SAND2022-3647C

Multi-Resolution Analysis Algorithm for Fast Fault Classification and Location in Distribution Systems

Miguel Jiménez Aparicio
Sandia National Laboratories,
Albuquerque, NM, 87185
mjimene@sandia.gov

Matthew J. Reno
Sandia National Laboratories,
Albuquerque, NM, 87185
mjreno@sandia.gov

Pedro Barba
Sandia National Laboratories,
Albuquerque, NM, 87185
pbarba@sandia.gov

Ali Bidram
University of New Mexico
Albuquerque, NM, 87131
bidram@unm.edu

Abstract— **This paper presents a new method for fault classification and location based on the Discrete Wavelet Transform decomposition and signal reconstruction - a type of Multi-Resolution Analysis. The designed signal-processing stage, which encompasses various signal transforms, plus the aforementioned decomposition in several frequency bands and the calculation of the signals' energy, provides a consistent generalization of the features that characterize the fault signal. Then, this data is fed into ensemble Machine Learning algorithms. The results show that this method is reasonably accurate while requiring a tiny amount of fault data, expanding the capabilities of Traveling Wave relays to achieve an accurate fault classification and location in just microseconds.**

Keywords— Discrete Wavelet Transform, Ensemble Machine Learning, Fault Classification, Fault Location, Distribution Systems.

I. Introduction

The evolution of relays and protection devices has advanced towards faster, more resilient, and more sensitive devices, with improved hardware and software capabilities. This has made fault detection and location faster and more accurate than ever. While traditional relays operate within 1 or 2 cycles, a few decades ago a new generation of Traveling Wave-based (TW) relays pushed this limit towards a few milliseconds. This was possible by the incorporation of techniques from others fields of knowledge, such as Digital Signal Processing, which provided more insights about how to take advantage of that phenomena for fault detection, classification, and location [1][2].

In recent years, Machine Learning (ML) methods are bringing new opportunities to Power Systems which could reduce even more the amount of data needed for this fault characterization. In this context, this paper aims to push beyond the limits provided by TW relays, and reduces the number of measurements up to the scale of microseconds. The approach described in this paper offers a re-visitation of a common approach on fault location and classification, as it is the usage of the Discrete Wavelet Transform decomposition, which is applied to a time window of 50 µs before and after the TW fault arrival. The signal is decomposed in six frequency bands, with an emphasis of frequencies over 100 kHz, and the signal energy is calculated over this time window using Parseval's Theorem. The post-fault energy is then used to summarize the response of the system to the fault. Finally, some state-of-the-art ensemble Machine Learning approaches (Random Forest and Tree Boosting) are employed to perform fault type classification and location using the previously generated data.

A. Literature review

Over the last couple of decades, there has been a growing interest in developing protective relays based on TWs, the initial part of a fault signal. Thanks to the TWs information, these devices are faster than conventional protection schemes based on post-fault measured impedances [3]. Since then, multiple approaches, most of them integrating signal processing techniques, have been developed to extract features of those fast transients. A few of them choose the Fast Fourier Transform (FFT) to get the frequency components of the measured signals [4][5] and use them as the input to an Artificial Neural Network (ANN) to calculate the distance to the fault or the fault.

However, as is indicated in many sources, Wavelet Transforms (WT) are usually preferred to FFT as they provide a time dimension to the frequency analysis. For example, the Continuous Wavelet Transform (CWT) has been employed in [6], on a distribution system for fault detection and classification, showing excellent performance. Another type of WT, the Discrete Wavelet Transform (DWT), is the choice of many algorithms because of its low computation time [7]. Most references tend to use either one or several decomposition levels given by the DWT. In the case of [8] and [9], the time differences on the peaks in the first level of decomposition are used to calculate the distance to the fault provided that the wave velocity is known.

