



Advanced methods of signal processing for Power Quality assessment.

Álvaro Jiménez Montero^{1,2}, Agustín Agüera Pérez^{1,2}, Juan José González de la Rosa^{1,2}, José Carlos Palomares Salas^{1,2}, José María Sierra Fernandez^{1,2} and Olivia Florencias Oliveros^{1,2}

¹ Department of Automation Engineering, Electronics and Computer Architecture, Cadiz University
Av. Ramón Puyol S/N., E-11202 Algeciras-Cádiz (Spain)
Phone/Fax number:+0034 9560280, e-mail: alvaro.montero@uca.es

² Research Group PAIDI-TIC-168: Computational Instrumentation and Industrial Electronics (ICEI), Av. Ramón Puyol S/N., E-11202 Algeciras-Cádiz, Spain

Abstract. The aim of this work is by using artificial neural networks (ANNs) compare six regression algorithms supported by 14 power-quality features, based on higher-order statistics (HOS). In addition, we have combined time and frequency domain estimators to deal with non-stationary measurement sequences; the final target is to implement the system in a smart grid to guarantee compatibility between all the equipment connected. The main results were based on spectral kurtosis measurements, which easily adapt to the impulsive nature of the power quality events. Through these results we have verified that the developed technique is capable of offering interesting results at classifying power quality (PQ) disturbance.

We can conclude that using radial basis networks, generalized regression and multilayer perceptron, we have obtained the best results mainly due to the non-linear nature of data.

Key words

Artificial neural networks (ANN), Power Quality (PQ), High-order statistics (HOS), Spectral kurtosis, Smart Grids (SG).

1. Introduction

As a result of the unstoppable growth of electronic equipments demanding electricity, consumers expect uninterrupted availability and quasi perfect power quality (PQ). Consequently, the worldwide interest in PQ is sustained in the basis that more sensible systems are constantly exposed to voltage disturbances, which are in turn caused by industrial equipment they are connected to. There is also the need for standardization and performance criteria for consumers and utilities, in order to develop measurement systems accessible to an average citizen in addition to its influence over and from the loads, and the recent potential inclusion in the modern smart grid (SG).

In this frame, an adequate PQ assures the necessary compatibility between all the equipment connected to the grid [1]. The terms of this compatibility gather several aspects: sustainable power with low losses, and high quality and security of supply and safety, being at the

same time economically efficient, reliable, and resilient [2].

Certainly, the future SG would introduce transformative technologies to meet these design requirements, integrating intelligence into end-use devices as the key to satisfy the demand response. In parallel, the industry would design ways to incorporate automatic end-use-load participation into the model so that customers are not bothered by these programs and decisions, and so their lifestyles are not inconvenienced [3].

Hence, the role of smart meters and sensors is the first revised aspect in the current and future SG. These automated meters (AM) use a two-way communicating infrastructure and centralized management, as well as new features such as the following: outage management, demand response, automatic load shedding, distribution automation, and the ability to enable and commute alternative energy sources.

Provided with this scenario, this research integrates artificial neural networks (ANNs) and advanced signal processing techniques based on higher-order statistics (HOS) to be implemented into an automated smart meter for PQ event detection and classification, within the frame of a SG with a high distribution penetration of renewable sources. Six regression algorithms were tested based on a hybrid time-frequency battery of characteristics, specially designed to deal with non-stationary measurement time series.

The feature extraction stage from PQ disturbances was based on HOS, which has proven to be efficient in several works. Indeed, since PQ events are sudden changes in the power line, HOS are potentially useful to characterize each type of electrical anomaly both in time and frequency domains. As a novelty, the present research exploits the combination of time and frequency domain features to deal with the inherent non-stationary associated to the electrical anomalies, with the goal of improving the performance of the ANN and making feasible the integration in a smart meter.

Regarding to backgrounds of HOS applications in the time domain, several notable works are worth mentioned e.g., Bollen et al. introduced new advanced statistical features

to PQ event detection [4]. Likewise, Gu and Bollen [5] found relevant characteristics associated to PQ events in the time and frequency domains.

The work by Ribeiro et al., also remarkable [6], extracted new time-domain features based on cumulants.

The same authors performed the classification of single and multiple disturbances using HOS in the time domain and Bayes' theory-based techniques [7]. HOS techniques and estimators have also been implemented to specifically detect sags and swells [8].

