

THE VARIATION OF POWER QUALITY INDICES DUE TO DATA ANALYSIS PROCEDURE

Sean Elphick and Vic Gosbell

Integral Energy Power Quality & Reliability Centre
School of Electrical, Computer and Telecommunications Engineering
University of Wollongong
Wollongong, NSW, 2522

Abstract

Power quality data is often reported using statistical confidence levels. This will exclude the most extreme data for a certain length of time depending on the interval over which the confidence level is applied. There is considerable conjecture as to the effect of applying statistical measures over different time intervals, e.g. several days, weeks or one year. If statistical confidence levels are applied over long intervals, the length of time not included in the statistical confidence interval is long. During such intervals disturbance levels may be continuously high and not be accounted for in the statistical parameter. This study investigates the effect different methods of aggregating data to a specific reporting period will have on the calculated index. Several data processing methods are trialled to evaluate the effect of using different aggregation intervals to produce an index to characterise disturbance levels for the whole year.

1. Introduction

There is a trend toward routine continuous monitoring of power quality to maintain a check on the "health" of the power system [1]. Commonly, measurements are taken at a representative sample of sites for voltage, unbalance, harmonics, sags and sometimes flicker. This type of routine monitoring results in copious amounts of data, particularly where numerous harmonic orders are monitored. For reporting purposes, this large amount of data needs to be reduced to a meaningful readable form without the loss of important detail. To achieve this it is necessary to aggregate data the many measured values to a representative single value or index, to characterise system behaviour over days, weeks or years depending on the specific reporting requirements.

There are two important concepts involved with the measurement and reporting of power quality data. The first is data sampling and aggregation methods. These are the methods used by power quality instrumentation to sample the raw data and aggregate it to acceptable time periods. The current IEC standard for power quality monitoring, IEC61000-4-30 [2] details the methods which power quality instrumentation should used to collect data.

Once raw data has been collected it is necessary to reduce the large amount of data to a form that will produce useful indices for reporting purposes; this is the second concept of power quality monitoring and reporting. Refs [3] and [4] describe methods that can be used to report power quality data. Power Quality levels are reported over common intervals such as weeks, months and years. How data is reduced from short time

periods to longer reporting periods and the variations that may be encountered depending on the methods used to achieve this are the subject of this study. Although [5] and [6] detail the interval over which data should be aggregated for assessment of harmonics and flicker respectively, no rationale is given and it is not clear how they should be chosen for assessing other common power quality disturbances such as steady state voltage.

Aggregation of power quality data from raw data sampled at short intervals such as 10 minutes to a predefined reporting period is usually achieved by way of statistical confidence levels. This technique is often used to avoid results being affected by the highly volatile maximum and minimum values which do not give a fair indication of the actual behaviour of the network for most of the time. The most commonly utilised statistical measure for reporting power quality data is the 95th percentile value. This value is recommended for use in [2] and is also used as the criteria for comparison of disturbance levels with limits in [5] and [6].

The concern that arises with this approach is that there is no way of knowing the behaviour of disturbance levels excluded by the statistical confidence level. For example, if the measure of interest was the disturbance level over one week, one may take the 95th percentile value of all the data recorded over the week to characterise the week. This leaves 5% of the week, or 8.4 hours worth of data excluded which will be higher than the reported value. It is possible that the site might have very high values of disturbance levels over 8.4 continuous hours with no effect on the statistical confidence level. Alternatively, the 95th percentile value

may be taken over each day of the week and these daily values aggregated to give the weekly measure. By taking the 95th percentile value daily the period of time every day now outside of the statistical analysis reduces to 5% of one day or 1.2 hours where the disturbance level will be higher than the reported value.

Analysis of data in this way poses two questions; the first is how much variation in calculated disturbance indices will be seen depending on the time interval over which the statistical confidence levels are applied; and the second is given that most power quality disturbances lead to continuous long term degradation of equipment so long as disturbance levels are not extreme how much difference does it really make?

This study addresses the first question posed above. Although [2], [5] and [6] give some indication of the methods that should be used to report power quality data, the reality is that many different methods are being used in industry and little is known about the effect that this will have on reported values. To investigate this, this study explores the variation in results when power quality levels are reported over 1 year using different methods. The characteristic statistical confidence level used to describe the year is the 95th percentile applied over various time intervals. Four common time intervals over which data may be aggregated to a yearly value are examined, namely, over the whole year, in monthly intervals, in weekly intervals and in daily intervals.

