HAL
open science

Data-driven approach for dip voltage fault detection and identification based on grid current vector trajectory analysis

Amel Adouni, Claude Delpha, Demba Diallo, Lassad Sbita

To cite this version:

Amel Adouni, Claude Delpha, Demba Diallo, Lassad Sbita. Data-driven approach for dip voltage fault detection and identification based on grid current vector trajectory analysis. 42nd Annual Conference of the IEEE Industrial Electronics Society (IECON 2016), IEEE, Oct 2016, Florence, Italy. pp.6971 - 6976, 10.1109/IECON.2016.7793167 . hal-01390878

HAL Id: hal-01390878

https://centralesupelec.hal.science/hal-01390878v1

Submitted on 12 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data-driven Approach for Dip Voltage Fault Detection and Identification based on Grid Current Vector Trajectory Analysis

Amel ADOUNI[1], Claude DELPHA[2], Demba DIALLO[3], Lassaad SBITA[1]

[1]Laboratoire Système Photovoltaïque, Eolien et Géothermique (SPEG)
Gabes, Tunisia

[2]Laboratoire des Signaux et des Systèmes (L2S), Université Paris Saclay
CNRS, CentraleSupélec, Univ. Paris-Sud, 91192 Gif Sur Yvette, France

[3]Group of Electrical Engineering - Paris (GeePs), Université Paris Saclay
CNRS, CentraleSupélec, UPMC, Univ. Paris-Sud, 91192 Gif Sur Yvette, France

Email : [1]Amel.Enig@gmail.com, [2]claude.delpha@lss. supelec.fr, [3]demba.diallo@centralesupelec.fr

Abstract— **This paper proposes a data driven approach for dip voltage fault detection and identification using the grid current vector trajectory in the stationary reference frame. Three features are extracted for the different operating conditions to build the database and analysed using Linear Discriminant Analysis to identify the fault type and subtype. In the subspaces spanned by the factorial components the four faults and eight out of nine faults subtype are successfully identified and isolated with an error rate less than 5%. Simulation results prove the efficiency of the proposed algorithm.**

Keywords*— *Dip Voltage, Grid current, fault detection, fault identification, Linear Discriminant Analysis.

I. INTRODUCTION

Electricity is increasingly essential in all human activities from industrial to transportation applications. The electricity flows into a grid from sources to end users through lines, circuit breakers, transformers, etc. The energy flow may be interrupted due to environmental disaster or inherent faults (faulty component). The energy unavailability can lead to catastrophic consequences or important economic losses.

The dip voltage is one of the crucial faults that can affect the electrical grid. Indeed, some equipment may trip when the voltage amplitude drops below 90% of the full scale for longer than one or two cycles [1]. The fault can be upstream or downstream the transformer of the power substation in the transmission line.

Therefore the grid maintenance is of utmost importance to fulfill the requirements of availability, reliability, safety and maintainability. Continuous health monitoring is one of the means to operate the grid efficiently. One of its components is the Fault Detection and Diagnosis (FDD) module that is designed to detect, isolate and if required estimate the fault severity. An efficient FDD method must be sensitive only to fault, robust to perturbations and the unknown inputs, simple, easy to implement and cost effective (in terms of number of sensors or quantity of information and computational burden).

There is an abundant literature on Fault Detection and Diagnosis (FDD) summarized in figure 1 [2]. Basically, there are four steps (modelling, pre-processing, features extraction and features analysis).

