

On-line Partial Discharge Analysis of Transmission and Distribution Assets

Paul L Lewin

The Tony Davies High Voltage Laboratory
University of Southampton
Southampton, Hampshire, UK
pll@ecs.soton.ac.uk

Abstract—It is becoming increasingly clear that methodologies for PD classification based on standard laboratory experimental data are not readily applicable when assessing on-line PD data measured in the field. It is not just that field data is corrupted by noise and disturbance, but also the significant differences between typical laboratory experiments to generate PD data and the generation of PDs in high voltage plant due to degradation of the insulation system. In this paper, the use of nonlinear time-series analysis on field data is shown to yield useful information, methods involving dimension reduction techniques are shown to allow identification of different sources and finally a method for designing standard finite impulse response filters that approximate the nonlinear analytical approach and are easy to implement in condition monitoring systems are discussed.

Keywords—*partial discharge, condition monitoring, clustering, time series analysis, stochastic analysis.*

I. INTRODUCTION

This paper outlines methodologies that are being developed specifically for the analysis of partial discharge (PD) measurements, obtained in the field for a range of transmission and distribution assets. Very specific solutions have been widely reported for particular problems (e.g. the use of UHF sensors incorporated into a transformer tank to locate PD activity [1]), however this paper is concerned with the development of approaches that can be used more generally, i.e. tools to analyse captured PD data from a range of sensor types mounted on different plant, over different frequency ranges. The need to know the true state of assets within power networks is becoming increasingly important as the change in generation mix, the impact of climate change, and the development of ‘smarter’ grids are having a significant effect on how power networks are operated. In turn changes in operation may mean significant changes in power flow and loading, such that circuits that have been operating reliably for some time under a certain set of conditions, start to have problems when these conditions change due to operational requirements. Early warning of a likely problem becomes imperative in order to guarantee operational security of supply.

It is generally assumed that one obvious indicator of asset health is the presence or absence of partial discharge activity. In the case of plant with polymeric insulation this is definitely true, but in the case of other insulation systems, research has

shown that on-line PD measurement data can give conflicting indications, not only of asset health but also imminent failure. One general approach for all forms of condition monitoring is to develop knowledge-bases for specific plant types that when used with field measurements can assist in interpretation of the measurement data [2]. However, this approach has to be used with care, as many suppliers of monitoring equipment see this information as proprietary and generally there is a commercial advantage in providing both the ability to detect PD activity as well as interpret the results on behalf of the asset owner/operator. The accepted graphical methods of presenting PD measurement data are very useful for the experienced PD engineer, but not necessarily for an asset manager. There have been many proposed measurands of PD data including time-based analyses, use of statistical moments, frequency-based approaches, combinations of time and frequency etc. These approaches try to characterise the features of the PD data, either by considering individual pulses or analysis of groups of PD pulses defined by their phase of occurrence on the ac cycle or by analysis of PD to PD behaviour (i.e. pulse sequence analysis, PSA). In some cases, features extracted from laboratory measurements have been used to train artificial neural networks or neuro fuzzy systems or other forms of ‘machine-learning’ algorithms [3]. However, it cannot be assumed that because these approaches can distinguish between different artificial PD sources under laboratory conditions, then they are applicable to data measured in the field. To date there is no supporting evidence that this is indeed the case.

A. Limitations of laboratory experiments in terms of analysis of field data

Obviously, there is a need to develop different analytical approaches for interpretation of on-line measurement of PD data. These methodologies need to consider the nature of field data and the differences when compared to that obtained from laboratory experiments. For example in the laboratory the tendency is to work at high fields and lower voltages, which ensures that it is possible to generate PDs from a test cell by exceeding the inception field whilst ensuring that the applied voltages are not excessive to the point where the test cell dimensions (to avoid flashover or breakdown) become too large. One facet of this approach is that generally signal to noise ratios are very high (almost noise free), sensitive measurements are possible and a lot of data can be generated

over a short time interval. However, the singular disadvantage may well be that the design of these experiments is such that the generated data is strongly deterministic, i.e the likelihood of a PD occurring once the inception field has been exceeded is very close to one, seed electrons are readily available and an avalanche very likely to occur. This can be compounded through the use of strongly divergent fields (e.g. point-plane). By designing experiments to create ‘good’ data there is a risk that the properties of the mechanisms leading to PD are different to that experienced in the field, where generally plant is operating at high voltages and low fields which in turn are designed to be uniform. Unlike within the laboratory, PD data measured in the field tends to appear to be far more stochastic, unless very close to the point of catastrophic failure. More recently these differences have been recognised and experiments designed in the laboratory to closely match operational conditions [4]. When it comes to analysing on-line PD data captured in the field there is a lot of information that is not known nor can be assumed, for example, the number of sources of PD, whether captured pulses are PD or disturbance and whether the captured pulses are internal or external to the plant. This may explain why initial analysis of field data using PSA has proven problematic.

