91. Institute of Electrical and Electronics Engineers (IEEE), doi: 10.1109/TDEI.2014.004815

[8] Seltzer-Grant M, Siew W, Corr E, Hu X, Zhu M, Judd M, Reid A and Neumann A 2015 Laboratory and Field Partial Discharge Measurement in HVDC Power Cables

[9] Mazzanti G 2017 Life and reliability models for high voltage DC extruded cables, IEEE Electrical Insulation Magazine., 33(4) 42-52. Institute of Electrical and Electronics Engineers (IEEE), doi: 10.1109/MEI.2017.7956632

[10] E. Gulski, P. H. F. Morshuis and F. H. Kreuger 1994 Conventional and time-resolved measurements of partial discharges as a tool for diagnosis of insulating materials, Proceedings of 1994 4th International Conference on Properties and Applications of Dielectric Materials (ICPADM), 2 666-669, doi: 10.1109/ICPADM.1994.414098

[11] F.H Kreuger 1990 Partial Discharge Detection in High Voltage Equipment

[12] IEEE Recommended Practice for the Detection of Partial Discharge and the Measurement of Apparent Charge in Dry -Type Transformers, in IEEE Std C57.124

[13] P. H. F. Morshuis and F. H. Kreuger 1992 A relation between time-resolved discharge parameters and ageing, Sixth International Conference on Dielectric Materials, Measurements and Applications, 37-40

[14] CIGRE Working Group Report 1969 Recognition of Discharges, Electra, No. 11

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

[16] Yongpeng X, Yong Q, Xiaoxin C, Aihuiping X, Gehao S and Xiuchen J 2016 Partial discharge feature extraction through contourlet transform for XLPE cable defect models classification International Conference on Condition Monitoring and Diagnosis 912-915, doi: 10.1109/CMD.2016.7757971

[17] Kunicki M, Cichoń A, Nagi L 2018 Statistics based method for partial discharge identification in oil paper insulation systems, Electric Power Systems Research, 163 B 559-571, https://doi.org/10.1016/j.epsr.2018.01.007

[18] Raymond W. J. K, Illias H. A and Bakar A. H. A 2017 High noise tolerance feature extraction for partial discharge classification in XLPE cable joints IEEE Trans. Electric. Insul. 24 66-74, doi: 10.1109/TDEI.2016.005864

[19] Evagorou D et al 2012 A lower dimensional feature vector for identification of partial discharges of different origin using time measurements Meas. Sci. Technol. 23 055606, doi: 10.1088/0957-0233/23/5/055606

[20] Mondal M and Kumbhar G 2017 Detection, Measurement, and Classification of Partial Discharge in a Power Transformer: Methods, Trends, and Future Research, IETE Technical Review., 35(5) 483-493, doi: https://doi.org/10.1080/02564602.2017.1335244

[21] Sarkar J and Patil R (2014), "Detection of Partial Discharges Occurring in HVDC Environment", 06, 2014

[22] Morshuis P and Smit J 2005 Partial discharges at dc voltage: their mechanism, detection and analysis IEEE Transactions on Dielectrics and Electrical Insulation., 12(2) 328-340. Institute of Electrical and Electronics Engineers (IEEE), doi: 10.1109/TDEI.2005.1430401

[23] Niemayer L 1995 A generalized approach to partial discharge modeling IEEE Trans. Electric. Insul. 2 510-528, doi: 10.1109/94.407017

[24] Morshuis P 1995 Assessment of dielectric degradation by ultrawide-band PD detection IEEE Trans. Electric. Insul. 2 744- 760, doi: 10.1109/94.469971

[25] Morette N, Daassi-Gnaba H, Ditchi T and Oussar Y 2017 Characterization of partial discharges in solid insulators under DC voltage using physical cavity properties IEEE Int. Sympos. Electr. Insul. Mat. 374-377, doi: 10.23919/ISEIM.2017.8088763

[26] Temmen K 2000 Evaluation of surface changes in flat cavities due to ageing by means of phase-angle resolved partial discharge measurement J. Phys. D Appl. Phys. 33 603, doi: 10.1088/0022-3727/33/6/303