2. Test Data

This study uses voltage, voltage unbalance and harmonic voltage (THD) data as the test data for comparing indices calculated using different aggregation methods. The data used for comparison in this study is sourced from data collected as part of the Long Term National Power Quality Survey (LTNPQS) project conducted by the University of Wollongong and detailed in [7]. This project has collected power quality data from over 400 sites predominately from the eastern seaboard of Australia for the past several years. Data from the 2004/2005 financial year is used in this study.

Sites to be used in this study were chosen based on data availability for the year. For the purposes of this study, the criterion for the data from a site to be included was that the site had data available for 99% of the year. Thirty-seven such sites were identified and make up the data set for this study. The data used in this study was monitored at intervals of between 10 minutes and 1 hour. The sites used in this study are a mixture of low voltage and medium voltage sites.

Although sites were selected that had 99% data availability for the year, this still leaves approximately 4 days for which the site may not report any data. If no data was present for a whole day, that day was simply excluded from the analysis. For example if one day was missing from a week, the weekly value was calculated using the remaining six days.

So that only data representing normal operating conditions was included in the study the data used in this study was first filtered. Data was excluded based on the following criteria:-

- If it was less than 0.8 per unit or greater than 1.2 per unit for voltage.
- If it was greater than 20% for unbalance.
- If it was greater than 20% for THD.

3. Test Reporting Periods

There are an infinite number of methods which may be used to aggregate and statistically analyse data to reduce it to produce a single index for the reporting period. One of the most common statistical analysis methods is use of the 95th percentile value as outlined in the introduction. In order to simplify this study, the 95th percentile value has been adopted as the statistical parameter that will be utilised in generating characteristic indices for the reporting intervals that are to be examined.

An added layer of complexity arises when it is necessary to characterise a longer time period using shorter time periods, for example evaluating a weekly value using daily values. The weekly value could be characterised by any statistical measure of the seven daily values. For example, the maximum of the daily values may be used to characterise the week, or it may be the average, or even the 95th percentile. The method used to perform this type of aggregation will have an effect on the reported value, however, the most appropriate of the measures to use and the benefits and drawbacks of each is beyond the scope of the study. For the purposes of this study when data for a longer reporting period is to be evaluated using shorter time periods the maximum of the values of the shorter time periods is taken to be characteristic. That is if a weekly value is to be evaluated from seven daily values, the maximum daily value is taken to characterise the week. This is common practice and is the technique used by the University of Wollongong when reporting power quality data.

As stated this paper focuses on methods of obtaining a yearly value for power quality disturbances. Four methods of obtaining a yearly value are examined in this paper namely:

- (i) The 95th percentile value of the disturbance over the whole year
- (ii) The maximum of the 12 95th percentile values calculated for each month
- (iii) The maximum of the 52 95th percentile values calculated for each week
- (iv) The maximum of the 365 95th percentile values calculated for each day

Given the above methods of calculating a yearly index the amount of data that will be excluded depending on the time interval over which the 95th percentile value is applied is as follows:

- For the 95th percentile value over the year, 18 days worth of data will be excluded.
- The monthly method of calculation of the 95th percentile value excludes between 1.4 and 1.55 days worth of data depending on the number of days in the month.
- The weekly method excludes 8.4 hours worth of data.
- The daily method excludes 1.2 hours worth of data.

Examining the above, it can clearly be seen how concerns over the use of statistical confidence levels and how they should be applied arises. In the case of reporting a yearly value 18 days worth of data will be excluded from the 95th percentile value. Using a single statistic it is impossible to know if these 18 days occur continuously and how high the levels actually are. At the same time use of a daily 95th percentile value excludes only 1.2 hours worth of data across the whole year which is very little. For data sampled at 10 minute intervals this represents only 8 intervals (out of 52560) that will be excluded for the whole year, for data sampled at 1 hour intervals only 2 (out of 8760) intervals will be excluded across the whole year. This may produce an index that is overly pessimistic and may be significantly higher than the disturbance levels for the vast majority of the year.

Appendix A shows sample yearly trends for each disturbance examined in this study showing how data is excluded depending on the time interval over which the 95th percentile is applied.