In this paper, analysis of field data through the use of dimension reduction techniques to identify PD signals from the same source is first detailed, before consideration of the application of nonlinear time series methods is discussed. Finally, a method is presented for developing ‘filters’ that approximately represent a non-linear analytical approach. The approximate filters lend themselves to implementation within standard PD monitoring equipment for on-line analysis of PD measurement data from transmission and distribution assets.

II. FIELD DATA

Figure 1 shows some typical data obtained from an 11 kV three phase belted cable, located in Central London. PD data are shown as a persistence plot in Fig.1a and as a ϕ -q-n plot in Fig.1b. Essentially these plots are the same, both plot the magnitude of a PD event (q) against occurrence in terms of the phase of the applied voltage (ϕ). In the case of the ϕ -q-n plot, the third dimension is the number of discharges of certain magnitude occurring at a point on phase. Previous work [5,6] has shown that it is possible to subdivide field data into subsets of data where each subset represents data from a different source. It is not readily apparent which are PD sources but this facilitates further more detailed analysis. For the data summarized in Fig.1, clustering reveals 4 separate sources (Fig.2.), each of which can then be represented in terms of a persistence or a ϕ -q-n plot, that may be more informative than trying to assess the whole data set. Various clustering techniques have been proposed from time-frequency mapping [5] to clusters determined by distribution of energy in time and frequency [6] and have been shown to generally perform well. Here, individual pulses are analysed using the wavelet transform, the energy distribution of wavelet detail coefficients at each level plus the energy of the approximation coefficients determined, and this vector is then represented as a single point in a 3 dimensional space. Pulses having similar structure in time and frequency will cluster together. These sub-sets of pulses can then be analysed separately.

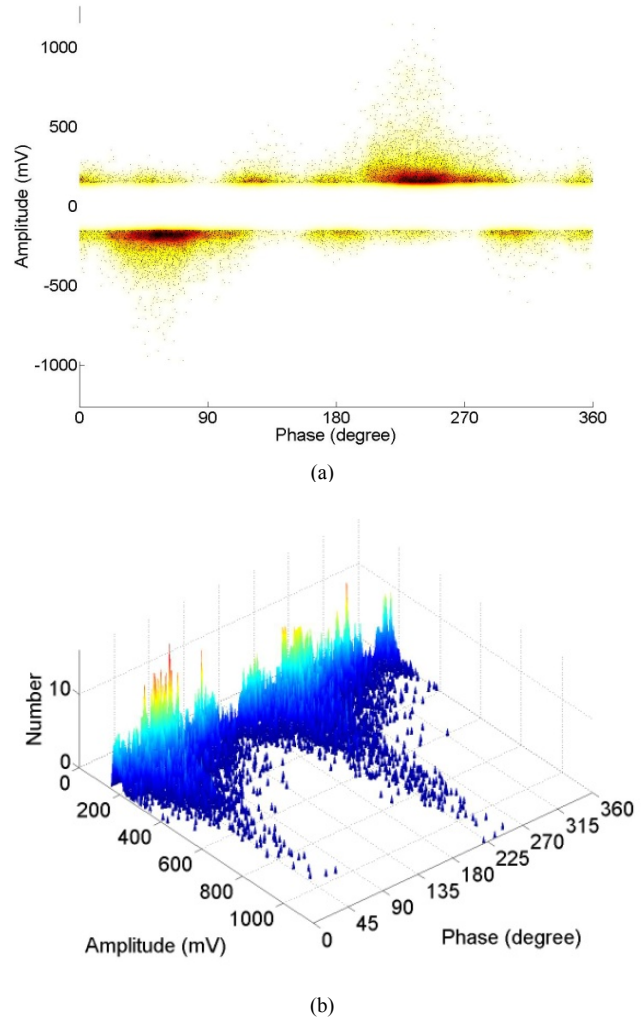


Fig. 1. (a) 2D and (b) 3D ϕ -q-n patterns for 157 cycles – data from a 3 phase belted 11 kV cable, Central London.