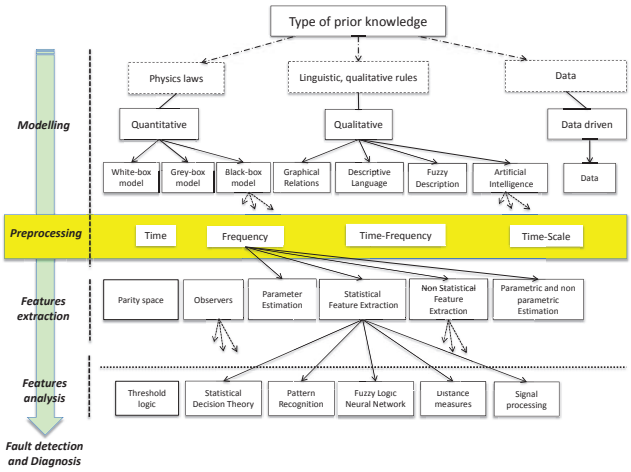


Figure 1. Fault Detection and Diagnosis Approaches.

For voltage grid fault diagnosis, the model-based approach (methods are derived from the physical laws and using analytical relations under different assumptions [3] such as observers like Luenberger or Kalman estimator [4]) would be tedious because of the complexity of the model due to the grid dimensions, the number of parameters and the different configurations.

For all the previous reasons, data driven approach is increasingly used. The idea is based on the analysis of the features extracted from pre-processed data in the time, frequency, time-frequency or time-scale domains. The extracted residuals (features) can be analyzed using different tools (signal processing methods like Hilbert-

Huang transform (HHT) [5-6], FFT, f-divergence [7]) depending on the desired performance, the user's expertise, the data dimensions and properties, etc.

There is a growing trend for the integration of renewable energy sources as power generation units to satisfy to the worldwide environmental issues of carbon-free society. The Wind Turbine (WT) is one of the leading renewable energy conversion systems. According to the International Energy Agency, the onshore wind leads the global renewable growth, accounting for over one-third of the renewable capacity and generation increase [8]. The WT is a dynamic, nonlinear and complex system characterized by coupled phenomena (mechanical, electrical, magnetic, thermal, etc).
The WT model and description can be found in [9]. The grid connected WT is displayed in figure 2.

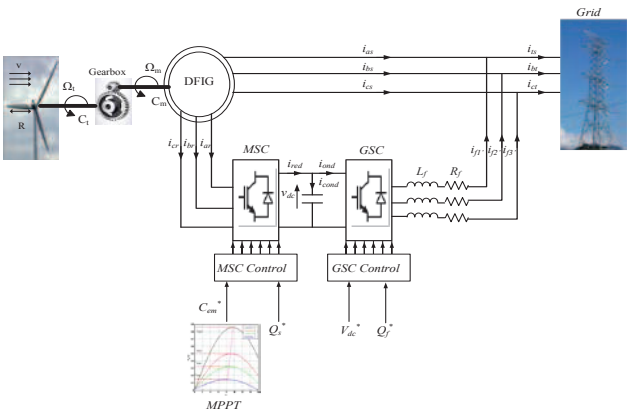


Figure 2. Topology of the wind energy conversion system

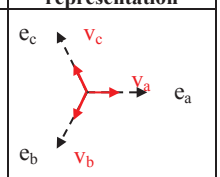
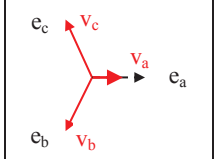
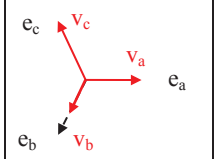
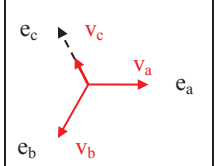
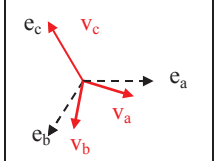
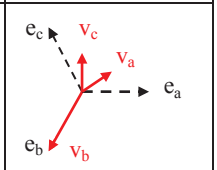
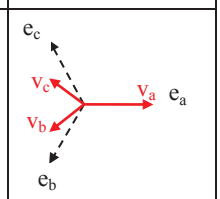
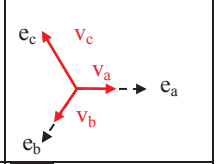
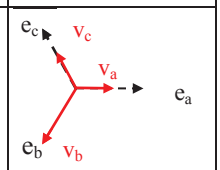
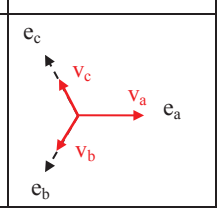
The paper is composed of 3 sections. Section 2 describes the fault modelling and the different steps of the data driven fault diagnosis approach. In section 4, the results of the use case are presented and discussed. The conclusion closes the paper.

