EFFICIENT POWER QUALITY ANALYSIS OF BIG DATA (CASE STUDY FOR A DISTRIBUTION NETWORK OPERATOR)

Etienne GASCH, Jan MEYER, Peter SCHEGNER
Technische Universitaet Dresden – Germany
etienne.gasch@tu-dresden.de

Karsten SCHMIDT
ENSO NETZ GmbH - Germany
karsten.schmidt@enso.de

ABSTRACT

Power Quality (PQ) monitoring plays an important role for a trouble-free operation of distribution grids. Prices for PQ monitoring equipment have continuously decreased during the last decade. Especially the integration into other equipment (IED – Intelligent electronic devices) like energy meters enables extremely affordable solutions. Due to this trend, network operators have intensified their PQ measurement activities and the number of monitored sites grows fast. As consequence the amount of PQ measurement data gets larger and larger and its efficient management and analysis becomes an increasing complex challenge. As also identified be the CIRED/CIGRE working group C4.112 "Guidelines for Power quality monitoring" [1], efficient algorithms to analyze this big data are a key issue in the future.

Starting with a short discussion of some major challenges of future PQ monitoring, the paper proposes a method for the calculation of an easy-to-interpret and flexible to aggregate PQ index. Next a brief description of the techniques developed for a device-independent handling of measurement data is provided. The main part of the paper addresses methods for efficient graphical visualization of large PQ data amounts. Two different approaches based on maps with fixed and flexible PQ zoom level are proposed. Finally the application of both visualization methods is illustrated in a case study with more than 80 measurement sites operated by a local distribution system operator (DSO).

INTRODUCTION

The software that comes bundled with PQ measurement devices is perfect for the detailed analysis of single sites and shorter time intervals. If the data amount increases, this software often becomes more and more impracticable and new approaches for data management and data analysis are required.

In terms of data management a new software has to deal with data coming from different vendors. The use of multiple software packages for different device types within one utility is not acceptable. Therefore special data filter applications have to convert proprietary file formats or databases from different vendors into a single, device-independent data stream. The filters can be implemented on file level (e.g. PQDIF according to IEEE Std. 1159.3), on database level (e.g. ORACLEs stored procedures) or on device level (approaches based on IEC 61850). More details can be found in [2].

In terms of data analysis the efficient and fast processing of large data amounts is an important issue. In Europe usually the PQ analysis is carried out according to the standard EN 50160. This means 82 PQ parameters[1] every ten minutes at each measurement site. If 100 sites are measured for 365 days this will result already in more than 400 million data points. A new analysis software has to present these data in a clear and easy-to-interpret way. Therefore new methods for a flexible data aggregation are essential.

The results of a survey about current practice of Power Quality assessment, which was carried out by the CIGRE/CIRED WG C4.112, also support the importance of the above requirements [3].

Figure 1 shows the principle schema of a measurement system that satisfies the above described requirements and which was implemented by the authors. The proprietary data of the individual measurement devices is converted and finally stored in a device-independent way. At present as storage system mostly databases like Oracle or MySQL are used. This central storage approach is only feasible, if the number of measurement devices does not exceed a critical value. If in the future millions of smart meters would be capable to measure PQ, hybrid approaches with an increased handling/storage and analysis functionality within the measurement devices could be required. More details about such approaches can be found in [2, 4].

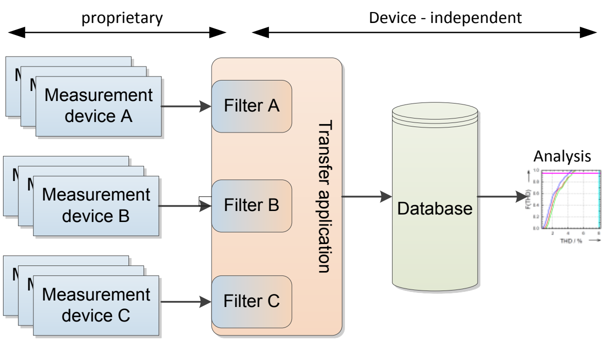

Figure 1 Principle schema of measurement system

With e.g. the UNIPEDE DISDIP table an accepted method exists for data aggregation of events, like voltage dips and swells or interruptions. For variations, like THD or flicker, such a generally accepted method is not yet available. One approach which is based on [5] and represents a further development of [6] introduces a system of normalized indices. The calculation of the