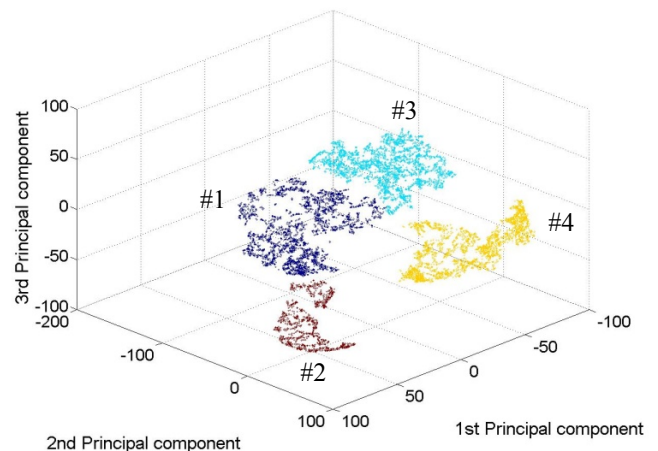


Fig. 2. Clusters of PD data from a 3 phase belted 11 kV cable, Central London.

A. Establishing the stochastic nature of on-line PD measurement data

It is clear that there are significant differences between laboratory data and on-line field data and it has been hypothesised that under field conditions, the PD process is one that exhibits deterministic chaos [7]. One way to investigate this is to consider the pulse to pulse behaviour, through the use of PSA plots. If it is assumed that the phase and real PD magnitude of a detected PD pulse are influenced by previous PD pulses, then the causal nature of the PD processes can be assessed through the stochastic generation of surrogate data sets. These surrogates are simulated time series of PD pulses that when plotted as a persistence plot will be identical to that of the original measurement data, but without the underlying relationships that are believed to exist in the real measurement data. Figure 3 shows PSA plots for a set of PD measurements made on a Central London 3 phase belted 11 kV belted cable circuit [8].

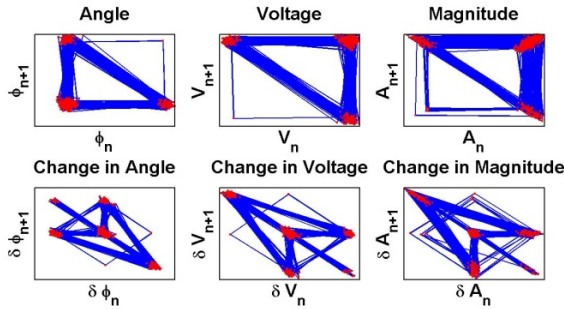


Fig. 3. PSA (vector) plot of on-line PD measurements from a Central London 11 kV cable circuit.

In Fig. 3 the red dots represent phase, applied voltage or measured PD magnitude relationships and the blue lines represent the transition from point to point in the plot. This can be compared with PSA plots of surrogate data, an example of which is shown in Fig. 4.

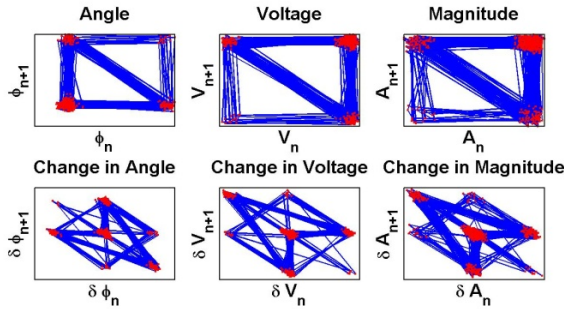


Fig. 4. PSA (vector) plot of a surrogate data set having the same persistence plot as that of the Central London 11 kV cable circuit data shown in Fig. 3.

On comparing the two sets of plots, clear differences can be observed. The lack of a cluster in the upper right hand corner of the angle plot of Fig. 3 indicates that two consecutive PD pulses occurring in the 3rd quadrant of the power cycle is rare and this is confirmed when looking at the magnitude plot of Fig 3., where the cluster in the lower left corner is

comparatively small, indicating the rarity of two consecutive PD events in the third quadrant with a negative magnitude. In the surrogate data (an example shown in Fig. 4), the same does not apply and there are consecutive events in the third quadrant of the power cycle.