The categorization of PQ anomalies had been formerly performed by Nezih and Ece in the work [9], where they proved that HOS and quadratic classifiers improved the second-order-based methods. The same authors previously had achieved performance in second-order computing, using 2-D wavelets and compression techniques [10]; finding, despite the promising results, the limits of the procedure and quantifying its heavy computational cost. Alienated to this work, the researches of Poisson et al. and Santoso et al. [11] also reported a wavelet-based method, finding the potential and the drawbacks of the technique in order to implement it in an intelligent meter.

The direct antecedent of the present research is the work by J.J.G. de la Rosa et al. [12], who performed a mixed study involving the time-domain variance, skewness, and kurtosis, and they obtained consequences combined with the spectral kurtosis (SK), and over a set of real-life measurements, some of them with mixed PQ perturbations. The same authors proposed preliminary criteria for six types of disturbances based on the former estimators [13]. In a previous work [14], they designed an offline case-based reasoner based on time-domain HOS estimators. Furthermore, the authors also used HOS features in classification techniques for a characterization of electrical PQ signals [15].

The present paper is designed as follows. The next section summarizes the main advances in the field, paying special attention at the applications of ANN for PQ analysis, and reasoning the contribution of HOS to the feature extraction stage and the architecture. Then, in the Higher-order statistics Section for PQ monitoring an enhancement proposal over the DWT is offered to expose the advantages of HOS for PQ monitoring and the state-of-the-art. In the Proposed methodology Section the HOS-based ANN procedure is detailed to later expose the results in Results Section; finally, the conclusions are drawn in the last section.

2. Integration of HOS and ANN

The evaluation of the integration of HOS and ANN is developed following the structure shown in Fig. 1, where the main stages of the pre-processing actions are also outlined within the whole computing process.

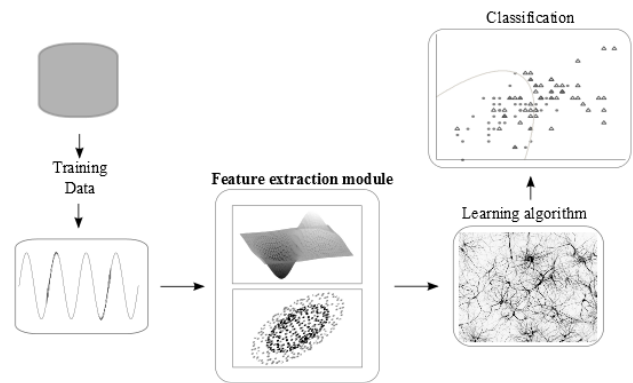


Fig. 1. Generic ANN structure which postulates the feature extraction stage previous to the ANN.

The performance of the potential PQ monitoring system and consequently the ANN is directly related to the pre-processing and feature extraction techniques used.

The main goal of the feature extraction is to represent the data set in a new feature space in which the probability to distinguish classes is higher than the one in the original space. Therefore, the identification of efficient pre-processing and feature extraction techniques is a key issue [16].

If the extracted features result in a feature space where the classes are well-enough separated, then it is said that there is a high probability that the designed classifier will demonstrate good performance. If not, the classifier will perform poorly no matter the technique used for the analysis.

Another requirement for a good feature extraction-based technique is the ability to provide the required separation in a low-dimensional space (medium complexity). In pattern recognition problems, we aim to find a reduced number of features in order to make the classifier design feasible. If the feature extraction technique results in a high-dimensional space, the feature selection will play an important role in the system design. In order to select a feature extraction technique, good knowledge of the classification issues is required.

Feature extraction tools most used in the literature are based in Effective Value (RMS), Discrete Fourier Transform (DFT), Wavelet Transform (WT), Cumulants of Higher-Order Statistics, and Principal Component Analysis (PCA).

As said before, in the present research, a non-stationary signal processing strategy based in higher-order statistics in the time and the frequency domains is postulated in order to extract a battery of features to be processed via ANNs with regression algorithms. ANNs have been used for classification purposes in myriads of works and have proved the utility for a long time [17]; Fig. 2 shows a generic architecture, indicating how the input pattern is processed via the neurons to provide the output vector. This basic figure is thought to introduce the concrete architecture of the NNT in the present work.

