Site Characterization Index for Continuous Power Quality Monitoring Based on Higher-order Statistics

Olivia Florencias-Oliveros, Juan-José González-de-la-Rosa, Jose-María Sierra-Fernández, Agustín Agüera-Pérez, Manuel-Jesús Espinosa-Gavira, and José-Carlos Palomares-Salas

***Abstract*—The high penetration of distributed generation (DG) has set up a challenge for energy management and consequently for the monitoring and assessment of power quality (PQ). Besides, there are new types of disturbances owing to the uncontrolled connections of non-linear loads. The stochastic behaviour triggers the need for new holistic indicators which also deal with big data of PQ in terms of compression and scalability so as to extract the useful information regarding different network states and the prevailing PQ disturbances for future risk assessment and energy management systems. Permanent and continuous monitoring would guarantee the report to claim for damages and to assess the risk of PQ distortions. In this context, we propose a measurement method that postulates the use of two-dimensional (2D) diagrams based on higher-order statistics (HOSs) and a previous voltage quality index that assesses the voltage supply waveform in a continous monitoring campaign. Being suitable for both PQ and reliability applications, the results conclude that the inclusion of HOS measurements in the industrial metrological reports helps characterize the deviations of the voltage supply waveform, extracting the individual customers' pattern fingerprint, and compressing the data from both time and spatial aspects. The method allows a continuous and robust performance needed in the SG framework. Consequently, the method can be used by an average consumer as a probabilistic method to assess the risk of PQ deviations in site characterization.**

***Index Terms*—Continuous statistical monitoring, big data, data compression, higher-order statistics (HOSs), power quality (PQ).**

I. Introduction

THE high penetration of renewable energy resources, as expected in the future smart grid (SG), establishes a challenging scenario for energy management. Consequently, it is driving and conducting the design of emerging monitoring equipment for power quality (PQ) and the strategies regarding the application and the associated reports. SG demands more efficient systems and monitoring algorithms deployed in advanced measurement infrastructure that creates a more flexible power system [1], [2]. Indeed, new PQ disturbances are appearing due to the uncontrolled connections of non-linear loads at different levels in the electrical network, e.g., from low voltage (LV) to high voltage (HV). Thus, it is a fundamental issue to create efficient monitoring strategy for different multiple disturbances in power system modernization. As a result, PQ is investigated in several areas, e.g., statistical signal processing, instrumentation and measurement for internet of things (IoT). New analytical tools should be developed to comply with a more complex online PQ analysis [3]-[5].

So far, European measurement campaigns have been traditionally based on widely accepted meters developed in accordance with IEC 61000-4-30 [6]. However, new measurement solutions are needed to fulfill the specific campaign requirements. PQ monitoring demands a more efficient data management strategy which ensures flexible reporting, e.g., time- and space-varying scales, which improves the site monitoring and allows measurement traceability and repeatability as well [7]. Moreover, current grid performance analysis conveys the characterization of multiple measurement locations, which produces huge volumes of data that require proper organization, i.e., big data. This in turn must be based on the elimination of redundant and erroneous information and the formulation of "ad hoc" indicators that bring together the significance of the measurements adapted to the requirements demanded by the customer [8], [9].

The reporting levels along with the measurement allocations are usually interpreted through the PQ triangle [10]. The graphical representation consists of a data framework in which the concepts of time and space compression are associated with each physical element within the entire electrical network, along with the magnitudes and the types of distur-

Manuscript received: February 24, 2020; revised: July 7, 2020; accepted: September 14, 2020. Date of CrossCheck: September 14, 2020. Date of online publication: November 20, 2020.

This work was supported by the Spanish Ministry of Science and Innovation (Statal Agency for Research), and the EU (AEI/FEDER/UE) via project PID2019-108953RB-C21 Strategies for Aggregated Generation of Photovoltaic Plants: Energy and Meteorological Operational Data (SAGPVEMOD), and the precedent TEC2016-77632-C3-3-R.

O. Florencias-Oliveros, J.-J. González-de-la-Rosa (corresponding author), J.-M. Sierra-Fernández, A. Agüera-Pérez, M.-J. Espinosa-Gavira, and J.-C. Palomares-Salas are with the Department of Automation Engineering, Electronics, Architecture and Computers Networks, Research Group TIC168, Higher Polytechnic School, University of Cádiz, Algeciras 11202, Spain (e-mail: olivia.florencias@uca. es; juanjose. delarosa@uca. es; josemaria. sierra@uca. es; agustin. aguera@uca.es; manuel.espinosa@uca.es; josecarlos.palomares@uca.es).

DOI: 10.35833/MPCE.2020.000041

bances. The new scheme of measurement campaigns is the result of more than a decade during which the scientific community has expressed serious concerns about the strategy adopted in the reports of the data that result from the PQ monitoring [11]. To tackle this issue, the method must be twofold. On one hand, it must be based on the standardized indices and standards. On the other hand, it should incorporate the elements that provide the monitoring campaigns with its own functionalities adapted to particular customer's requirements.

In fact, along with the deregulation of the market that comes with the renewable energy, specific PQ measurements are called to contribute to improvements in the compatibility between consumers and grid operator's solutions. The idea of using more understandable indicators lies not only for energy suppliers but also for the end users, producers and consumers (prosumers). The EN 50160 [12] does not differentiate the responsibilities between suppliers and end-consumers in the point of common coupling (PCC), and the physical connections between the power grid and the end-user [13]. Indeed, this standard describes the voltage supply and deviations in the PCC. From the consumer side, the user must fulfill the harmonized electromagnetic compatibility (EMC) standards accomplishing the immunity and emission [14]. Thus, in the power grid context, a two-flow energy takes place at PCC. Besides, current standards have been developed when linear consumers dominate the distribution networks and they must consider the end-consumer's influence in the SG context [15].

A PQ index should assess the performance of continuous and discrete electrical disturbances. These two strategies are usually based on the techniques that compress the acquired time-series cycle by cycle, extracting the information both in the time and frequency domains. For instance, the site indices may unify the measurements and compute individual weekly percentiles at different physical layers along the entire network. Nevertheless, far from being updated, PQ norms and standards still do not gather sufficiently flexible standardized measurement methods, e. g., implemented in new meters. Therefore, for continuous monitoring, it is necessary to incorporate new indices and ensure that their compliance remains within 95% of confidence interval within a week [6]. Current power measurement strategies are definitely insufficient to achieve a full characterization and cope with the current electrical disturbances and consequently bound its causes [8], [16], [17].

Based on normal operation conditions, we develop a long-term PQ campaign on how to characterize the network, looking out its deviation from the ideal steady state. The proposal makes use of an index based on higher-order statistics (HOSs). The performance is introduced in class *S* of the instruments according to IEC 61000-4-30 [6], and the flexibility is added to the measurements and the reporting which match the current SG context [17] for possible class *M* instruments [8]. Four new levels of PQ monitoring functionality are then covered. The main highlights of the research are as follows. First, HOSs are capable of characterizing the waveform distortion by monitoring their probability density function (PDF) as well as reporting the symmetry, amplitude and tail deviations. Thus, for continuous analysis, we propose a global deviation index of PQ based on HOSs. Second, these measurements help characterize the waveform in a consumer installation under normal operation conditions in different measurement scenarios in a permanent regime, i.e., daily or weekly. This contributes to improving future prediction tools based on artificial intelligence. The authors' previous work [18] settled down the basis of the index and showed a preliminary controlled experience that has been tested and validated in the current work.