In summary, through the use of clustering techniques, identified subsets of on-line PD measurement data have been shown to exhibit short-term correlated activity that is not evident in surrogate data sets having the same persistence plot. This is a useful observation, and although the approach does not lend itself to the development of tools and automatic processes, it is more important that some measure of the short-term correlated behaviour be used.

One potential tool is the Generalised Hurst Exponent (GHE) [9]. The GHE, $H(q)$ gives an indication as to whether a process (in this case a sequence of pulses, $p(t)$) has an underlying Gaussian distribution. Assuming that the pulse series can be defined by the scaling properties, S_q of its structure, such that

$$S_q = \langle |p(t + \tau) - p(t)|^q \rangle_t, \quad (1)$$

where τ is a time lag that is small compared to the duration of the captured pulse sequence. Then it is possible to approximate the scaling property in terms of the GHE, i.e.

$$S_q \sim \tau^{qH(q)}, \quad q = 1, 2 \quad (2)$$

Estimating $H(q)$ for $p(t)$ reveals information about the underlying behaviour of the pulse sequence. If H_1 (i.e. $q=1$) is larger than 0.5 then increases in values are likely to be followed by further increases (behaviour is persistent) and if H_1 is less than 0.5 then there is anti-persistent behaviour and any increase is likely to be followed by a decrease. In both cases the further away the exponent is from 0.5 the greater the persistence of behaviour within the pulse sequence.

Analysis has been undertaken for a range of on-line measurement data for circuits in London, UK and Cyprus as well as for measurement data from artificial PD sources. To allow comparison and avoid issues with scaling, the PSA vectors for magnitude and angle were determined for each data set and then respective Hurst components (in terms of mean and standard variation) determined. Some of the obtained results are shown in Table I.

TABLE I. SELECTED HURST EXPONENT VALUES

Source	Data Type	Mean	Standard Deviation
London	Magnitude	0.62	0.007
	Angle	0.78	0.004
Cyprus	Magnitude	0.44	0.010
	Angle	0.44	0.009
Void Source	Magnitude	0.70	0.005
	Angle	0.69	0.007
Surface Discharge	Magnitude	0.83	0.006
	Angle	0.82	0.006

The mean values closest to 0.5 (nearest behaviour equivalent to Brownian motion) are for the Cyprus data, which is also anti-persistent. All other values are greater than 0.5, so it can be concluded that either the pre-processing approach with

the Cyprus data did not isolate PD pulse-like information or that the original data does not contain any PD measurement information. Comparison of the Hurst Exponents of the magnitude data, reveals that the artificial PD sources generate PD magnitudes that are more persistent than the field data, whereas the angle Hurst Exponent for a void discharge source is the least persistent. Other non-linear time series tools can be used to investigate the structure of field PD data and this is an on-going area of research [8].

B. Moving forward, developing approaches that can be implemented in Condition Monitoring plant.

Previous work at the Tony Davies High Voltage Laboratory, University of Southampton has concentrated on the development of methodologies to locate PD sources in large transformers [10]. The basic approach is to identify internal PD activity, locate the source and then consider the severity of the problem. Initial work was based on using the theory initially proposed by L. V. Bewley [10], which was modified to give a description of the response seen at the end of the windings for an impulse event occurring at some point along the winding itself [11]. The derived response is the solution of a high order linear ordinary differential equation whose constant coefficients were modelled based on the classic lumped parameter model of a winding. As an approximate method, this approach does work, but was found to be limited.

Fundamentally, the problem is that the assumptions made to create a linear model of a high voltage transformer winding that is useful for PD analysis do not hold and this is best illustrated by consideration of winding frequency response measurements. The issue is that if a winding arrangement is linear, the frequency response must be positive real, in other words the range of phase of the response over the frequency range measured must be within $\pm 90^\circ$. This is not a problem near power frequencies (say 10 to 100s of Hertz) but as a PD event is broadband, which at source may have UHF components, consequently the winding response over a range of up to several hundred MHz is required. At higher frequencies the phase compared with the winding phase at power frequency is substantially greater than $\pm 90^\circ$

More recently a non-linear approach, that relies on the data alone and does not require a priori knowledge of the transformer winding construction has been investigated [12]. The selected approach has the same basis as the use of clustering for source location discussed earlier in Section II. Simultaneous measurement of PD at either end of the winding, for example by using broadband radio frequency current transformers (RFCTs) to measure earth currents at the bushing tap point and the neutral to earth connection will give two related sets of PD pulse data. The relationship between these two data sets can be investigated using the same clustering approach as used on other field data such as the London cable circuit (Fig. 1 and Fig. 2). However, in this case the relationship between the clusters is more important.

