

## DETECTION OF EMERGING FAULTS IN POWER TRANSFORMERS USING SELF-ORGANIZING MAPS.

Francisco Martin<sup>1</sup>, Olga Duran<sup>2</sup>, Jose A. Aguado<sup>1</sup>

<sup>1</sup>Dep. of Electrical Engineering. University of Malaga, Spain

<sup>2</sup>SEC Faculty, Kingston University, UK

### Abstract

Power transformers are a crucial part of the power system, one of the largest infrastructures in industrialised countries. In particular, wind turbine transformers are subjected to frequent thermal cycling as a function of varying turbine loads. Thus transformers are prone to developing faults and defects that can involve high repair costs for instance due to the repeated thermal stress on the winding. Faults develop mainly when the insulation produces small leakage currents between turns, which if not detected early, might become short circuits that can result in interruptions in electricity supply, and difficult and costly repairs. An optimum overhaul of damaged transformers is not accomplished often because of lack of appropriate inspection tools. Detailed assessment and preventive maintenance work, which will allow the detection and repair of failures at early stages, is believed to be the only suitable way to cope with power transformer degradation at low cost. This paper presents a methodology based on the analysis of current signals converted by the S transform for the detection of incipient faults in transformers. The procedure is based on calculating the energy of the zones of the time-frequency spectrum. Its main advantage is its possible real time implementation that can be applied while the transformer is in use. Experimental results with PSCAD are presented.

**Keywords:** Power transformer, S transform, self organizing maps, incipient fault.

### 1. Introduction

Power systems are mainly used to convert natural resources into electrical energy. There are two types of electric energy resources; conventional such as gas, diesel, coal and nuclear, and non conventional that include renewable such as solar, wind, fuel cells or biogas. Once transformed, the electric power energy is transported to the load centres through power network. Power transformers are a fundamental part of power systems. The faults in transformers can completely destroy or cause serious damage and costly repairs, along with long interruptions. The faults in transformers are caused by small leaks (incipient faults) that damage the insulation. The detection of incipient faults on time is vital to prevent that failures occur and destroy the transformer.

In particular, wind turbine transformers are subjected to thermal phases as that depend on the turbine loads [1]. This causes repeated thermal stress on the windings that can create hot spots and partial discharges which can damage insulation. This produces small leakage currents between turns, which if not detected early, might become short circuits [2].

The main objective here is to detect those small leaks through analysis of current and voltage signals while the transformer operates. Other objectives are to analyse the behaviour of the transformer to all types of disturbances and classification of them, designing a method to distinguish magnetizing currents of faults and other disturbances.

At present, the voltage and current signals are not pure sinusoidal. In these quasi-stationary states clearly non periodic components (transient electromagnetic) and other disturbances coexist, as well as random noise. Therefore, the classical method of the Fourier transform of power in the analysis of periodic signals, is not suited for this type of signals. Time-frequency transforms are more suitable for non-stationary signals. Among them, wavelet transforms have shown their efficiency in the non-stationary signal analysis [3,4]. More recently, the S transform [5] has been used by some researchers in Power Quality, giving promising results [6]. These transformations can be combined with artificial intelligence techniques like neural networks and fuzzy logic [4,7] for the construction of an automated detection of incipient faults before they develop into a short circuit and produce a serious breakdown and costly disruptions.

However, there are some very challenging problems that need to be solved. On the one hand, small defects in the insulation or incipient faults lead to very small magnitude currents. These currents cause short reach, and if not detected early, can cause serious damage. In the first case, the currents can not be detected by the residual current or overcurrent. Differential protection does not detect these currents, being much lower than the current percentage spread. Nor overcurrent protection, which is set for large fault currents. The same goes for thermal protection, as the temperature increase produced by the leakage current is negligible.

On the other hand, ones needs to identify magnetizing currents, which are often mistaken with ground faults, causing nuisance tripping of protection devices, with consequent unnecessary interruptions of supply. The method that is

usually employed to distinguish the magnetizing currents of internal faults is the protection of differential percentage with retention by harmonics, but this method has not completely solved the problem, and sometimes does not prevent nuisance tripping to occur [1].

This paper presents a method for the detection of incipient faults in transformers, and distinguish them from different types of disturbances (faults external, internal, inrush currents, etc.). The method is based on the analysis of current and voltage signals using time-frequency transform (S transform) combined with artificial self-organizing maps (SOM) for the classification stage.

