CHALMERS
UNIVERSITY OF TECHNOLOGY

VISUALIZING THE RESULTS FROM UNSUPERVISED DEEP LEARNING FOR THE ANALYSIS OF POWER-QUALITY DATA

VISUALIZING THE RESULTS FROM UNSUPERVISED DEEP LEARNING FOR THE ANALYSIS OF POWER-QUALITY DATA

Roger A. de Oliveira[1], Chenjie Ge[2], Irene Y.H. Gu[2], Math H.J. Bollen[1]*

[1]*Luleå University of Technology, Skellefteå, Sweden*
[2]*Chalmers University of Technology, Gothenburg, Sweden*
**math.bollen@ltu.se*

Keywords: POWER-QUALITY, POWER-QUALITY MONITORING, POWER-SYSTEM HARMONICS, MACHINE LEARNING, DEEP LEARNING.

Abstract

This paper presents a visualisation method, based on deep learning, to assist power engineers in the analysis of large amounts of power-quality data. The method assists in extracting and understanding daily, weekly and seasonal variations in harmonic voltage. Measurements from 10 kV and 0.4 kV in a Swedish distribution network are applied to the deep learning method to obtain daily harmonic patterns and their distribution over the week and the year. The results are presented in graphs that allow interpretation of the results without having to understand the mathematical details of the method. The inferences given by the results demonstrate that the method can become a new tool that compresses power quality big data in a form that is easier to interpret.

1 Introduction

Utilities and industrial consumers perform continuous power quality (PQ) monitoring to obtain information on the supply and equipment performance [1]. Long-term PQ measurements result in a large amount of data. For instance, one year of continuous monitoring of 39 harmonics and 40 interharmonics (3 voltages, 3 currents, 10-minute values) results in about 31 million data points per location. Manual analysis of this data type is possible; however, it is too time-consuming for multiple location measurements. Existing automatic methods to handle PQ data obtain indices such as the 95th percentile, average, or maximum values of the Total Harmonic Distortion (THD) over a given period [2, 3]. The representation of the PQ data in such indices is a very limited approach, as it does not allow obtaining typical variation patterns versus time. In order to assist the experts, new automatic tools are still needed to compress the PQ big data in an easier-interpretation form.

Artificial intelligence (AI) tools have been developed since the ’90s to identify patterns in power system data. Before 2015, most of the tools were based on supervised learning which requires pre-labelled training data to predict the labels of the unseen data. The application of supervised learning was focused on the classification and recognition of events [4]. Since 2015, there is a growth in deep learning (DL) algorithms that can handle power system data without the need to pre-define features. Even with the possibility of automatic feature extraction, none of the approaches has provided a solution for the extraction of the harmonics patterns. Faced with that, we provided in [5] an unsupervised method, briefly summarized in Section 2, based on DL to extract automatically patterns in daily variations of harmonics.

In Section 3, we show how the results from the method in [5] become a tool that assists power engineers in the analysis of PQ measurements. The automatic results given by DL are presented, in Section 4, for harmonics measurements in a Swedish distribution network. We concentrate the analyses on the visualization of the automatic results from a PQ viewpoint. The main objectives of this paper are: (1) to illustrate the inferences that can be obtained by the application of the method to PQ data; (2) to provide a guide for the power engineers in the interpretation of the automatic results given by DL.

2 The Deep Learning Method

The DL method consists of an autoencoder (AE) followed by a clustering algorithm to obtain the principal features and main patterns of variations in PQ data as shown in Fig 1. The following describes each stage of the method. Mathematical descriptions can be found in [5].

The pre-processing consists of representing the time-series of a long period (input data) in individual daily samples. For instance, Fig. 1 shows a harmonic voltage time-series for an entire year, which results in 52560 samples (365 days x 144 samples per day). The pre-processed data is obtained by reshaping the entire time-series in 365 data sequences with 144 samples (10 min-values).

The AE consists of an encoder that maps the original data to a compressed format that contains the principal features. The decoder is a reverse structure of the encoder, which can reconstruct a representation of the original data by the principal features. Both encoder and decoder contain several layers of convolutional neural networks.