4. Results

4.1 Variation of Voltage Data

Assessment of voltage data is made using the Absolute Voltage Deviation, an index developed by the University of Wollongong. The AVD is a method of calculating a voltage index that describes the absolute deviation of voltage levels around the centre of the voltage range. Methods of calculating the AVD are fully detailed in [8].

Figure 4.1 shows the variation between a yearly AVD value calculated by taking the 95th percentile of the AVD values across the whole year and a yearly AVD value calculated by taking the 95th percentile of the AVD values over each month and using the maximum of these to characterise the year. Figure 4.2 shows the variation between a yearly AVD value calculated by taking the 95th percentile of the AVD values across the whole year and a yearly AVD value calculated by taking the 95th percentile of the AVD values over each week of the year and using the maximum of these to characterise the year. Figure 4.3 shows the variation between a yearly AVD value calculated by taking the 95th percentile of the AVD values across the whole year and a yearly AVD value calculated by taking the 95th percentile of the AVD values for each day of the year

and using the maximum of these to characterise the year.

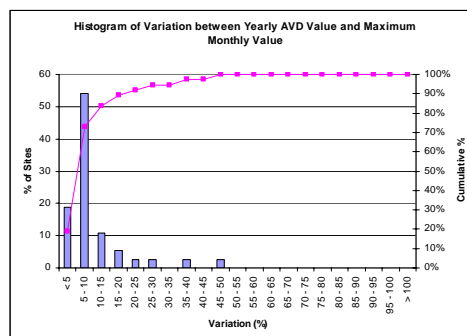


Figure 4.1: Variation between AVD values calculated Yearly and Monthly

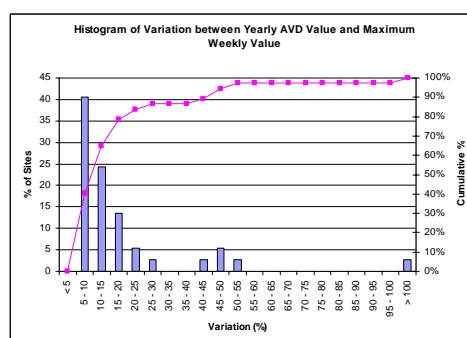


Figure 4.2: Variation between AVD values calculated Yearly and Weekly

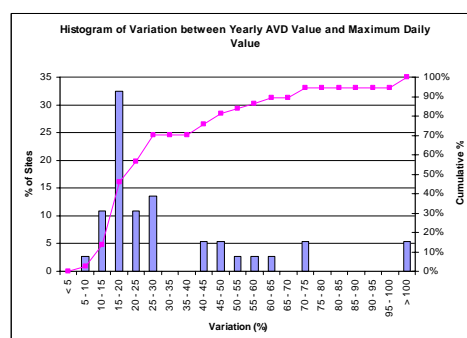


Figure 4.3: Variation between AVD values calculated Yearly and Daily

Figures 4.1, 4.2 and 4.3 show that there will be significant variation in the reported value of AVD depending on how the 95th percentile value is calculated.

Figure 4.1 shows that majority of variations between yearly AVD values calculated across the whole year and yearly AVD values calculated using the maximum monthly method are less than 10%. That is, a yearly AVD value calculated using the maximum of the monthly values method will be, on average, 10% higher than a yearly AVD value calculated by taking the 95th percentile of the AVD values across the whole year. Figures 4.2 and 4.3 show that as the period over which the AVD value is calculated is reduced larger variation in the reported values is observed. The weekly reporting period is the one which is recommended for use by IEC61000-4-30. The average variation between yearly

AVD values calculated using the weekly approach and the whole year approach is found to be 19% with the maximum 119%.

For yearly AVD values calculated using the maximum daily 95th percentile value, the average variation is 35%, that is, the yearly AVD value calculated using the maximum day approach will be on average 35% larger than the yearly AVD value calculated by taking the 95th percentile value across the whole year. This is a significant figure. The maximum variation seen between a yearly value calculated using these two approaches is 174%.

4.2 Variation of Unbalance Data

Assessment of unbalance data is achieved through calculation of the 95th percentile level over the time interval of interest. Figures 4.4, 4.5 and 4.6 show the variation between yearly unbalance values calculated using the 9th percentile value over the whole year and (i) the maximum monthly 95th percentile value, (ii) the maximum weekly 95th percentile value and (iii) the maximum daily 95th percentile value respectively.