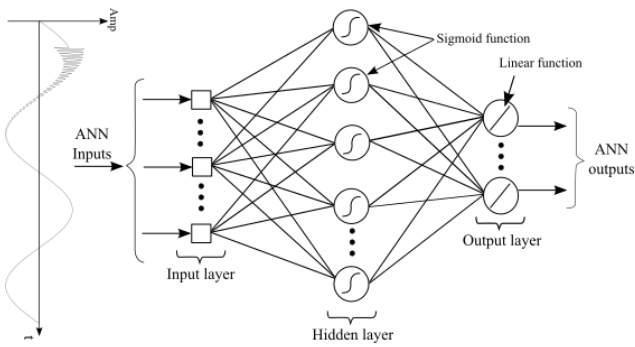


Fig. 2. Generic ANN structure which postulates the feature extraction stage previous to the ANN.

The performance of the potential PQ monitoring system and consequently the ANN is directly related to the pre-processing and feature extraction techniques used. Apart from the difficulties in implementing a real-time processor in the smart grid frame, the main disadvantage of the above procedures lies in the fact that accuracy and repeatability are highly compromised by the second-order estimators (e.g., WT and VRMS) used in the feature extraction stage. Indeed, drawbacks arise when the data are corrupted by noise; especially when the number of samples of the signal window is reduced, and the resolution and repeatability are degraded. These facts are dramatically increased if the tests are performed over synthetic signals of controlled-lab experiences, where predictability is tacitly supposed. In the real-world experiences, the system should be prepared for unpredictable phenomena, both in the time and in the frequency domains. As explained hereinafter, this goal is accomplished by statistical parameters of an order higher than two.

3. The proposed methodology: HOS-based ANN

The final goal of this work is the implementation in the future SG to guarantee compatibility between all the equipment connected. The Data used in this work was simulated using MATLAB software. The data set generated by the simulation consists of 550 samples including the different studied disturbance kinds, which cover the following disturbances: oscillatory transient, impulsive transient, interruption, harmonic permanent distortion, harmonic temporal distortion, sag, sag plus oscillatory transient, and swell. Each signal comprises a 20K-point synthetic time-domain register with 1 s duration (20 KHz sampling frequency). An additive normal noise process (1% of the amplitude of the signal) has been added to achieve a more realistic behaviour. Fig. 3 shows an example of the signals, as well as healthy signals, employed in this paper. The classification techniques used to classify the PQ disturbances are based on regression algorithms.

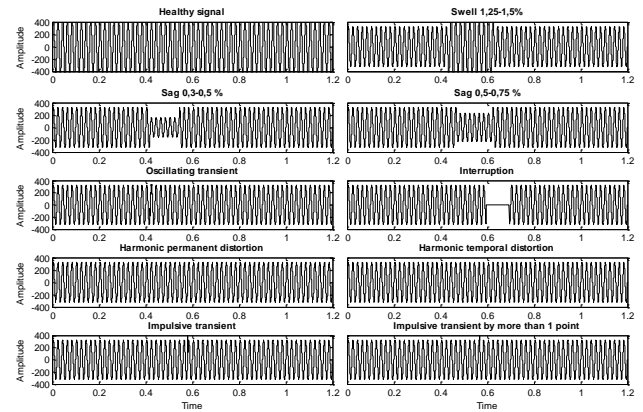


Fig. 3. Example of the healthy signal and different disturbances studied.

Data used to realize the classification were based on representative coefficients obtained from the PQ disturbances referred above. These coefficients are acquired by a process of feature extraction based on the combination of higher-order statistics in time and frequency domains. The HOS have been computed using a 400-points sliding window (which corresponds to a signal period), with a shift of 10 points over a vector of 20,000 points. After the extraction stage, a total of 14 characteristic features are selected, nine of which correspond to time domain and the remaining five to frequency domain. The coefficients selected in the first one correspond to the maxima and minima, and stable in the second-, third-, and fourth-order cumulants at zero lags (directly related to the variance, skewness, and kurtosis).

On the other hand, the coefficients selected in the second one correspond to the frequency of extreme value of SK, bandwidth of dome, extreme value of SK, number of peaks in SK, and very targeted dome (between 0 and 1). Fig. 4 shows the modular graphic corresponding to the feature extraction module.