This paper is organised as follows. In Section 2 background is described where the development of the basic theory of the S transform and Self Organising Maps are described, respectively. Results using simulated signals are presented in Section 3. Finally, conclusions are drawn in Section 4.

## 2. Background

### 2.1. The S transform

The S transform[5] is being used in recent years by some researchers in the time-frequency analysis of signals in power systems, especially in Power Quality, with very good results. The time-frequency transform decomposed into its spectrum signals in both frequency and time dimensions, with good resolution in time and frequency. Its use permits the utilization of information contained in transients with the advantages of being able to analyze the signals in real time, using fast algorithms [8], making them very effective in prevention efforts and fault protection.

The ST was introduced as an alternative to the STFT for localization of time-frequency spectra. The ST gives time and frequency information, as does the STFT, but it uses a variable window length that provides information at different resolutions, such as in the case of the WT.

The ST can be derived from the WT by modifying the phase of the window function or mother wavelet.

Given a time-dependent signal,  $x(t)$ , the ST can be derived as the product of the signal and a phase correction function  $e^{(-i2\pi f\tau)}$ .

The S-transform of  $x(t)$  is defined as:

$$S(\tau, f) = \int_{-\infty}^{\infty} x(t) g(t) e^{-i2\pi f t} dt \quad (1)$$

where  $g(t)$  is the Gaussian modulation function, defined as:

$$g(t) = \frac{|f|}{k\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2k^2}} \quad (2)$$

The expression becomes:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{k\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2k^2}} e^{-i2\pi f t} dt \quad (3)$$

The discrete version of (3) is calculated, taking advantage of the efficiency of the fast Fourier transform.

The discrete Fourier transform of the time series  $x(t)$  is obtained as:

$$H\left[\frac{n}{NT}\right] = \frac{1}{N} \sum_{k=0}^{N-1} h(kT) e^{-i\left(\frac{2\pi nk}{N}\right)} \quad (4)$$

where  $n, k=0,1,\dots, N-1$  ( $N$  is the number of samples).

The discrete S-transform is obtained by allowing  $f \rightarrow n/NT$  and  $\tau \rightarrow jT$ :

$$S\left(jT, \frac{n}{NT}\right) = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] G(m, n) e^{\frac{i2\pi m j}{N}} \quad j, n=0, 1, \dots, N-1 \quad (5)$$

where

$$G(m, n) = e^{-\frac{2\pi^2 m^2 \alpha^2}{n^2}} \quad (6)$$

and  $j, m, n = 0, 1, \dots, N-1$ .

The discrete inverse of the S-transform can be obtained as:

$$x[kT] = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{j=0}^{N-1} S \left[ \frac{n}{NT}, jT \right] \right\} e^{i2\pi nk} \quad (7)$$

The output from ST analysis is a complex matrix whose rows and columns represent frequency and time, respectively. Each column represents the local spectrum in time. Frequency-time contours with the same amplitude spectrum are also obtained. This information is used to detect and characterize power disturbance events.

## 2.2. Self Organising Maps

Self-organising maps (SOM) [9] is a method to represent large amounts of high dimensionality data by a much lower complexity topological map. The SOM may be used for data visualisation, clustering, classification and many other applications.

The SOM consists of a regular low-dimensionality grid (usually a 2D grid), each node of which is associated with a weight vector  $m_i$  (also called prototype or codebook vector) with the same dimension as that of the data. The map units are connected to adjacent ones by a neighbourhood relation. Each map unit may then be considered as having two sets of coordinates: the weight vectors, in the input space and the position on the map in the output space [].

The number of map units varies but typically goes from a few dozen to several thousand. The selection of the number of map units depends on the application, taking into account that there is a trade-off between accuracy and generalisation capabilities of the SOM. The larger the number of map units, the more detail is produced. However, the input vectors being spread out on a large number of units, may lead to poor generalisation.

The SOM is trained iteratively by a two-layered neural network. The first layer is the input layer, with as many nodes as the dimension of the data. The second layer is the competitive layer and corresponds to the map units in the output space. The two layers are fully connected with the weight vectors.

The SOM training algorithm resembles the k-means: at each training stage, data is presented to the network and the map unit whose weight is closer to the input vectors (best-matching unit, (BMU)) is updated towards the inputs. The distinction is that not only the BMU weight vector is updated but also its topological neighbours on the map. The units in the map become then ordered in regions with similar weights.