Other references put the location/classification tasks in the hands of ML algorithms. Same as before, they feed the detail coefficients (or the energy) given by the DWT into algorithms such as SVM [10], Radial Basis Function (RBF) neural network [11], or Bayes classifiers [12]. [13] makes a comparison between Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and ANNs, highlighting the superiority of the first one on that particular case study. [14] provides an ultra-fast method, based as well on ANN, that performs fault classification using just one-eighth of a cycle of post-fault data. Some other papers investigate ANN for protecting DC systems [15]. Finally, [16] uses the energy associated with the DWT decomposition coefficients in DC systems, and then uses an SVM for fault classification and a Gaussian Process Regression Engine to calculate the fault location. This method reports needing about 200 µs of data. The other algorithms in literature need at least several milliseconds.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

B. *Contributions of the paper*

As it has been introduced before, the main contributions of this paper are: First, developing a curated (and more robust) signal processing method with comparatively lower windowing times (which makes the algorithm faster) and higher sampling frequency (which makes the algorithm more capable of detecting high-frequency patterns), decreasing the amount of required post-fault data to 50 microseconds. As far as the authors know, this is the first proposed method that claims to achieve a good performance given that small amount of measurements. Second, using powerful state-of-the-art ML methods for both classification and location tasks, which allows higher accuracies using less temporal information.

II. The System

The system used in this paper, created by the authors, is shown in Figure 1. It consists of one 12.47 kV variable-length distribution line that is connected to a load of 300 kVA per phase. Faults occur on the load bus. Simulations were performed in PSCAD for 160 fault locations (increasing the length in steps of 25 meters), for 3 fault types (Single-Line-to-Ground, Line-to-Line, and 3-Phase), and for 7 resistance values ranging from 0.01 to 10 Ω. In total, there were 3360 simulations. Three-phase voltages and currents are measured on the secondary side of the transformer at a sampling frequency of 10 MHz.

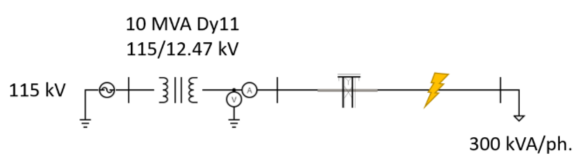

Figure 1. Schematic of the system.

III. The Method

The method described in this paper can be separated into two main stages: the signal processing block that aims to process the measured signals (+/- 50 µs since TW time of arrival) and extract useful information, and the Machine Learning stage where several algorithms are trained to perform the actual fault location and classification. The workflow is shown in Figure 2.

As it can be seen, the first part of the signal processing block consists of applying some transforms to both the 3-Phase voltage and current signals. Here, the Clarke, Karrenbauer, and DQ0 transforms are employed. Numerical results showed that the larger amount of features, the better performance. Each transform performs a different operation on the measured data, which is later used to create additional features to train the ML algorithms. Figure 3 shows the results of applying these transforms to a set of signals containing information on TWs. All faults occur at 30 ms and the TW arrives just a few microseconds later.

The second step, conceptually, is a high-pass filtering of the transformed signals, which is implemented using the DWT. In particular, it uses the reconstructed coefficients of the DWT decomposition of the input signals for a user-defined number of

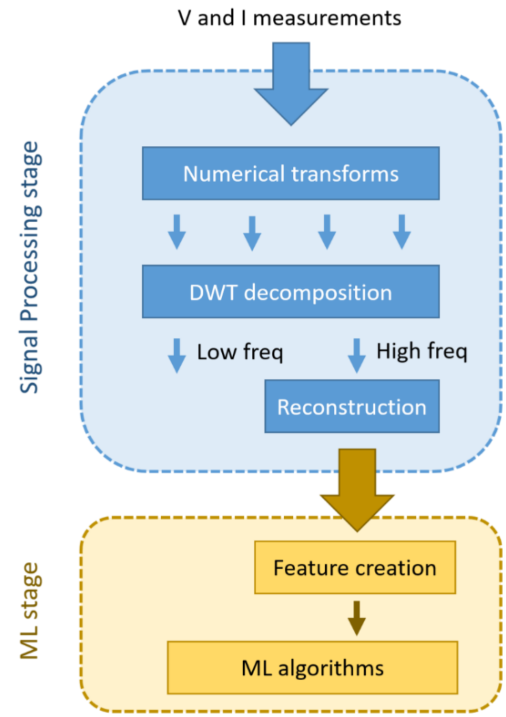

Figure 2. Workflow of the algorithm.