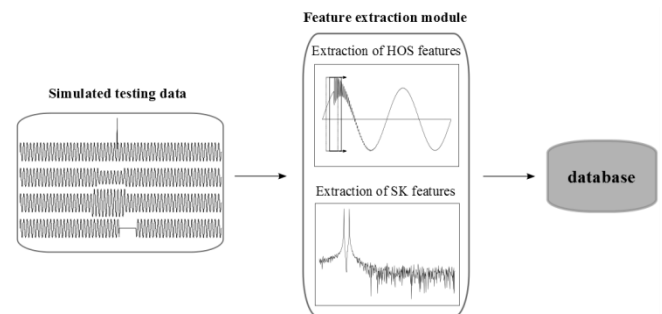


Fig. 4. Procedure of the feature extraction stage. The lower subfigure forwards the reader the use of the spectral kurtosis, extracting the kurtosis of each frequency component.

In order to illustrate the capability of the SK to discriminate PQ disturbances, we have selected a practical example consisting of an oscillatory transient coupled to the power sine wave. The analysis result is depicted in Fig. 5. The time-domain variance increases when it bumps into the transient; this behaviour is independent of the transient frequency. Similarly, the time-domain skewness and kurtosis detect slight variations. The real detection took place in the frequency domain; the SK produces a

real enhancement in 2,000 Hz, along with the high-resolution bump (narrow peak). The resulting data after carrying out the feature extraction is a matrix of dimension $n_s \times f$, where n_s is the number of samples and f is the number of features selected. Then, the classification algorithms are used to classify the PQ disturbances.

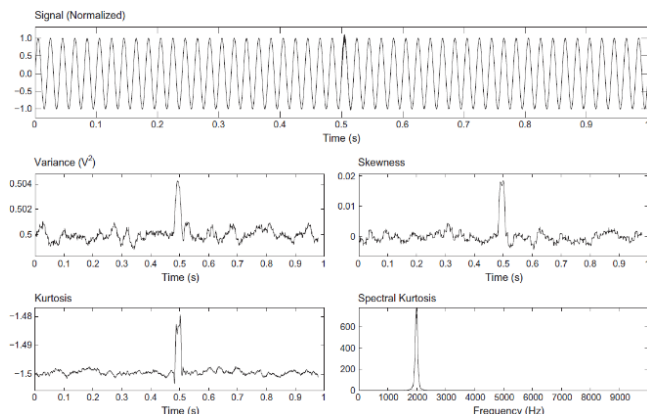


Fig. 5. An example of mixed analysis in the time and frequency domains.

The classification techniques used in this work were based on regression algorithms. These techniques were adaptive linear neuron (LIN), multilayer ANNs (BP1 and BP2), radial basis function (RBF), exact radial basis (ERB), and generalized regression network (GRN). Each model was intelligently adapted to meet the objective of PQ classification because each had different characteristics. Table I shows the selected parameters corresponding to the architecture and activation function of each model used. The final design of each model was obtained by the optimization of the parameters shown in this table.

Table I. - Parameters of the proposed ANNs.

ANN	LIN	BP1	BP2	RBF	ERB	GRN
Hidden layers	-	1	2	1	1	1
Neurons hidden layer 1	-	[4 to 10]	[4 to 10]	[1 to 150]	[1 to 150]	[1 to 150]
Neurons hidden layer 2	-	-	[2 to 5]	-	-	-
Spread	-	-	-	[1 to 20]	[1 to 20]	[1 to 20]
TF	-	S	S	G	G	G
TF output	L	L	L	L	L	L
Training algorithm	WH	LM	LM	k	k	k
Performance function	MSE	MSE	MSE	MSE	MSE	MSE

Acronyms: LM: Levenberg Marquardt; WH: Widrow Hoff; S: Sigmoid; G: Gaussian; L: Linear; k : k-means; TF: Transfer function; MSE: Mean square error.

The building of the models was performed by following two steps so as to efficiently classify the disturbances.

1) **Normalization:** Data were normalized so that they were in the interval $[-1, 1]$, and ensuring that all selected features represent the same dynamic range to achieve faster computation, which is by the way a desirable property when considering implementation issues [18].