The rest of paper is organized as follows. Section II exposes the need for continuous monitoring based on HOS. The measurement and analysis are presented in Section III. The analysis results are provided in Section IV. Section V presents the discussion of the results and the contributions of the proposed methods. Finally, Section VI concludes the paper.

II. The Need for Continuous Monitoring Based on HOS

A. PQ Index

Recently, managing information is proposed in measurement solutions by converting big data from each smart meter into a series of probability distributions, calculating the pairwise distance between load profiles. Each long series of demand data are transformed into a single two-dimensional (2D) diagram [19], and become more traceable and understandable for both customers and suppliers. We propose a 2D diagram of HOS (2D-HOS) that characterizes the voltage, compresses the time-domain information, and computes the evolution of the quality vector of the waveform shape cycle by cycle. It contributes to the main PQ issues including data compression, an scalable index, and continuous monitoring based on a more feasible solution. The proposed method differs from the conventional PQ measurements which are based on the second-order magnitudes, e.g., the power spectrum and the total harmonic distortion (THD). Considering a complementary resource that contributes to improve the PQ with a global indicator that informs about the waveform, the proposed method helps characterize the point in test and monitor the network in an holistic way. It is allowed to trigger a more specific off line in case the parameters fall outside the normal operation conditions and overcome the contractual requirements.

In [18], the index is primarily tested during a short-term one-week campaign, which is an effective tool for voltage characterization using HOS. Two strategies are extracted: the first strategy detects the extreme values of the PQ events; and the second one consists of the continuous monitoring patterns of working and non-working days. Based on a long-term horizon, the method consists of a site characterization, which compares the performance in a public building with an isolated consumer. Indeed, the voltage is characterized during six weeks considering the individual indices that integrate the aggregated PQ index as the variance, the skewness and the kurtosis. Extracting histograms of the different weekly indices based on cumulative density function (CDF) and

PDF helps characterize the voltage behaviour under normal operation conditions. In addition, data compression is explored, and the PQ index proves global effectiveness, which detects the network state in terms of voltage waveform deterioration with less computation load. The PQ fingerprint in the HOS space, e. g., clusters regions, along with the time evolution of the PQ index manages to characterize the network at different time stamps with different resolutions. Thus, the flexibility and the potential of the method are shown to incorporate solutions on future instrumentation. For instance, as a direct result, consumer's fingerprints related to waveform characteristics and day ahead are characterized.

The HOS-based indices introduce shape parameters of the signal which are not traditionally included in the norms. They usually deal with second-order measurements. As stated in the following sections, the variance detects the changes in the amplitude as a result of power change, which is indicative of sags and swells [20]. In addition, the variance exhibits a linear behaviour with the root mean squrare (RMS) value and characterizes the power supply under continuous normal operation conditions, which is the main goal of the proposed method. The skewness can be positive or negative depending on the sizes of the right and left tails of the PDF, as shown in Fig. 1. If Fig. 1(a) and (c) is more evident than Fig. 1(b) and (d), the skewness is negative, vice versa.

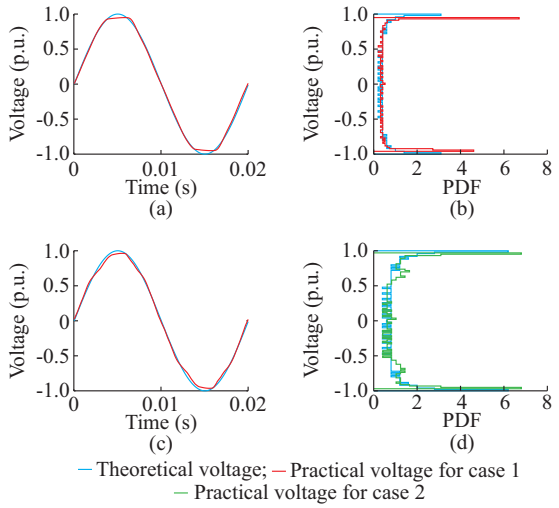

Fig. 1. Voltage supply and PDF of two steady-state voltage signals compared with the theoretical for cases 1 and 2. (a) Waveform for case 1. (b) PDF for case 1. (c) Waveform for case 2. (d) PDF for case 2.

In PQ events, the non-symmetry behaviour generally indicates the half-cycle in which the deviation takes place. The skewness detects transients and the non-symmetry of the initial and end cycles of events such as sag and swells. The kurtosis characterizes the tails of the statistical distribution. In the bimodal distribution of a voltage sinusoidal cycle of 50 Hz as shown in Fig. 1, the tails of the distribution are the maximum and minimum values. Indeed, the kurtosis measures the outliers in the form of heavy tails, mainly in the region of the maximum and minimum waveform values. In the context of voltage waveforms, the flatter the top and down regions are, the lower the kurtosis will be [21], [15].

Regarding the feature extraction stage, the importance lies in the previous preprocessing that guarantees the maximum probability of detecting not only the events when performing permanent monitoring but also the network behaviours under normal operation condition. The information is crucial for the power recognition systems [22], [19] in terms of statistical behaviours. Also, the PQ campaigns must produce reliable data that accomplish the metrological requirements from the Bayesian perspective and reduce the uncertainty in the process control [23]. In literature, the detection and characterization of real-time events are not improved [24] and examples on the characterization of multiple events in real-time monitoring can not be found frequently [25]. Thus, the current PQ technology lacks a generalized solution to the problem that handles all different events, i.e., single and multiple events [25], [26]. Other signal processing tools such as wavelets [27] and estimators based on echo state networks [28] have provided good results in shape identifications, but the noise influence is still noticeable. Also, there is a need for targeting new types of disturbances, specially hybrids. Considering the type of disturbances, it seems that the harmonics are the most necessary family of disturbances that requires special attention in continuous campaigns [21], [27], [28].

The literature review has revealed the need for new analyzing tools, which track the waveform continuously, rather than for the power change, e. g., THD. Thus, an alternative tool is needed for the traditional second-order time-domain indicators in permanent PQ surveillance applications [15]. The use of HOS is motivated hereinafter by comparing its performance with an indicator of the signal power. In Table I, the HOS performance is simulated and compared with one of the traditional waveform indicators as the crest factor (CF) for different simulated time-domain signals.

