



Article

# Unsupervised Monitoring System for Predictive Maintenance of High Voltage Apparatus †

Christian Gianoglio \*, Edoardo Ragusa , Andrea Bruzzone , Paolo Gastaldo ,  
Rodolfo Zunino  and Francesco Guastavino 

Electrical, Electronics and Telecommunication Engineering and Naval Architecture Department (DITEN), University of Genoa, 16145 Genova, Italy; edoardo.ragusa@edu.unige.it (E.R.); andrea.bruzzone@edu.unige.it (A.B.); paolo.gastaldo@unige.it (P.G.); rodolfo.zunino@unige.it (R.Z.); francesco.guastavino@unige.it (F.G.)

\* Correspondence: christian.gianoglio@edu.unige.it

† This paper is an extended version of our paper published in International Conference on Applications in Electronics Pervading Industry, Environment and Society. (Pisa, Italy, 11–13 September 2019).

Received: 30 January 2020; Accepted: 26 February 2020; Published: 2 March 2020

**Abstract:** The online monitoring of a high voltage apparatus is a crucial aspect for a predictive maintenance program. Partial discharges (PDs) phenomena affect the insulation system of an electrical machine and—in the long term—can lead to a breakdown, with a consequent, significant economic loss; wind turbines provide an excellent example. Embedded solutions are therefore required to monitor the insulation status. The paper presents an online system that adopts unsupervised methodologies for assessing the condition of the monitored machine in real time. The monitoring process does not rely on any prior knowledge about the apparatus; nonetheless, the method can identify the relevant drifts in the machine status. In addition, the system is specifically designed to run on low-cost embedded devices.

**Keywords:** predictive maintenance; embedded systems; partial discharges

## 1. Introduction

A predictive maintenance apparatus monitors an electrical machine continuously to prevent a breakdown in the insulation system [1–4]. Electrical, mechanical and thermal stress sources systematically contribute to deteriorate motors. In the absence of a scheduled maintenance, this may bring about sudden breakdowns, with the consequent significant economic losses. Thus, predictive maintenance is replacing reactive maintenance, as the latter does not prevent motor failures and downtime costs. Predictive analytics relies on the ever-growing availability of reliable Internet-of-Thing (IoT) technologies, to drive the real-time scheduling of specific maintenance programs for the target machine. As a result, an effective predictive maintenance system should prompt alerts whenever the electrical machine needs fixing to ensure operation continuity.

In the case of powered electrical machines, partial discharges (PDs) are both the symptoms and the causes of defects in the insulation systems. According to IEC 60270 [5], a partial discharge (PD) is a localized, electrical discharge that only partially bridges the insulation between conductors; it can (or cannot) occur adjacent to a conductor. PDs take place when the amplitude of the electric field locally exceeds the breakdown strength limit of the insulating material. At the beginning of the phenomenon, a local concentration of electrons induces an ionization and a consequent streamer, known as electron avalanche; the availability of free electrons eventually starts the discharge process. PDs belong to four main categories: those occurring at the surface of insulating material (namely surface discharge), PDs that take place within gas bubbles in liquid insulation, PDs occurring in vacuum within solid insulation (voids or delaminations), and those that affect an electrode in gas (corona activity). PD activity—if

ignored—could damage the insulating system until a total breakdown, i.e., a complete short circuit between two conductors. Detecting PDs and recognizing the possible sources of defects is therefore crucial for (1) estimating the status of the insulation system and (2) scheduling a maintenance program that can prevent sudden breakdowns [6–8].

In the literature, many works deal with the online monitoring of PDs. Non-invasive sensors (i.e., conducted or irradiated sensors) typically acquire the original, analog signals, which are sampled by an oscilloscope. The resulting time series, in a digital format, undergo a feature-extraction process and ultimately feed a classification system. The classifier identifies the various categories of defects that may affect the observed apparatus. Machine-learning techniques prove especially useful to classify the sources of PDs. In [9], the authors first collected data with an ultra high frequency (UHF) antenna, then extracted a set of features from raw data. The application of the K-means algorithm in the feature space allowed to identify homogeneous clusters and to distinguish different kinds of defects. In [10] a high-frequency current transformer (HFCT) captured the signals; after feature extraction, a neural network classified the defects. A HFCT sensor was used in [11] to collect signals, whereas features resulted from wavelet decomposition; the DBSCAN clustering algorithm [12] was applied to group the sources of defects. Recently, both antenna and HFCT sensors monitored a gas insulated substation (GIS) [13], and clustering techniques again performed the PD sources categorization. In [14] authors proposed a tensor-based algorithm, filtering noise, to classify different PD sources. These method relies on geometric information of PD pattern. In recent years, powerful Graphic Processing Units (GPUs) and the associate distributed-computing paradigms supported solutions based on deep neural networks (DNNs), capable of addressing very complex problems. Convolutional neural networks (CNNs), in particular, were employed in partial discharges recognition and classification. The approach presented in [15] classified six different PD defects constructed in oil by using a deep neural network, which processed phase-resolved PD (PRPD) pattern as images. In [16], the authors used a CNN to classify PD sources in a gas insulated system (GIS). That paper showed that the deep network improved on state-of-art algorithms. Several works adopted CNNs to classify PD sources in a variety of applications [17–22]. As a matter of fact, a drawback of deep learning techniques is that their deployment on low-resources devices is very challenging, as they require both a considerable computational capability and large memories. In addition, deep networks involve a lot of parameters, hence the training process calls for a huge amount of data, which might not be easily available in some applications.

Overall, the above-mentioned approaches all involve a data-driven classification systems, which in turn requires a training procedure. In addition, those methods do not provide either a run-time monitoring of the aging phenomena nor an estimation of the remaining life of a machine, a pair of essential requirements for a cost-effective maintenance scheduling.

Similar issues affected the methods proposed in [23,24]. In the former paper, an UHF antenna monitored a power transformer for 65 days, and yielded phase-resolved PD patterns. Those patterns were matched to three reference templates, by computing the associate correlation coefficients. Thus template matching allowed the system both to classify the PD sources and to monitor changes over time. This method, however, only applies if one knows in advance all possible defects (i.e., the PD pattern shapes) that may affect the observed device. In [24] the authors introduced a histogram-similarity criterion to support a supervised classification approach. Histograms represented PD patterns, hence the various defects formed a database of histograms. At run time, a PD pattern is acquired and yields a histogram that is compared, according to a statistical hypothesis-test procedure, with each reference pattern included in the database. This approach also requires a large dataset of labeled histograms and cannot predict the remaining life time of the insulation system. Multi-defects PD patterns also represent an issue. A similar technique was proposed in [25].

An empirical model enabling the assessment of the lifetime of a specimen was proposed in [26]. By applying a pulse voltage waveform to a set of twisted pairs specimens, the authors induced the aging process, and showed that by collecting some PD patterns for a sufficient time, one could compute

some quantities and model the lifetime of any twisted pairs specimen accordingly. Likewise, in [27] a sinusoidal voltage at a 500 Hz frequency was applied to twisted pairs. Then, the same features adopted in [26] characterized the PD patterns. The results proved that after an initial transient period, a linear relation existed between a specific feature and the degradation time of the specimens. Thus, these approaches both showed that one could collect data from a set of available specimens, build an according model, and predict the lifespan of a new, unseen specimen. The same technique was adopted recently in [28] to predict the life span of medium-voltage XLPE cables.

In [29] the authors proposed a continuous monitoring of PD phenomena in medium voltage grids. The system acquires and processed PD data automatically by rejecting noise [30] and separating PD patterns according to the defects [31]. Clustering methods allowed to attain that goal. On-line monitoring was implemented by tracking the evolution of the clusters as aging progressed, whereas a training phase supported the criteria to trigger automated warnings.