2) **Datasets:** In order to design and test a classification algorithm, sets of samples from the patterns are required. Usually three datasets are used: training, validation and testing. The training dataset is used for system design during the classifier learning phase. The validation dataset is used in order to verify the classifier generalization performance. Finally, the test dataset is used to assess on the out-of-sample set the classification power of a model.

It is crucial for generalization purposes that the test dataset must not be used during the training phase of the classification algorithm, but for a limited-size dataset this is not always possible.

Finally, the complete design of a classification system must include an evaluation of its performance, an important step which could lead to the complete redesign of the system. The goal is to estimate the classification error of the designed system with a finite dataset available for the system design.

In this work, the classification criterion is based on the parameter hit rate (HR), which is defined as follows:

$$HR = \frac{N_C}{N_T} \times 100 \% \quad (1)$$

where N_C is the number of test samples and N_T is the number correct disturbance recognition.

This criterion was used to assess the capacity of classification to each model, providing a measure of comparison between them.

4. Results

Once the assessed models were configured and optimized, they were used in the out-of-sample set. Because the database was small, we choose randomly two test sets, and in each of them, 100 experiments are launched by model. This is realized to achieve statistically meaningful results which rule out the random factors influencing the ANNs.

For each model, we performed three analysis in function of the features used: HOS in the time domain (HOST), the SK, and the mixed (time and frequency domain) analysis (HOST+SK). In all of them, the HR was calculated for each model on the 200 offline tests, observing the arithmetic mean. The obtained results showing the percentage of effectiveness for classifying disturbances are presented in Table II-IV. In most models, results obtained in the mixed analysis are better than in the other analyses except isolated results. These exceptions represent the 27.27% of the cases presented in these

Tables, where the 14.28% of them correspond to HOS_t features, and the remaining 12.99% correspond to SK features.

Table II. - Percentages of effectiveness considering HOS features in time and frequency domains (HOS_t+SK).

PQ disturbance	LIN	BP1	BP2	RBF	ERB	GRN
Harmonic permanent distortion	0.00	85.8	78.0	54.9	100.0	69.3
Harmonic temporal distortion	10.5	74.5	70.0	42.1	93.0	48.4
Impulsive transient	53.4	82.8	81.4	38.5	96.3	100.0
Impulsive transient by more than one point	94.3	83.7	81.7	68.0	95.9	100.0
Oscillatory transient	5.99	75.2	77.2	83.6	96.9	90.1
Interruption	18.6	84.2	81.3	95.6	91.0	96.3
Sag 0.3% to 0.5%	40.5	67.4	69.7	38.9	90.1	64.8
Sag 0.5% to 0.75%	44.8	53.0	58.8	33.4	89.9	87.7
Swell 1.25% to 1.5%	18.8	84.0	78.6	40.0	99.6	94.9
Sag plus oscillatory	0.00	43.9	51.6	18.6	88.6	27.4
Healthy signal	0.00	80.8	74.9	99.9	100.0	96.3

Table III. - Percentages of effectiveness considering HOS features in time domain (HOS_t).

PQ disturbance	LIN	BP1	BP2	RBF	ERB	GRN
Harmonic permanent distortion	0.0	59.8	49.6	52.8	83.2	1.8
Harmonic temporal distortion	0.0	60.7	59.9	53.0	79.7	33.7
Impulsive transient	0.0	53.2	56.3	19.8	80.3	10.0
Impulsive transient by more than one point	17.6	49.2	55.2	16.9	81.1	12.3
Oscillatory transient	98.5	24.1	25.8	45.0	79.5	95.3
Interruption	14.6	88.8	88.2	100.0	80.2	86.7
Sag 0.3% to 0.5%	38.8	54.4	59.0	60.6	79.7	14.5
Sag 0.5% to 0.75%	46.7	42.5	35.8	28.3	79.6	88.1
Swell 1.25% to 1.5%	17.9	76.8	84.3	39.4	80.3	50.4
Sag plus oscillatory	0.0	14.1	11.9	8.8	80.2	0.0
Healthy signal	0.0	12.0	25.2	0.0	95.2	0.0

Table IV. - Percentages of effectiveness considering HOS features in frequency domain (SK)