Figure 5 shows cluster analysis for a PD source near the centre of the winding and a source near to the neutral to earth point. The distance between the clusters is related to the differences in transmission path from the source to the

measurement points. Consequently this approach could be used for source location in the field.

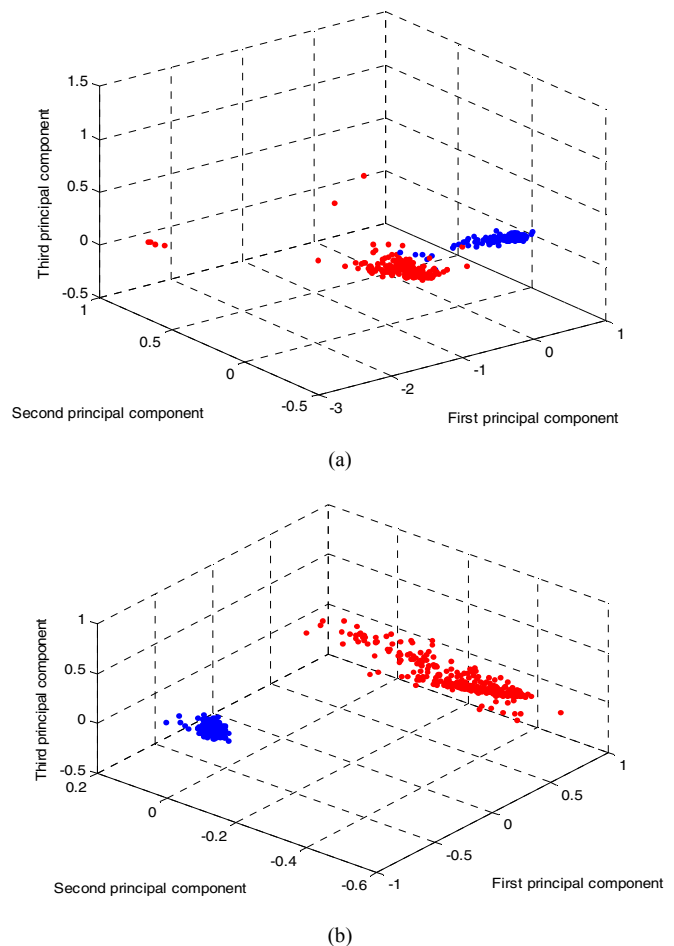


Fig. 5 Cluster analysis for (a) a PD source near the centre of a high voltage interleaved winding and (b) a PD source near the Neutral to Earth point. Measurements were made with RFCTs at either end of the winding the red dots corresponding to PDs measured at the neutral to earth point and the blue dots PDs measured at the Bushing tap point.

The method used is to identify each detected PD pulse, then use wavelet decomposition to determine the distribution of pulse energy, then to use t-Distributed Stochastic Neighbor Embedding (t-SNE) for dimension reduction in order to represent each energy distribution as a single point in three dimensional space. Developing this into a real-time algorithm for use in CM equipment is possible but likely to require significant hardware at the measurement point. A better solution is to consider the possibility to design a ‘filter’ that represents each principle component axis, so that a measured PD signal can be analyzed by passing it through three filters in parallel and then computing the share of energy of each filter output. Initial results from this work are very promising and are presented in another paper at this conference [13]. The supporting theory is based on the use of principle component analysis rather than t-SNE and the resulting finite impulse response filters need to be of relatively high order (c30-40) in order to produce reasonably accurate results, but in terms of

development of a general methodology for analysis of on-line PD data this has significant potential.

III. CONCLUSIONS

It is increasingly clear that in order to develop general tools and techniques for analysis of on-line PD measurement data of transmission and distribution assets, a different approach needs to be adopted. In terms of laboratory-based experimentation, there is a need to move away from the 'traditional' approach of studying behaviors of specific artificial PD sources to developing experiments that closely relate to field conditions, e.g. [4], and contain known sources of PD as a result of deliberate faulty construction or third party damage or overstressing during operation.