The learning using a sequential training algorithm may be summarised as follows:

- The weight vectors are randomly initialised. The initial learning rate,  $\alpha_0$ , and the size of the neighbourhood,  $N_0$ , are initialised.
- The prototype vectors are randomly initialised
- A sample input vector,  $x_j = [x_1, x_2, x_3, \dots, x_n]$ , is randomly chosen from the input data set.
- The distances between the input vector and the weight vectors are computed.
- Weights are updated for all units that are in the neighbourhood of the BMU. The updating equation is

$$\Omega_{ij}(t+1) = \Omega_{ij}(t) + \alpha(t) \lambda(m_i - \text{BMU}) (x_j - \Omega_{ij})$$

where  $\alpha(t)$  is the learning rate which decreases with the iteration number  $t$  and  $\lambda(m_i - \text{BMU})$  is the window function, which is used to ensure that neighbouring points in the target space have similar weights. At the latest stage, when the neighbourhood gets restricted to a single unit, the algorithm becomes similar to the k-means. The number of iterations is usually 30 to 50 times the number of map units [9].

With the SOM, high dimensionality data are visualised on a low-dimensionality display, where the relationships between the high dimensionality data are converted into simple geometric relationships on the image points. The visualisation may only be used to obtain qualitative information. Further clustering has to be done if quantitative results are to be obtained. By clustering the SOM, instead of directly the data, the computing time is reduced. Moreover, the clustering quality is improved since outliers are reduced in the first stage. In order to find the clusters, the number of map units in the SOM should be much larger than the expected number of classes. This is because the neighbourhood function in the SOM training brings neighbouring units closer to each other rather than cluster them. Different approaches have been proposed for clustering, such as clustering based on distance matrices (U-matrix) [10]. The U-matrix is a tool to visualise the distances between each map unit and its neighbours. These distances are inversely proportional to the density of the map prototypes. High values of the distance matrix indicate where the cluster borders lie, while local minima of that matrix indicate cluster centres. Local minima of the distance matrix

may be found and then used as cluster centres of the SOM. The rest of the map units are assigned to the cluster whose centre is closest.

### 3. Experimental results

#### 3.1. Simulated data

The PSCAD/EMTDC package has been used for simulating the following operating conditions in power transformers:

- Normal condition
- Inrush currents
- Internal incipient faults
- External faults

Normal condition has been simulated at different load levels. Inrush currents are produced when the transformer switches off and on with remanent flux. Different states of remanent flux and different reconnection times and load states have been considered.

Internal incipient faults have been simulated as a short-circuit of a small number of turns in the primary or secondary winding. 1%, 2%, 3%, 4% and 5% of the turns have been considered. Fig. 1 shows the PSCAD schematics of a transformer with an incipient internal fault. A switch has been included to allow randomly switch on and off in order to produce incipient faulty currents.

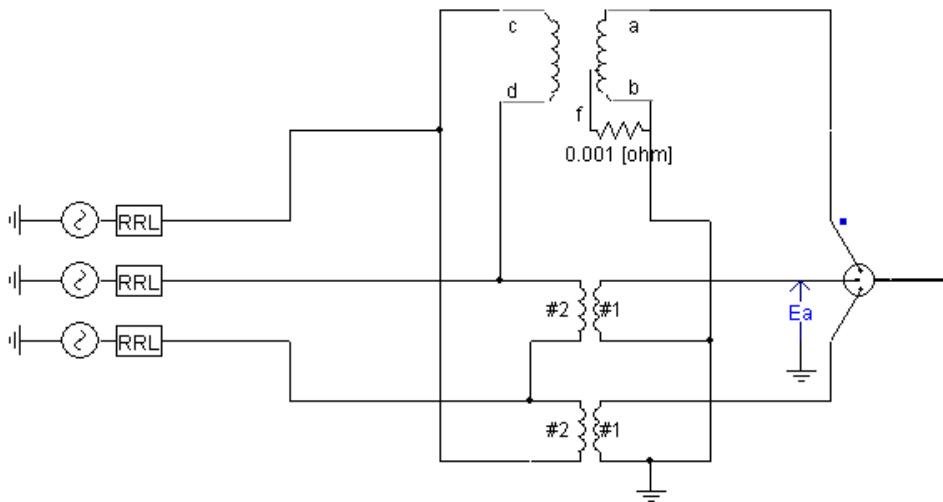


Fig. 1. PSCAD schematic of internal incipient faults

External faults have been simulated in the power lines next to the transformer. Phase-ground, two phase, two phase-ground and tri-phase faults, at different times of occurrence, at different load conditions and in different points of the line been considered. Fig. 2 shows signals corresponding to the four operating conditions.