PQ disturbance	LIN	BP1	BP2	RBF	ERB	GRN
Harmonic permanent distortion	0.0	88.0	70.0	100.0	100.0	0.0
Harmonic temporal distortion	1.3	17.1	24.8	1.7	80.5	0.0
Impulsive transient	3.8	56.5	55.4	42.2	82.5	0.0
Impulsive transient by more than one point	57.4	68.2	73.4	77.5	79.4	100
Oscillatory transient	3.2	52.1	58.2	39.0	79.8	2.9
Interruption	56.5	33.2	33.2	31.7	79.9	58.1
Sag 0.3% to 0.5%	21.2	29.6	33.3	23.4	78.3	39.8
Sag 0.5% to 0.75%	47.7	47.4	45.6	41.6	79.0	59.2
Swell 1.25% to 1.5%	8.5	43.5	45.1	27.8	79.5	0.0
Sag plus oscillatory	0.0	24.3	25.6	100.0	80.7	0.0
Healthy signal	0.0	78.5	63.0	51.0	100.0	0.0

Figure 7 and Table V show the global percentages associated to the different models attending to the three different analyses (and their respective sets of features) commented above. These percentages are obtained for each model as the average of the hit rates associated to each PQ disturbance.

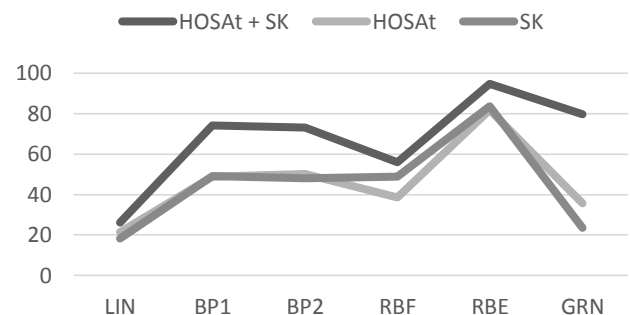


Fig. 7. Global percentages considering different features

Table V. - Global percentages considering different features

Features	LIN	BP1	BP2	RBF	ERB	GRN
HOS _t + SK	26,1	74,1	73,0	55,8	94,7	79,5
HOS _t	21,3	48,7	50,1	38,6	81,7 8	35,7
SK	18,1	48,9	48,0	48,7	83,6	23,6

Results show that the algorithms based on linear models are worse than those based on non-linear schemes. The best models both individually and collectively are ERB, GRN, and BP1.

5. Conclusion

The complexity and variability of electrical signals demand a new signal processing methods in order to distinguish each type of perturbation. This fact is even more precious in modern smart grid, where distributed energy resources are connected asynchronously, in high penetration scenarios. For that reason pattern recognition becomes an essential enabling tool for the identification and control of the upcoming electric smart-grid environment, along with the physical localization of the fault.

In this paper, six regression algorithms were applied and compared for PQ disturbances classification.

The novel aspect is the introduction of new representative coefficients based on HOS in time and frequency domains. These coefficients are the inputs used in the classification algorithms to verify the occurrence of single or multiple disturbances in the electric signals.

The data used to test the proposed method were generated by the MATLAB software. The PQ disturbances considered are the most common on the supply. The best models both individually and collectively are obtained employing radial basis networks, generalized regression, and multilayer perceptron. The overall hit rates obtained are 94.70%, 79.59%, and 74.17%, respectively. This is consistent with non-linearity of the used data and emphases of the non-linearity that provide the HOS.

Once the obtained results demonstrate that the proposed method can effectively classify different kinds of PQ disturbances, it is necessary to do more training and tests to find an ANN classifier with better characteristics than the one obtained in this research.

In summary, it might be concluded that the proposed method can help to satisfactorily classify different types of PQ disturbances. Then, once their usefulness has been established, it is necessary to expand the experiment to find an ANN classifier with better characteristics than obtained in this research.

Acknowledgements

The authors would like to thank the Spanish Government for funding the research project TEC2010-19242-C03-03 (SIDER-HOSAPQ). This work is newly supported by the Spanish Ministry of Economy and Competitiveness in the frame of the Statal Plan of Excellency for Research, via the project TEC2013-47316-C3-2-P (SCEMS-AD-TED-PQR). Our unforgettable thanks to the trust we have from the Andalusian Government for funding the Research Group PAIDI-TIC-168 in Computational Instrumentation and Industrial Electronics (ICEI).