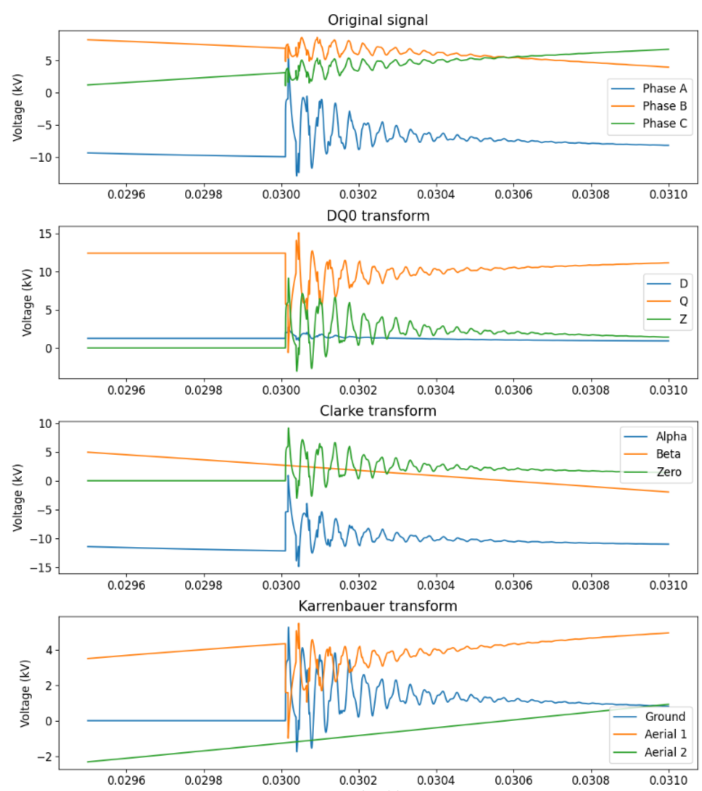

Figure 3. Applied transforms during TW arrival (SLG fault at 3100 meters, for 10 ohms resistance).

frequency bands. For this project, a window of 100 µs (50 µs before and 50 µs after the fault detection) is used for performing the DWT. This corresponds to 1000 samples. Note that this algorithm is only triggered after the fault has been detected by other means. The energy of those reconstructed signals is calculated and is employed as input for the ML models. The usage of just the medium and high-frequency components of the signals is motivated by the theory of TWs, which states that as the wave propagates across the system, there is an attenuation in both the frequency and velocity [17]. Therefore, the information retrieved from the frequency spectrum of the wave is key for both fault location and classification.

Given the importance of the DWT and the ML stage, the next two sections provide a detailed explanation of these parts of the proposed method.

A. The Discrete Wavelet Transform (DWT)

The DWT is a widely used tool for time-frequency analysis. As a more detailed explanation about how it works can be accessed in a lot of superb references - as in [7] - just a few key concepts will be included here in order to provide a more concise explanation. The method used for this paper, the Multi-Resolution Analysis (MRA), consists of splitting the input signal into several signals of different frequency ranges. This is efficiently implemented through the DWT.

The DWT algorithm resembles a pyramidal structure, known as subband coding, in which the signal goes through multiple decomposition levels. Each level contains a set of high-pass and low-pass filters (determined by the mother wavelet), which return the so-called detail and approximation coefficients, respectively. The inputs for the following levels are the approximation coefficients of the previous level. It is important to note that the output of each filter is downsampled by 2. This is explained by the fact that, in each new decomposition level, the cut-off frequency has been divided by 2 in comparison to the previous level, as well as the expected maximum frequency in the output on the high-pass filter. Therefore, following the Nyquist criteria, the number of samples representing the decomposition level (directly related to the sampling frequency) can be halved, and the frequency information would still be the same.

In summary, each decomposition level is defined by two filters and a down-sampling by 2 of the signal. This can be compactly expressed as the following set of convolutions:

$$y_{detail}[n] = \sum_n x[n] \cdot h[2k - n] \tag{1}$$

$$y_{approx}[n] = \sum_n x[n] \cdot g[2k - n] \tag{2}$$

Where $x[n]$ is the input signal for the given decomposition level, $h[n]$ and $g[n]$ are the high-pass and low-pass filter coefficients, and $y_{detail}[n]$ and $y_{approx}[n]$, the detail and approximation coefficients for that level, respectively. Note that we are just interested in the detail coefficients, which contain the frequency data in that particular frequency band.