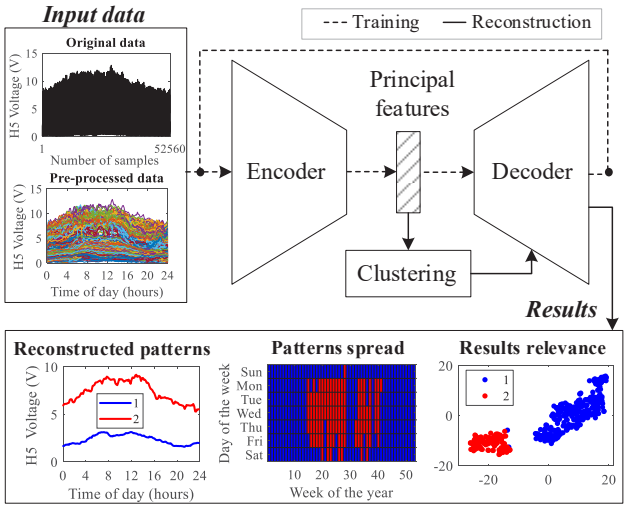


Fig. 1 Block diagram of the DL method [5].

The AE training is performed to minimize the reconstruction error between the original and the decoded data. Through the training process, the principal features of the input data are obtained. To analyse the possible underlying patterns of variation from the input, clustering is applied to group the principal feature vectors of data. The reconstructed patterns are obtained by applying the cluster centres of the principal features on the trained decoder.

3. The Visualisation Method

The method results in three plots, shown in the lower part of Fig 1: reconstructed patterns, patterns spread, and results relevance. It is these plots that assist the power engineers.

3.1 Reconstructed patterns

This plot shows the two daily patterns that are most dominant in the data. In engineering terms, these could be described as "typical patterns". In the method as applied here, only two patterns are assumed, but there is no limitation in the number of clusters. Our experience is, however, that interpretation generally becomes more complicated with an increasing number of clusters. The differences and similarities between intraday variations of the two reconstructed patterns allow making inferences for the magnitude and origin of a PQ disturbance over a day. For instance, in Fig. 1 the reconstructed pattern 2 is higher than pattern 1 in terms of harmonic voltage during the whole day. Then, pattern 2 represents high distortion days, while pattern 1 represents low distortion days. Even though the voltage magnitude is distinct for each pattern, the intraday variations are similar. For example, both patterns present steps at 8:00 and 16:00, and the highest values occur at the same time (between 6:00 and 16:00). The similarity in the intraday variations indicates that the disturbance cause is the same.

3.2 Patterns spread

This plot indicates, for each day of the year, which of the two typical daily patterns best fits the observed pattern for the day. The plot shows how the patterns are spread over the year and the occurrence of any weekly or seasonal variations. For instance, in Fig.1 the high-distortion days are in red (pattern 2) and occur mostly on weekdays during spring and summer.

3.3 Results relevance

This graph is the result of an advanced mathematical transformation, the details of which are not relevant here, which shows how strong the clusters are. It can show, for example, if the patterns show a continuous change or if they form two or more distinct groups. For instance, in Fig. 1 the two clustered patterns are reasonably well separated. It indicates that the data features obtained are effective to reveal distinctive patterns.

4 Study case and Results

PQ data sets from a continuous PQ measurement in a Northern Swedish distribution system over the year 2017 are applied to the proposed DL method. In each feeder, class A PQ monitors are installed in both MV and LV sides of the distribution transformers as shown in Fig. 2.

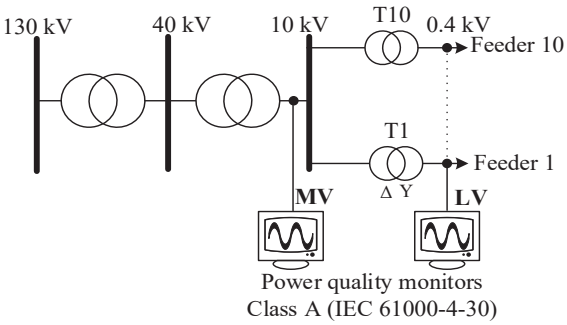
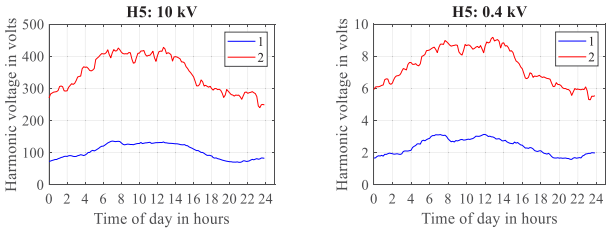


Fig. 2 Measurements setup in a Northern Swedish distribution system.