[27] Chen S, Billings S.A, Luo W 1989 Orthogonal least squares methods and their application to non-linear system identification Int. J. Control 50 1873–1896, doi: 10.1080/00207178908953472

[28] Guyon I and Elisseeff A 2003 An introduction to variable and feature selection Journal of Machine Learning Research 3 1157- 1182

[29] Cristianini N and Shawe-Taylor J 2006 Support Vector Machines and other Kernel-based Learning Methods (Cambridge: Cambridge University Press)

[30] B. T. Phung, Z. Liu, T. R. Blackburn, R. E. James. Recent Developments for On-line Partial Discharge Detection in Cables. University of New South Wales, Australia

[31] Guide for electrical Partial Discharge Measurements in compliance to IEC 60270

[32] 2010 CIGRE 444 WG D1.33 Guidelines for Unconventional Partial Discharge Measurements

[33] Liu Z, Phung B.T, Blackburn T. R 1999 The propagation of partial discharges pulses in a high voltage cable Proc. Of AUPEC/EECON eds 287 - 292

[34] Sriram S, Nitin S, Prabhu K and Bastiaans M 2005 Signal denoising techniques for partial discharge measurements, IEEE Transactions on Dielectrics and Electrical Insulation., 1182-1191. Institute of Electrical and Electronics Engineers (IEEE), doi: 10.1109/TDEI.2005.1561798

[35] Lei L, Wang C and Liu X 2013 Discrete Wavelet Transform Decomposition Level Determination Exploiting Sparseness Measurement IJECE 7 691–694

[36] Cunha C, Carvalho A, Petraglia M, Lima A 2015 A new wavelet selection method for partial discharge denoising, Electric Power Systems Research, 125 184-195, https://doi.org/10.1016/j.epsr.2015.04.005

[37] Zhou X, Zhou C and Kemp I. J 2005 An improved methodology for application of wavelet transform to partial discharge measurement denoising IEEE Trans. Electric. Insul. 12 586-594, doi: 10.1109/TDEI.2005.1453464

[38] H de Oliveira Mota, L Dutra da Rocha, T Cunha de Moura Salles, F Henrique Vasconcelos 2011 Partial discharge signal denoising with spatially adaptive wavelet thresholding and support vector machines Electric Power Systems Research, 81 2 644-659, https://doi.org/10.1016/j.epsr.2010.10.030

[39] Evagorou D et al 2010 Feature extraction of partial discharge signals using the wavelet packet transform and classification with a probabilistic neural network IET Science, Measurement & Technology 4 177-192, doi: 10.1088/0957-0233/17/9/001

[40] L Hao et al 2008 Improving detection sensitivity for partial discharge monitoring of high voltage equipment Meas. Sci. Technol. 19 055707, doi: 10.1049/iet-smt.2009.0023

[41] Ma X, Zhou C and Kemp I. J 2002 Interpretation of wavelet analysis and its application in partial discharge detection IEEE Trans. Electric. Insul. 9 446-457, doi: 10.1109/TDEI.2002.1007709

[42] Morette N, Ditchi T and Oussar Y 2018 Partial Discharges Measurements and Analysis as an Evaluation Tool for the Reliability of Polymeric-Insulated Cables used under HVDC Conditions ICD 1-4 8514708, doi: 10.1109/ICD.2018.8514708

[43] Hastie T, Tibshirani R and Friedman J 2009 The Elements of Statistical Learning (New-York: Springer)

[44] Luo Y, Li Z and Wang H (2017), "A Review of Online Partial Discharge Measurement of Large Generators", Energies., oct, 2017. Vol. 10(11), pp. 1694. MDPI AG

[45] N. Oussalah, Y. Zebboudj, S.A. Boggs 2007 Partial discharge pulse propagation in shielded power cable and implications for detection sensitivity, IEEE Electrical Insulation Magazine, 23(6) 5-10, doi: 10.1109/MEI.2007.4389974

[46] Wang Q, Michau G and Fink O (2019), "Domain Adaptive Transfer Learning for Fault Diagnosis", In 2019 Prognostics and System Health Management Conference (PHM-Paris)., may, 2019. IEEE