HIGHER-ORDER STATISTICS-BASED METHOD FOR TRANSMISSION LINE FAULT CLASSIFICATION

Pedro Henrique Galdino Bouzon
Federal University of Lavras, Automation Department, Campus, ZIP Code 37.200-900, Lavras, MG, Brazil

Gabriel Aparecido Fonseca
Federal University of Lavras, Automation Department, Campus, ZIP Code 37.200-900, Lavras, MG, Brazil

Mônica Maria Leal
Federal University of Rio Grande do Norte (UFRN), ECT, Campus, P. O. Box 1524, ZIP Code 59,078-900, Natal, RN, Brazil

Alysson Alves Fernandes
Federal University of Lavras, Automation Department, Campus, ZIP Code 37.200-900, Lavras, MG, Brazil

Leandro Rodrigues Manso Silva
Federal University of Juíz de Fora (UFJF), Engineering Department, Campus, P. O. Box 20010, ZIP Code 36,036-110, Juíz de Fora, MG, Brazil

Belisario Nina Huallpa
Federal University of Lavras, Automation Department, Campus, ZIP Code 37.200-900, Lavras, MG, Brazil

Flavio Bezerra Costa
Michigan Tech, 1400 Townsend Drive, Houghton, Michigan 49931-1295

Danton Diego Ferreira
Federal University of Lavras, Automation Department, Campus, ZIP Code 37.200-900, Lavras, MG, Brazil
Corresponding author: danton@ufla.br

Abstract: Research on power transmission lines has been the subject of several studies aiming to provide relevant information to electrical system users. One of the focuses of this study area is transmission line fault classification. This paper presents an approach for fault classification using higher-order statistics and an artificial neural network-based classifier. A detector based on Euclidean distance was implemented to reduce classifier complexity. The proposed method takes advantage of requiring only ½ cycles of postfault data to perform the classification; therefore, it is suitable for real-time processing. The proposed method classified 10 classes of faults with global efficiency above 97%.

Keywords: Higher-order Statistics, Fault Classification, Transmission Lines.

Received: June 12, 2023 - Accepted: July 15, 2023

INTRODUCTION

Power transmission lines are long elements exposed to natural and human events that can lead to malfunctions. A fault is defined as any failure that interferes with normal current flow (GRAINGER; STEVENSON, 1994). Faults occur mainly when electrical discharges damage the line insulators. In this context, fast fault detection and classification have become essential so that ultrafast relays can trip to minimize the harmful effects of a fault.

The types of transmission line faults that can occur can be categorized into eleven classes. According to Avagaddi, Edward, and Ravi (2017), single line-to-ground (SLG) faults are the most common and least dangerous, occurring in just one phase (AG, BG, CG). Two-phase faults (AB, BC, CA) are slightly more rare and severe. Double line-to-ground (DLG) faults, which involve two phases and a ground (ABG, BCG, CAG), are much less common and more harmful.

Finally, three-phase faults (ABC and ABCG) rarely occur; however, when they do, they can lead to electric system collapse. Three-phase faults are usually known as symmetric faults, while the other faults are known as asymmetric faults.

Several fault detection and classification methodologies can be found in the literature. In general, they differ in the feature extraction technique and classification method. The wavelet transform (WT) was used together with an artificial neural network (ANN) in Silva, Souza, and Brito (2006) and with linear discriminant analysis (LDA) in Yadav and Swetapadma (2015). The half-cycle discrete Fourier transform (HCDFT) was adopted in the works reported in Arash Jamehbozorg and S Mohammad Shahrtash (2010) and Jamehbozorg and SM Shahrtash (2010) in conjunction with a decision tree to classify faults. The concept of the Stockwell transform (ST) was used in Samantaray, Dash, and Panda (2006), Samantaray and Dash (2008) and Samantaray (2013).

In Thukaram, Khincha, and Vijaynarasimha (2005), principal component analysis (PCA) was applied directly to current and voltage signals, and in Cheng, Wang, and Gao (2015), a feature extraction stage based on random projections was applied. In Godse and Bhat (2020), the feature extraction and feature selection stages were conducted using a morphological median filter (MMF) and information gain, respectively. Three methods were used for fault classification (a decision tree, a neural network and a support vector machine (SVM)), and their performances were compared to analyze which one had better accuracy. Almeida et al. (2017) used independent component analysis (ICA) as a preprocessing stage and signal filtering for an SVM-based classifier to compare the classification of the filtered signal and noisy signal.

In this study, voltage signals are continuously monitored by a Euclidean distance-based detector (RIBEIRO, E. G. et al., 2018). If a fault is detected, higher-order statistics (HOS) are used to extract features of current and voltage signals in the form of cumulants. Fault classification is performed with a multilayer neural network (MNN). The novelties of the method are the use of cumulants in a form proposed by Danton D Ferreira et al. (2011) to classify faults in transmission lines. HOS is used in the preprocessing phase to extract meaningful characteristics from the signal, after which Pearson's correlation coefficient and Fisher's linear discriminant are applied to the data to select the most relevant features for classification. The advantage of this approach is that HOS is immune to Gaussian noise and has a good ability to represent nonlinear processes.