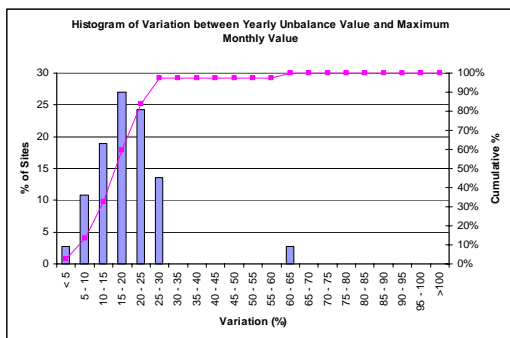


Figure 4.4: Variation between Unbalance values calculated Yearly and Monthly

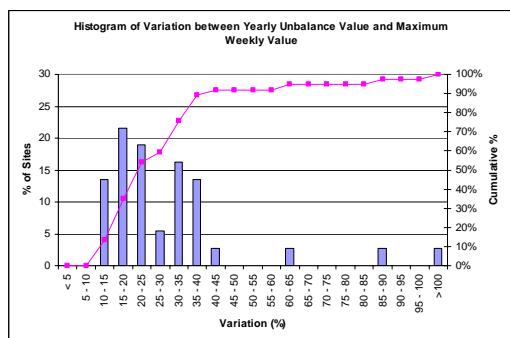


Figure 4.5: Variation between Unbalance values calculated Yearly and Weekly

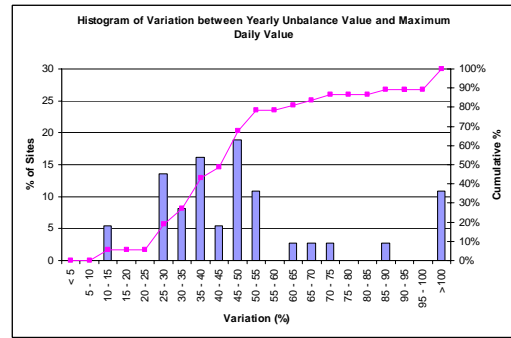


Figure 4.6: Variation between Unbalance values calculated Yearly and Daily

Figures 4.4, 4.5 and 4.6 show similar results are observed for unbalance as were observed for voltage with variation in unbalance being slightly larger than that seen for voltage. Once again there is significant variation in the reported values dependant on the time interval over which the 95th percentile value is applied. The average variation between an unbalance value calculated over the whole year and an unbalance value calculated using the maximum monthly value was found to be 19% and the maximum variation was 64%.

For the unbalance calculated over weekly and daily intervals the average variation between these values and one calculated over the whole year was found to be 31% and 56% respectively. This variation was the highest observed for the 3 disturbances considered in this study. The maximum variation for the same methods of calculation was found to be 176% and 264% respectively.

4.3 Variation of Harmonic (THD) Data

Assessment of harmonics data is achieved through calculation of the 95th percentile level of the Total Harmonic Distortion (THD) over the time interval of interest. Figures 4.7, 4.8 and 4.9 show the variation between yearly unbalance values calculated using the 9th percentile value over the whole year and (i) the maximum monthly 95th percentile value, (ii) the maximum weekly 95th percentile value and (iii) the maximum daily 95th percentile value respectively.