#### 3.2. Classification

The normal operating frequency is 50 Hz. The sampling frequency is 4 KHz. Thus the Nyquist frequency is 2 KHz. 200 samples or two and half cycles of each signal have been analysed in each simulation. Fig. 3 shows the corresponding S-transforms for the signals depicted in Fig. 2.

In order to reduce the number and complexity of the input features for the SOM, three steps are performed as follows:

- 1) The signals are filtered so the information corresponding to the 50Hz is not included.
- 2) The energy of the S transform is used instead of the raw signals.
- 3) Moreover, the signals are normalised so the total energy of the signal equals one.

The number of samples used for classification where 1200 (300 samples for each operating condition). Fig. 4 shows examples of the feature vectors that are fed to the SOM.

The results show good discrimination between the signals corresponding to normal operation, internal and external faults. Signals corresponding to inrush current and normal operation are not discriminated. This is acceptable since both signals correspond to non faulty scenarios and the main objective is to distinguish between faulty and non faulty signals.

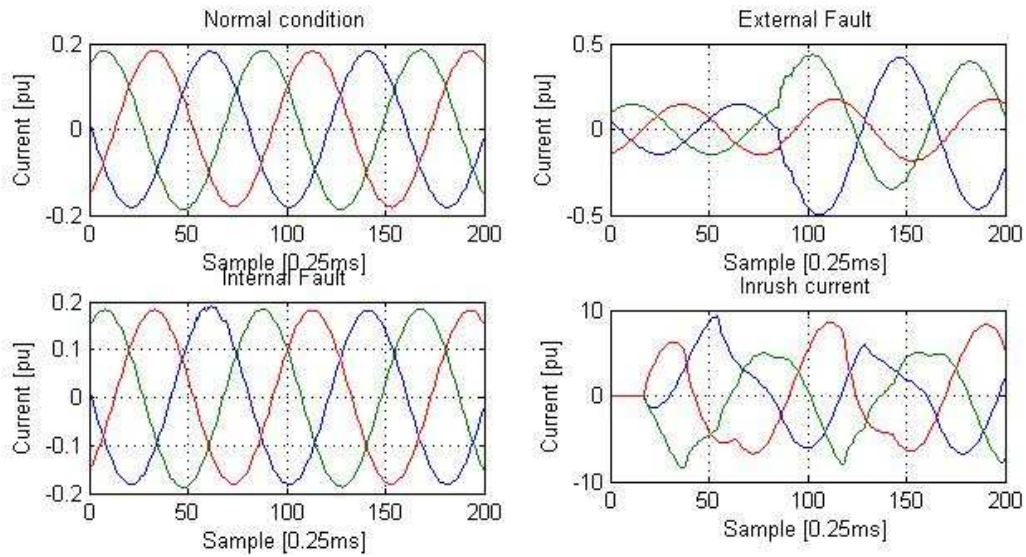


Fig. 2. Typical simulated signals for the different operating conditions.

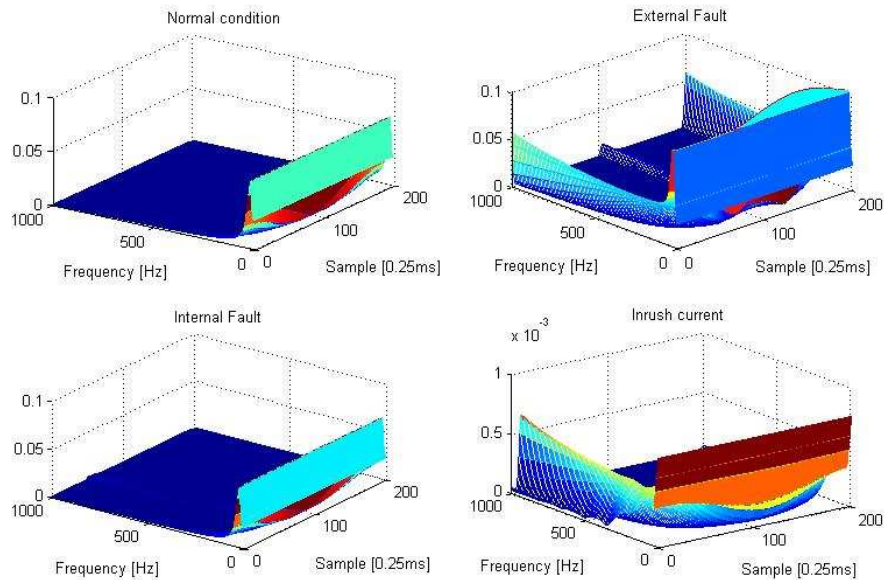


Fig. 3. Corresponding S transforms for the signals depicted in Figure 2.