1 Power frequency is not considered.

indices is based on EN 50160, but other standards can be used as well. In general four different indices are distinguished:

- PQ index (QI)
  (assessment of the 95th percentile)
- Risk index (RI)
  (assessment of the data above the 95th percentile)
- Trend index (TI)
  (assessment of possible long-term trends)
- Variation index (VI)
  (assessment of e.g. seasonal variations)

While the first two indices are analyzed on weekly basis, the last two are applied to time series of the first two indices.

In general the visualization of the indices has to follow a top-down approach. Starting with a few highly aggregated indices, which could e.g. be part of a network control system, the user can obtain a quick and comprehensive overview of the PQ status of the whole system. More details can be obtained by “jumping” down to levels with lower aggregation. The indices can either be presented in tables and charts or by geographical maps.

This paper focuses on ways for the geographical presentation of the PQ index QI. It starts with a short explanation, how QI is obtained and introduces two methods for graphical visualization. Finally both methods are applied to the measurement data of 86 sites of a MV distribution network. The data is provided by a regional DSO.

THEORY OF POWER QUALITY INDEX

Measuring PQ according to EN 50160 means analysing 82 PQ parameters for every measurement site and every 10-minute time interval: 3x THD, 3x voltage magnitude, 1x unbalance, 3x longterm-flicker, 3x24 voltage harmonics. To ensure comparability between the different PQ parameters, the 95th percentiles (m) of the 10-minute-values of one week where normalized by relating them to their limits (g).

$$QI = \left(1 - \frac{m}{g}\right) \cdot 100 \qquad (1)$$

The resulting PQ index represents the percentage reserve of the considered PQ parameter to the respective limit. The PQ index QI is divided into four categories, which are indicated by different colors:

Table 1 Proposed categories for PQ index

Cat.	QI value range	Description
A	QI > 50	High reserve (green)
B	25 > QI ≥ 50	Sufficient reserve (yellow)
C	0 > QI ≥ 25	Low reserve (orange)
D	QI ≤ 0	No reserve/limit violation (red)

Colors and values are illustrated in Fig. 2 by using THD and longterm-flicker as example. The limits according to EN 50160 are $P_{lt} = 1$ and THD = 8 %.

P_{lt}	>1	1	0,75	0,5	0
THD / %	>8	8	6	4	0
QI	<0	0	25	50	100

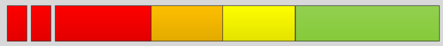

Figure 2 Colored PQ index QI

An individual PQ index QI_i is calculated for each PQ parameter, each week and each site. These can be considered as three dimensions and can be presented by a cuboid as shown in Fig. 3, where each small cube represents one individual PQ index QI_i.