The IEEE 62.2 standard [32] includes diagnostic methods that can provide information about the insulating system and materials. The most common approach relies on the electrical direct current (DC) resistance measurement. The involved quantities are: the value of the resistance  $R1$  measured one minute after the application of the test voltage, and the polarization index  $PI$ , which is the ratio of the resistance value measured after ten minutes to the resistance value detected after one minute. Low  $R1$  and  $PI$  values indicate possible issues due to conductive currents. Another widely adopted method involves the measurement of the power factor, usually denoted as  $Tan\delta$ . This is an alternate current (AC) voltage-based technique, giving information about the active power absorbed by the insulation. High  $Tan\delta$  values indicate the presence of electric conductive phenomena inside the insulating system. As far as PD measuring is concerned, those diagnostic methods give information about the conditions of the insulating system at run time, but can hardly yield some hint about the behavioral trend, even if the measurements repeat over time during degradation.

The research presented in this paper shows that a monitoring system only relying on an unsupervised approach can support predictive maintenance effectively. As a result, the monitoring process does not require any prior knowledge about the observed apparatus. The anomaly detection paradigm drives the surveillance of aging in the insulation system. The system periodically extracts features from PD patterns, and prompts an alert whenever the machine exhibits an anomalous behavior with respect to past history. By the resulting approach one can characterize the status of the insulation system in real time and schedule maintenance operations accordingly.

That monitoring approach presents some crucial features. First, it does not depend on either a knowledge base or some training process. Secondly, it is computationally light, since standard statistical tools (such as the Chi-Square test and the Kolmogorov-Smirnov test) carry out anomaly detection. In this regard, the paper focuses on the implementations of the overall approach in low-cost embedded systems. Third, the method can provide a basic module within a comprehensive IoT-based predictive maintenance apparatus. The signals acquired in real time by local monitoring apparatuses can all be collected by a remote warehouse. Advanced (and computationally heavy) data analysis by means of Artificial Intelligence tools may then apply to further improve the prediction ability of each local monitoring system. The paper illustrates an implementation of such strategy by applying the empirical model presented in [27].

### 1.1. Contributions

The main novel contribution of the present work is a fully unsupervised strategy, that adopts an anomaly detection paradigm for the real-time monitoring of insulation systems. To the best of the authors' knowledge, state-of-the-art approaches that assess the lifetime of insulation systems typically rely on either knowledge bases or Machine-Learning (ML) models, which should be supported by a training process. By contrast, the method described in this paper is specifically designed to detect aging phenomena by only using local information, collected from the monitored machine.

Furthermore, the paper tackles the hard problem of effective implementation on embedded, low-cost devices. Toward that end, the paper presents both an algorithm to support the deployment of that strategy on low-power, low-cost embedded devices, and the design of an IoT-based predictive maintenance system embedding the proposed strategy.

The rest of the paper is organized as follows. Section 2 reviews the empirical model for the electrical aging assessment presented in [27]. Section 3 shows how anomaly detection paradigms can drive the recognition of a change in status within an insulation system. Section 4 introduces and analyzes the proposed on-line monitoring system. Section 5 details an IoT-based predictive system based on the proposed strategy. Section 6 discusses the outcomes of the experimental session. Finally, Section 7 provides a few concluding remarks.

## 2. Background: An Empirical Model for Electrical Aging Assessment

Only a very few works in the proposed methodologies in the literature can support on-line monitoring of the aging of an insulation system. The method proposed in [26] and further developed in [27] provides a state-of-the-art approach for building an empirical model enabling the assessment of the lifespan of a specimen (in particular, twisted-pair specimens).

The process starts by recording the aging of a set of specimens until their breakdown. Aging is induced by applying a pulse voltage [26] or a sinusoidal voltage [27]. To monitor that progress, PD patterns are periodically extracted; in the following,  $T_a$  will denote the acquisition time for a single PD pattern, the index  $i$  marks the PD pattern matrix.

For each specimen, the total average apparent charge per cycle  $q_a(t_j)$  is computed for the generic PD pattern extracted at time  $t_j$ . For this purpose, the positive average current  $I^+$  and the negative average current  $I^-$  are first computed:

$$I^+ = \frac{\sum_i N_i q_i^+}{T_a}; \quad (1)$$

$$I^- = \frac{\sum_i N_i q_i^-}{T_a}. \quad (2)$$

here,  $N_i$  is the total number of measured partial discharges with an amplitude corresponding to an apparent positive  $q_i^+$  or negative  $q_i^-$  charge. Eventually,  $q_a(t)$  is obtained as follows:

$$q_a(t_j) = \frac{I^+ + |I^-|}{f} \quad (3)$$

where  $f$  is the frequency of the power supply voltage applied to the specimen.

Each monitored specimen is then characterized by the linear transfer function  $q_a(t)$  that is worked out by exploiting the empirical measurements  $q_a(t_j)$ . As a result, one has

$$q_a(t) = I_a * t + Q_0 \quad (4)$$

where  $I_a$  is the average current per cycle and  $Q_0$  is the offset charge corresponding to the theoretical average charge per cycle in  $t = 0$ , i.e., the charge measured in a cycle just after the inception of the PD activity. The coefficients  $Q_0$  and  $I_a$  are determined via linear regression. Indeed, as reported in [27], linear regression is computed only on the empirical measurements collected after the instant  $\hat{T}$ , which correspond to the end of the transient zone that characterizes any specimen.

In [26], the breakdown time  $D$  of a twisted pair is expressed as a function of the two quantities  $I_a$  and  $Q_0$ :

$$D = K * Q_0^{-n_1} * I_a^{-n_2} \quad (5)$$

where  $K$ ,  $n_1$  and  $n_2$  are parameters to be tuned. Equation (5) can be written in a linear form by switching to a logarithm representation:

$$\log D = \log K - n_1 * \log Q_0 - n_2 * \log I_a \quad (6)$$

the availability of a set of specimens with the corresponding triplets  $(D, I_a, Q_0)$  allows one to estimate the value of  $K$ ,  $n_1$  and  $n_2$  in (6) via linear regression.

Eventually, after setting  $K$ ,  $n_1$  and  $n_2$ , the expression (5) makes it possible to infer the lifetime of a new, unseen specimen. The inference process requires first to monitor the specimen until its  $Q_0$  and  $I_a$  can be reliably estimated. Then,  $D$  can be assessed by using  $Q_0$  and  $I_a$  as input in (5).

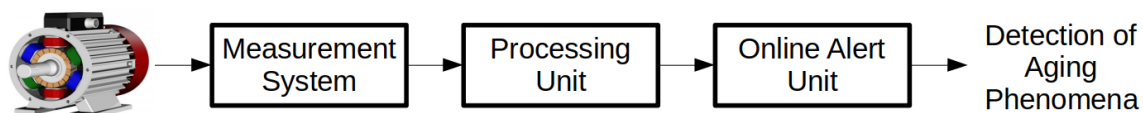
### 3. Detecting Aging Phenomena via Anomaly Detection

The proposed framework assesses in real time the aging of the insulation system by adopting an unsupervised approach. The goal is to implement a monitoring system that does not rely on the availability of a knowledge base. This kind of characteristic demonstrates the main difference between this work and the research presented in [26,27,29]. For this purpose, the strategy adopted is to check at any time,  $t$ , whether the electrical machine presents some anomalous behavior with respect to the past.

In general, anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior [33]. Network intrusion detection, health anomaly detection, spam detection, fraud detection, and industrial damage detection represent only a few examples of applications relying on the anomaly detection paradigm. Actually, anomaly detection can be implemented by exploiting statistical techniques or ML methods, such as clustering, random forest, SVM, and neural networks. In practice, one chooses among the available approaches by taking into account aspects such as the amount of data, the cost of data labeling, the data type (e.g., time series, images, timestamps, etc.), the hardware resources, and the real-time constraints.

The proposed strategy uses statistical methods to implement anomaly detection [34–38]. The underlying rationale is to avoid methodologies that involve training processes. Statistical methods rely on the consideration that normal data instances occur in high probability regions of a stochastic model, while anomalies occur in the low probability regions of the stochastic model. Hence, hypothesis testing is exploited: a normal behavior is fit to the data and then a statistical inference test is applied to determine if new data belongs to the normal model or not.