II. FAULT DETECTION AND ISOLATION

A. Fault description

In this paper, the four principal dip voltage faults are under study. They are classified to symmetrical fault (fault type A), asymmetrical fault (fault type B, C and E). The fault B is dip voltage in one phase while the other two phases remain unchanged (three possibilities). The fault type C is a dip voltage in two phases with angle deviation while the third phase remains unchanged (three possibilities also). The Fault type E is a just dip voltage in two phases while the third remains unchanged (three possibilities also). All the possible configurations are described in Table I. In each case, the expression is given and its corresponding representation in the vector space is drawn both in healthy and faulty case. In the healthy case all the phases have the same amplitude $|V_i| = E$. If the fault occurs in phase i, $V_i = (1-d)V$ where d is the voltage dip amplitude.

TABLE I. PHASOR REPRESENTATION OF THE DIP VOLTAGE.

Type	Sub type	Voltages expressions	Phasor representation
A		$v_a(t) = (1-d)V\sqrt{2}\sin(wt)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
B	Phase a	$v_a(t) = (1-d)V\sqrt{2}\sin(wt)$ $v_b(t) = V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
	Phase b	$v_a(t) = V\sqrt{2}\sin(wt)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
	Phase c	$v_a(t) = V\sqrt{2}\sin(wt)$ $v_b(t) = V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
C	Phases a&b	$v_a(t) = (1-d)V\sqrt{2}\sin(wt - \alpha)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3} + \alpha)$ $v_c(t) = V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
	Phases a&c	$v_a(t) = (1-d)V\sqrt{2}\sin(wt + \alpha)$ $v_b(t) = V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3} - \alpha)$	
	Phases b&c	$v_a(t) = V\sqrt{2}\sin(wt)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3} - \alpha)$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3} + \alpha)$	
E	Phases a&b	$v_a(t) = (1-d)V\sqrt{2}\sin(wt)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
	Phases a&c	$v_a(t) = (1-d)V\sqrt{2}\sin(wt)$ $v_b(t) = V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	
	Phases b&c	$v_a(t) = V\sqrt{2}\sin(wt)$ $v_b(t) = (1-d)V\sqrt{2}\sin(wt - \frac{2\pi}{3})$ $v_c(t) = (1-d)V\sqrt{2}\sin(wt + \frac{2\pi}{3})$	

B. Data-driven approach

Hereafter a data-driven approach is preferred because of its simplicity as it is only based on the measured line (grid) currents. Figure 3 displays the flowchart.

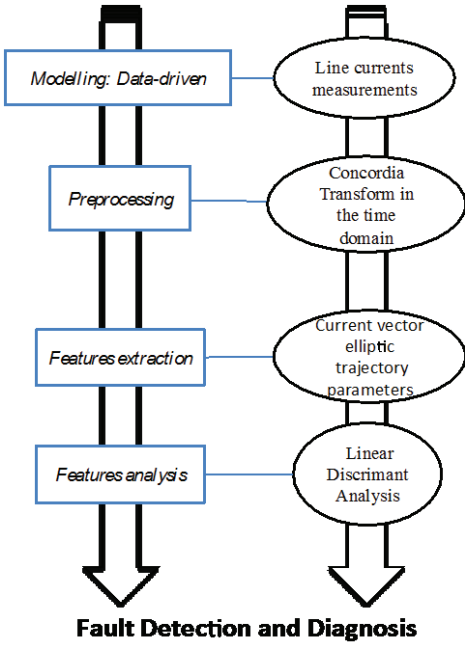


Figure 3. Flowchart of the Fault Detection and Diagnosis

1) Preprocessing and extraction of the features

The pre-processing consists of computing the Concordia Transform of the measured currents under the assumption that they are balanced. This is not conservative even if due to inherent imperfections, there is always a slight unbalance. The current vector trajectory in this stationary reference frame for the different situations (healthy and all the faulty cases) is computed. Three main characteristics (features) are extracted from the current trajectory:

- D_1 represents the value of the current i_β when i_α is equal to 0. It is the intersection between the current trajectory and the vertical axis,
- D_2 represents the value of the current i_α when i_β is equal to 0,
- The shift angle denoted θ between the major axis of the current trajectory and the horizontal axis $\theta = [OC_1, OC_2]$.

Examples can be observed in Fig. 4 that displays the current trajectory in the considered frame both in the healthy and faulty cases. It shows the extracted features.

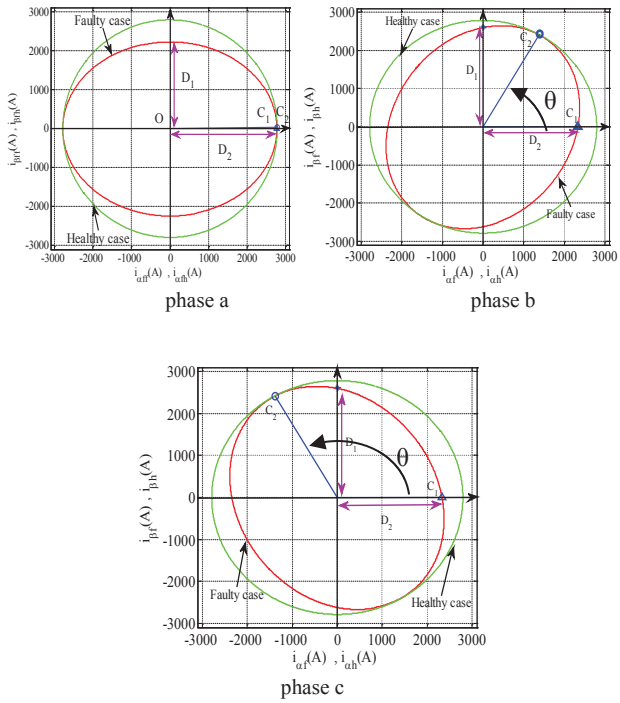


Figure 4. Current trajectory in the stationary (α,β) frame for healthy and faulty cases

For each cycle (20ms) we calculate the three features aforementioned for different operating conditions representing the healthy and faulty cases. Afterwards the information is arranged in a matrix X used as the database for the analysis. Fig. 5 summarizes the proposed procedure.

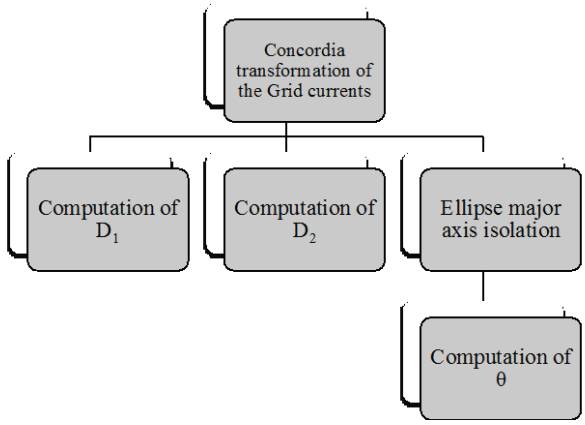


Figure 5. Computation of the features

2) Analysis of the Features

Once the aforementioned features are computed, they have to be analysed to make the diagnosis (fault detection and fault identification).