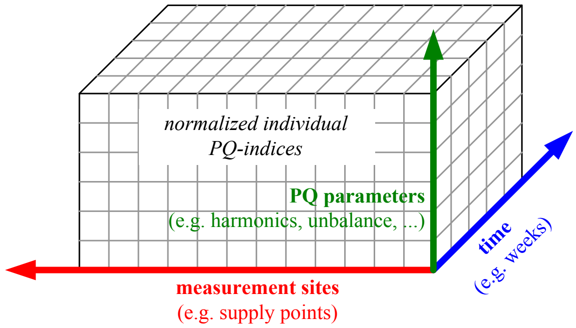

Figure 3 PQ index visualization by a cuboid

Aggregation can be best visualized as squeezing the cuboid in one or multiple dimensions. Aggregation in dimension of PQ parameter (green line) e.g. results for each site and each week in a single, so-called site PQ index QI_s. The cuboid becomes a rectangle and can be easily presented e.g. as colored contour plot (see [2] for an example). Multiple site indices can be further aggregated to region PQ indices QI_r representing different supply areas. If knowledge about the minimum reserve for a particular time period is required, aggregation is done along the time axis (blue). An aggregation in all three dimensions (time, site, PQ parameter) results in one single index, which is called network PQ index QI_n and represents the minimum reserve for all considered sites, PQ parameters and measurement weeks. A network index of $QI_n = 42$ e.g. means that no PQ parameter at any site and any week has less than 42 % reserve, which is usually enough information e.g. for management staff. Details about how the aggregation is performed can be found in [6].

The described aggregation flexibly reduces the amount of presented data for a better overview of the PQ status of a network or network region in any type of presentation. By disaggregating the network PQ index, it is possible to find the respective individual PQ index, which is responsible for a low network PQ index. This way the user can easily find the cause of any low reserve.

The PQ index QI can be presented in tables, charts and maps. This paper will focus on techniques for visualization by geographical maps.

DATABASE IMPLEMENTATION ISSUES

Each PQ measurement survey needs a database to store the large amount of measurement data. Each vendor of measurement devices uses its own methodology of data handling, storage and analysis. This makes it virtually impossible to perform overall analyses for measurement data of different vendors. Therefore the application of the PQ index requires a device-independent database.

During conversion and transfer of the individual data formats into the device-independent database, the software can already calculate selected PQ indices. This increases the performance of the PQ index visualization software considerably. It also enables additional querying possibilities, e.g. weeks and/or sites with PQ indices below a certain threshold or within a particular PQ category (cf. Tab. 1).

In addition to the PQ parameters itself, the database is designed to store further data describing the measurement site. These can be divided into electrical parameters, like short circuit power or consumer topology and non-electrical parameters, like the geographical coordinates. The benefit of these additional parameters is an extension and increased flexibility while querying the database. It enables e.g. to search for sites, which have a short circuit power below a certain value. By geographical coordinates it is possible to find measurement sites in a selected area or at a defined altitude. If e.g. a particular site has a PQ reserve of only 10 %, it is easy to find close by sites in order to assess the impact of the poor quality on neighboring networks. Geographical coordinates have a benefit compared to the postal address when sites are moved. Especially in MV networks it happens quite often that a distribution station is moved short distances (e.g. to the other site of a crossing). While it can get a completely new address, the geographical coordinates remain almost the same.

METHODS FOR GEOGRAPHICAL VISUALIZATION

Using the geographical coordinates, the PQ index for each measurement site can be plotted in geographical maps or included in an existing GIS (geo information system). Depending on the aggregation level of PQ index, which is referred to as “PQ zoom level” in this paper, two kinds of maps are considered for representation: maps with variable PQ zoom level and maps with fixed PQ zoom level. Any available data filter can of course also be applied to the graphical presentation (e.g. only THD, only sites with low short circuit power, …).

Maps with flexible PQ zoom level

The major idea of different PQ zoom levels is the representation of different PQ index aggregation levels. Fig. 4 exemplarily shows a low PQ zoom level of a particular region of a supply area. The PQ zoom level is high enough to make all individual site PQ indices for this particular region visible. Zooming out will further aggregate the site PQ indices of the region into a region PQ index (Fig. 5). In order to keep the map simple but comprehensive, the aggregation is carried out individually for the four PQ categories according to Tab. 1. Finally the best and worst PQ categories with their respective PQ index are presented. The total number of aggregated sites is indicated by an additional small box.

As example Fig. 4 shows 7 measurement sites with site PQ indices ranging from 17 to 87 and covering the three PQ categories A, B and C. Zooming out (Fig. 5) will merge/aggregate the site PQ indices into the region PQ index. For the worst category (cat. C) the minimum value $QI_s = 17$ is kept, for the best category (cat. A) the minimum value $QI_s = 62$ is kept. The small box with white background indicates the total number of aggregated sites.