There are features of the DWT that make this approach very convenient. First, halving the size of the coefficients implies less processing time than other WT, such as the CWT. This fast processing speed is what is needed for fault protection in power systems. Also, the DWT decomposition provides the minimal amount of information that is needed for reconstruction, which implies a more efficient usage of the resources.

The selected mother wavelet is Daubechies 7 (db7). This family is commonly used in power systems (particularly for TW applications) because their sharp shape is perfect for detecting low amplitude, short duration, and fast decay signals [18]. Different Daubechies wavelets were considered, but db7 gave slightly better numerical results. This makes sense with other research papers, as lower order wavelets (as db3) have less cut-off frequencies (so its suitability for TWs is diminished), while too large orders incur excessive computation time [19]. As such, db7 offers a perfect balance.

As a side note, the Wavelet Transforms have a trade-off between resolution on time and frequency domains. That means that higher-frequency components can be identified with accurate temporal resolution but poorer frequency resolution. For lower-frequency components, it is the other way around: higher-frequency resolution but lower temporal resolution. In the case of the DWT, this can be easily observed by taking a look at the implementation: For the used sampling rate (10 MHz), the maximum frequency that can be effectively sampled is one-half (5 MHz) following Nyquist's rule. The first decomposition level will encompass the frequency components on the range 2.5 – 5 MHz. Due to the down-sampling by 2, this first level would need 500 coefficients to accurately describe the frequencies of the signal. The next decomposition on the approximation coefficients has a frequency range from 1.25 to 2.5 MHz. Therefore, as the maximum frequency dropped in half, the required number of coefficients also halved to 250. The same process will be repeated for the next levels. As the level increases, fewer coefficients are used for the signal representation in the time domain, which could lead to inaccuracies regarding when the frequencies appeared in the wave. This decomposition can be continued up to the maximum level (where the down-sampling by 2 is no longer possible), which depends on the length of the window and the employed mother wavelet. In this project, the decomposition goes just up to level 6. The ranges of analyzed frequencies goes from 5 MHz to 78.125 kHz.

This loss of coefficients (and the physical meaning that they have) is not desired on MRA, where decomposition of the signal by frequency ranges is the goal. For this reason, the decomposition levels are reconstructed back to voltage and current signals of length 1000 samples. The reconstruction process, following the same procedure, consists of successive convolutions of the detail coefficients and the synthesis filters. The fact that the detail coefficients are the only ones that are reconstructed is what gives the similarity to a high-pass filtering (in frequency bands) of the measured signal. In Figure 4, one of those reconstructed voltage signals for some specific fault conditions over a window of 100 µs can be observed.

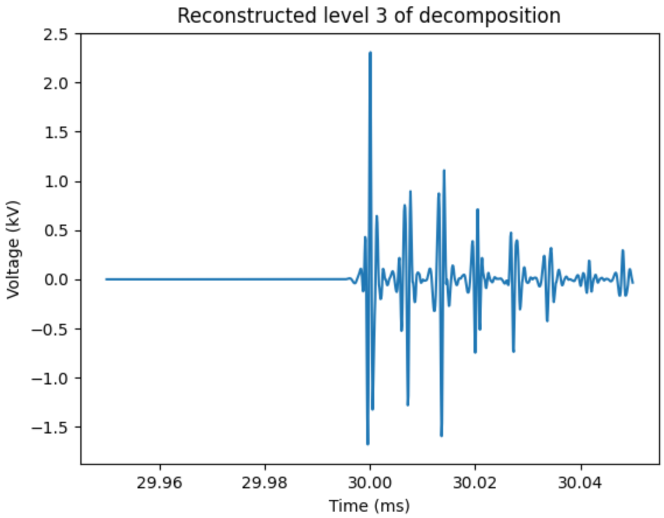

Figure 4. Reconstructed third level of decomposition (SLG fault at 25 meters, for 0.01 ohms resistance and Clarke transform and the alpha component)