Because there are 4 fault types denoted A, B, C and E and 6 sub-types (see Table I), the features' analysis is conducted in two steps as shown in fig.6.

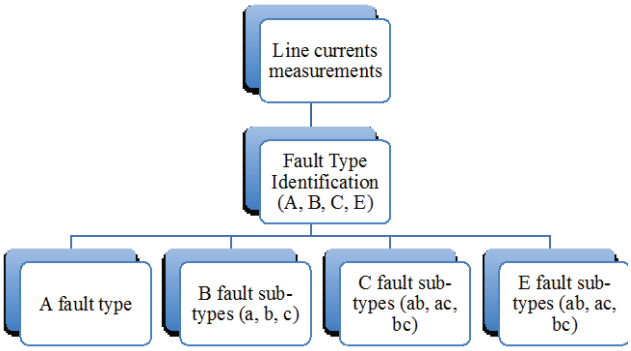


Figure 6. Fault identification and isolation flowchart

Linear Discriminant Analysis (LDA) Method

Linear Discriminant Analysis (LDA) is a discriminant factorial analysis related to both multivariate analysis of variance and multiple regressions [10]. For this method, all the subjects (measurements) in the database are affected to *a priori* groups (healthy and faulty for example). The aim of this method is first descriptive and then decisive. In the descriptive part this method is used to reduce the dimensionality of a problem. In the decisive part, the discriminant procedure identifies a linear combination of quantitative predictor variables which best characterize the differences among the chosen a priori groups. For this procedure the differences between each group are maximized and the differences inside each group are minimized. With such a procedure, a linear decisive law corresponding to the obtained discrimination is produced. The efficiency of this law can be controlled by the one-leave-out cross validation method on the original data base. In most cases, the created law is used to classify or identify new cases as unknown test samples.

If we consider our original dataset $\boldsymbol{X}$ with $\boldsymbol{n}$ observations described by $\boldsymbol{m}$ features and displayed in $\boldsymbol{K}$ a priori groups (n_k denotes the observations belonging to the kth a priori group). As explained before, we are looking for the new representation space that best represents the $\boldsymbol{K}$ groups. We need to create new factorial axis $\boldsymbol{Z_h}$ uncorrelated 2-by-2 for which the data belonging to the same group projected on these axes are as close as possible and the data belonging to different groups are as far as possible. For LDA, only $\boldsymbol{K-1}$ factorial axis can be obtained where the different groups are best separated in the chosen groups.

If we denote the variance-covariance matrix by $\boldsymbol{\Omega_k}$ representing the data dispersion inside the kth group we can define the Intra-Group dispersion as:

$$\boldsymbol{\Omega} = \sum_{\boldsymbol{k}} \boldsymbol{n_k} \times \boldsymbol{\Omega_k}$$

The distances between the groups will be defined by $\boldsymbol{B}$ the variance-covariance matrix Inter-Groups considering gravity centres for the database μ and for each group μ_k:

$$B = \sum_k n_k (\mu_k - \mu)(\mu_k - \mu)^T$$

Thus the total variance-covariance matrix $\boldsymbol{R}$ can be written as $\boldsymbol{R} = \boldsymbol{B} + \boldsymbol{\Omega}$ according to the Huyghens' theorem. For computing the new factorial axis $\boldsymbol{Z_h}$ we need to solve the equation $R^{-1}Bu = \lambda u$ and calculate the eigenvectors u with the eigenvalues $\boldsymbol{\lambda}$.

The second part of LDA consists of trying to identify new unknown cases by classifying them in the chosen groups. For that purpose, the linear combination function f_k of features which best separate the different classes is then used as a linear classifier function.