Maps with fixed PQ zoom level

Maps with fixed PQ zoom level tries to find regions with similar PQ index values in an automatic way. As example Fig. 6 shows the same map as Fig. 5. All sites are clustered into groups with similar site PQ indices. The clustering is performed autonomous and is repeated with each update of the PQ indices. It tries to combine as much as possible sites that are close to each other and

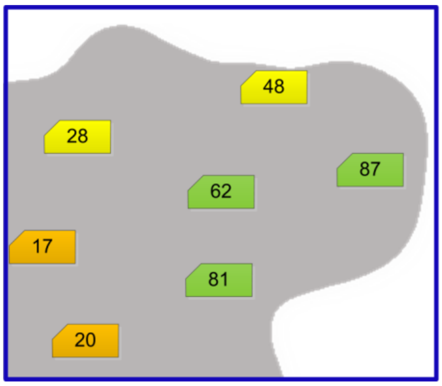

Figure 4 Flexible map of fictive values at high PQ zoom level (aggregation level at site indices)

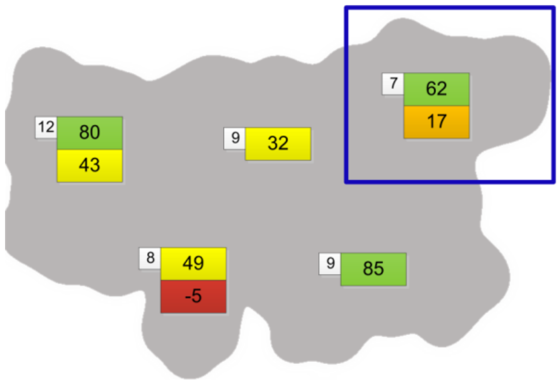

Figure 5 Flexible map of fictive values at low PQ zoom level - (aggregation level at regional indices; blue box indicates zoomed region in Fig. 4)

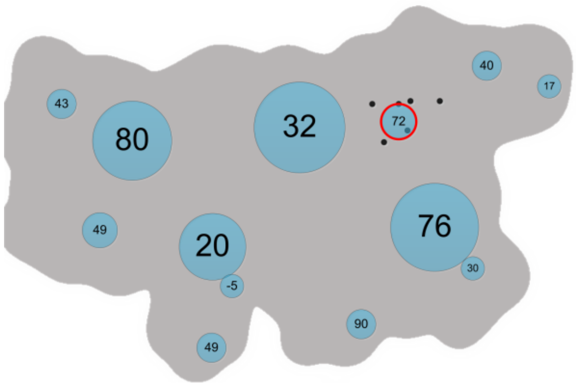

Figure 6 Map with fixed PQ zoom level (fictive values)

have similar site PQ indices. This clustering uses a distance measure based on site PQ index values and geographical distances and has the target to minimize the number of clusters. The size of the circles corresponds directly to the number of sites in the cluster. The circle is centered in the geographical center of the clustered sites. By selecting a circle, the associated sites are shown. This is exemplarily illustrated by the black dots around the selected circle, which is highlighted with a red edge (Fig. 7).

The value in each circle represents the minimum site PQ index of the cluster. It has a similar meaning than the region PQ index in Fig. 5, but without the grouping by PQ categories. If different clusters would overlap in the map, the smaller cluster is moved to the nearest edge of the bigger cluster.

Another possibility is to define a set of network regions (fixed site clusters) and to allocate the measured sites to these regions. In this case each region could be colored based on the PQ category of the region PQ index.