Figure 1 schematizes the proposed framework. First a measurement system samples—at a given frequency—the amplitudes of the signals sensed by an HFCT sensor in a time window  $\delta$ . In the following,  $\mathbf{v}_t$  will denote the vector of raw data extracted in a time window  $\delta$  starting at instant  $t$ . Second, a *Processing Unit* processes  $\mathbf{v}_t$  to extract the signals that characterize aging phenomena. Then, an *Online Alert Unit* uses anomaly detection to identify, in real-time, sudden changes in the status of the monitored insulation system.



**Figure 1.** The proposed framework for the real-time detection of aging phenomena

The *Processing Unit* (Figure 2) receives as input  $\mathbf{v}_t$  and works out a pair of vectors:  $\mathbf{w}$  and  $\mathbf{x}$ . Vector  $\mathbf{w}$  is computed by the same algorithm that characterizes a PD pattern, but phase information is discarded in this case. Toward that end, first a PD pattern is extracted from  $\mathbf{v}_t$ ; it is represented as a matrix, whose columns correspond to the power supply phases, and rows mark the maximum amplitudes of the discharges. Thus, each element in the matrix identifies an amplitude-phase pair; the contents of that cell give the occurrences of discharges at that amplitude-phase pair in the time window  $\delta$ . The resolution of the analog-to-digital converter (ADCs) actually determines

the quantization levels in both the amplitude range and the phase range. Figure 3 illustrates this mechanism in a toy example, where a 2-bit ADC is employed. As a result, the PD pattern has four levels of amplitudes ( $A_1, A_2, A_3, A_4$ ) and as many levels of phases ( $P_1, P_2, P_3, P_4$ ). In the example, only one discharge with a maximum amplitude  $A_1$  occurred at phase  $P_2$ , whereas two discharges with the same amplitude occurred at phase  $P_4$  phase. Figure 3 gives the associate vector  $w$ , which includes three elements with value  $A_1$ , since phase information is discarded. Likewise, vector  $w$  includes two elements with value  $A_2$ , five elements with value  $A_3$ , and one element with value  $A_4$ . The resulting size of  $w$ , in general, will depend on the non-zero occurrences of the PD pattern. The choice of not using phase information is related to the fact that such information is useful when the goal is to recognize the category of defect that generated the PD. However, the proposed strategy only focuses on the detection of aging phenomena.

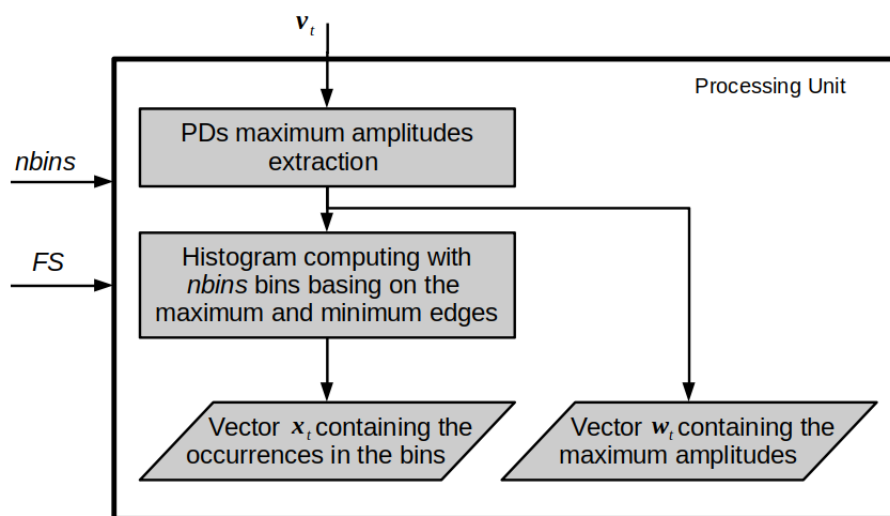


Figure 2. Processing unit entitled to compute both vector of the amplitudes and vector of occurrences in the bins from the raw data.

A1	0	1	0	2
A2	1	0	1	0
A3	1	2	1	1
A4	0	0	1	0
	P1	P2	P3	P4

$$w = [A_1, A_1, A_1, A_2, A_2, A_3, A_3, A_3, A_3, A_3, A_4]$$

Figure 3. PD pattern matrix example with 2 bits ADCs and the amplitudes fullscale set on 1 V.

Vector  $x$  holds the histogram of the amplitudes of raw data. In that histogram, each bin covers a range of amplitudes. Given the fullscale  $FS$  adopted to sample the signal, the bins are obtained by uniform quantization of such interval. The number of bins  $nbins$  and  $FS$  are two input parameters.

Anomaly detection proceeds by applying statistical hypothesis testing to  $w$  and  $x$ . A pair of well-known statistical tests, namely Chi-Square (Chi2) and Kolmogorov–Smirnov (KS) support this task. Chi2 and KS have been selected as 1) they can support non-parametric tests (i.e., no assumptions on the specific probability distributions are involved) and 2) they allow a real-time implementation of hypothesis testing even on low-cost, low power and resource-constrained embedded devices. The null

hypothesis is that the signal measured at time  $T^*$  and any signal measured before  $T^*$  come from the same population. In other words, an anomaly occurs when the measurement at time  $T^*$  is not consistent with previous measurements. The underlying hypothesis is that aging phenomena lead progressively to significant changes in the distribution of PDs. Remarkably, such discontinuities can be detected even in the absence of trained classifiers or any knowledge base.

Hypothesis testing involves a pair of vectors: one carries the measurement at time  $T^*$ , the second one holds the reference, i.e., a representative sample of the population that characterizes the apparatus before time  $T^*$ . The latter vector should be updated whenever an anomaly (i.e., a major discontinuity in the observed status) is detected. The Chi2 test compares two discrete distributions that are expressed in the form of histograms, whereas the (non parametric) KS test carries out hypothesis testing by observing the empirical cumulative distribution functions (ECDF) of data. Thus, as per Figure 2, Chi2 and KS support the detection of anomalies in  $\mathbf{x}$  and  $\mathbf{w}$ , respectively.

In the case of Chi2 test, let  $n_{bins}$  be the number of bins in the two discrete distributions  $\mathbf{x}$  (measured at time  $T^*$ ) and  $\tilde{\mathbf{x}}$  (measured before time  $T^*$ ). Then, let  $O_i$  be the value of the histogram of  $\mathbf{x}$  for the  $i$ th bin, whereas  $\tilde{R}_i$  is the value of the histogram of  $\tilde{\mathbf{x}}$  in the  $i$ th bin. Finally, the quantity  $\chi^2$  is worked out as

$$\chi^2 = \sum_{i=1}^{n_{bins}} \frac{(K_1 O_i - K_2 \tilde{R}_i)^2}{O_i + \tilde{R}_i} \quad (7)$$

where  $K_1$  and  $K_2$  are scaling factors to make up for unbalanced samples. Given  $\chi^2$  and the degrees of freedom  $DF$ , which corresponds to the number of non-empty bins, the p-value can be computed from the *Chi2* distribution. The null hypothesis is accepted if  $p\text{-value} < \alpha$ , where  $\alpha$  is the significance level set a-priori (usually  $\alpha=0.05$ ).

The KS test compares the empirical cumulative distributions (ECDF) of  $w_T$  and  $\tilde{w}$ . Given  $N$  points  $p_n$  sorted in increasing order of values, the ECDF is defined as

$$FO_n = i_n / N \quad (8)$$

where  $i_n$  is the number of elements smaller than  $p_i$  in  $w_T$ . Likewise, FE is the ECDF of  $\tilde{w}$ . As a result, the KS test statistic relies on

$$D = \sup_n |FO_n - FE_n| \quad (9)$$

Given  $D$ , the  $p$ -value can be computed as in [39]. Eventually, the null hypothesis is rejected if the significance level  $\alpha$  is lower than the  $p$ -value.