The measurement data consist of 10-min values according to IEC 61000-4-30 [6] for harmonics 2 through 50. To illustrate the visualization of the results, data sets from harmonics measurements in a feeder for both MV and LV are analysed in this paper. The forthcoming sections present the automatic plots for some odd harmonic orders in power networks. In order to guide the interpretation of the plots, analyses for the 5th harmonics are presented in detail. The inferences for 3th, 7th, and 9th are summarized. The choice for these orders was based on the meaningfulness of them in power systems analysis.

4.1 Results and analysis for 5th harmonics

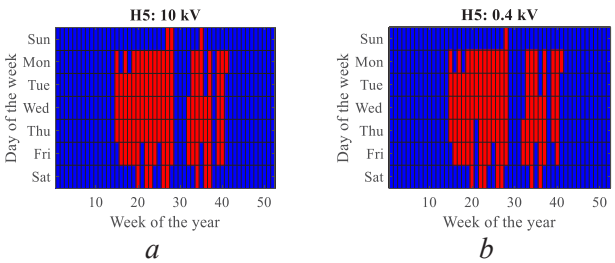
To analyze the daily variations of the 5th harmonics (H5), Fig. 3 shows the reconstructed patterns for 10 kV (a) and 0.4 kV (b). For both voltage levels, there are two distinct patterns represented as '1' (blue) and '2' (red). Inferences can be made by visualizing the variation of the harmonic voltage during the 24-hour interval. For instance, one can note that pattern 2 has higher harmonic voltage than pattern 1, in both voltage levels. A possible inference is that the intraday variations are the same in the 10 kV and 0.4 kV, i.e. the patterns share a similar shape in both voltage levels. The highest values in patterns 1 and 2 occur during the day between 6:00 and 18:00. The steps around 6:00 and 18:00 are higher in pattern 2 than in pattern 1.



a *b*

Fig. 3 Reconstructed patterns for H5: (a) 10 kV; (b) 0.4 kV.

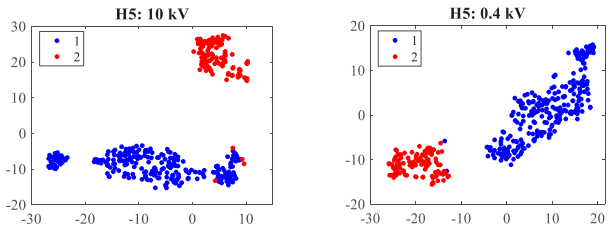
To further analyze the distribution of original data sequences in these two patterns over the entire 1-year period, Fig. 4 shows how the patterns are spread in terms of weekdays versus week numbers over 2017 for both voltage levels. The blue cells correspond to days classified in pattern 1, and the red cells correspond to data in pattern 2 (higher distortion). By the visualization of the plots, one can infer that the high-distortion days occur mostly on weekdays during spring and summer for both voltage levels. The exception is from weeks 29 through 32 and week 38 when only low-distortion days occur. An important conclusion that the same underlying phenomenon causes the 5th harmonics for both voltage levels as the daily variations of the patterns (Fig. 3) and spread over the year (Fig. 4) share the same trends.



a *b*

Fig. 4 Patterns spread for H5: (a) 10 kV; (b) 0.4 kV. Blue cells correspond to pattern 1, red cells to pattern 2.

Finally, Fig. 5 shows the 2D scatter plot of high-dimensional clustered feature vectors to analyse the relevance of the results. For both voltage levels, one can see that the two clustered patterns are reasonably well separated. It indicates that the data features obtained from AE are effective to reveal distinctive patterns.

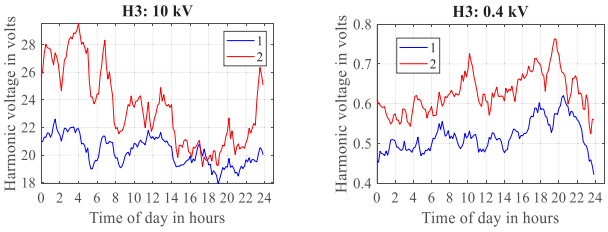


a *B*

Fig. 5 Results relevance for H5: (a) 10 kV; (b) 0.4 kV.