TABLE I
COMPARISON OF HOS PERFORMANCE WITH CF FOR DIFFERENT SIMULATED TIME-DOMAIN SIGNALS UNDER HARMONIC DISTORTION

Simulation condition			Statistical PQ index				CF
Interval (s)	Harmonic	Waveform behaviour	Variance	Skewness	Kurtosis	PQ index	
0-0.4	Ideal	Symmetric-sinusoidal	0.5000	0	1.5000	0	1.00
0.4-0.8	Second order	Symmetric-nonsinusoid	0.4565	0	1.6109	0.1546	1.04
0.8-1.2	Third order	Symmetric-nonsinusoid	0.6452	0	1.2222	0.4229	1.02
0.2-1.6	Fifth order	Symmetric-nonsinusoid	0.2785	0	1.9800	0.7017	1.01
0.2-1.6	Seventh order	Symmetric-nonsinusoid	0.2905	0	2.1621	0.8718	1.01
0.6-2.0	Nineth order	Symmetric-nonsinusoid	0.2513	0	2.2417	0.9905	1.00

In each time interval, a new harmonic is incorporated to the simulation [29]. It is observed that as the number or harmonics increases, the PQ index, which is a combination of variance, skewness, and kurtosis in absolute values, is capable of detecting harmonic additions. However, CF remains unchanged [29]. The skewness is zero because all signals are symmetric. Reference [30] introduces statistical waveform similarity metrics, using successive cycles of the signal and detecting local outliers. The proposed method is similar but using HOSs, which also provides noise immunity and non-Gaussian characterization.

B. HOS-based PQ Index

HOS estimation has been proposed through the last decade to infer new statistical characteristics associated with the data from non-Gaussian time-series in predominant Gaussian background, which can be theoretically considered as a result of the summation of different noise processes. Within the context of PQ disturbance detection, the targeted electrical disturbance is always considered as non-Gaussian, while the floor is assumed to be a stationary Gaussian signal [20].

With an r^{th}-order real-valued stationary random process (original analogue time-series) $x(t)$, we define a set of random variables (time-series) given by:

$$x(t), x(t+\tau_1), x(t+\tau_2), \ldots, x(t+\tau_{r-1}) \tag{1}$$

where t is the discrete time; and $\tau_r = rT_s$ is the r^{th}-order time shift applied to the original data time-series, and T_s is the sampling period. The joint (compacted notation) r^{th}-order cumulant of the random variables $C_{r,x}(\cdot)$ is given by:

$$C_{r,x}(\tau_1, \tau_2, \ldots, \tau_{r-1}) \equiv Cum\left[x(t), x(t+\tau_1), \ldots, x(t+\tau_{r-1})\right] = E\left\{x(t)x(t+\tau_1)x(t+\tau_2)\ldots x(t+\tau_{r-1})\right\} \tag{2}$$

where $Cum[\cdot]$ denotes the cumulant function; and $E\{\cdot\}$ is the expected value. The concept of cumulant is defined in (2) as the autocorrelation between the original time-series and their time-shifted versions. In other words, the cumulants quantify the mathematical similitude between two of more time-series. Depending on the cumulants' orders, the interdependency will lead to specific state of the system under test, and will enable the inference of some properties related to the system behaviour. By using (2), the most common cases of the cumulants are the second-, third- and fourth-order versions defined as:

$$C_{2,x}(\tau) = E\{x(t)x(t+\tau)\} \tag{3}$$

$$C_{3,x}(\tau_1, \tau_2) = E\{x(t)x(t+\tau_1)x(t+\tau_2)\} \tag{4}$$

$$C_{4,x}(\tau_1, \tau_2, \tau_3) = E\left\{x(t)x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)\right\} - C_{2,x}(\tau_1)\cdot C_{2,x}(\tau_2-\tau_3) - C_{2,x}(\tau_2)C_{2,x}(\tau_3-\tau_1) - C_{2,x}(\tau_3)C_{2,x}(\tau_1-\tau_2) \tag{5}$$

For a non-zero mean process, with no time shifting, we have the well-known statistics:

$$\gamma_{2,x} = E\{x^2(t)\} = C_{2,x}(0) \tag{6}$$

$$\gamma_{3,x} = E\{x^3(t)\} = C_{3,x}(0,0) \tag{7}$$

$$\gamma_{4,x} = E\{x^4(t)\} - 3\gamma_{2,x}^2 = C_{4,x}(0,0,0) \tag{8}$$

Since HOSs have succeeded in other fields, e.g., vibration mechanics, acoustic detection of insects, they are suitable for the new power grid. It contributes significantly to the classification of electrical disturbances since it addresses not only the instantaneous power, but also those associated with waveform [20], [31]. The circumstance confers HOSs with suitable characteristics to analyze the new type of multiple disturbances that occur in the electrical network with distributed generation.

The strategy consists of calculating three statistics: the variance, the skewness and the kurtosis. In the case of an ideal voltage supply of 50 Hz, this triplet takes the reference values of 0, 0.5, 1.5. This is assumed as the steady state from which deviations are measured in the HOS planes. For practical purposes, the absolute deviation index is not null, since the power line is not pure at all. Likewise, by gathering the measurements of each record in three statistical parameters, the memory savings are notable, which indicate that an index of these characteristics is suitable for dealing with big data.

In order to define the generic index, the following magnitudes are introduced: Δt is the measurement interval; and Δt contains M periods of the power signal; s_{ij} is the j^{th} sample statistic associated with the i^{th} period; $\hat{s}_j$ is the expected j^{th} statistic; and N is the number of statistics; the PQ deviation index $PQ_{\Delta t}$ is a function of the specific deviations of each individual statistic with respect to their expected values, and it is given by the general expression as:

$$PQ_{\Delta t} = f(|s_{11} - \hat{s}_1|, \ldots, |s_{ij} - \hat{s}_j|, \ldots, |s_{MN} - \hat{s}_N|) \tag{9}$$

While the theoretical value for the index is zero, i.e., each statistic equals its estimates, in practice it has to be calibrated depending on the location of the point under test and the specific operation conditions. A particular case of (9) consists of using the summation of each individual deviation:

$$PQ_{\Delta t} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|s_{ij} - \hat{s}_j|}{M} \tag{10}$$

The deviations of each statistic from its ideal value assess the waveform. Three deviations terms are used in (10). The final expression for the PQ deviation index is described as:

$$PQ_{\Delta t} = \frac{\sum_{i=1}^{M}|var_i - \widehat{var}| + |sk_i - \widehat{sk}| + |kur_i - \widehat{kur}|}{M} \tag{11}$$

where var_i, sk_i, and kur_i are the variance, the skewness, and the kurtosis of the i^{th} period, respectively; and the symbol $\widehat{\ }$ denotes the estimated value. The indices measure the quality of the voltage in terms of the waveform from a statistical point of view. To illustrate the concept of statistical distribution, i.e., a steady-state voltage supply, understand their deviations, and introduce the definition of compression, two different cases under normal power delivery conditions are compared in Fig. 1. Case 1 corresponds to a public building in the University of Cádiz, Spain, and case 2 corresponds to a household user. The purpose of time compression is to reduce all the data within a cycle into a single triplet, which is composed of three statistics compared with the ideal values

within the expression of the index.