Compared with other works in this area, the contribution of the paper is that the proposed method classifies fault occurrences in the transmission line when little information is available. Few projects explore the possibility of dealing with fewer than ¼ cycles. Therefore, for this project, a model was projected by varying the signal sampling after a fault occurrence. Tests were performed considering ¼ prefault cycles and 1, ½, ¼, ⅛, 1/16, 1/32 postfault cycles. For each variation, a model was built, and their performances were compared, which allowed us to measure the impact of the lack of information for the proposed method. It is important to note that a fairly good result was obtained even when signals with 1/32 cycles were used.

The rest of the paper is divided as follows: Section 2 describes the proposed method, where Higher-order statistics are explained, and the data set used. In Section 3, the results are presented and discussed. Finally, the conclusions are presented in Section 4.

MATERIAL AND METHODS

The proposed method

The proposed method is organized as follows: Voltage signals are monitored by a detector based on Euclidean distance (RIBEIRO, E. G. et al., 2018); if an abnormal condition is detected, current and voltage signals are segmented and features based on higher-order statistics are extracted; The most significant characteristics are selected based on Fisher's linear discriminant, and the redundance among them is eliminated using Pearson's coefficient; Finally, a multilayer perceptron performs the classification in one of the following fault categories: $AT, BT, CT, AB, AC, BC, ABT, ACT, BCT$ and

ABT . Figure 1 shows the block diagram for the proposed method. Fig 1 shows the block diagram for the proposed method.

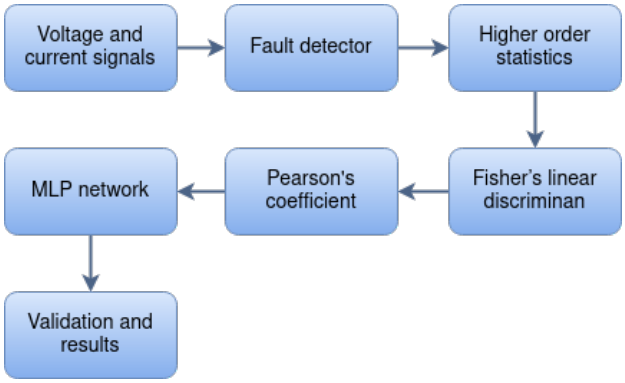

Figure 1: Proposed method block diagram.

Table 1: Parameters of the electrical system model.

Source - S1	Source - S2
$V = 500.0\angle 0°kV$	$V = 500.0\angle 20°kV$
$Z_{S10} = 1.0140 + j18.754\Omega/km$	$Z_{S20} = 1.1268 + j20.838\Omega/km$
$Z_{S11} = 0.8710 + j25.661\Omega/km$	$Z_{S21} = 0.9681 + j28.513\Omega/km$
Transmission Lines - TL1 & TL2	
$Z_{TL1} = 0.3996 + j0.9921\Omega/km$	$Z_{TL2} = 0.0333 + j0.3170\Omega/km$
$Y_{TL1} = 3.0839\mu\mho/km$	$Y_{TL2} = 5.2033\mu\mho/km$

The data was divided into design data (70%) and test data (30%).

Table 2: Simulation parameters.

Resistance (Ohms)	Distance to Local Bar (km)
1	20
1	150
1	280
50	150
100	150

Dataset

The data used in this project were generated using the Bergeron model with Simulink/MATLAB and ATP software. The model used was presented by Costa, Souza, and Brito (2010); however, some of the fault characteristics were changed for this project. Its diagram can be seen in Figure 2, and Table 1 presents its parameters. The transmission system is composed by two transmission lines of 300 km each, the rated voltage is 500 kV and a noise with 60 dB signal-to-noise ratio (SNR) was added. The fundamental frequency adopted was 60 Hz, and the sampling frequency was 15,360 Hz, resulting in 256 points per cycle. Ten fault types were simulated: AT, BT, CT, AB, AC, BC, ABT, ACT, BCT and ABC.

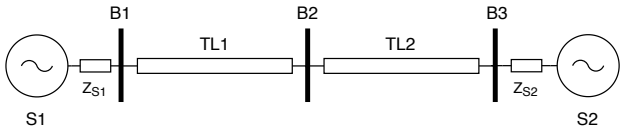

Figure 2: Diagram for the electrical system used to generate the data.

Altogether, 950 faults were generated. Table 2 describes the five parameters' combinations used in the simulation. In each combination, 190 signals of each fault type were generated with incidence angles from 0° to 180° with steps of 10°.

Fault detection

The fault detector used in this work was proposed by Eduardo G Ribeiro et al. (2018), where the monitoring is made with a sliding window (sample by sample) equal to one cycle of the fundamental signal. In this method, the voltage signal is considered a d-dimensional point, where d is the size of the monitored window. The distance from this point to the center of the space, defined as $\mathbf{c} = [0\,0\,0\cdots 0_d]$, is then calculated. The fault detection (distance calculation) operation is described as follows:

$$r = \|\mathbf{v} - \mathbf{c}\|^2 \tag{1}$$

where $\mathbf{v}$ represents the monitored window and is the center of the space. This analysis derives a hyper spherical region on the d-dimensional space that models the nondisturbed signals (RIBEIRO, E. G. et al., 2018).