4.2 Results and analysis for 3rd harmonics

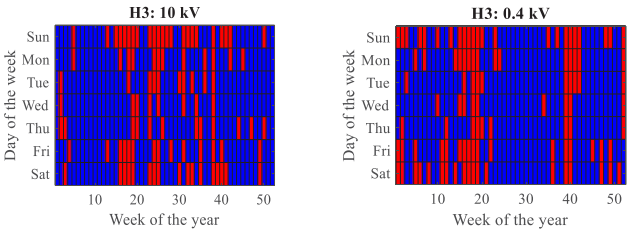
The two reconstructed patterns in 3rd harmonics at 10 kV present similar variations for some periods of the day, as shown in Fig. 6 (a). For instance, there are similar trends in the harmonic voltage between 4:00 and 7:00 and also between 19:00 and 24:00. Pattern 2 is higher than pattern 1 during the day, the only exception is for a period around 17:00 and 18:00. Fig. 6 (b) shows that the two patterns also show similar variations during the day in 0.4 kV, for example, the highest peak in the evening and some lesser peaks during daytime. However, these intraday variations in 0.4 kV are not the same as in 10 kV.



a *b*

Fig. 6 Reconstructed patterns for H3: (a) 10 kV; (b) 0.4 kV.

The spread of the patterns is randomly and distinctly distributed in the two voltage levels as shown in Fig. 7. Similarly, the scatter plots in Fig. 8 do not show any separation in pattern clusters.



a *b*

Fig. 7 Patterns spread for H3: (a) 10 kV; (b) 0.4 kV.

By Fig.7, one can see that summer days present more instances classified as low distortion days (pattern 1) in 0.4 kV than 10 kV. In contrast, high distortion days (pattern 2) are more incident in 0.4 kV than in 10 kV for winter days

between weeks 1 and 15. The only similarity in the voltage levels for the patterns spread is between the weeks 16 and 20, where the days were classified in majority to high-distortion days (pattern 2). However, the reconstructed patterns for pattern 2 in 10 kV and 0.4 kV do not show the same intraday variations in Fig. 6. Due to these differences in the patterns spread and reconstructed patterns, one could infer that the 3rd harmonic cause is not the same in both voltage levels, moreover that there is no propagation of the 3rd harmonics through MV/LV transformers.

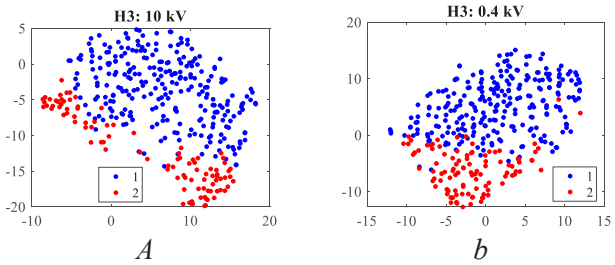


Fig. 8 Results relevance for H3: (a) 10 kV; (b) 0.4 kV.

4.3 Results and analysis for 7th harmonics

The reconstructed patterns for 10 kV and 0.4 kV of 7th harmonic (H7) (Fig. 9) present similar daily variation as H5, the dependence on the time of the week and the time of year is also very similar (Fig. 10). This indicates the same origin for both harmonics in both voltage levels. The separation between the clusters for H7 (Fig. 11) is as clear as for H5.

4.4 Results and analysis for 9th harmonics

The two reconstructed patterns in 10 kV present distinct variations during the day, as shown in Fig. 12 (a). For instance, there is a slight increase in pattern 1 between 6:00 and 18:00 while in pattern 2 there is a decrease during the same period. In contrast, the two patterns in 0.4 kV present similar intraday variations, for instance, both present an evening peak.