In Fig. 1, the real measurements of the two waveforms are computed through the HOS, in case 1, $var=0.4826$, $sk=-0.0067$, $kur=1.5092$, and in case 2, $var=0.4720$, $sk=-0.0039$, $kur=1.4726$. Finally, the PQ index computes each of the individual waveform deviations. In case 1, PQ is 0.0333 p.u., and in case 2, PQ is 0.0593 p.u.. The most deteriorated signal with a higher PQ comes from the household. In fact, two different degrees of the deviation are detected. In Fig. 1(b), the amplitude fluctuations are detected by the variance, because the signal exhibits a flat top, while the skewness and kurtosis remain around their ideal values. In Fig. 1(d), the deviations in the top and bottom regions stand out over the rest of the cycle, and are detected by the variance. Additionally, a presence of a certain harmonic distortion is revealed by the kurtosis. Also, as both cycles are symmetric, the deviation in the skewness is negligible. Although this is a simulation based on two different cycles, the PQ confirms the behaviour of each cycle. Table II presents the time-domain compression using HOS and the aggregated index, which elicits that the method is capable of maintaining the resolution in terms of signal compression and represents the information of a cycle (0.02 s). It also inform the network behaviour using less computation resources needed in continuous measurement campaigns.

TABLE II
TIME-DOMAIN COMPRESSION FOR CASES 1 AND 2

Scenario	Variable size	Signal compression (byte)	Class
Case 1	500×1	4000	Double
Case 2	500×1	4000	Double
PQ for case 1	1×1	8	Double
PQ for case 2	1×1	8	Double
HOS for case 1	1×3	24	Double
HOS for case 2	1×3	24	Double

In [21] and [29], daily PQ patterns are identified, and differences between the trends of hourly PQ within working and non-working days are presented. The following performances of HOS are found.

1) The performance of the indices according to the impact of instantaneous fundamental frequency changes on voltage waveforms is measured according to EN 50160.

2) A minimum of 5 kHz sampling frequency is needed. In field measurements, a 25 kHz sampling frequency is adopted.

3) HOSs are immune to noise.

4) The sliding window is used to sweep the waveform and extract the indices. Theoretically, the length of window from 1 cycle up to 10 cycles would exhibit similar indices. No overlapping is used as it sweeps one period.

Besides, we show the potential of HOS in improving predictions. The strategy could be based on the learning of daily, weekly, and monthly 2D patterns. However, more research work should be done in the PQ trend pattern for long-term campaigns. We report the PQ index time-series and the 2D-HOS during three time scales, i.e., day, week, and month, and aim to establish different strategies in long-term measurement campaigns that would use the PQ index based on HOS.

III. MEASUREMENT AND ANALYSIS

The measurements are conducted at the University of Cádiz, Spain, during a six-week campaign. The goal is to monitor the 50 Hz LV at a sampling rate of 25 kHz (500 samples per cycle). The devices used in the acquisition system are the chassis NI cDAQ-9188 of National Instruments™, using an analogue input module NI-9225 C-series. It is connected via the ethernet to a PC in which a LabVIEW-based program developes continuous analysis. With the information generated using the sample frequency of 25 kHz, for each period (20 ms), the dimension of the data vector is $20/25\times10^6=800000$. The algorithm calculates the three statistics, i.e., variance, skewness, and kurtosis, using a cycle-by-cycle window without overlapping. Each window computes 8000 data, which are reduced to four parameters. In a subsequent stage, the values are subtracted to the ideal quantities; the absolute values are calculated; and the final differences are added to the PQ index in (11). In order to establish a procedure to study different patterns, all the measurements are analyzed offline using MATLAB. The monitoring procedure through HOS-based deviation index is shown in Fig. 2, where u_o is the output variable.

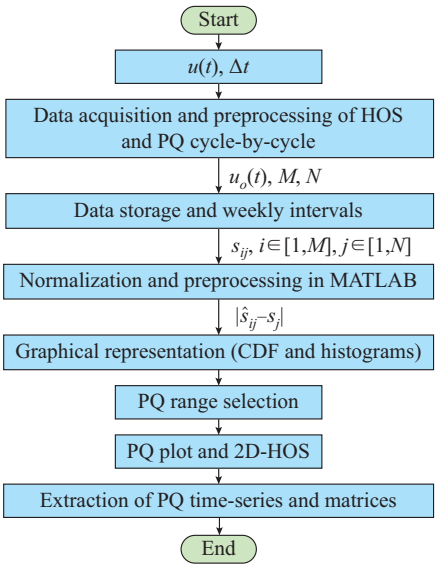

Fig. 2. Monitoring procedure through HOS-based deviation index.

An N-statistic $u(t)$ signal is processed over a preselected time interval Δt. M period s_{ij} are calculated and compared with their nominal values $\hat{s}_i$. The first compression consists of the time compression during the feature extraction stage. The second takes place in the space, averaging different 2D diagrams.

IV. ANALYSIS RESULTS

Figure 3 shows the histograms for different weekly indices based on CDF and PDF. All the measurements have been obtained during a six-week campaign in the same con-

nection point of the building. The histograms are normalized. Each value in the horizontal axis is equal to the relative accumulated number of observations in the current and former bins. The proposed method suggests the idea of studying probability distribution disturbances in order to have a profile of the distribution line [32].

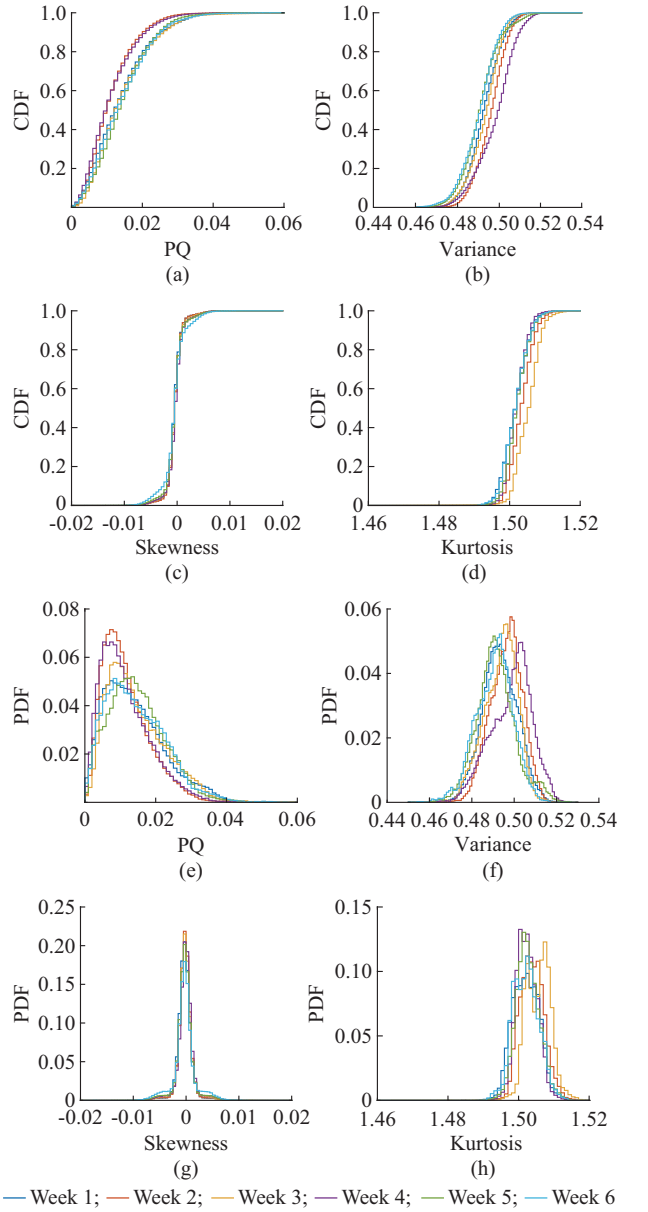

Fig. 3. Histograms for different weekly indices based on CDF and PDF. (a) CDF of PQ. (b) CDF of variance. (c) CDF of skewness. (d) CDF of kurtosis. (e) PDF of PQ. (f) PDF of variance. (g) PDF of of skewness. (h) PDF of kurtosis.