#### 4. Online Monitoring for Predictive Maintenance

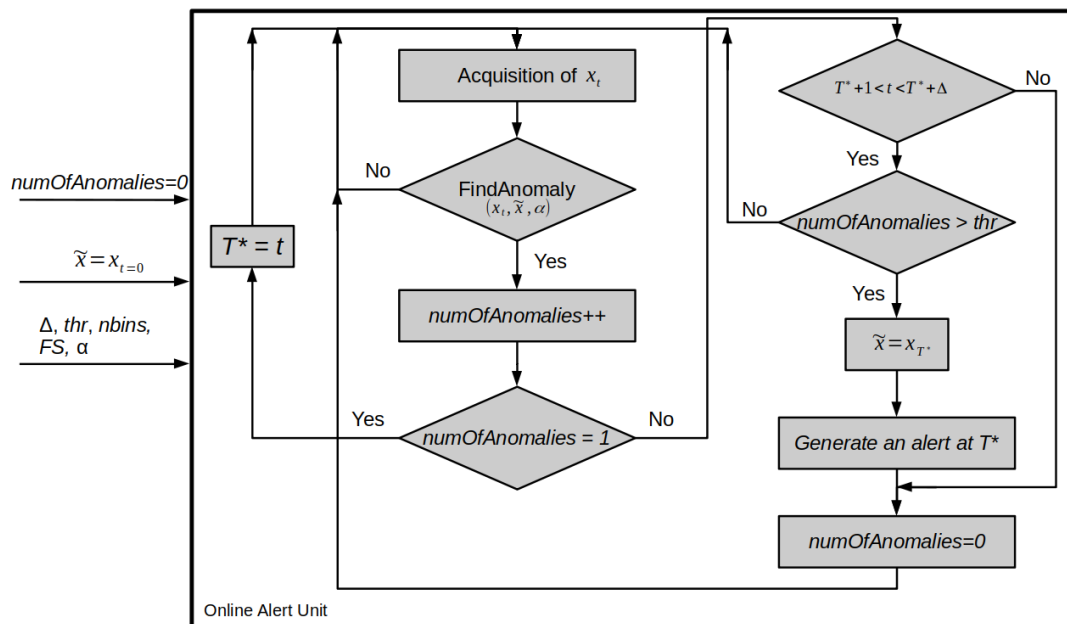
The online monitoring system identifies, at run time, abrupt discontinuities in the status of the electrical machine, which are reported as alerts. At time  $t = 0$  the system is switched on, and the pair of vectors,  $\mathbf{w}_{t=0}$  and  $\mathbf{x}_{t=0}$ , are initialized as per Section 3.

Figure 4 outlines the anomaly detection monitoring system; the flowchart illustrates the configuration for the Chi2 test. First, the reference vector  $\tilde{\mathbf{x}}$  is set to  $\mathbf{x}_{t=0}$  and the number of anomalies *numOfAnomalies* is initialized to 0. The parameters  $\Delta$ , *thr*, *nbins*, *FS* and  $\alpha$  characterize the *Online Alert Unit*.

In normal operation, the *Online Alert Unit* continuously gets the vector  $\mathbf{x}_t$  measured at time  $t$ , and verifies the occurrence of an anomaly (as per Section 3) by computing the function *FindAnomaly*( $\mathbf{x}_t, \tilde{\mathbf{x}}, \alpha$ ). The latter algorithm may implement either the Chi2 test or the KS test, and returns 1 when the null hypothesis is rejected, and 0 otherwise. The detection of an anomaly triggers a second procedure, that verifies whether an alert should be actually prompted. Let  $T^*$  denote the time mark of the first anomaly detection. The procedure propagates an alert only if—in the time window between  $T^*$  and  $T^* + \Delta$ —the anomaly (*numberOfAnomalies*) again shows up at least *thr* times.

If the system generates an alert, the reference  $\tilde{x}$  is updated and set to  $x_{T^*}$ , i.e., to the vector acquired at the instant in which the system triggered the first anomaly. If instead the system does not detect the required number,  $thr$ , of anomalies within the time interval  $[T^*, T^* + \Delta]$ , the quantity  $numOfAnomalies$  is reset and a new acquisition starts.

In practice, the monitoring system only prompts an alert when it detects a sequence of consecutive anomalies. Thus, significant discontinuities in the status of the insulation system only generate an alert. The approach is fully unsupervised and does not require any previous knowledge about the apparatus. Moreover, the computational complexity of the whole monitoring procedure is negligible. This in turn means that the monitoring system can be deployed in low-cost, resource-limited embedded systems.



**Figure 4.** Flowchart of the online alert unit entitled to output an alert if a significant change is detected.

## 5. IoT-Based Predictive System

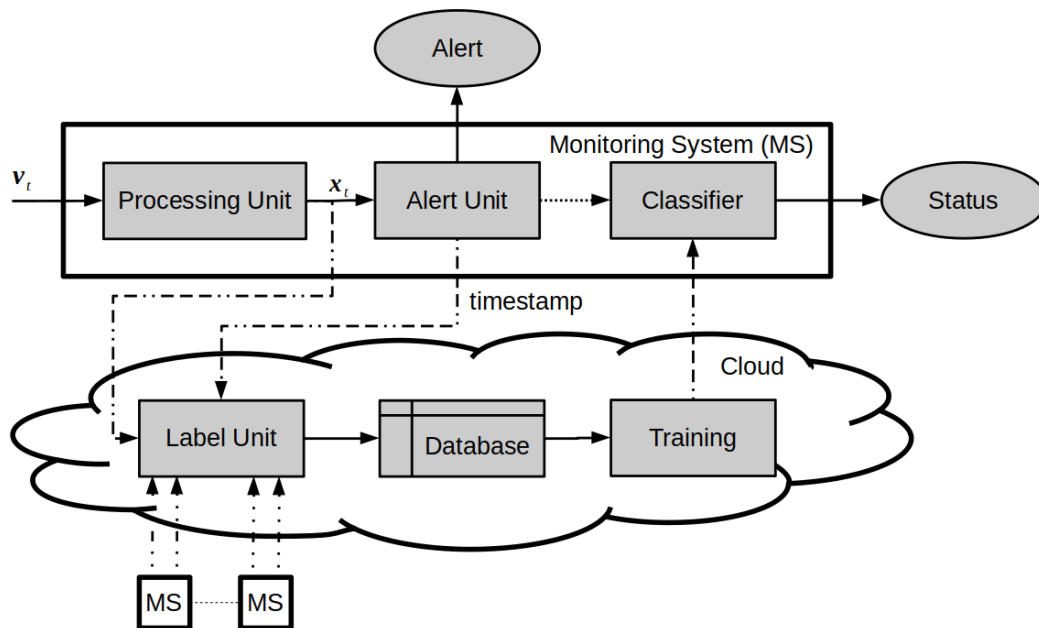
The monitoring system can be regarded as a module of a comprehensive IoT-based architecture for predictive maintenance, as per Figure 5. The framework illustrated in Figure 4 runs in an embedded system positioned in proximity of the monitored electrical machine, and provides alerts related to aging phenomena as described. Moreover, an additional, classifier module relying on Machine-Learning paradigms can be deployed to automatically predict the status of the machine at each time  $t$ . This allows estimating the proper time for maintenance.

This scenario envisions a remote data-warehouse to analyze the data collected from a variety of maintenance apparatuses, and to adjust the inference functions to be supported by local classifiers. As a major result (1) the training set can progressively collect all the available information from the connected maintenance apparatuses; and (2) the time-consuming, computationally demanding training procedure can be carried out by a central computing unit. This in turn means that the local monitoring system can take advantage of an hardware-friendly inference function that is updated as long as the training set grows.