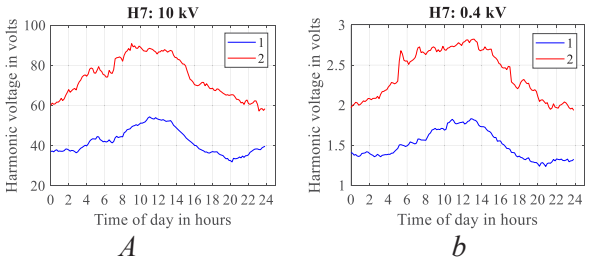


Fig. 9 Reconstructed patterns for H7: (a) 10 kV; (b) 0.4 kV.

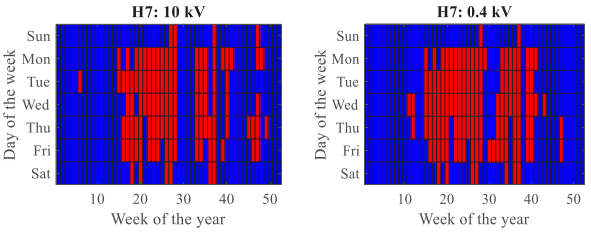


Fig. 10 Patterns spread for H7: (a) 10 kV; (b) 0.4 kV.

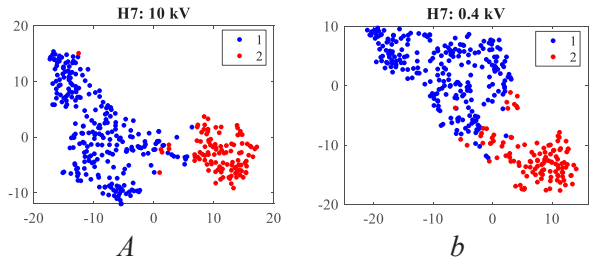


Fig. 11 Results relevance for H7: (a) 10 kV; (b) 0.4 kV.

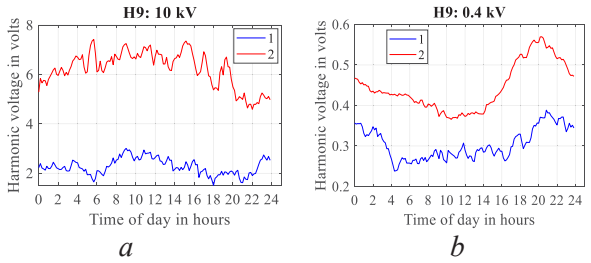


Fig. 12 Reconstructed patterns for H9: (a) 10 kV; (b) 0.4 kV.

Fig. 13 also points out an opposite spread of the patterns between 10 kV and 0.4 kV; the high distortion days (pattern 2) are more prominent in the summer and spring for 10 kV and the winter for 0.4 kV. The spread of the patterns is more homogenous in 0.4 kV; for instance, there are only two days classified as pattern 2 between weeks 20 and 26, while 10 kV shows both pattern 1 and pattern 2 during that period. Fig. 14 also points out that the patterns are more separated in 0.4 kV than in 10 kV.

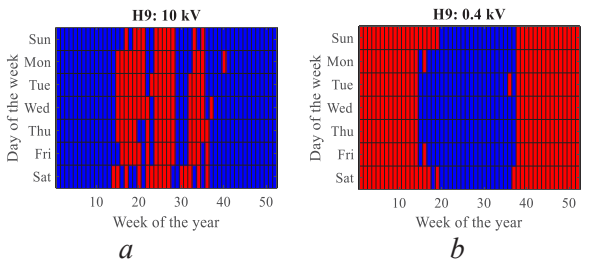
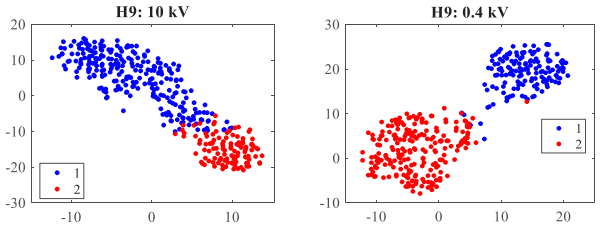


Fig. 13 Patterns spread for H9: (a) 10 kV; (b) 0.4 kV.



a *b*

Fig. 14 Results relevance for H9: (a) 10 kV; (b) 0.4 kV.