So, if we consider a new test case x_a defined for the same initial variables, we can compute the distance $d(x_a, \mu_k)$ between the new case and the centers of gravity of the groups:

$$d(x_a, \mu_k) = (x_a - \mu_k)^T R^{-1} (x_a - \mu_k)$$

The linear combination function of the features f_k for each group used as a linear classifier is also called discriminant function:

$$f_k(x_a) = d(x_a, \mu_k) x_a^T R^{-1} x_a$$

The new sample x_a will be affected to the kth group with the lowest linear result:

$$x_a \in k \text{ if } f(x_a) = \min\{f_k(x_a)\}$$

In this application, the discriminant function was computed but was not applied to unknown samples.

III. Results and Discussion

In order to evaluate the efficiency of the method, the algorithm is implemented in Matlab®. We inject a 30% dip voltage in the electrical grid in arbitrary different time intervals. Consequently, the three phase voltages are affected according to the faulty phase(s) considered. Data are collected in databases for the 3 mentioned features (d_1, d_2, θ).

For our analysis the number of samples according to the considered situation is displayed in Table II.

TABLE II. Number of samples in the databases

Fault type	Fault subtype			Number of samples			Total
A	-			40			40
B	a	b	c	44	42	44	130
C	ab	bc	ac	42	42	43	127
E	ab	bc	ac	86	43	43	172

Using the LDA, we have first tried to discriminate the 4 fault types A, B, C, E whatever the fault subtypes. The database is then composed of 469 samples to be separated in 4 classes. So we have 3 factorial axes containing all the

information to be separated. As the first two axes retain most of the information (99,97%), we present in Figure 7 the LDA result for the 4 fault types discrimination.

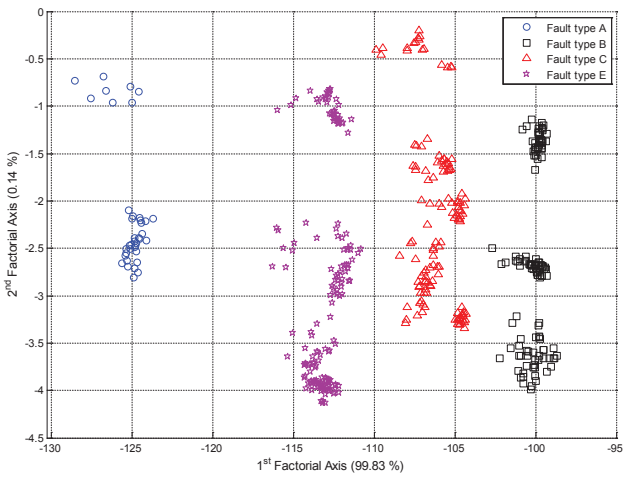


Figure 7. Fault type discrimination with LDA

As it can be observed, the four fault types can be correctly separated whatever the fault subtype considered. This result is confirmed by the confusion matrix given in table III. 100% of the data corresponding to the fault type A, B or E are correctly classified while 99,21% are considered to be well classified for the C fault type. In fact 0,78% of the data is classified as E fault type corresponding to only 2 misclassified samples. This result is confirmed by the one-leave-out cross validation approach, which validates the created discrimination function. It could then be used for the discrimination of new test samples corresponding to the mentioned fault types without errors.

TABLE III. Confusion Matrix for Fault type analysis

		A posteriori			
		Fault type A	Fault type B	Fault type C	Fault type E
A priori	Fault type A	100	0	0	0
	Fault type B	0	100	0	0
	Fault type C	0	0	99,21	0,78
	Fault type E	0	0	0	100

As we can properly discriminate the fault type, we can consider each of these types in separate databases and try to discriminate the subtype for each of these faults. The result for the fault type B is displayed in Figure 8. In this case the 2 factorial axes represent 100% of the information. We obtain in this case a perfect separation between the subtypes a and b. Only one sample initially belonging to the subtype c is misclassified in the subtype b.