In this project, detection was conducted using only the voltage signals. When a fault was detected in a given phase, the same cut point was used for the other two. This point was also used in the current signals without the need to use the detector again.

To define the region where there is no fault, nominal signals comprising phases from $-180°$ to 180° and SNR (*signal-to-noise ratio*) of 60 dB are

used to determine the values of r in this case. From these values, the lower and upper limits are adopted as:

$$r_{max} = \bar{r} + 3 * \sigma \tag{2}$$

and

$$r_{min} = \bar{r} - 3 * \sigma \tag{3}$$

where $\bar{r}$ is the mean value and σ is the standard deviation of r.

When a fault is detected, a window with a quarter of prefault cycles and ½ postfault cycles is segmented for feature extraction. The quantity of pre- and postfault data was determined experimentally. Those values were initially set to the size of one cycle and then reduced to a point where the results were still acceptable. The prefault data were introduced to add information about the transition from a normal condition to a fault condition.

Higher-order statistics

In the past, signal processing, signal analysis, system identification, and signal estimation problems were primarily based on second-order statistical information. Autocorrelations and cross-correlations are examples of second-order statistics (SOS). The power spectrum, which is widely used and contains useful information, is also based on the SOS in that the power spectrum is the one-dimensional Fourier transform of the autocorrelation function. As Gaussian processes exist, and a Gaussian probability density function (pdf) is completely characterized by its first two moments, the analysis of linear systems and signals has so far been quite effective in many circumstances. It has nevertheless been limited by the assumptions of Gaussianity, minimum phase systems and linear systems. When signals are non-Gaussian, the first two moments do not define their pdf, and consequently, higher-order statistics (HOS), namely, of order greater than two, can reveal information about them other than SOS alone.

Higher-order statistics (HOS) can be described by means of cumulants or moments. The first applies to random signals, while the second applies to deterministic signals. This technique is better explored when used on non-Gaussian processes and nonlinear systems (FERREIRA, Danton D et al., 2011), (FERREIRA, D. et al., 2009), which makes the approach promising for transmission line fault classification.

The cumulants of the second, third, and fourth orders of a random signal $x[n]$ with $Ex[n] = 0$ are respectively given by Mendel (1991):

$$c_{2,x}[i] = E\{x[n]x[n+1]\}, \tag{4}$$

$$c_{3,x}[i] = E\{x[n]x^2[n+1]\}, \tag{5}$$

$$c_{4,x}[i] = E\{x[n]x^3[n+1]\} - 3c_{2,x}[i]c_{2,x}[0] \tag{6}$$

Considering a finite vector with size N, the stochastic approximations result in (KUSHNER; YIN, 2003):

$$\hat{c}_{2,x}[i] = \frac{2}{N}\sum_{n=0}^{N/2-1} x[n]x[n+i], \tag{7}$$

$$\hat{c}_{3,x}[i] = \frac{2}{N}\sum_{n=0}^{N/2-1} x[n]x^2[n+i] \tag{8}$$

and

$$\hat{c}_{4,x}[i] = \frac{2}{N}\sum_{n=0}^{N/2-1} x[n]x^3[n+i] - \frac{2}{N^2}\sum_{n=0}^{N/2-1} x[n]x[n+i]\sum_{n=0}^{N/2-1} x^2[n], \tag{9}$$

where $i = 0, 1, \cdots, \frac{N}{2} - 1$. (7), (8), and (9) cannot be used for $i > \frac{N}{2} - 1$ because the index $n + i$ exceeds the size of the vector. Due to this limitation, alternative approximations were introduced by Moisés V Ribeiro et al. (2007). In this new approach, the cumulants are estimated as follows:

$$\hat{c}_{2,x}[i] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x[\mathrm{mod}(n+i, N)], \tag{10}$$

$$\hat{c}_{3,x}[i] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x^2[\mathrm{mod}(n+i,N)] \tag{11}$$

and

$$\hat{c}_{4,x}[i] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x^3[\mathrm{mod}(n+i,N)] - \frac{3}{N^2}\sum_{n=0}^{N-1} x[n]x[\mathrm{mod}(n+i,N)]\sum_{n=0}^{N-1} x^2[n], \tag{12}$$

where $\mathrm{mod}(n+1,N)$ is the integer remainder of the division of $n+i$ by N. The use of the mod operator implies in the assumption that the vector $x[n]$ is periodic.

The use of HOS to obtain relevant information about electrical signals in power systems has been shown to be a good alternative for feature extraction (ROSA; MORENO-MUNOZ, 2009). For a finite feature vector, such statistics can be approximated by (10), (11) and (12). In this work, such approximations were used to extract relevant higher-order features from the monitored three phase voltage and current signals.