4.5 Discussion of the results

In order to confirm mathematically the inferences provided by the visualization of the DL results, Table 1 shows the correlations between the patterns for each harmonic and between the spread over the year. H3 presents a moderate correlation between patterns 1 and 2 in 0.4 kV and in 10 kV, indicating a common cause. However, the correlation is negative between the low-distortion patterns (pattern 1) in 0.4 kV and in 10 kV. The correlation is also negative between the high-distortion patterns (pattern 2). The correlation of patterns spread is also small (12.32 %). It points out that H3 does not spread from one voltage level to the other. H5 and H7 presented a high correlation between the patterns, and between the voltage levels. The correlation is also high for the spread of the patterns over the year. It confirms that these harmonics have the same source and that they propagate through MV/LV transformers. H9 presents higher correlation of the patterns in 0.4 kV than in 10 kV, which points out that the patterns in 10 kV are not related to the same cause. The correlation is negative between the pattern in 0.4 kV and 10 kV. The same holds for the spread of the patterns. This way, the inferences given by DL are aligned with PQ knowledge that harmonics 3 and 9 do not spread through MV/LV transformers. While H5 and H7 spread through a large part of the distribution.

Table 1 Correlations between the patterns for each harmonic

Correlation (%)	H3	H5	H7	H9
Pattern 1 and 2 in 0.4 kV	67	93	92	74
Pattern 1 and 2 in 10 kV	78	94	91	42
Pattern 1 in 0.4 kV and 10 kV	-51	97	93	-72
Pattern 2 in 0.4 kV and 10 kV	-67	98	96	-35
Spread in 0.4 kV and 10 kV	12	96	78	-68

4.6 Comparison between DL and traditional methods

Existing methods are limited to an expert to direct the analysis. For instance, an expert can decided to average the harmonic values according to temperature over a year: negative temperatures (autumn and winter), and positive temperature. By our method, the spread of patterns is obtained automatically. For instance, for the 5th harmonic, the plot of patterns spread (Fig. 4) revealed three weeks around week 30 with a pattern not related to weather conditions. Moreover, Fig. 15 shows that the quality of the daily patterns is also improved by DL for the 5th harmonic. DL can lead better with more severe days and reconstruct intraday variations in more detail then using simple values averaging.

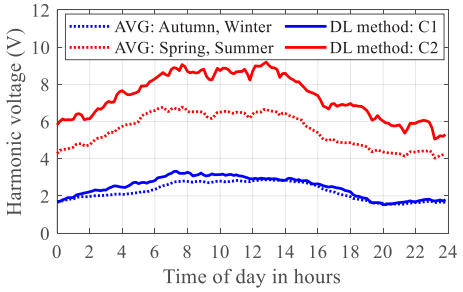


Fig. 15 Comparison between averaging and DL method.

5 Conclusion

This paper presented a method to extract automatically patterns from PQ data through DL. Through harmonics measurements in a distribution system, a guide for the interpretation of the automatic results given by DL was provided. It was shown that, even with without any pre-knowledge of DL, inferences related to the daily variations, seasonality, origin, and propagation of harmonics could be obtained simply by the visualization of the DL results. The inferences given by the visualization were confirmed by the correlation between the patterns. Besides, the inferences given by DL were aligned with PQ theory. Overall, this paper showed that our method could become an additional tool to assist power engineers, which compresses big data from PQ measurements in a form that is easier to interpret.

6 Acknowledgements

This work was financially supported by the Swedish Energy Agency.

7 References

[1] Vlahinić, T., Brnobić, D., Vučetić, D.: 'Measurement and analysis of harmonic distortion in power distribution systems', Electric Power Systems Research, 2009, 79, (7), pp. 1121–1126

[2] IEC 61000-4-30: 'Testing and measurement techniques-power quality measurement methods', 2009

[3] Caramia, P., Capinelli, G., and Verde, P.: 'Power quality indices in liberalized markets' (John Wiley & Sons, 2009)

[4] Santoso, S., Powers, E., Grady, W., Parsons, A.: 'Power quality disturbance waveform recognition using 29 wavelet-based neural classifier', IEEE Transactions on Power Delivery, 2000, 15, (1), pp. 222–228

[5] Ge, C., Oliveira, R., Gu, I., Bollen, M.: ' Deep Feature Clustering for Seeking Patterns in Daily Harmonic Variations', IEEE Transaction on Instrumentation and Measurement, 2021, 70, (1), pp. 1-10.