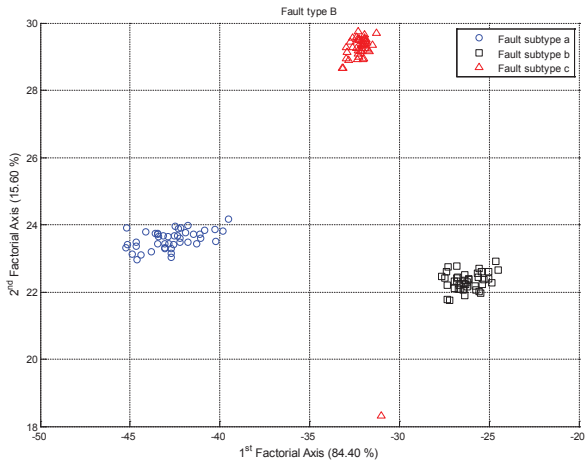


Figure 7. Fault subtype discrimination for the fault type B with LDA

The corresponding confusion matrix summarizing the results is given in table IV.

TABLE IV. Confusion Matrix for Fault type B and subtypes analysis

		A posteriori		
		Fault sub type a	Fault sub type b	Fault sub type c
A priori	Fault sub type a	100	0	0
	Fault sub type b	0	100	0
	Fault sub type c	0	2,27	97,73

The results obtained in this table IV are mainly confirmed by the one-leave-out cross validation technique. That confirms the very good separation result.

The same analysis is done by considering the database of the fault type C and its subtypes ab, ac and bc. In this case the two factorial axes obtained with LDA are considered and represent 100% of the information in the database.

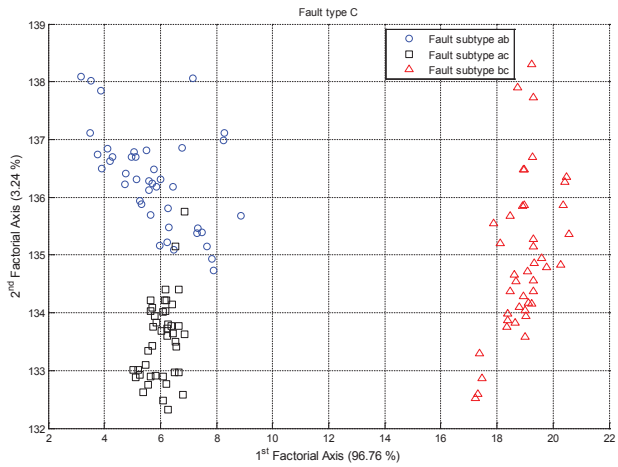


Figure 8. Fault subtype discrimination for the fault type C with LDA

As shown in Figure 8, the discrimination is efficient. The confusion matrix in table V widely confirms this result. For this fault type only the subtype bc is perfectly separated. For the subtypes ab and ac, there are some classification errors (2 for subtype ac and 4 for subtype ab).

TABLE V. Confusion Matrix for Fault type C and subtypes analysis

		A posteriori		
		Fault sub type ab	Fault sub type ac	Fault sub type bc
A priori	Fault sub type ab	95,24	4,76	0
	Fault sub type ac	4,65	95,35	0
	Fault sub type bc	0	0	100

The good separation result results obtained and displayed in this table V are also confirmed by the leave one out cross validation technique. As for the previous case the two factorial axes obtained with LDA are considered and represent 100% of the information in the database.

The last analysis done here concerns the database of the fault type E and its subtypes ab, ac and bc.