In practice, the remote database stores the data provided by all the monitored systems (MS in Figure 5). The *Label Unit* characterizes those data along with the timestamps of the alerts, thus building a labeled dataset. In other words, the timestamps, which stem from an unsupervised process, can be used to label the data provided by online measurements. The remote data-warehouse applies machine learning methodologies to that database and computes inferences about the status of the insulation systems under investigation, given a measure  $x_t$  (or  $w_t$ ) for an apparatus. Within the paradigm of edge computing, the local embedded system just implements the inference



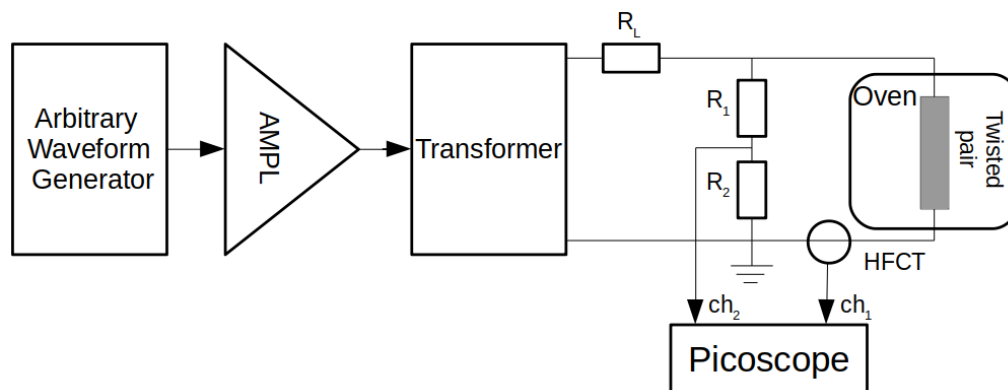
function, which receives its parameters from the cloud. The literature provides some examples of hardware-friendly implementations of inference functions [40–42] that could be integrated in our low resources system. Such works exploited the paradigm of random basis neural networks to implement an inference function supported by a single hidden-layer feedforward neural network. As a major result, the eventual classifier can achieve a satisfactory trade-off between generalization performance and computational complexity. This in turn enables the deployment of the classifier on resource-constrained devices.



**Figure 5.** Monitoring system for alert detection and classification, and cloud infrastructure collecting data from several monitoring system .

## 6. Experimental Results

The experimental session involved a set of twisted pair specimens that underwent aging tests according to standard IEC 60851-5. Figure 6 shows the measurement circuit. The sinusoidal power supply, having a frequency of 500 Hz, was provided by an arbitrary waveform generator and then amplified to reach a 3.74 KV peak to peak amplitude. A gained HFCT with a band-pass behavior in the range  $3 \div 20$  MHz, placed around the ground cable, served as sensor. The signals were sampled by a Picoscope with a bandwidth in the range  $0 \div 200$  MHz and a maximum sample frequency of 1 GSamples/s. The Picoscope collected both the signal from the twisted pair specimens and the signal from a resistive divider as phase reference. All the specimens were placed in an oven at  $150$  °C.



**Figure 6.** Power supply and measurement system.

In the following, Section 6.1 will presents the outcomes of a set of preliminary experiments, which were designed to properly adjust the resolution in the ADC. Section 6.2 will discuss the outcomes of the experiments involving the proposed online monitoring system. Finally, Section 6.3 will analyze the computational complexity of the monitoring system.

### 6.1. Preliminary Experiments: Setting the Resolution

The first experimental session aimed at evaluating the role played by the analog-to-digital conversion in the monitoring system. Thus, given a specimen under test, the signals were acquired in parallel by two Picoscope adopting a common fullscale (20 V) and different resolutions in the ADC: 8 bit and 12 bit, respectively. The 8-bit resolution actually corresponds to the most adopted configuration in the literature for sampling of PD signals. The 12-bit resolution has been included in the experiment to analyze the behavior of the anomaly detection strategy when higher sampling resolutions are available. The status of the specimen was monitored every two minutes, with a time window  $\delta = 0.5$  s.

Figure 7 shows—on a time scale—the alerts generated by the monitoring system for one of the specimens under test. In this case, the specimen breakdown occurred after 23 h. Figure 7a refers to the monitoring system adopting a 8-bit resolution. The red bars tag the alerts generated by the KS test, while the blue bars tag the alerts generated by the Chi2 test. Analogously, Figure 7b gives the alerts generated by the monitoring system adopting a 12-bit resolution. Here, the green bars correspond to KS test and the black bars correspond to the Chi2 test. In the 8-bit setup, the anomaly detection unit based on Chi2 test generated the first alert after one hour. The second alert occurred after 5 h; in the following three hours, the alerts occurred almost periodically. Then, the gap between successive alerts progressively increased as aging proceeded. The anomaly detection unit based on KS test produced fewer alerts. However, it is interesting to note that also in this case the gap between successive alerts progressively widened as the status of the specimen worsen. Possibly the higher sensitivity of the Chi2 test may depend on the setup of parameter  $nbins$ , which has been set empirically to  $nbins = 25$ . In the 12-bit setup the sequence of alerts almost follows the same trend. Again, the Chi2 test provides more sensitivity.

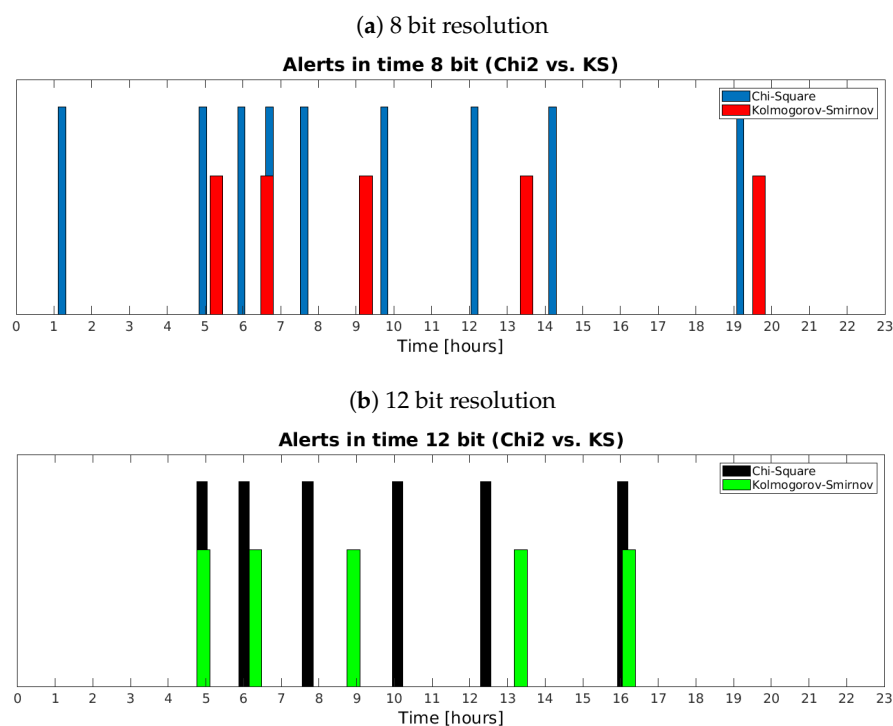


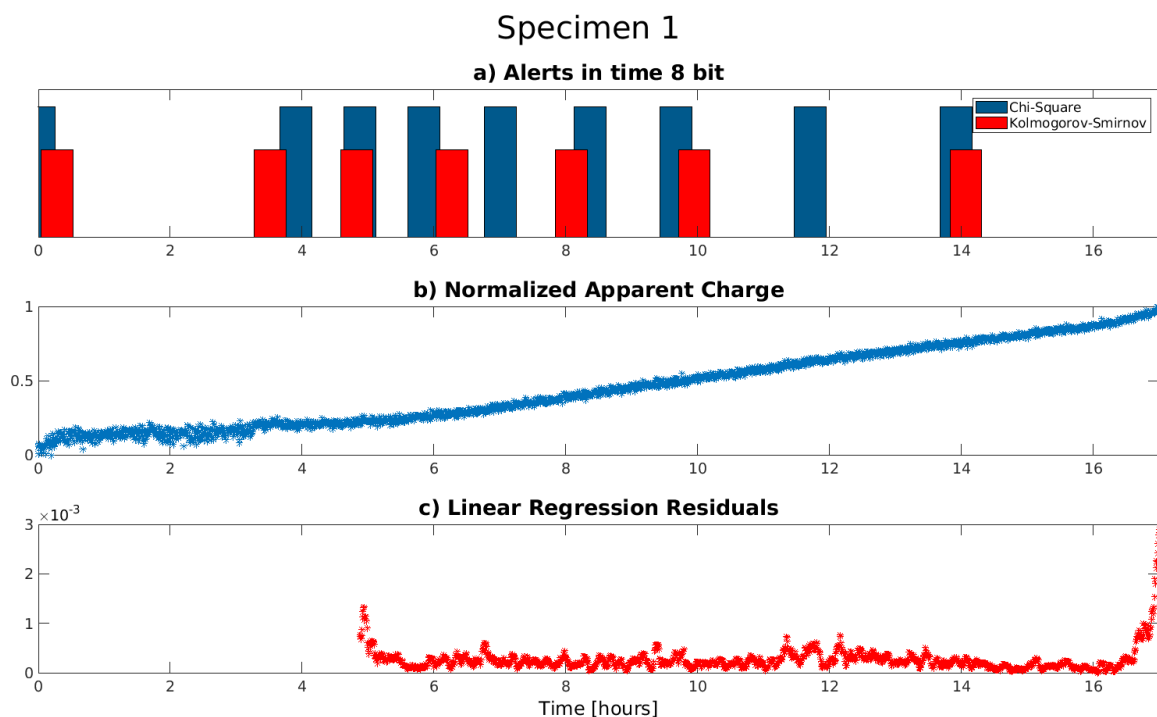
Figure 7. Alerts produced by Chi2 and KS test with resolutions of 8 and 12 bit.