According to Mendel (1991), the second-order cumulant $C_{2,x}(i)$ is just the autocorrelation of $x[n]$. The third-order cumulant $C_{3,x}(i)$ measures if a random process (in this case the signal x[n]) is symmetrically distributed, and if so, its third-order cumulant equals zero. Thus, the third-order cumulant is a measure of how different a random process is from a symmetric distribution. On the other hand, the fourth-order cumulant$C_{4,x}(i)$ provides a measure of the distance of a random process ($x[n]$) from Gaussianity. If signal $x[n]$ presents a Gaussian distribution, the fourth-order cumulant is zero. It is worth mentioning that equations (4) to (6) give the variance, skewness and kurtosis measures in terms of cumulants at zero lags (i = 0). For different lags, further higher-order information may be captured with the cumulants from the transmission line faults that are useful for classification purposes.

Fisher's discriminant ratio

Since the set of features generated by this approach is very large, Fisher's discriminant ratio (FDR) was used to select the best cumulants that discriminate between classes. FDR has been applied in classification problems to select and reduce input data (BARBOSA et al., 2016), (NAVES; BARBOSA; FERREIRA, 2016), (GUEDES et al., 2015). The FDR cost function for a problem with M classes is given by (THEODORIDIS et al., 2010):

$$\mathbf{J} = \sum_{i}^{M}\sum_{j\neq i}^{M}(\mu_i - \mu_j)^2 \odot \frac{1}{(\sigma_i^2 + \sigma_j^2)} \tag{13}$$

where μ_i, σ_i, μ_j and σ_j are the mean and variance vectors in classes i and j, respectively. Symbol $\odot$ implements the Hadamard product. Considering a feature space with $l = 1, 2, ..., L$ features, the higher is the value of J_l, more relevant is the l-st feature.

Pearson's Correlation Coefficient

Since FDR does not remove redundancy, another technique must be used for this purpose. In this work, an algorithm based on Pearson's correlation coefficient was implemented to limit the correlation between the selected features. The calculated Pearson coefficient for the features c_i and c_j with autocorrelation and cross-correlation matrices equal to σ_{ii}, σ_{jj} and σ_{ij} is given by:

$$\mathbf{R_{ij}} = \frac{\sigma_{ij}}{\sigma_{ii} * \sigma_{jj}} \tag{14}$$

where the closer to the unit the absolute value of R_{ij} is, the greater the correlation of features.

From (14), the developed algorithm calculates the R_{ij} between all features selected by the FDR, and for each pair of feature with $R_{ij} > 0.8$, the feature with the lowest J value (see Equation (13)) is discarded.

Multilayer Neural Network

The multilayer neural network (*multilayer perceptron* - MLP) consists of a system of simple interconnected neurons and represents a nonlinear mapping between an input vector and an output vector. The neurons are connected by weights and output signals that are a function

of the sum of the neuron inputs modified by a nonlinear activation function (GARDNER; DORLING, 1998).

According to Murtagh (1991), to define a multilayer artificial neural network, it is necessary to determine its configuration or architecture (number of layers, number of neurons per layer, etc.), the activation function to be used in neurons, the training method to be used, and how to update weights (real time or *offline*).

During network training, the *dropout* technique was used, which prevents the occurrence of *overfitting* and provides a method for efficiently combining different neural networks architectures. For this purpose, neurons are temporarily removed from the network along with all their input and output connections (SRIVASTAVA et al., 2014). The choice of units to be removed is made at random and for this project a rate of 20% was chosen for the removal of neurons.

Another technique used was *early stopping*, which also avoids *overfitting* and is based on monitoring a certain metric during training. A constant check is carried out between the training and test sets to determine whether there is an improvement, interrupting the process at the right time and preventing the network from memorizing the input data.

In this project, the hyperbolic tangent function was used in the hidden layers, and softmax was used in the output. In addition, the Adam optimizer was chosen for the search process of minimizing the classification error. The next section presents more details about the implemented fault classifier.

Fault classification

At this stage, the selected cumulants serve as the basis for training a multilayer neural network. This classification method was chosen because it has a low computational burden and good capability of solving nonlinear problems. Initially, the model layout was set to 48 input nodes, 24 hidden neurons and 10 output neurons (being one neuron for each class of fault), and then it was pruned by gradually removing neurons and observing the change in the model's accuracy. The final architecture of the classifier is presented in Table 3. Such architecture consists of two layers. The hidden layer has eight neurons with sigmoidal hyperbolic tangent activation function, described as follows:

Table 3: Classifier Architecture.

Layer	Neurons	Activation Function
Hidden Layer	8	Hyperbolic Tangent
Output Layer	10	Softmax

$$g_1(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{15}$$

The output layer has ten neurons, representing the possible faults, with a softmax activation function, described as follows:

$$g_2(x_i) = \frac{e^{-x_i}}{\sum_{j=1}^{10} e^{-x_j}} \tag{16}$$

For network training, the Adam optimizer proposed by Kingma and Ba (2014) was applied. This optimizer was chosen because it has shown a faster convergence than others, such as stochastic gradient descent (KINGMA; BA, 2014). The hyperparameters of this algorithm were kept at values suggested by the authors of the method. The used learning rate was 0.01.