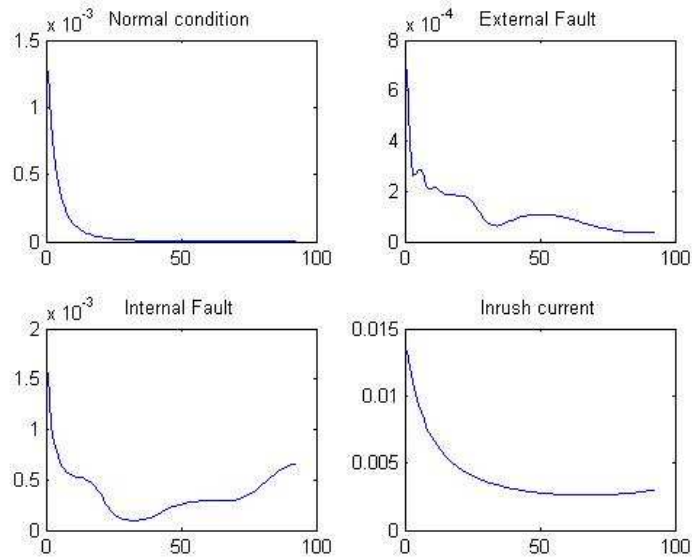


Fig. 4. Examples of feature vectors fed to the SOM for the different types of operating conditions.

## 5. Conclusions

This paper presents a novel approach that combines S transform and SOM methods to classify signals corresponding to faulty and non faulty transformers, respectively. The S transform is used to extract the time-frequency response of the signal. Further, the normalised energy of the S transform is compute and fed to a SOM to enhance discrimination between transformer internal fault, inrush currents, external faults and normal operation. Signals corresponding to non faulty cases (i.e. normal operation and inrush currents) are classified as a single class. This is acceptable since both signals correspond to non faulty scenarios and the main objective is to distinguish between faulty and non faulty signals. Moreover, the proposed approach allows discrimination between faulty and not faulty transformer with high success rate.

SOM has the advantage of being faster than the classical ANNs and easier to design. Further work will be dedicated to improve the success rate of fault classification (internal and external) and to classify different types of faults.

## References

- [1] K. Remington "Why do wind turbine transformers fail so often", Available online: <http://www.windpowerengineering.com/featured/business-news-projects/why-do-wind-turbine-transformers-fail-so-often/>
- [2] P. M. Anderson, Power System Protection, IEEE Press, 1999.
- [3] Heydt, A.W. Galli "Transient Power Quality Problems Analyzed using Wavelets" IEEE Trans. on Power Delivery, Vol.12, N°2, April 1996.
- [4] F. Martín, J.A. Aguado. Wavelet-based ANN approach for transmisión line protection. IEEE Trans. on Power Delivery, Vol. 18, No. 4, October 2003.
- [5] Stockwell, RG, L Mansinha, and RP Lowe (1996). Localization of the complex spectrum: the S transform, IEEE Transactions on Signal Processing 44 (4), p 998-1001.
- [6] S Santoso, EJ Powers, WM Grady; Power quality assessment via wavelet transform analysis; Power Delivery, IEEE, May 1996.
- [7] Mao, P.L.; Aggarwal, R.K. A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network. IEEE Trans on Power Delivery.
- [8] Kelly Sansom, "Fast S Transform", University of Calgary, <http://www.ucalgary.ca/news/utoday/may31-2011/computing>
- [9] T. Kohonen. Self-Organizing Maps. Springer Series in Information Sciences, Berlin, Heidelberg, 1995. (Third Extended Edition, 2001)
- [10] J Vesanto and M Sulkava. Distance matrix based clustering of the self-organizing map. In Jos R Dorronsoro, editor, International Conference on Artificial Neural Networks- ICANN- Lecture Notes in Computer Science, volume 2415, pages 951-956, madrid, Spain, Aug 2002, Springer.