There is an extensive range of approaches for identifying PD signals that have similar characteristics; examples include time-frequency mapping and wavelet energy distribution with dimension reduction. In essence these approaches are similar in that inherent features of captured PD pulses are used to create sub-sets of PD pulses having similar features. This in turn facilitates the use of non-linear time series analysis on the identified sub-sets of PD data. Initial work has generated surrogate data sets for comparison with real measurement data and results reveal there is short term correlated activity in the field measurement data. Investigations into this relationship have used the Generalized Hurst Exponent to identify the persistence of this behavior and again results reveal differences between laboratory data from artificial sources (more persistent) and field data (less persistent but not random).

Through the development of an understanding of the non-linear nature of field PD data, it has been possible to develop data driven approaches that could be used to locate sources of PD in large transformer windings. The disadvantage is that any algorithm is likely to be computationally and memory intensive. To overcome this, linear approximations (in the form of specifically designed filters) can be used to represent the non-linear analytical methodology.

IV. ACKNOWLEDGMENT

The author would like to thank the condition monitoring research team of the Tony Davies High Voltage Laboratory, University of Southampton, UK and colleagues around the

world who have been willing to collaborate; sharing both ideas and data.

REFERENCES

- [1] M. D. Judd, L. Yang, I. B. B. Hunter "Partial discharge monitoring of power transformers using UHF sensors. Part 1: sensors and signal interpretation," IEEE Electrical Insulation Magazine, 21(2), March-April 2005, pp 5-14.
- [2] J. Lapworth "A novel approach (scoring system) for integrating dissolved gas analysis results into a life management system" 2002 IEEE International Symposium on Electrical Insulation, 7-10 April 2002, Boston, USA, pp 137-144.
- [3] H. Ma, J. C. Chan, T. K. Saha and C. Ekanayake "Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources," IEEE Trans DEI, 20(2), April 2013, pp 468-478.
- [4] J. A. Hunter, P. L. Lewin, L. Hao, C. Walton and M. Michel "Autonomous classification of PD sources within three-phase 11 kV PILC cables," IEEE Trans DEI, 20(6), December 2013, pp 2117-2124.
- [5] G. C. Montanari, A. Cavallini, F. Puletti "A new approach to partial discharge testing of HV cable systems," IEEE Electrical Insulation Magazine, 22(1), January-February 2006, pp 14-23.
- [6] L. Hao, P. L. Lewin, J. A. Hunter "Discrimination of multiple PD sources using wavelet decomposition and principle component analysis," IEEE Trans DEI, 18(5), October 2011, pp 1702-1711.
- [7] S. Azizi-Ghannad, J. K. Nelson and M. Embrechts "An investigation of chaotic behaviour of PD activity in power generator systems," 1995 IEEE Conference on Electrical Insulation and Dielectric Phenomena, 22-25 October 1995, Virginia Beach, VA, USA, pp 353-356.
- [8] L. Petrov, P. L. Lewin and C. Tadeusz "On the Applicability of Nonlinear Time series Methods for Partial Discharge," IEEE Trans DEI, 21(1), February 2014, pp 284-293.
- [9] H. E. Hurst "Long-term storage capacity of reservoirs" Transactions of the American Society of Civil Engineers, 1951, 116, pp 770-808.
- [10] L. V. Bewley, "Travelling Waves on Transmission Systems," 2nd ed, John Wiley and Sons, New York, 1951.
- [11] P. L. Lewin, I. O. Golosnoy and R. Mohammad, "Locating Partial Discharge Sources in High Voltage Transformer Windings," 2011 IEEE Electrical Insulation Conference, 5-8 June 2011, Annapolis, MD, USA, pp 196-200.
- [12] M. S. Abd Rahman, P. Rapisarda and P.L.Lewin, "The study of the effect in variation of the AC applied voltage to the partial discharge localization inside transformer windings," 2013 IEEE Conference on Electrical Insulation and Dielectric Phenomena, 20-23 October 2013, Shenzhen, China, pp 1290-1293.
- [13] M. S. Abd Rahman, P. Rapisarda and P.L.Lewin, "The use of three dimensional filters for on-line partial discharge localisation in large transformers." To be published 2014 IEEE Electrical Insulation Conference, 8-12 June 2014, Philadelphia, PA, USA.