RESULTS AND DISCUSSION

Feature Extraction

Second-, third- and fourth-order cumulants were extracted from phases A, B, C and Z of the voltage and current signals. For each order of cumulant, a set of N features was generated, where N is the size of the window segmented by the detector, and represents the lags applied in the cumulant computation. Due to the symmetry of the second-order cumulants, the first half of them were discarded. The quadratic mean value and the absolute maximum value of the signals were also extracted, leading to a total of 1,456 features.

To reduce the number of features and thus the computational complexity of the method, FDR was applied to the total set of features. Figure 3 shows the result of the FDR cost function in terms of the feature set. Each point in this figure

represents the value of importance given by Fisher's criterion to the corresponding feature.

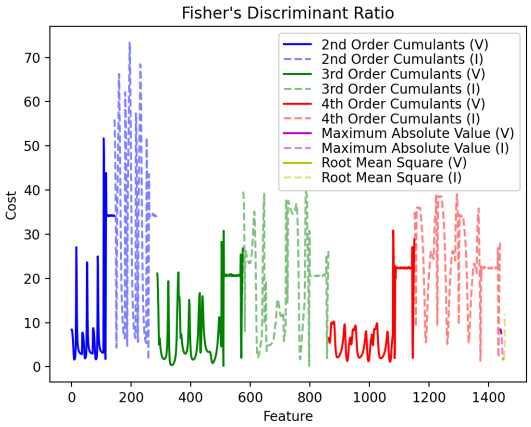

Figure 3: Fisher's discriminant ratio cost function in terms of the feature set.

After applying the FDR, 80 features with less than 50% correlation between them were selected. Then, this number was further reduced by removing one feature at a time and performing the classification. Since the accuracy achieved using all 80 features and the 25^{th} best were the same, the number of selected features was reduced to 25.

Figure 4 shows the number of selected features for each phase. The maximum absolute values were selected only for voltages and the root mean squares for currents. The fourth-order cumulants were selected from phases A and B of the current signals and from phase B of the voltage signals. The third-order cumulants were selected from all phases of the current and phases A, B and C of the voltage. The second-order cumulants were selected from all phases of the voltage signals. This shows that all phases are important and that the use of both current and voltage signals is not redundant as all of them contributed to the classification task, according to both selection criteria used (FDR and linear correlation).

The number of features extracted from phase Z is smaller than the others because not all faults include the ground. The feature selection method succeeded in capturing this information from the data.

Observing the results shown in Figure 4, it is

Figure 4: Amount of selected features for each phase.

possible to note that the RMS value is significant for all phases of current, except phase Z, it is due to the fact that the RMS value of the currents could increase significantly during the fault in the faulty phase. For the voltage signals, the Maximum Absolute Value (MVA) shows to be significant for all phases, including phase Z that could be justified by the fact that the MVA will decrease in faulty phases due to a voltage unbalance affecting all phases, including phase Z. Also, it can be observed that the at least one type of cumulant is extracted from each phase, except for phase Z, which is interesting to make the classification system more robust to noise.

To check whether the cumulants provide useful information regarding the presence of faults, the sample distribution of a cumulant was approximated by a Gaussian distribution using Parzen windows (PARZEN, 1962), as shown in Figure 5. These cumulants were extracted from phase C of the current signal. The distributions of the classes containing phase C are located to the left, while the others are located to the right showing a certain distinction of fault C from the others.

Figure 6 shows the feature space generated by the three best-selected features (C1, C2 and C3) for different quantities of postfault data. With one cycle duration of postfault, there is a linear separation among the classes. Using at least ⅛ cycles we still have a good separation among classes. In other cases, the overlaps become more evident. Nevertheless, when all 25 features were used, the classifier showed good results in all cases.

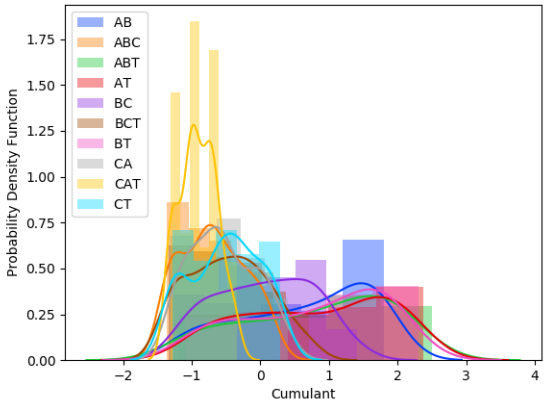

Figure 5: Probability density estimation of the cumulants extracted from phase C of the voltage signal.