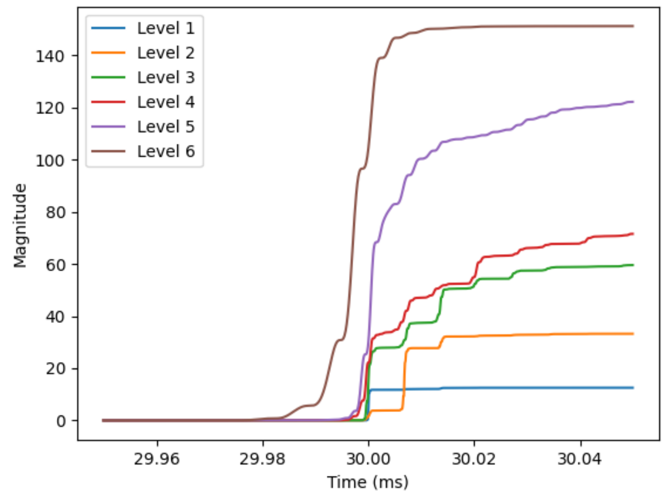

Figure 5. Energy of the reconstructed level signals (SLG fault at 25 meters, for 0.01 ohms resistance and Clarke transform and the alpha component)

The next step is to calculate the energy associated with each reconstructed level using the Parseval's Theorem, which for a discrete and finite signal x is calculated as:

$$E_k(j) = \sum_{n=0}^{j} x_n^2 \quad for\ j = 1, \ldots, N \tag{3}$$

Where k is the level of energy, j is the sample index and N is the total number of samples in the reconstructed signal. The equation is the same for both voltage and current signals. Using the Parseval Energy, the abrupt arrival of the traveling wave after the fault occurs will be more evident. During normal operation, the voltage and current values are really small (signals should not have those frequency components). When the traveling wave arrives, the magnitudes of the signals at those levels become much larger for a few instants, before coming back to almost null once the TW is attenuated enough. When the energy of the signal is calculated over the defined time window, a step-wise shape appears, as it is shown in Figure 5. Note that the first oscillations lead to a large energy increase, while the next oscillations have a relatively smaller effect, approximately reaching a "steady-state". Also, the trade-off between frequency and time resolution can be appreciated on level 6, where DWT reconstructed coefficients start to increase some microseconds before the known arrival of the TW.

The feature creation part of the method just takes the "steady-state" values for the training and testing datasets. The data is organized as a table, in which the columns are the steady-state values for both voltage and current measurements, the 3 transforms (which have 3 components each), and the 6 reconstructed levels of decomposition. Altogether, they sum up to 108 values per fault simulation. Afterward, 2730 of those faults form the training set, while the other 630 comprise the testing set. Note that those final values are highly dependent on the selected transform, the fault resistance value, the distance to the fault, and the type of fault. This is what the ML algorithms will use for classification and location. These variations for distance can be observed in Figure 6. The shadowed areas correspond to variations for several resistance values.

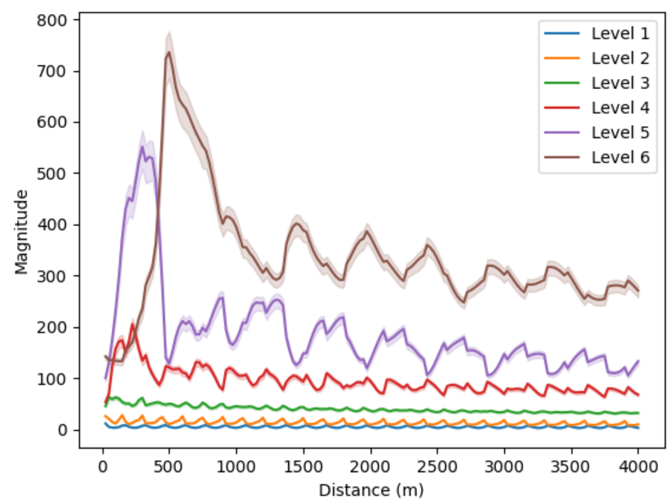

Figure 6. Steady-state energy variation due to distance (SLG fault and Clarke transform)

B. *Machine Learning/ Deep Learning approaches*

A comparison between several algorithms has been made in this paper, with special emphasis on ensemble learning, which consists of combining multiple ML models in order to create a more accurate and powerful model. Currently, it is one of the most used techniques to improve the performance. There are several techniques to implement ensemble models, such as bagging, boosting, or stacking. For comparison, a Random Forest (RF) algorithm (example of bagging), implemented on scikit-learn library, has been selected, along with Google's TensorFlow Boosted Trees (BT) estimator (Gradient Tree Boosting). Some notions of these techniques will be given later. These particular algorithms have been selected because of their excellent accuracy, computational efficiency, and scalability for a larger number of samples/ larger systems (which is where outperforms other common ML algorithms, such as SVMs). Finally, some of the state-of-the-art Machine Learning algorithms integrate different types of models into one ensemble model that takes advantage of the individual skills of each of the components (known as stacking). In this line, one model that groups the two aforementioned individual methods (RF and BT) has been implemented.