References

[1] MHJ Bollen, S Bahramirad, A Khodaei, in *Proceedings of the 2014 IEEE 16th International Conference on Harmonics and Quality of Power (ICHQP)*. Is there a place for power quality in the smart grid? (University Politehnica of Bucharest Romania, 2014), pp. 713–717 2.

- [2] Y Xiao, *Communication and networking in smart grids*. (CRC Press, Broken Sound Parkway NW, Suite 300, 2012)
- [3] WK Reder, IEEE smart grid. Part 2: a grand vision for smart grid. Technical report. IEEE (2014). <http://smartgrid.ieee.org/education>
- [4] MHJ Bollen, IY-H Gu, PGV Axelberg, E Styvaktakis, Classification of underlying causes of power quality disturbances: deterministic versus statistical methods. *EURASIP J. Adv. Signal Process.* **2007**(1), 1–17 (2007)
- [5] YH Gu, MHJ Bollen, Time-frequency and time-scale domain analysis of voltage disturbances. *IEEE Trans. Power Deliv.* **15**(4), 1279–1283 (2000)
- [6] MV Ribeiro, CAG Marques, CA Duque, AS Cerqueira, JLR Pereira, Detection of disturbances in voltage signals for power quality analysis using HOS. *EURASIP J. Adv. Signal Process.* **2007**(1), 1–13 (2007)
- [7] MV Ribeiro, JLR Pereira, Classification of single and multiple disturbances in electric signals. *EURASIP J. Adv. Signal Process.* **2007**(1), 1–18 (2007)
- [8] A Agüera-Pérez, JC Palomares-Salas, JJG de la Rosa, JM Sierra-Fernández, D Ayora-Sedeño, A Moreno-Muñoz, Characterization of electrical sags and swells using higher-order statistical estimators. *Measurement.* **44**(Issue 8), 1453–1460 (2011)
- [9] ON Gerek, DG Ece, Power-quality event analysis using higher order cumulants and quadratic classifiers. *IEEE Trans. Power Deliv.* **21**(2), 883–889 (2006)
- [10] DG Ece, ON Gerek, Power quality event detection using joint 2D wavelet subspaces. *IEEE Trans. Instrumentation Meas.* **53**(4), 1040–1046 (2004)
- [11] S Santoso, WM Grady, EJ Powers, J Lamoree, SC Bhatt, Characterization of distribution power quality events with fourier and wavelet transforms. *IEEE Trans. Power Deliv.* **15**(1), 247–254 (2000)
- [12] JJG de la Rosa, JM Sierra-Fernández, A Agüera-Pérez, JC Palomares-Salas, A Moreno-Muñoz, An application of the spectral kurtosis to characterize power quality events. *Electrical Power Energy Syst.* **49**, 386–398 (2013)
- [13] JJG de la Rosa, JM Sierra-Fernández, A Agüera-Pérez, JC Palomares-Salas, A Jiménez-Montero, A Moreno-Muñoz, in *Proceedings on the IEEE International Workshop on the Applied Measurements for Power Systems (AMPS)*, vol. 1. Power quality events' measurement criteria based in higher-order statistics: towards new measurement indices (IEEE Aachen, Germany, 2013), pp. 73–79
- [14] JJG de la Rosa, A Agüera-Pérez, JC Palomares-Salas, JM Sierra-Fernández, A Moreno-Muñoz, A novel virtual instrument for power quality surveillance based in higher-order statistics and case-based reasoning. *Measurement.* **45**(7), 1824–1835 (2012)
- [15] JC Palomares-Salas, JJG de la Rosa, A Agüera-Pérez, A Moreno-Muñoz, Intelligent methods for characterization of electrical power quality signals using higher order statistical features. *Przeglad Elektrotechniczny.* **2012**(8), 236–243 (2012)
- [16] DD Ferreira, CAG Marques, JM de Seixas, AS Cerqueira, MV Ribeiro, CA Duque, Exploiting Higher-Order Statistics Information for Power Quality Monitoring. (InTech Open Science, 2011). <http://intechopen.com>
- [17] S Haykin, *Neural Networks*. (Englewood Cliffs, NJ, 1994)
- [18] J Sola, J Sevilla, Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. Nuclear Sci.* **44**(3), 1464–1468 (1997)