In Fig. 3(a), 90% of the weekly values achieve 0.026 or below, and the maximum values reach the range of 0.04-0.06. In Fig. 3(e), the values of the supply quality index are skewed to the left, which confirms the trend to a high probability of reaching the value of 0.01. It is relevant because the theoretical value for the PQ is 0.

The patterns of week 2 and week 4 seem to adopt similar trends, and are different from the rest of the weeks. In addition, high values of up to 0.12 denote the presence of the events. To achieve a better interpretation of the graphs, a representation range of 0-0.06 has been selected.

Figure 3(b) shows a wider range of variation. Ninety percent of the values are included in the interval [0.49, 0.51], which includes the theoretical value of the variance as 0.5. With regards to Fig. 3(f), except for those corresponding to week 4, they all exhibit a symmetrical distribution pattern with a very high probability within the interval of [0.49, 0.51].

Figure 3(c) confirms that for the voltage signal, the skewness denotes a symmetrical behaviour (around the ideal value of 0) for more than 90% of the cycles analyzed, with a very small deviation and admissible value.

Finally, based on Fig. 3(d), more than 90% of the values fall below 1.51. Note that the theoretical value of kurtosis is 1.5. For Fig. 3(h), the highest values of kurtosis fall within the interval [1.50, 1.51], and the general interval for this index is [1.495, 1.552].

Therefore, according to the weekly histograms, while the skewness and the kurtosis seem to have the most stable ranges, the variance exhibits a wider one. The results obtained, mainly those related to variance, help understand that the majority of fluctuations that occur during a week are associated with the changes in variance, i.e., amplitude changes of the waveform.

Also, the changes in the tails present a probability of occurrence evidenced by the deviation in the kurtosis index. This situation is different from the less frequent changes in the symmetry of the distribution, which constitutes a working hypothesis for future experiments. However, the skewness is another term of the PQ index having specific weight in certain measuring points with deviations in the symmetry.

Next, in order to obtain daily patterns of 2D diagrams, two weeks have been compared based on the variance versus the kurtosis in Fig. 3. The values with the maximum temporal resolution of 0.02 s (cycle by cycle) have been considered. For each day of the considered week, its intensity graph of variance versus the kurtosis has been considered. The skewness is obviated for the reasons stated above. Similarities have been found between the patterns of different days within a week and those of the days corresponding to different weeks.

Figure 4 shows the representation of PQ index time-series along the first two weeks of the monitoring campaign and their trend, and different 2D diagrams of day-to-day patterns of HOS.

Measurements took place from Monday to Sunday, from November 13th to 28th, 2017. During these two weeks, all the cycles of the measured signals were processed except a lack of 7.2% of data from week 1, and 0.6% of data from week 2 (missing 12-hour and 1-hour monitoring data, respectively) because of a connection loose between the acquisition unit and the connection point during the monitoring campaign. However, in order to carry out a robust characterization with the least number of data, only a representative part was selected. The criterion adopted (as forwarded in pre-

vious sections) was to take a measurement in every 1000 data points for each container in the histograms (bins). The goal was to eliminate statistically redundant information. Before carrying out the later compression, the time-series of PQ measurements seemed to have coupled noise, which was the visual effect produced by the accumulation of data. Even so, it was observed that the cycle-by-cycle PQ index time-series exhibited a trend, which was easily reproducible by a simple mathematical model.

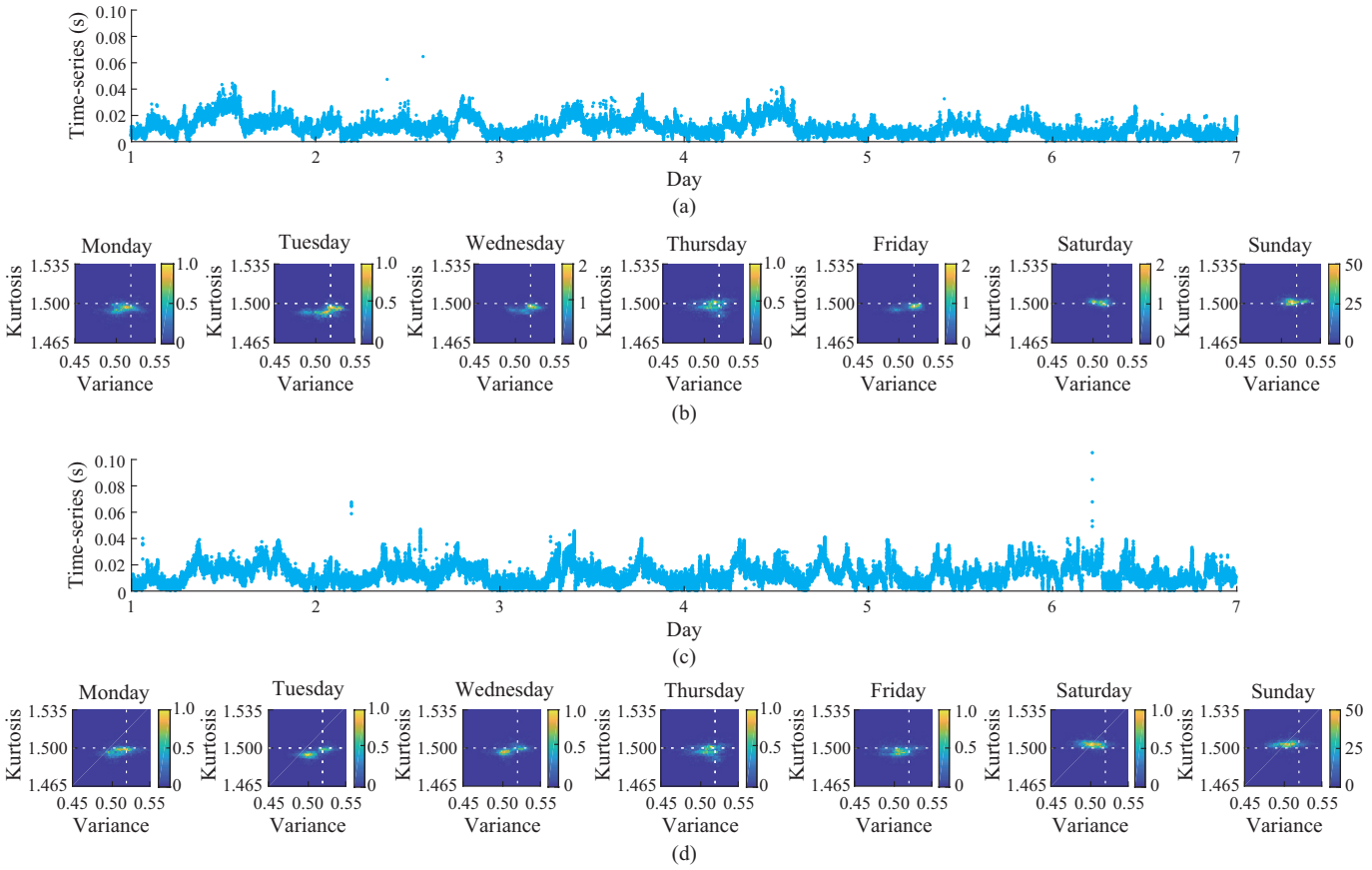

Fig. 4. Respresentation of PQ index time-series. (a) Monitoring campaign and trend for week 1. (b) 2D diagram of day-to-day pattern of HOS for week 1. (c) Monitoring campaign and trend for week 2. (d) 2D diagram of day-to-day pattern of HOS for week 2.