a) Random Forest (RF) estimator: This method splits the dataset on N independent datasets, randomly choosing $\sqrt{F}$ of F available features for each dataset, which is considered as a form of bagging. Then, a total of N decision trees are trained. The prediction would be either the result of voting (for classification) or the average of the individual predictions (for regression). This approach provides huge advantages, which are aligned with the goals of the project. First, RF provides variance reduction and overfitting avoidance. The method that is presented in this paper displays a strong signal processing part that reduces the features to a "small" set of final energy values with no temporal evolution. This implies a huge reduction in training data that could easily lead to overfitting. This is a problem because the testing dataset (and a real fault) would present slightly different energy values. Therefore, it is necessary to keep a good generalization capability. Training trees on random subsets of features ensures that. Also, this algorithm brings in other desirable features such as its fast training. Although the dataset has more than one hundred columns, as only a few are used at the same time, training is comparatively quicker. In addition, the trees are trained in parallel.

b) Boosted Trees (BT) estimator: This algorithm also combines multiple decision trees, but the boosting approach follows a different procedure. This time, N decision trees are trained as weak learners, which means that every one of them is able to outline a very basic rule on the dataset. The trees are trained sequentially in this case, and each added tree has to reduce the loss function. The resulting algorithm is a strong learner that can cope with the complexity of the training dataset.

c) Stacking: As explained before, this approach looks for combining the predictions (and averaging out the errors) of already-trained models using what is called a "meta-model". In this case, a linear estimator is chosen for this role. The predictions of the RF and the BT on the training set are used as input for training this method. This method is especially useful when the errors in the predictions are uncorrelated (otherwise the meta-learner would mainly have the same behavior as the individual estimators). Note that due to the good results on classification, this approach is used only for fault location.

IV. Results

A. Fault type classification

For the fault type classification task, the algorithms must be able to discern whether the measurements are of a Single-Line-to-Ground fault, Line-to-Line fault, or 3-Phase fault. Accuracy is defined as the number of correctly predicted fault types over the total number of faults. Table I summarizes the individual results of both RF and BT algorithms. As it can be observed, accuracy is close to 100% in both of them. The reason for this excellent performance is the large energy magnitude variations among each type of fault, which makes this task easy.

TABLE I. Fault Type Classification Accuracy

Classifier	Accuracy
Boosted Trees	99.84%
Random Forest	100%

B. Fault location

The results for the BT and RF regressors, along with the Stacking method can be observed in Figure 7. The average and the standard deviation of the errors are shown in Table II. Most of the faults are located with an error well below 200 m.

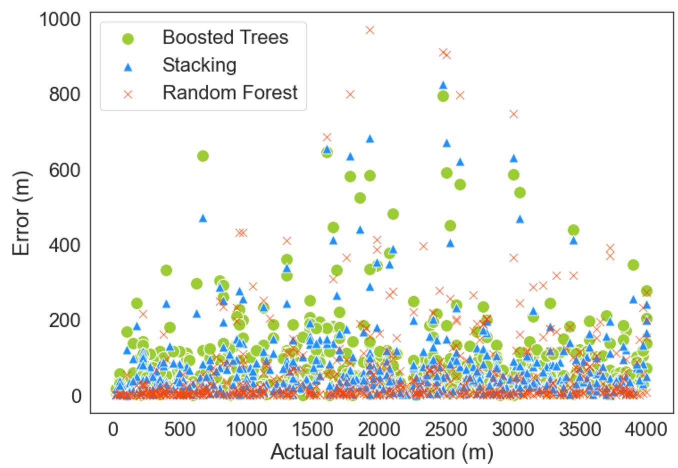

Figure 7. Prediction errors relative to fault distance

TABLE II. Mean and STD of Prediction Errors

Regressor	Mean (m)	Standard Deviation (m)
Boosted Trees	72.64	99.14
Random Forest	53.64	113.01
Stacking	**62.95**	**94.77**

As it can be seen in Figure 7, the Random Forest model tends to have lower errors in most of the cases, but there are a few outliers. However, the BT algorithm is much more consistent. The stacking of both algorithms leads to a middle point, lowering the prediction errors (respect to BT) and the number and magnitude of outliers (respect to RF). This can be observed in Figure 8.