APPLICATION EXAMPLE

Background

During the regular exchange of their energy meters in the MV distribution system, the ENSO NETZ GmbH decided to replace the old meters by new ones with integrated PQ functionality. The meters are exclusively used for internal metering purposes and not for billing. Even if the implementation does not cover all PQ parameters according to EN 50160 (only voltage magnitude, THD, selected harmonics, flicker) and does not fully comply with IEC 61000-4-30, the data is useful for an informative overview of the PQ situation in the grid. One important advantage is the low additional price of the PQ functionality and the already existing infrastructure for data management. Anyway meter manufacturers, who include PQ measurement in their meters, are encouraged to comply at least with IEC 61000-4-30 class S. For full comparability between different instrument types and/or vendors, instruments complying with class A must be used. More details about cost aspects of PQ monitoring can be found in [7].

Until mid of 2013 about 120 new meters had been installed. Due to differences in configuration for most meters THD, flicker and selected harmonics are available. RMS values are available for all meters, but as phase to ground values only. These are only of limited use, because in MV grids the phase to phase values of voltage magnitude must be assessed. The whole data transfer is realized by the software of the meter manufacturer. An individual converter was developed to scan the incoming folder for new PQ measurement data and to transfer it into the device-independent database. In general any analysis software can access the data by SQL queries. In this case study a specific MATLAB program was developed for the analyses.

Results

To illustrate the application of the different maps, one week in 2011 with 86 available measurement sites is selected. As map a stylized map of the whole supply area of the DSO is used. Only THD and flicker of every phase are selected for calculating the site PQ indices.

In order to provide a simple overview about the PQ situation, all measurement sites are displayed in the stylized map at their geographical position using the color code according to Tab. 1 (Fig. 7). Most of the sites are in green color (high reserve). Every considered PQ parameter of the “green” sites has more than 50% reserve to the limit of EN 50160, which means that the 95th percentile is $P_{lt} < 0,5$ for flicker and THD < 4 % for total harmonic distortion. The red sites in the upper left corner represent industrial customers of the DSO, which are not directly connected to the public MV network. The benefit of such a map is the easy identification of regions, where the limits of EN 50160 are not met.

A map with flexible PQ zoom is used to quantify the PQ index. Fig. 8 presents the map at its lowest PQ zoom level. It was decided in the project to keep at minimum PQ zoom level seven regional PQ indices. This allows more information than the single network PQ index only, but is still comprehensive. The 86 site PQ indices range from QI = -18 to QI = 71. The network PQ index corresponds to the minimum site PQ index, which is QI = -18. Excluding the industrial customers the network PQ index amounts QI = 55. Only one region, namely the one with the industrial customers does not comply with

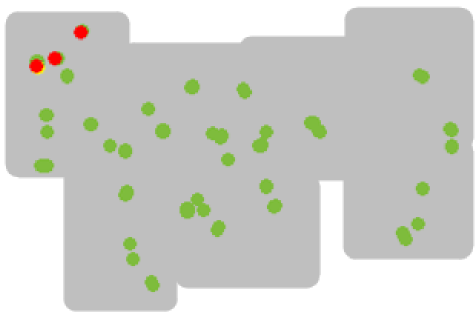

Figure 7 Stylized map with all sites colored by PQ category of the site index

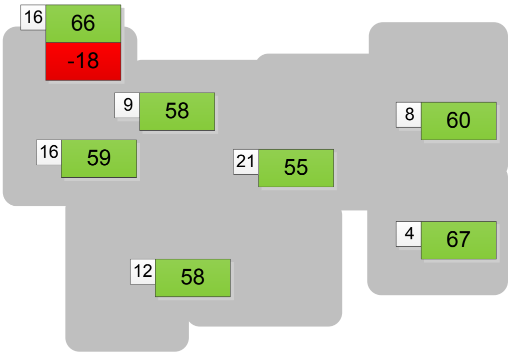

Figure 8 Map with flexible PQ zoom level at lowest level (regional indices shown)

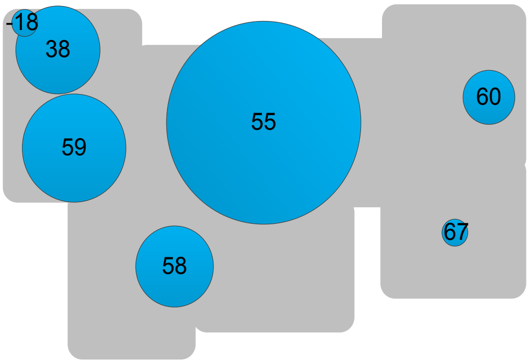

Figure 9 Map with fixed PQ zoom level

the limits according to EN 50160. All other sites are "green" with a PQ reserve of more than 50 % for the considered PQ parameters.