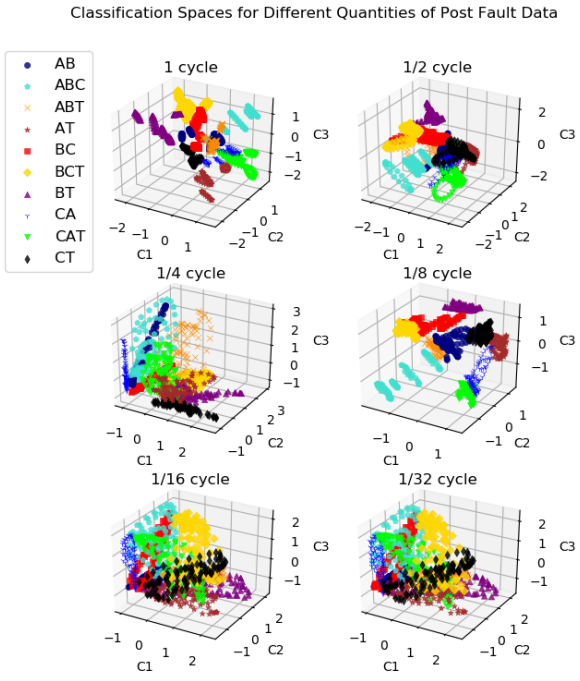

Figure 6: Feature spaces for the considered classes built by the three most relevant features considering different postfault signal length (1, 1/2, 1/4, 1/8, 1/16 and 1/32 cycles.

Classification

To validate the classifier, the k-fold cross-validation with 10 divisions has been used considering the design data set. Fig 7 presents the boxplot of the achieved accuracies for different postfault values in the window segmented by the detector, in terms of percentage, considering the 10 folds. To evaluate the performance of the proposed method for reduced postfault windows length, we varied them from 1 to 1/32 cycles of the fundamental component. It obtained 100% accuracy for ½ and ⅛ postfault cycles, 99.78% for 1, ¼ and 1/16 postfault cycles and 97.66% for 1/32 postfault cycles.

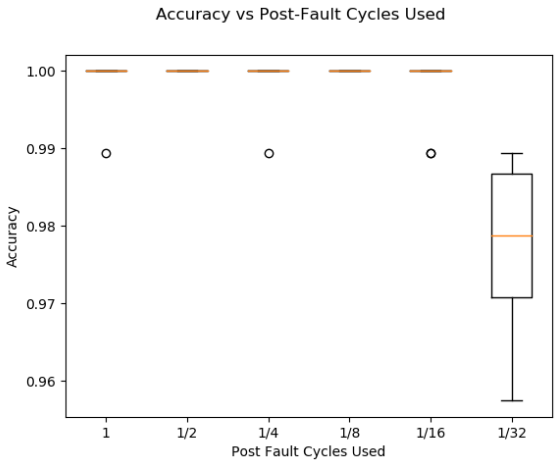

Figure 7: Influence of the number of postfault cycles used on the model's accuracy (10-fold cross-validation).

Fig 8 shows the confusion matrix for the detector configured with 1/32 of porstfalt cycles, considering the testing data set, which consisted of 30% of the data. Note that for single-phase faults and AC and ABT faults the classification result was 100%. The remaining faults presented hit percentages between 96% and 97%.

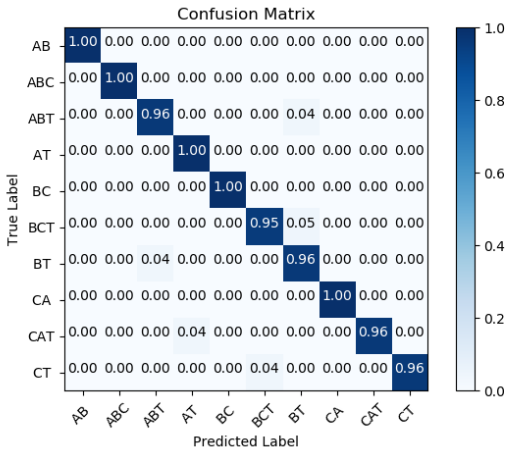

Figure 8: Confusion matrix obtained from a test set made with 30% of the available data (randomly selected).

Table 4: Projects comparison.

Scheme	Processing technique	Classification method	Data window	SNR	Accuracy
Proposed	HOS, FDR and Pearson's coefficient	Artificial neural network (MLP)	½ cycle	60 dB	97.66%
Yadav and Swetapadma (2015)	Wavelet transform	Linear discriminant analysis	¼ cycle	Not informed	100%
Chatterjee and Debnath (2020)	Cross correlation and Butterworth filter	Fuzzy inference system	½ cycle	20-40 dB	99.34%
Fathabadi (2016)	Finite impulse response filter	Support vector machine	1 cycle	Not informed	100%
Swetapadma and Yadav (2016)	Discrete Fourier Transform	Naive Bayes Classifier	Not informed	Not informed	99.99%
Mishra, Yadav, and Pazoki (2018)	Singular value decomposition and S-transform	Bagged tree ensemble classifier	½ cycle	20-40 dB	100%

Table 4 shows the characteristics and accuracy of current methods that were developed in the area of fault classification. Although these methods were not replicated with the same database described in this paper, analyzing their results, it is possible to verify that the proposed method is the only one that can obtain good accuracy even though it uses less fault information (1⁄32 postfault cycles). The results when using more fault information (1, ½, ¼, ⅛ and 1⁄16 post fault cycles) are similar to what is found in the literature. However, the proposed method takes advantage of being more suitable for real-time processing and protection purposes, since it is able to process reduced signal windows (with less postfault information) with good performance. It is important to emphasize that the main advantage of using less samples than the other methods is the capability of detecting the fault events and tripping the relay faster. The usage of few samples also contributes to the reducing of the computational complexity making the proposed method suitable to be implemented in real time in a digital relay. Another point that differs the proposed method from the others is the usage of Higher-Order Statistics information as input to the classifier. The others use only deterministic information or cross-correlation, that are both second-order statistic features.