During the first week, there was a lack of data on Monday 13th, because the acquisition started from noon. Indeed, during the nights, the PQ index was nearer to the null ideal value, and reached the maximum of 0.04 at noon. Also, there was a second maximum around 0.03 corresponding to the network behaviour during the afternoons, which was shown on Tuesday 14th, Wednesday 15th, and Thursday 16th. On Wednesday 15th, some outliers indicated a PQ of 0.07. In addition, time-series on Friday 17th showed that PQ was lower during the afternoon, which was similar to that during the night. Moreover, Saturday 19th and Sunday 20th exhibited completely different PQ patterns and a trend within the interval of [0.01, 0.02].

Focusing on the one-day color maps (represent the variance versus kurtosis), the regions with shades were close to yellow. As was the case in histograms, the variance reached the greatest elongations, which were represented in the yellow pattern of the graph along the horizontal axis. For some days, two centroid-type regions were observed in the 2D diagram corresponding to different states of the network during the same day. This fact was even more visible on Tuesday 14th, Wednesday 15th, and Thursday 16th. On Saturday 19th and Sunday 20th, the pattern was more diffuse at the periphery of the graph center and more intense in the graph center. Precisely, the greatest number of measurements occured on Sunday 20th, which was nearest to the ideal supply value.

Furthermore, during the second week, the PQ index time-series exhibited a similar behaviour to week 1. There was a lack of PQ data on Monday 13th during the noon as a result of the monitoring campaign. From Monday 21st to Friday 25th, two maximum regions were detected, and there were less behaviours during the nights. A more unstable pattern can be observed during the weekend compared with the same period during week 1. However, in general terms during the weekends, the PQ trend was, most of the time, near to the ideal values located in the graph center. Some outliers can be observed on Wednesday 23th and on Sunday 27th that reached a PQ of 0.08 and 0.1, respectively. Besides, the day-to-day color maps confirmed the patterns in two different clusters from Monday to Friday and the weekends with a centered behaviour according to the PQ trend. The minimum PQ was always over zero, which was indicative of a non-ideal behaviour of the voltage supply waveform under normal operation conditions.

V. Discussion of Results and of Proposed Method

A. General Discussion of Results

The analyses of CDF and PDF help establish a characterization of the waveform in the point under test. For a weekly/monthly campaign, the most representative indices seem to be the variance and the kurtosis. Nevertheless, it is important to mention that based on the authors' experience, skewness can be useful in the strategies more focused on small-length windows and the event detection strategies. The main contributions include establishing the more realistic measurements to the individual ranges of indices and detecting their region of the maximum probability within the whole campaign duration.

Time-domain analysis helps detect the waveform deviation computed by the statistics. Also, the individual contribution of each indicator to the PQ is identified. Indeed, the daily PQ fluctuates depending on the day of the week, the hourly trend of the network, and the energy usage during working or non-working hours. During the night, the PQ cycle-by-cycle can fluctuate between 0.01 and 0.02. During the morning, the deviation of PQ increases, and during the noon, it reaches to the maximum of 0.04. Also, between 13:00 and 16:00, there is a drop because of the lunch time. A second increase of the index occurs in the evening, since the University under study in this paper is open until 22:00. Finally, during the midnight, the PQ decreases again, recovering the minimum values. Indeed, 2D diagrams allow visualizing such behaviours by emphasizing the areas of signal persistence throughout the hours, days and weeks.

In order to develop the site characterization through the PQ features, the time-series can be scaled to different average windows, as shown in Fig. 5. Note that the average of the PQ values help compress the size of time-series and extract the information about the general trend of the index with less computation load. Regarding the permanent monitoring strategy, the final representation of PQ smoothes the cycle-by-cycle waveform, avoiding transient effects but providing a general trend information. To set a compliance limit for the values, following the average strategy, the PQ index should be set to be 0.03 at 95% of the full scale. Nevertheless, as the indicator is averaged, the index looses the event detection resolution but the measurements help train the proposed method for more efficient performance of the PQ control. PQ based on HOS changes from one network to another depending on the quality of the waveform in the point influenced by multiple factors, which include the changes in voltage admitted by the actual PQ standards. Depending on the requirements of customer and network, this non-ideal behaviour can be assumed or not. In general, both the utility and the end-users can facilitate that information, introducing the continuous PQ monitoring of the waveform and characterizing their deviation under normal operation conditions. In practical applications, the minimum requirements in terms of computation are resumed in the Table III. 10-min PQ demands of 474720 bytes per month (15824 plus 30 days) in a year (474720 plus 12 months) equals 0.00569664 gigabytes. The PQ strategy deployed in a thousand of smart meters would demand a storage capacity of 5.69 gigabytes per year. The procedure can be incorporated in future campaigns and the solution accomplishes the monitoring challenges of the next generation of advanced metering infrastructure in terms of compression and reporting efficiency.

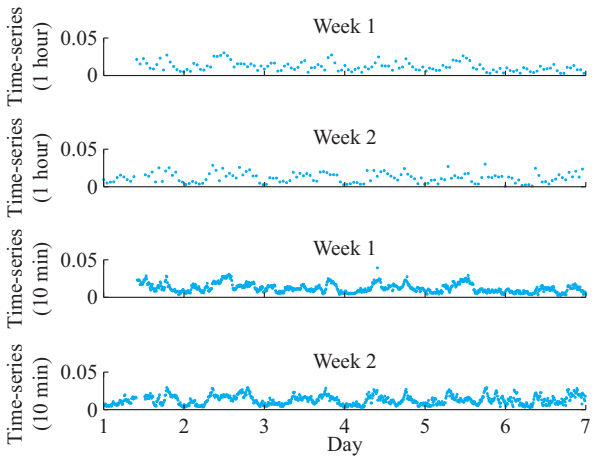

Fig. 5. Different PQ monitoring strategies informing about hourly-averaged PQ index based on HOS and PQ.

TABLE III
Data Compression Results Using Different Strategies

Parameter	Variable size	Data compression (byte)	Class
Week data based on HOS	29685000×6	1424880000	Double
Week data based on PQ	29685000×2	474960000	Double
Variance versus kurtosis	102×102	83232	Double
PQ (1 hour)	164×2	2624	Double
PQ (10 min)	990×2	15824	Double

B. Contributions

All in all, the objective of the proposed method does not reside in the specific events, but in characterizing a long-term time-series. Indeed, the proposed method characterizes different power signal states from a global point of view, but not focusing on specific events, as compared with other similar researches. However, we have carried out a comparison based on the objectives of the proposed method on the data set using low complexity event classification (LCEC) [33] and optimal time-frequency response (OTFR) [34], [35], which are good classifiers in voltage signals. Results are similar to those in [36], where HOSs have good performance and are efficient. However, in a long-term measurement campaign, the proposed method computes and exhibits voltage behaviour and the fluctuations between weeks.