In general, as expected, the experiments revealed that the 12-bit setup is more robust to the quantization noise. In this sense, the early alert generated by the Chi2 unit in the 8-bit setup of the experiment showed in Figure 7b might be the effect of a coarser quantization. Indeed, KS test proved less sensitive to changes in resolution. Overall, one should also consider that the 8-bit setup represents a more cost effective solution. In the following, only the 8-bit setup will be considered.

## 6.2. Experimental Session

The second experimental session aimed to evaluate the ability of the proposed monitoring system to support a maintenance apparatus. The session involved a total of six specimens and the measurement setup schematized in Figure 6. The PD signals have been sampled using a 8-bit resolution. The status of a specimen was monitored every minute, with a time window  $\delta = 0.5$  s.

Figure 8 gives the results obtained in the experiments involving specimen #1. The figure compares the outcomes of the proposed monitoring system with those of the state-of-the-art method describe in Section 2. The plot in Figure 8a shows the alert generated by the proposed monitoring system by adopting the same format of Figure 7a. The plot in Figure 8b gives—on the same time scale of Figure 8a—the average apparent charge per cycle  $q_a(t)$  computed as per (3); here, the values of  $q_a(t)$  have been normalized in the range  $[0, 1]$ . The plot in Figure 8c provides—on the same time scale of Figure 8a—the residuals referred to the linear regression computed on the data plotted in Figure 8 as per (4). As specified in Section 2 the linear regression is computed after removing the data belonging to the transient zone. Figure 8b proves that in this experiment the transient ended after about four hours and a half.

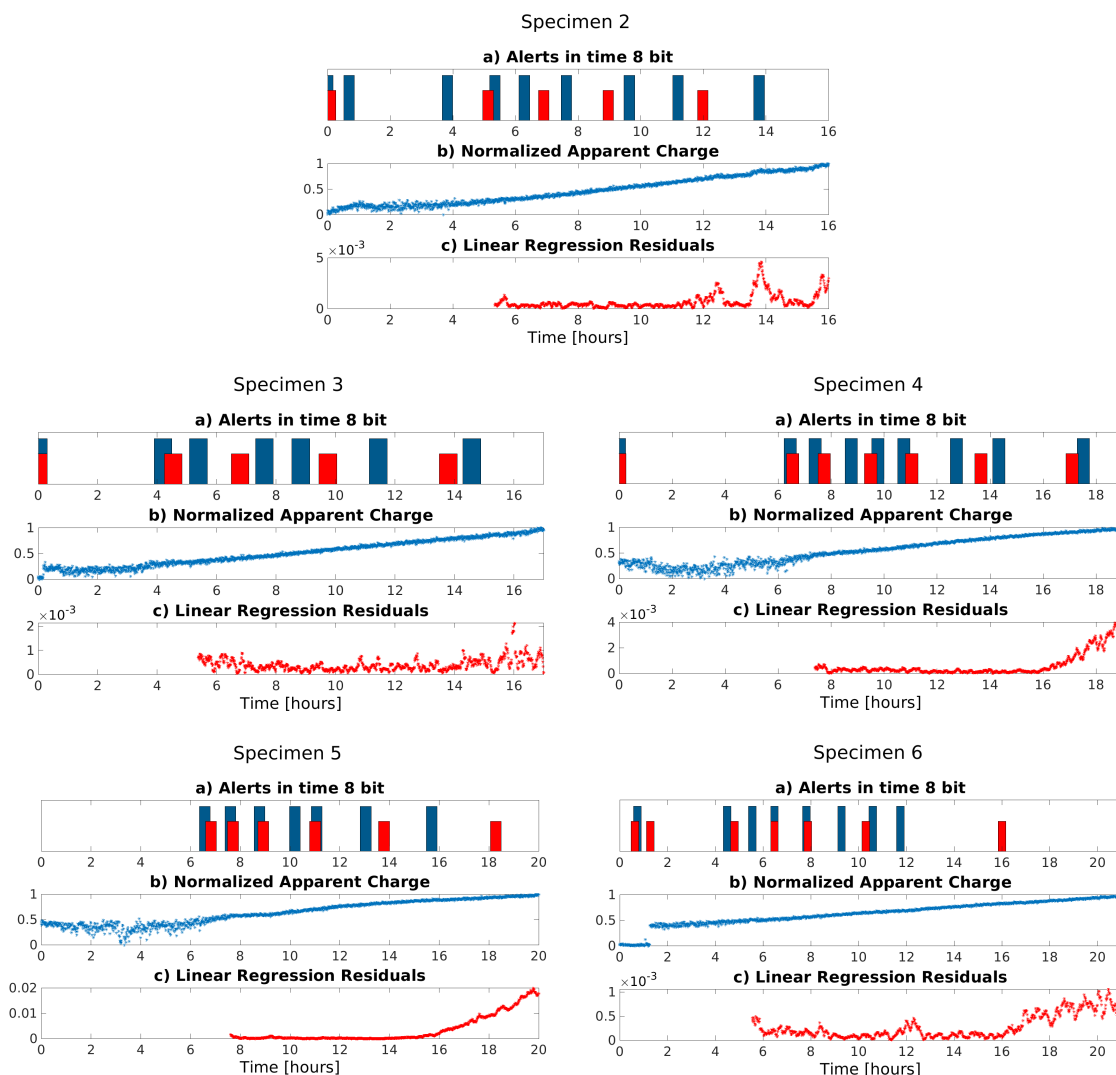


**Figure 8.** Specimen 1: (a) alerts generated by the 8 bit system; (b) normalized apparent charge computed on PD Pattern and (c) linear regression residuals on normalized apparent charge.

Figure 8 stimulates a few interesting observations. First, the sequence of periodical alerts in Figure 8a starts exactly when the average apparent charge per cycle shows that the transient has ended (as per Figure 8b). Second, the proposed monitoring system warns the observer of the incipient critical aging of the specimen well before the breakdown. Figure 8a shows that the alerts lose the periodical pattern after the 10th hour. That is, the monitoring system indicates that the specimen has

entered in a new phase, which in practice corresponds to the incipient critical aging. Actually, the same information is conveyed by the average apparent charge per cycle a few hours later. Figure 8c shows that  $q_a(t)$  deviates significantly from the linear trend only when the breakdown is very close. Overall, one may conclude that the proposed monitoring system seems able to effectively support a predictive maintenance procedure by adopting a fully unsupervised strategy.

Figure 9 shows the outcomes of the experiments involving the remaining five specimens. For each specimen, the results have been presented by adopting the same format of Figure 8. In general, the observations made for specimen #1 still hold. The sequence of periodic alerts in the proposed monitoring system marks the end of the first part of the specimen's lifetime. Indeed, the gap between subsequent alerts starts increasing when the third part of the specimen's lifetime begins. It is easy to note that this event always anticipate the alarm conveyed by the trend of the average apparent charge per cycle, which departs from the expected linear progression when the breakdown is approaching.