Zooming into the map provides further details until the site PQ index level. Further information can be obtained by an additional message box, which appears if the mouse is moved over a PQ index. While zooming out, a cluster method based on geographical distances is used to aggregate the site PQ indices stepwise to regional PQ indices until only the initially mentioned seven PQ indices are visible. In all regions the number in the colored box shows the worst PQ index (e.g. QI=67% in the lower right corner) of the aggregated sites. The region in the upper left corner shows two PQ indices (QI = -18 and QI = 66). As described in the section before, this indicates that the PQ index in this region covers more than one PQ category. Therefore the best and the worst PQ category are shown together with their respective lowest PQ index.

The presentation of the PQ index based on fixed PQ zoom level maps is shown in Fig. 9. Again the stylized map is used. In contrast to the clusters in the map with flexible PQ zoom level, the clusters are not only obtained by geographical distance, but also by similarity of PQ index values. Therefore even without the ability to zoom, number and size of circles can change over time. The number in each circle shows the minimum PQ index of the respective region. The center of a circle is located mostly in the center of the aggregated region. If two circles overlap, the smaller circle is shifted to the nearest edge of the bigger circle (see upper left corner in Fig. 9).

SUMMARY

Due to the increasing integration of PQ functionality into intelligent devices, like meters or protection relays, the amount of PQ measurement data is expected to growth very fast. In order to cope with this big data new approaches for data management and data analysis, especially for continuous PQ parameters like THD, are required.

Beside design aspects for the implementation of a device-independent database the major part of the paper addresses methods for a comprehensive graphical visualization of large PQ measurement data. It is based on a PQ index, which is easy to interpret and flexible to aggregate and represents the remaining reserve compared to a set of limits. At the top level (highest aggregation) only a few values are provided, which are useful for reporting to e.g. the management staff or in a network control center. The most detailed level is intended for PQ experts. The graphical presentation is based on maps with flexible or fixed PQ zoom level. The maps provide a quick and comprehensive overview about the PQ status of a whole supply area and allow an easy identification of network regions, which are problematic in terms of Power Quality. Currently the authors develop a Web application, which allows an interactive PQ analysis making printed reports almost obsolete.

REFERENCES

[1] CIGRE Report 596: CIGRE/CIRED Joint Working Group C4.112, "Guidelines for Power Quality Monitoring", October 2014.

[2] J.M.R. Gordon, et. al., 2011, "Design Aspects for large PQ Monitoring Systems in future Smart Grids", *IEEE PES General Meeting,* 2011.

[3] J. Milanovic et.al., 2014, "International Industry Practice on Power-Quality Monitoring", IEEE Transactions on Power Delivery 29/2, 2014.

[4] F. Zavoda et. al., 2015, "CIGRE/CIRED/IEEE Working Group C4.24 – Power Quality in the Future Grid – Status Report", *this conference*

[5] V.J. Gosbell, B.S.P. Perrera, H.M.S.C. Herath, 2002, "Unified power quality index (UPQI) for continuous disturbances" *10th International Conference on Harmonics and Quality of Power (ICHQP)*, vol. 1, p.312-316

[6] J. Meyer et.al., 2005,"Efficient Method for Power Quality Surveying in Distribution Networks", *18th CIRED*, Italy 2005.

[7] J. Meyer et. al., 2012, "Contemporary and future aspects of cost effective power quality monitoring — Position paper of CIGRE WG C4.112", *Electric Power Quality and Supply Reliability Conference*, Estonia, 2012.