We are not only differentiating between two signal states [32]. The proposed method extracts all the patterns and helps monitor the main zones within the 2D traces. Moreover, the aforementioned results show different behaviour zones within a day.

Based on our experience and the measurements, we can assure that the PQ index based on HOS can tackle the limitations on the data base, e.g. a data permutation or any other incidence such as sampling errors, because we have charac-

terized the previous pattern and it is easier to check any anomaly in the database. On the contrary, the methods based on supervized learning on the extreme values are unable to target previously unregistered waveforms [30]. We have used the Euclidian distance [37] to obtain the 2D traces, residing the contribution in targeting different behaviour zones, even within the dynamics of the same week [29].

With respect to [15], this is conceived to detect disturbances and not thought for a mid-term campaign. Thus, it is a global index without imposing a previous classification. Indeed, it is an open topic in the PQ research. Regarding our method, visualizing the network state in a longer-term measurement campaign would be the next goal.

VI. Conclusion

The contribution of the proposed method lies in the site characterization of the consumer's behaviour based on HOS monitoring, and the strategy of extracting the individual customers' pattern fingerprint. A PQ index is utilized which computes the voltage operation conditions and their deviations that come from both sides of the network, i.e., the utility and the customer, when we measure on the client's side. The proposed method assumes customer's deviations as intrinsic characteristics of the network, which is not quantified in the traditional analysis.

While traditional indices inform only about the power fluctuations, the HOS estimators provide the information regarding the waveform and constitute the new terms of the PQ index to assess the nature of PQ in the SG. The statistical features that are incorporated in the strategy are the symmetry and the tailedness of the signal under test, characterizing their PDF. The analysis is valid for continuous disturbances and event detection, or continuous and permanent monitoring. Indeed, the proposed method is scalable both in time and space, and can be deployed downstream and upstream depending on the measurement campaign objectives.

The PQ index in the HOS space provides the patterns which are more aligned to the instantaneous state of the network, considering time and the waveform characteristics, and determines the percentage of the data which are convenient to be stored according to the PQ monitoring objectives.

Additionally, the proposed method helps reduce the data computation and reports about the statistical features of the waveform more intuitively. The highest compression is made through the detection of the PQ hourly pattern. Thus, it manages to accomplish monitoring strategies and objectives related to smart meters with new PQ functionalities. As a final idea, our method would contribute to characterize the instantaneous fundamental frequency fluctuations, which is a limitation of the current time-based methods.

Considering potential usages of the PQ index based on HOS, the followings can be adopted for site characterization: PQ data compression in different intervals, e.g., journey, day-night, measurements for each hour, each 15 min, each 1 min, and each 1 s.

In performance analysis, two highlights are considered.

1) Different patterns can detect the outliers in the waveform with an origin in the PQ events.

2) We recommend an analysis window of a PQ each 10 min, 1 min or less.

Within the context of troubleshooting, we recommend an analysis window of PQ for each 10 min, 1 min or less. In advanced applications that evolve to PQ analysis based on artificial networks, waveform feature extraction in order to establish a more accurate artificial network and a reliability or PQ classification reflects the behaviour of the point under test.

To establish a relationship between HOS and the consumer's energy pattern, climatic forecasting is convenient. Indeed, the PQ pattern is related to the time with high or low energy demands. Thus, the proposed method would satisfy the requirements of the modern power grid to carry out permanent monitoring in future networks.

References

[1] S. Pukhrem, M. Basu, and M. F. Conlon, "Probabilistic risk assessment of power quality variations and events under temporal and spatial characteristic of increased PV integration in low-voltage distribution networks," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3246-3254, Mar. 2018.

[2] M. H. J. Bollen, S. Bahramirad, and S. Bahramirad, "Is there a place for power quality in the smart grid?" in *Proceedings of 2014 16th International Conference on Harmonics and Quality of Power (ICHQP)*, Bucharest, Romania, May 2014, pp. 713-717.

[3] M. H. J. Bollen and I. Gu, *Signal Processing of Power Quality Disturbances*. New York: Wiley-IEEE Press, 2006.

[4] D. S. P. M. de P. F. Ribeiro and C. A. Duque, *Power Systems Signal Processing for Smart Grids*. New York: Wiley-IEEE, 2013.

[5] O. N. Gerek and D. G. Ece, "Compression of power quality event data using 2D representation," *Electric Power Systems Research*, vol. 78, pp. 1047-1052, Jun. 2008.

[6] International Electrotechnical Commission, *Electromagnetic Compatibility (EMC)-Part 4-30: Testing and Measurement Techniques-Power quality Measurement Methods*, IEC standard, 2009.

[7] M. J. Bollen, P. Baumann, Y. Beyer *et al*., "Guidelines for good practice on voltage quality monitoring," in *Proceedings of 22nd IEEE International Conference on Electricity Distribution (CIRED)*, Stockholm, Sweden, Jun. 2013, pp. 1-4.

[8] J. M. Romero-Gordon, J. Meyer, and P. Schegner, "Design aspects for large PQ monitoring systems in future smart grids," in *Proceedings of 2011 IEEE PES General Meeting*, Detroit, USA, Jul. 2011, pp. 1-8.

[9] J. Meyer, M. Klatt, and P. Schegner, "Power quality challenges in future distribution networks," in *Proceedings of 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies*, Manchester, UK, Dec. 2011, pp. 1-6.

[10] J. Braun, V. J. Gosbell, and D. Robinson, "XML schema for power quality data," in *Proceedings of 17th IEEE International Conference on Electricity Distribution (CIRED)*, Barcelona, Spain, May 20, pp. 1-5.

[11] V. J. Gosbell, B. Perera, and H. Herath, "Unified power quality index (UPQI) for continuous disturbances," in *Proceedings of 10th IEEE International Conference on Harmonics and Quality of Power (ICHQP)*, Rio de Janeiro, Brazil, Oct. 2002, pp. 316-321.

[12] European Committee for Electrotechnical Standardization. *Voltage Characteristics of Electricity Supplied by Public Electricity Networks.* EN Standard 50160, 2010.

[13] J. Meyer, M. Klatt, and P. Schegner, "Power quality challenges in future distribution networks," in *Proceedings of 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies*, Manchester, UK, Dec. 2011, pp. 1-6.

[14] The-European-Parliament and the Council of the European Union, "Directive 2014/30/EU of the european parliament and of the council of 26 february 2014 on the harmonisation of the laws of the member states relating to electromagnetic compatibility (recast)," *Official Journal of the European Union*, vol. 1, pp. 96-79, Feb. 2014.

[15] D. D. Ferreira, J. M. de Seixas, A. S. Cerqueira *et al*., "A new power quality deviation index based on principal curves," *Electric Power Systems Research*, vol. 125, pp. 8-14, Aug. 2015.