**Figure 9.** Specimen 2-6: (a) alerts generated by the 8 bit system and (b) normalized apparent charge computed on PD Pattern and (c) linear regression residuals on normalized apparent charge.

It is worth noting that the sequence of alerts generated by the proposed monitoring system depends on the adopted anomaly detection criterion. Thus, one may get different outcomes when adopting KS test rather than Chi2 test. In general, the Chi2 test produces more alerts than the KS test. As discussed above, such behavior might be related to the setup of the parameter  $nbins$  in the Chi2 test. However, all the experiments confirm that the monitoring system can reliably assess aging in the

specimens. Such result is very interesting in that the proposed approach does not need to rely on any knowledge base.

The experimental session also inspires further developments. The proposed monitoring system based on alerts may act as an automated trigger for the monitoring system based on the average apparent charge per cycle proposed in [26,27]. In practice, as the sequence of periodical alerts corresponds to the end of the transient, one can automatically identify the instant  $\hat{T}$  from which linearity is expected in  $q_a(t)$ . Thus, an augmented monitoring system combining the two strategies could be able to start collecting  $q_a(t)$  only when the transient has ended. Such data may be sent in real time to a central unit entitled to estimate the value of parameters  $K$ ,  $n_1$  and  $n_2$  in (6) by exploiting the data acquired from different electrical machines.

### 6.3. Computational Cost

The proposed monitoring system has been deployed on a Raspberry Pi 3 Model B+ [43] by using the Python programming language. The tests aimed at evaluating the computational cost of the two main blocks, i.e., *Processing Unit* and *Online Alert Unit*. Thus, the execution times of the two blocks have been assessed by running the monitoring system on the Raspberry Pi, which represent a suitable example of low-cost embedded computing unit. In the adopted setup, the PD signals were sampled using a 8-bit resolution. The status of a specimen was monitored every minute, with a time window  $\delta = 0.5$  s.

Table 1 gives the execution times for, respectively, the *Processing Unit*, the *Online Alert Unit* adopting the Chi2 test, and the *Online Alert Unit* adopting the KS test. For each block the average execution time and the maximum execution time are showed. As expected, the KS test brings about a larger execution time than the Chi2 test. Actually, in the proposed system the KS test processes  $\mathbf{w}$ , which in general is an high-dimensional vector; conversely, the Chi2 test processes  $\mathbf{x}$ , which store the occurrences of an histogram with  $n_{bins} = 25$ .

**Table 1.** Estimated execution times for the computing blocks involved in the proposed monitoring system.

	Processing Unit [ms]	Chi2 [ms]	KS [ms]
<b>Average Time</b>	2.3	1.7	60.3
<b>Maximum Time</b>	4.4	42.4	132.0

Overall, the outcomes of this experiment proves that the monitoring system can achieve real-time performance on a low-cost embedded computer. In this sense, one should consider that in a real case the insulation systems aging is a very slow process. While in the proposed experiment the time interval between two acquisitions was set to one minute, in general one would set this time to a few hours. Accordingly, the execution times achieved on the Raspberry Pi can be considered insignificant on that time scale.

## 7. Conclusions

This paper shows that anomaly detection paradigms can support a fully unsupervised monitoring system for predictive maintenance of high voltage apparatus. The proposed method is computationally light and fits IoT solutions that rely on edge computing. The monitoring system can identify in real time the significant changes in the status of the apparatus, thus revealing aging effects. At the same time, the system enables the automated labeling of acquired data, which become structured information to be stored and processed by a central unit collecting the outcomes of several apparatuses. Eventually, the unsupervised strategy can be locally augmented by a supervised inference model that is driven by such central unit.

The experimental session led to the following observations:

- the unsupervised system properly detects the changes in the status of the specimens, thus enabling a maintenance schedule that can avoid unwanted breakdowns.
- When adopting the Chi2 test to implement anomaly detection a systematic pattern can be identified. After a transient period corresponding to the first part of the specimen's life, the alerts assumes a periodical trend, which ends when the aging starts to severely affect the specimen.
- The adoption of a coarser resolution in the ADC may somewhat increase the sensitivity of Chi2 test, which in turn may be adjusted by properly tuning the *nbins* parameter.
- The effectiveness of the proposed fully unsupervised strategy has been confirmed by the comparison with the state-of-the art supervised strategy based on the average apparent charge per cycle [26,27].
- The monitoring systems has real-time performance on a low-resources device.

Future directions mostly involve the employment of the online unsupervised monitoring system to assess the aging of more complex apparatuses. Moreover the IoT-based maintenance apparatus will be developed. As discussed in Section 6.2, the proposed strategy may combined with the strategy presented in [26,27] to further improve the effectiveness of the aging assessment. Thus, local information collected via the unsupervised system can be exploited at a central level to properly set the parameters in (6). Nonetheless, the general scheme of Figure 5 could also exploit a different inference model.

**Author Contributions:** Conceptualization, C.G., E.R. and P.G.; data curation, A.B. and C.G.; formal analysis, C.G. and E.R.; investigation, C.G. and A.B.; methodology, C.G., E.R. and P.G.; project administration, F.G.; resources, F.G., P.G. and R.Z.; software, C.G. and E.R.; validation, C.G. and E.R.; visualization, all authors; writing—original draft preparation, all authors; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

**Funding:** Authors received no specific funding for this work.

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

## Abbreviations

The following abbreviations are used in this manuscript:

AC	Alternate Current
ADC	Analog Digital Converter
Chi2	Chi-Square
CNN	Convolutional Neural Network
DC	Direct Current
DNN	Deep Neural Network
GIS	Gas Insulated System
HFCT	High Frequency Current Transformer
IoT	Internet of Things
KS	Kolmogorov–Smirnov
PD	Partial Discharge
UHF	Ultra High Frequency

## References

1. De Faria, H., Jr.; Costa, J.G.S.; Olivas, J.L.M. A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis. *Renew. Sustain. Energy Rev.* **2015**, *46*, 201–209. [[CrossRef](#)]
2. Selcuk, S. Predictive maintenance, its implementation and latest trends. *Proc. Inst. Mech. Eng. Part J. Eng. Manuf.* **2017**, *231*, 1670–1679. [[CrossRef](#)]
3. Stone, G.C. A perspective on online partial discharge monitoring for assessment of the condition of rotating machine stator winding insulation. *IEEE Electr. Insul. Mag.* **2012**, *28*, 8–13. [[CrossRef](#)]
4. Luo, Y.; Li, Z.; Wang, H. A review of online partial discharge measurement of large generators. *Energies* **2017**, *10*, 1694. [[CrossRef](#)]