CONCLUSION

This paper proposed a classification method based on higher-order statistics that uses only 1⁄32 postfault cycle. In this sense, it was possible to classify 10 fault types, including single-phase, two-phase and three-phase faults. To reduce the complexity of the classifier, an Euclidean distance-based detector was implemented.

It should be mentioned that the third- and fourth-order cumulants, used as input to the classifier, are less sensitive to Gaussian noise, which gives the method robustness.

Despite the proposed method has not yet been tested in real fault situations, the experiments made via simulated signals were built to imitate real situations, since we have considered the most important characteristics, and statistical tests were carried out by considering realistic variations in the fault parameters. One of the disadvantages of the proposed classification method is that it requires a training stage that depends on the available data set. However, as the focus of the work is on transmission systems, which are usually well documented, it is possible to perform realistic simulations to generate a complete data set for training. For future works, the authors intend to implement the method in a Field-programmable gate array (FPGA) hardware to evaluate the performance of the method in a real-time configuration. In addition, we intend to build a visual interface via LabVIEW software to communicate with the FPGA systems for monitoring purposes.

ACKNOWLEDGEMENT

This work was carried out with support from the National Council for Scientific and Technological Development (CNPq) and the Minas Gerais Research Support Foundation (FAPEMIG).

REFERENCES

ALMEIDA, A. R. et al. **ICA feature extraction for the location and classification of faults in high-voltage transmission lines**. v. 148. [S.l.]: Elsevier Ltd, July 2017. P. 254–263. DOI: 10.1016/j.epsr.2017.03.030.

AVAGADDI, Prasad; EDWARD, Belwin; RAVI, Kuppan. **A Review on Fault Classification Methodologies in Power Transmission Systems: Part I**. v. 5. [S.l.: s.n.], 2017. DOI: 10.1016/j.jesit.2017.01.004.

BARBOSA, Tássio S et al. Fault detection and classification in cantilever beams through vibration signal analysis and higher-order statistics. **Journal of Control, Automation and Electrical Systems**, Springer, v. 27, n. 5, p. 535–541, 2016.

CHATTERJEE, Biswapriya; DEBNATH, Sudipta. Cross correlation aided fuzzy based relaying scheme for fault classification in transmission lines. **Engineering Science and Technology, an International Journal**, v. 23, n. 3, p. 534–543, 2020. ISSN 2215-0986. DOI: https://doi.org/10.1016/j.jestch.2019.07.002.

CHENG, Long; WANG, Lingyun; GAO, Feng. Power system fault classification method based on sparse representation and random dimensionality reduction projection. In: IEEE. 2015 IEEE Power & Energy Society General Meeting. [S.l.: s.n.], 2015. P. 1–5.

COSTA, Flavio; SOUZA, B.A.; BRITO, N.S.D. Real-time detection of fault-induced transients in transmission lines. **Electronics Letters**, v. 46, p. 753–755, June 2010. DOI: 10.1049/el.2010.0812.

FATHABADI, Hassan. Novel filter based ANN approach for short-circuit faults detection, classification and location in power transmission lines. **International Journal of Electrical Power and Energy Systems**, v. 74, p. 374–383, 2016. ISSN 0142-0615. DOI: https://doi.org/10.1016/j.ijepes.2015.08.005.

FERREIRA, Danton D et al. Exploiting Higher-Order Statistics Information for Power Quality Monitoring. **Power Quality: Intech Open Access Publisher**, p. 345–362, 2011.

FERREIRA, DD et al. HOS-based method for classification of power quality disturbances. **Electronics letters**, IET, v. 45, n. 3, p. 183–185, 2009.

GARDNER, M.W; DORLING, S.R. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. **Atmospheric Environment**, v. 32, n. 14, p. 2627–2636, 1998. ISSN 1352-2310. DOI: https://doi.org/10.1016/S1352-2310(97)00447-0.

GODSE, R.; BHAT, S. Mathematical Morphology-Based Feature-Extraction Technique for Detection and Classification of Faults on Power Transmission Line. **IEEE Access**, v. 8, p. 38459–38471, 2020.

GRAINGER, J.J.; STEVENSON, W.D. **Power System Analysis**. [S.l.]: McGraw-Hill, 1994. (Electrical engineering series). ISBN 9780071133388.