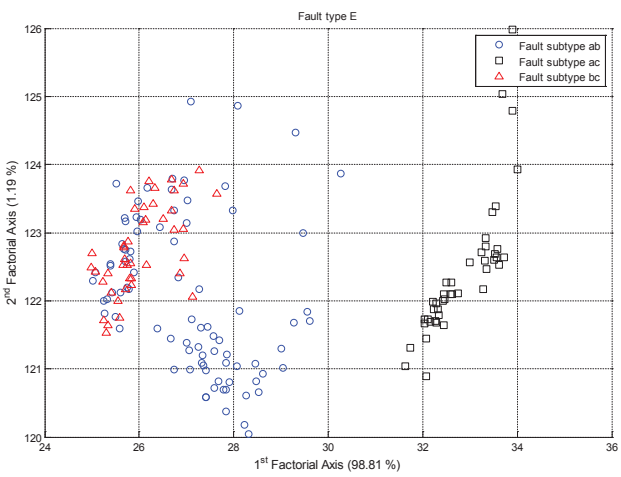


Figure 9. Fault subtype discrimination for the fault type E with LDA

The classification results obtained for this analysis are less good. For the subtypes ab and bc several overlappings are observed (see Figure 9).

The confusion matrix in table VI confirms these results. In this last case, the linear classification is not sufficient for discriminating the subtypes with no errors.

TABLE VI. Confusion Matrix for Fault type E and subtypes analysis

		A posteriori		
		Fault sub type ab	Fault sub type ac	Fault sub type bc
A priori	Fault sub type ab	68,60	1,16	30,23
	Fault sub type ac	0	100	0
	Fault sub type bc	41,86	0	58,14

IV. CONCLUSION

This paper proposes a procedure for dip voltage fault detection and identification based on the grid current measurements. After the Concordia transformation, three features are extracted from the current vector trajectory analysis. Using the Linear Discriminant Analysis of these features, the four fault types and the 8 among the 9 fault subtypes have been identified and isolated with less than 5% error. For the last fault subtype, the linear classification failed with an error rate of 40%.

REFERENCES

[1] AK. Goswami, CP. Gupta, GK. Singh, "An analytical approach for assessment of voltage sags". Int J Electr Power Energy Syst, vol. 31, pp. 418–426, 2009.

[2] J. Harmouche, C. Delpha, D. Diallo, Y. Le Bihan, "Statistical approach for non-destructive incipient damage detection and characterisation using Kullback-Leibler Divergence", IEEE Transactions on Reliability, Accepted for publication.

[3] A. Baggini. Handbook of power quality. John Wiley & Sons; 2008.

[4] U. Andreas, N. Kilian, W. Ralf, S. Stefan, "SOGI based grid fault detection for feeding asymmetrical reactive currents to fulfil LVRT requirements", The IEEE AFRICON conference, Mauritius, 9-12 September 2013, pp. 1-5.

[5] D., Solomatine, L. M. See, and R. J. Abrahart, "Data-driven modelling: concepts, approaches and experiences. Practical hydroinformatics". Springer Berlin Heidelberg, pp. 17-30, 2009.

[6] L. Lovisolo, J.A. Moor Neto, K. Figueiredo, L. de Menezes Laporte, J.C. Dos Santos Rocha, "Location of faults generating short-duration voltage variations in distribution systems regions from records captured at one point and decomposed into damped sinusoids", IET Generation, Transmission & Distribution, Vol. 6, Issue. 12, pp. 1225–1234, 2012.

[7] J. Harmouche, C. Delpha, D. Diallo, "Improved Fault diagnosis of ball bearings based on the global spectrum of vibration signals" IEEE Trans. on Energy Conversion, vol. 30, n°1, pp. 376-383, March 2015

[8] P. Thakur, A. K. Singh, S.B. Singh, "Type Detection of Voltage Sags through Voltage Unbalance Factor", 15th IEEE International Conference on Harmonics and Quality of Power (ICHQP), Hong Kong, China, 17-20 June 2012, pp. 892-896.

[9] A. Adouni, D. Diallo, C. Delpha, L. Sbita, "Voltage Dip fault Detection and Identification based on Principal Component Analysis : application to Wind Energy Conversion System", IEEE International Symposium on Industrial Electronics (ISIE), Buzios, Rio de Janeiro, Brazil, 2015.

[10] D.F. Morrisson, *Multivariate statistical methods*, Singapore : Mc Graw Hill, 2nd edn., 1988, pp. 415