5. *High-Voltage Test Techniques—Partial Discharge Measurements*; International Standard IEC 60270:2000; International Electrotechnical Commission: Geneva, Switzerland, 2000.
6. Wu, M.; Cao, H.; Cao, J.; Nguyen, H.L.; Gomes, J.B.; Krishnaswamy, S.P. An overview of state-of-the-art partial discharge analysis techniques for condition monitoring. *IEEE Electr. Insul. Mag.* **2015**, *31*, 22–35. [[CrossRef](#)]
7. Fofana, I.; Hadjadj, Y. Electrical-based diagnostic techniques for assessing insulation condition in aged transformers. *Energies* **2016**, *9*, 679. [[CrossRef](#)]
8. Li, S.; Li, J. Condition monitoring and diagnosis of power equipment: review and prospective. *High Voltage* **2017**, *2*, 82–91. [[CrossRef](#)]
9. Gao, W.; Ding, D.; Liu, W. Research on the typical partial discharge using the UHF detection method for GIS. *IEEE Trans. Power Deliv.* **2011**, *26*, 2621–2629. [[CrossRef](#)]
10. Wang, M.H. Partial discharge pattern recognition of current transformers using an ENN. *IEEE Trans. Power Deliv.* **2005**, *20*, 1984–1990. [[CrossRef](#)]
11. Hao, L.; Lewin, P.; Hunter, J.; Swaffield, D.; Contin, A.; Walton, C.; Michel, M. Discrimination of multiple PD sources using wavelet decomposition and principal component analysis. *IEEE Trans. Dielectr. Electr. Insul.* **2011**, *18*, 1702–1711. [[CrossRef](#)]
12. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd* **1996**, *96*, 226–231.
13. Álvarez, F.; Garnacho, F.; Ortego, J.; Sánchez-Urán, M. Application of HFCT and UHF sensors in on-line partial discharge measurements for insulation diagnosis of high voltage equipment. *Sensors* **2015**, *15*, 7360–7387. [[CrossRef](#)]
14. Gianoglio, C.; Ragusa, E.; Bruzzone, A.; Gastaldo, P.; Torello, E.; Guastavino, F. Tensor Based Algorithm for Automatic Partial Discharges Pattern Classification. In Proceedings of the 2019 21st European Conference on Power Electronics and Applications (EPE'19 ECCE Europe), Genova, Italy, 3–5 September 2019; pp. P.1–P.8.
15. Catterson, V.; Sheng, B. Deep neural networks for understanding and diagnosing partial discharge data. In Proceedings of the 2015 IEEE Electrical Insulation Conference (EIC), Seattle, WA, USA, 7–10 June 2015, pp. 218–221.
16. Song, H.; Dai, J.; Sheng, G.; Jiang, X. GIS partial discharge pattern recognition via deep convolutional neural network under complex data source. *IEEE Trans. Dielectr. Electr. Insul.* **2018**, *25*, 678–685. [[CrossRef](#)]
17. Li, G.; Rong, M.; Wang, X.; Li, X.; Li, Y. Partial discharge patterns recognition with deep Convolutional Neural Networks. In Proceedings of the 2016 International Conference on Condition Monitoring and Diagnosis (CMD), Xi'an, China, 25–28 September 2016; pp. 324–327.
18. Dey, D.; Chatterjee, B.; Dalai, S.; Munshi, S.; Chakravorti, S. A deep learning framework using convolution neural network for classification of impulse fault patterns in transformers with increased accuracy. *IEEE Trans. Dielectr. Electr. Insul.* **2017**, *24*, 3894–3897. [[CrossRef](#)]
19. Li, G.; Wang, X.; Li, X.; Yang, A.; Rong, M. Partial discharge recognition with a multi-resolution convolutional neural network. *Sensors* **2018**, *18*, 3512. [[CrossRef](#)]
20. Wan, X.; Song, H.; Luo, L.; Li, Z.; Sheng, G.; Jiang, X. Pattern recognition of partial discharge image based on one-dimensional convolutional neural network. In Proceedings of the IEEE 2018 Condition Monitoring and Diagnosis (CMD), Perth, WA, Australia, 23–26 September 2018; pp. 1–4.
21. Wang, Y.; Yan, J.; Sun, Q.; Li, J.; Yang, Z. A MobileNets Convolutional Neural Network for GIS Partial Discharge Pattern Recognition in the Ubiquitous Power Internet of Things Context: Optimization, Comparison, and Application. *IEEE Access* **2019**, *7*, 150226–150236. [[CrossRef](#)]
22. Banno, K.; Nakamura, Y.; Fujii, Y.; Takano, T. Partial Discharge Source Classification for Switchgears with Transient Earth Voltage Sensor Using Convolutional Neural Network. In Proceedings of the IEEE 2018 Condition Monitoring and Diagnosis (CMD), Perth, WA, Australia, 23–26 September 2018; pp. 1–5.
23. Tenbohlen, S.; Coenen, S.; Djamali, M.; Müller, A.; Samimi, M.H.; Siegel, M. Diagnostic measurements for power transformers. *Energies* **2016**, *9*, 347. [[CrossRef](#)]
24. Yazici, B. Statistical pattern analysis of partial discharge measurements for quality assessment of insulation systems in high-voltage electrical machinery. *IEEE Trans. Ind. Appl.* **2004**, *40*, 1579–1594. [[CrossRef](#)]
25. Jannah, R.R.; Khayam, U. Design, implementation, and testing of partial discharge signal pattern recognition and judgment system application using statistical method. In Proceedings of the IEEE ICEVT & IMECE, Surakarta, Indonesia, 4–5 November 2015; pp. 314–318.

26. Guastavino, F.; Dardano, A.; Torello, E. Measuring partial discharges under pulsed voltage conditions. *IEEE Trans. Dielectr. Electr. Insul.* **2008**, *15*, 1640–1648. [[CrossRef](#)]
27. Guastavino, F.; Gianoglio, C.; Torello, E.; Ferraris, M.; Gianelli, W. Electrical Aging Tests on Conventional and Nanofilled Impregnation Resins. In Proceedings of the 2018 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), Cancun, Mexico, 21–24 October 2018; pp. 156–158.
28. Ahmed, Z.; Kluss, J.; Wallace, D. Partial discharge measurements and techniques for pattern recognition and life prediction of medium voltage XLPE cables. In Proceedings of the 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Perth, WA, Australia, 23–26 September 2018; pp. 1–4.
29. Caprara, A.; Cavallini, A.; Garagnani, L.; Guo, J. A novel approach for continuous monitoring of Partial Discharge phenomena on Medium Voltage equipments. In Proceedings of the 2018 IEEE Electrical Insulation Conference (EIC), San Antonio, TX, USA, 17–20 June 2018; pp. 495–498.
30. Cavallini, A.; Contin, A.; Montanari, G.C.; Puletti, F. Advanced PD inference in on-field measurements. I. Noise rejection. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 216–224. [[CrossRef](#)]
31. Cavallini, A.; Conti, M.; Contin, A.; Montanari, G. Advanced PD inference in on-field measurements. II. Identification of defects in solid insulation systems. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 528–538. [[CrossRef](#)]
32. *IEEE Guide for Diagnostic Field Testing of Electric Power Apparatus—Electrical Machinery*; International Standard IEEE 62.2:2004; IEEE Power and Energy Society: Piscataway, NJ, USA, 2004.
33. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. *ACM Comput. Surv. (CSUR)* **2009**, *41*, 1–58. [[CrossRef](#)]
34. Aissa, N.B.; Guerroumi, M. Semi-supervised statistical approach for network anomaly detection. *Procedia Comput. Sci.* **2016**, *83*, 1090–1095. [[CrossRef](#)]
35. Ye, N.; Chen, Q. An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems. *Qual. Reliab. Eng. Int.* **2001**, *17*, 105–112. [[CrossRef](#)]
36. Yu, M. A nonparametric adaptive CUSUM method and its application in network anomaly detection. *Int. Adv. Comput. Technol.* **2012**, *4*, 280–288.
37. Manikopoulos, C.; Papavassiliou, S. Network intrusion and fault detection: A statistical anomaly approach. *IEEE Commun. Mag.* **2002**, *40*, 76–82. [[CrossRef](#)]
38. Goldstein, M.; Dengel, A. *Histogram-Based Outlier Score (Hbos): A Fast Unsupervised Anomaly Detection Algorithm*; German Research Center for Artificial Intelligence (DFKI): Kaiserslautern, Germany, 2012; pp. 59–63.
39. Massey, F.J., Jr. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [[CrossRef](#)]
40. Gianoglio, C.; Guastavino, F.; Ragusa, E.; Bruzzone, A.; Torello, E. Hardware Friendly Neural Network for the PD Classification. In Proceedings of the 2018 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), Cancun, Mexico, 21–24 October 2018; pp. 538–541.
41. Ragusa, E.; Gianoglio, C.; Gastaldo, P.; Zunino, R. A digital implementation of extreme learning machines for resource-constrained devices. *IEEE Trans Circuits Syst. II: Express Briefs* **2018**, *65*, 1104–1108. [[CrossRef](#)]
42. Ragusa, E.; Gianoglio, C.; Zunino, R.; Gastaldo, P. A Design Strategy for the Efficient Implementation of Random Basis Neural Networks on Resource-Constrained Devices. *Neural Process. Lett.* **2019**, 1–19, doi:10.1007/s11063-019-10165-y. [[CrossRef](#)]
43. Raspberry Pi Foundation. Raspberry Pi 3 Model B+. Available online: <https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/> (accessed on 28 February 2020).