[16] J. V. Milanovic, M. H. J. Bollen, and N. Cukalevski, "Guidelines for

monitoring power quality in contemporary and future power networks-results from CIGRE/CIRED JWG C4.112," in *Proceedings of Host Publication CIRED 2014*, Paris, France, Aug. 2014, pp. 1-5.
[17] D. Maheswaran, V. Selvaraj, and D. P. Manjaly, "Power quality monitoring systems for future smart grids," in *Proceedings of 23rd International Conference on Electricity Distribution CIRED 2015*, Lyon, France, Jun. 2015, pp. 1-5.
[18] O. Florencias-Oliveros, J. J. G. de-la Rosa, and A. Aguera-Pérez, "Reliability monitoring based on higher-order statistics: a scalable proposal for the smart grid," *Energies*, vol. 12, no. 55, pp. 1-14, Dec. 2019.
[19] R. J. Hyndman, X. Liu, and P. Pinson, "Visualizing big energy data: solutions for this crucial component of data analysis," *IEEE Power and Energy Magazine*, vol. 16, pp. 18-25, May 2018.
[20] A. Agüera-Pérez, J. C. Palomares-Salas, J. J. G. de-la Rosa *et al*., "Characterization of electrical sags and swells using higher-order statistical estimators," *Measurement*, vol. 44, pp. 1453-1460, Oct. 2011.
[21] O. Florencias-Oliveros, J. J. G. de-la Rosa, A. Agüera-Pérez *et al*., "Power quality event dynamics characterization via 2D trajectories using deviations of higher-order statistics," *Measurement*, vol. 125, pp. 350-359, Sept. 2018.
[22] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2000.
[23] J. M. Pou and L. Leblond, "ISO/IEC guide 98-4: a copernican revolution for metrology," *IEEE Instrumentation & Measurement Magazine*, vol. 21, pp. 6-10, Oct. 2018.
[24] S. Khokhar, A. A. B. M. Zin, A. S. B. Mokhtar *et al*., "A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 1650-1663, Nov. 2015.
[25] M. K. Saini and R. Kapoor, "Classification of power quality events-a review," *International Journal of Electrical Power & Energy Systems*, vol. 43, pp. 11-19, Dec. 2012.
[26] O. P. Mahela, A. G. Shaik, and N. Gupta, "A critical review of detection and classification of power quality events," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 495-505, Jan. 2015.
[27] A. Rahmani and A. Deihimi, "Reduction of harmonic monitors and estimation of voltage harmonics in distribution networks using wavelet analysis and narx," *Electric Power Systems Research*, vol. 178, pp. 1-9, Jan. 2020.
[28] A. Deihimi and A. Rahmani, "Application of echo state networks for estimating voltage harmonic waveforms in power systems considering a photovoltaic system," *IET Renewable Power Generation*, vol. 11, no. 13, pp. 1688-1694, Jul. 2017.
[29] O. Florencias-Oliveros, "Instrumental techniques for power quality monitoring," Ph.D dissertation, Higher Polytechnic School of Algeciras, University of Cádiz, Cádiz, Spain, 2020.
[30] A. F. Bastos and S. Santoso, "Universal waveshape-based disturbance detection in power quality data using similarity metrics," *IEEE Transactions on Power Delivery*, vol. 35, no. 4, pp. 1779-1787, Nov. 2019.
[31] J. J. G. de la Rosa, A. Aguera-Pérez, J. C. Palomares-Salas *et al*., "A novel virtual instrument for power quality surveillance based in higher-order statistics and case-based reasoning," *Measurement*, vol. 45, pp. 1824-1835, Aug. 2012.
[32] B. Li, Y. Jing, and W. Xu, "A generic waveform abnormality detection method for utility equipment condition monitoring," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 162-171, Jan. 2017.
[33] A. S. Cerqueira, C. A. Duque, R. M. Trindade *et al*., "Digital system for detection and classification of electrical events," in *Proceedings of 2005 IEEE International Symposium on Circuits and Systems*, Kobe, Japan, May 2005, pp. 5417-5420.
[34] M. Wang and A. V. Mamishev, "Classification of power quality events using optimal time-frequency representations-part 1: theory," *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1488-1495, Feb. 2004.
[35] M. Wang, G. I. Rowe, and A. V. Mamishev, "Classification of power quality events using optimal time-frequency representations-part 2: application," *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1496-1503, Feb. 2004.
[36] D. D. Ferreira, A. S. Cerqueira, C. A. Duque *et al*., "HOS-based method for classification of power quality disturbances," *Electronics Letters*, vol. 45, no. 3, pp. 183-185, Feb. 2009.
[37] E. G. Ribeiro, T. M. Mendes, and G. L. Dias, "Real-time system for automatic detection and classification of single and multiple power quality disturbances," *Measurement*, vol. 128, pp. 276-283, Jun. 2018.

Olivia Florencias-Oliveros received the Ph.D. degree in energy engineering in 2020 at the University of Cádiz, Cádiz, Spain. She is now a Lecturer at University of Cádiz, in the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). Her research interests include smart grid and building monitoring techniques, power quality, instrumentation technologies, virtual instruments, smart metering and statistics.

Juan-José González-de-la-Rosa received the M.Sc. degree in physics-electronics in 1992 at the University of Granada, Granada, Spain, and the Ph.D. degree in industrial engineering in 1999 at the University of Cádiz, Cádiz, Spain. He has four recognitions in the field of engineering and communication engineering, computation and electronics by the Spanish Commission for the research quality. He was awarded on Knowledge Transfer Recognition by Spanish Government. He is a Full Professor in electronics and founder of the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). His research interests include HOSs, power quality and computational intelligence techonolgies for measurement systems.

José-María Sierra-Fenández received the Ph.D. degree in industrial engineering in 1998 at the University of Cádiz, Cádiz, Spain. He is a Lecturer at the same University in the the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). His research interests include energy technologies to manage energy efficiency and renewable energies, SG, power quality, instrumentation technologies, smart metering and HOSs.

Agustín Agüera-Pérez received the M.S. degree in physics in 2004 at the University of Seville, Seville, Spain, and the Ph.D. in industrial engineering in 2013 at the University of Cádiz, Cádiz, Spain. He is now a Professor of electronics in the Department of Automation Engineering, Electronics, Architecture and Computer Networks at the University of Cádiz and a Researcher of the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). His research interests include energy meteorology, power quality and virtual instruments.

Manuel-Jesús Espinosa-Gavira received the M.S degree in 2018 at the University of Cádiz, Cádiz, Spain. He is now a Ph.D. student in energy and sustainable engineering at the University of Cádiz from 2017 and Member of the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). His research interests include power quality, time-series analysis, sensor networks, meteorology applied to renewable energies and energy efficiency.

José-Carlos Palomares-Salas received the M.S. degree in industrial engineering in 2008, and the Ph.D. degree in industrial engineering in 2013 (summa cum laude), both at the University of Cádiz, Cádiz, Spain. Currently, he is a Professor at the University of Cádiz, in the Department of Automation Engineering, Electronics, Architecture and Computer Networks and also a Member of the Research Group in Computational Instrumentation and Industrial Electronics (PAIDI-TIC-168). His research interests include intelligent systems, forecasting and wind speed modeling.