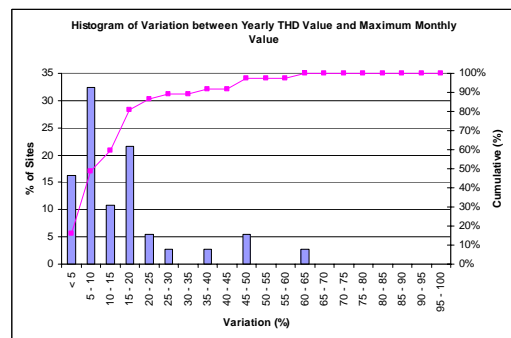


Figure 4.7: Variation between THD values calculated Yearly and Monthly

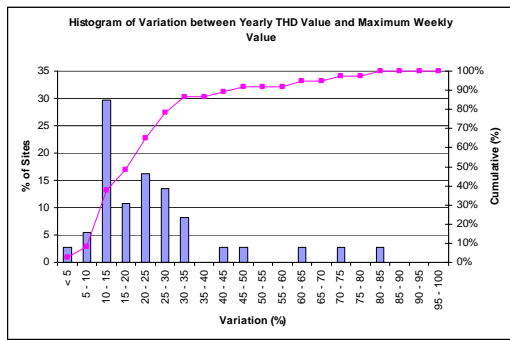


Figure 4.8: Variation between THD values calculated Yearly and Weekly

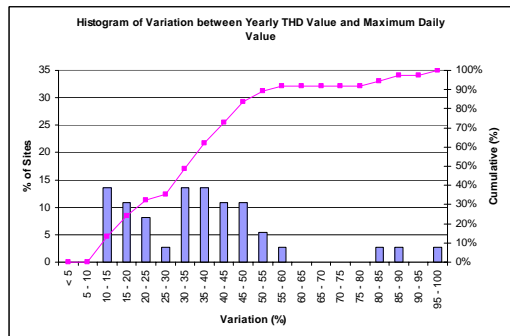


Figure 4.9: Variation between AVD values calculated Yearly and Daily

Figures 4.7, 4.8 and 4.9 again indicate that THD values show significant variation based on the method used for index calculation. For the yearly value calculated using the whole year method as opposed to the maximum monthly method, the average variation was found to be 15% and the maximum variation was found to be 60%.

The average variation between the yearly method and the maximum weekly method was found to be 24% and the maximum 80%.

Figure 4.9 shows that there is large variation between yearly THD indices calculated over the whole year and a yearly THD index calculated using the maximum daily value. Analysis shows that the average variation between these two indices is 37% and the maximum is 96%.

4.4 Summary of Variation

Table 4.1 conveniently summarises the information found previously. Using the value calculated across the whole year as a benchmark the normalised values of the indices corresponding to the shorter intervals have been calculated. The average value of these across all sites is given in Table 4.1 as the parameter Avg. To show the range of variation across the sites the coefficient of variation shown as the parameter CV in Table 4.1 has also been determined. It can be seen from Table 4.1 that the variation between the indices calculated to give a yearly value increase as the interval over which the statistical confidence level is applied decreases.

The coefficient of variation allows conclusions to be made as to the predictability of a value for a shorter time

interval based on a value calculated using the 95th percentile value across the whole year. Table 4.1 shows that there is a high level of confidence between the value calculated over the whole year and the maximum monthly value. For example, it can be concluded that it is highly likely that for voltage, the maximum monthly value will be 10% higher than a value calculated by applying the 95th percentile across the whole year. It can be seen that as the interval over which the statistical confidence levels is applied decrease, the coefficient of variation increases indicating that values for shorter term intervals can be predicted with less confidence.

Table 4.1: Summary of Variations

Disturbance	Calculation Method					
	Maximum Month		Maximum Week		Maximum Day	
	Avg (%)	CV (%)	Avg (%)	CV (%)	Avg (%)	CV (%)
Voltage	110	8.2	119	17.5	135	25
Unbalance	119	8.3	131	21.7	156	29.8
Harmonics	115	11.6	124	14.1	137	14.9

5. Conclusions

This study focussed on calculation of an index to characterise disturbance levels across a whole year. To this end four methods of calculating such an index, have been examined in this paper. These methods are:

- (i) The 95th percentile over the whole year
- (ii) The maximum of the 12 95th percentile values calculated for each month
- (iii) The maximum of the 52 95th percentile values calculated for each week
- (iv) The maximum of the 365 95th percentile values calculated for each day

Analysis of the variation between indices calculated using the reporting methods outlined above indicates that significant variation will be seen depending on the reporting method used. As expected the maximum variation is seen between indices calculated across the whole year which excludes a significant amount of data and indices calculated using the maximum daily method which excludes very little of the data. In all cases the average variation between these two methods is between 35% (for voltage) and 55% (for unbalance). That is, a yearly index calculated using the maximum day approach will, on average, be more than 35% larger than the corresponding index calculated using the whole year approach. This clearly indicates that if indices are to be compared across sites the method used in calculating the indices will have a significant impact on the outcomes of such comparisons. The same comments apply if indices are compared across time or to limits.

IEC61000-4-30 recommends one week as the measurement and reporting interval to be used for the disturbances examined in this study. Analysis of the variation between indices calculated using weekly values and indices calculated across the whole year