GUEDES, Juan Diego Silva et al. Non-intrusive appliance load identification based on higher-order statistics. **IEEE Latin America Transactions**, IEEE, v. 13, n. 10, p. 3343–3349, 2015.

JAMEHBOZORG, A; SHAHRTASH, SM. A decision tree-based method for fault classification in double-circuit transmission lines. **IEEE Transactions on Power Delivery**, IEEE, v. 25, n. 4, p. 2184–2189, 2010.

JAMEHBOZORG, Arash; SHAHRTASH, S Mohammad. A decision-tree-based method for fault classification in single-circuit transmission lines. **IEEE Transactions on Power Delivery**, IEEE, v. 25, n. 4, p. 2190–2196, 2010.

KINGMA, Diederik P; BA, Jimmy. **Adam: A method for stochastic optimization**. [S.l.: s.n.], 2014.

KUSHNER, Harold; YIN, G George. **Stochastic approximation and recursive algorithms and applications**. [S.l.]: Springer Science & Business Media, 2003. v. 35.

MENDEL, Jerry M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. **Proceedings of the IEEE**, Ieee, v. 79, n. 3, p. 278–305, 1991.

MISHRA, P. K.; YADAV, A.; PAZOKI, M. A Novel Fault Classification Scheme for Series Capacitor Compensated Transmission Line Based on Bagged Tree Ensemble Classifier. **IEEE Access**, v. 6, p. 27373–27382, 2018.

MURTAGH, Fionn. Multilayer perceptrons for classification and regression. **Neurocomputing**, v. 2, n. 5, p. 183–197, 1991. ISSN 0925-2312. DOI: https://doi.org/10.1016/0925-2312(91)90023-5.

NAVES, Raphael; BARBOSA, Bruno HG; FERREIRA, Danton D. Classification of lung sounds using higher-order statistics: A divide-and-conquer approach. **Computer methods and programs in biomedicine**, Elsevier, v. 129, p. 12–20, 2016.

PARZEN, Emanuel. On Estimation of a Probability Density Function and Mode. **Ann. Math. Statist.**, The Institute of Mathematical Statistics, v. 33, n. 3, p. 1065–1076, Sept. 1962. DOI: 10.1214/aoms/

1177704472. Available from: <https://doi.org/10.1214/aoms/1177704472>.

RIBEIRO, Eduardo G et al. Real-time system for automatic detection and classification of single and multiple power quality disturbances. **Measurement**, Elsevier, v. 128, p. 276–283, 2018.

RIBEIRO, Moisés V et al. Detection of disturbances in voltage signals for power quality analysis using HOS. **EURASIP Journal on Applied Signal Processing**, Hindawi Publishing Corp., v. 2007, n. 1, p. 177–177, 2007.

ROSA, Juan José González de la; MORENO-MUNOZ, Antonio. Higher-order characterization of power quality transients and their classification using competitive layers. In: IEEE. 2009 Compatibility and Power Electronics. [S.l.: s.n.], 2009. P. 390–395.

SAMANTARAY, SR. A systematic fuzzy rule based approach for fault classification in transmission lines. **Applied soft computing**, Elsevier, v. 13, n. 2, p. 928–938, 2013.

SAMANTARAY, SR; DASH, PK. Pattern recognition based digital relaying for advanced series compensated line. **International Journal of Electrical Power & Energy Systems**, Elsevier, v. 30, n. 2, p. 102–112, 2008.

SAMANTARAY, SR; DASH, PK; PANDA, G. Fault classification and location using HS-transform and radial basis function neural network. **Electric Power Systems Research**, Elsevier, v. 76, n. 9-10, p. 897–905, 2006.

SILVA, KM; SOUZA, Benemar A; BRITO, Nubia SD. Fault detection and classification in transmission lines based on wavelet transform and ANN. **IEEE Transactions on Power Delivery**, IEEE, v. 21, n. 4, p. 2058–2063, 2006.

SRIVASTAVA, Nitish et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929–1958, June 2014.

SWETAPADMA, Aleena; YADAV, Anamika. Protection of parallel transmission lines including inter-circuit faults using Naïve Bayes classifier. **Alexandria Engineering Journal**, v. 55, n. 2, p. 1411–1419, 2016. ISSN 1110-0168. DOI: https://doi.org/10.1016/j.aej.2016.03.029.

THEODORIDIS, Sergios et al. **Introduction to pattern recognition: a matlab approach**. [S.l.]: Academic Press, 2010.

THUKARAM, D; KHINCHA, HP; VIJAYNARASIMHA, HP. Artificial neural network and support vector machine approach for locating faults in radial distribution systems. **IEEE Transactions on Power Delivery**, IEEE, v. 20, n. 2, p. 710–721, 2005.

YADAV, Anamika; SWETAPADMA, Aleena. A novel transmission line relaying scheme for fault detection and classification using wavelet transform and linear discriminant analysis. **Ain Shams Engineering Journal**, v. 6, n. 1, p. 199–209, 2015. ISSN 2090-4479. DOI: https://doi.org/10.1016/j.asej.2014.10.005.