The fact that the Stacking approach gives a little bit of extra overall accuracy in the tested fault locations is motivated by the fact that the correlation of the Boosted Trees and Random Forest algorithms errors is low and no major trend is observed, as it can be seen in Figure 9.

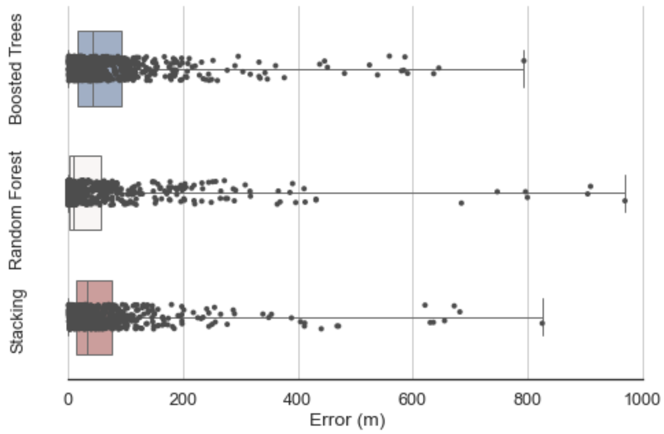

Figure 8. Distribution of prediction errors

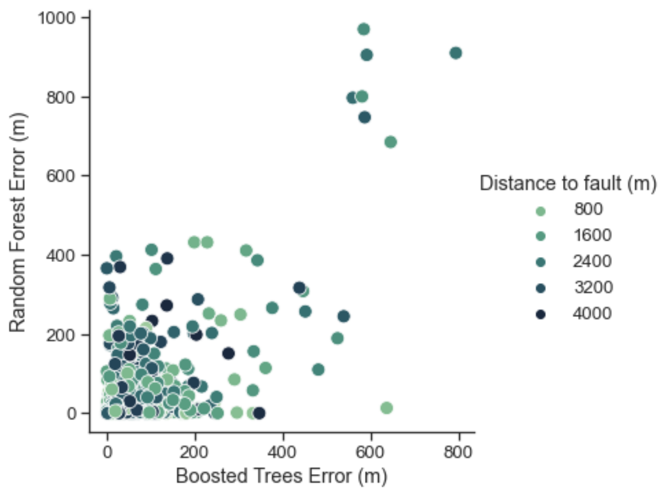

Figure 9. Correlation between BT and RF errors

A closer look at the distribution of errors (for Stacking) with respect to the fault type is shown in Figure 10. It can be observed that Line-to-Line faults are more prone to problems than other faults, while SLG and LLL have comparatively lower prediction errors and fewer outliers.

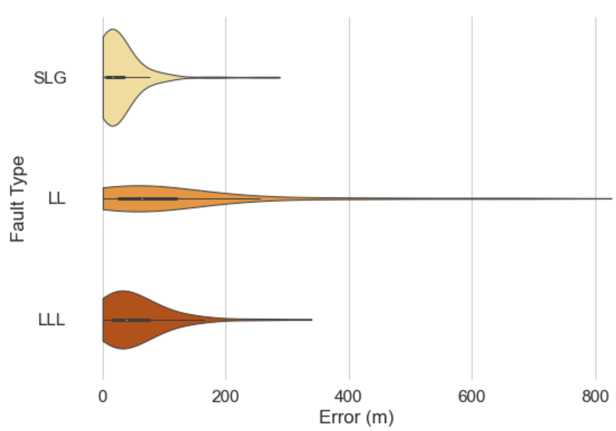

Figure 10. Distribution of location errors per fault types

Regarding the distribution of errors in relation to the fault resistance, it can be observed in Figure 11 that, first, using the Stacking approach some fault simulation predictions are driven into lower errors and, second, that the faults that still have larger errors are the ones for higher fault impedances.

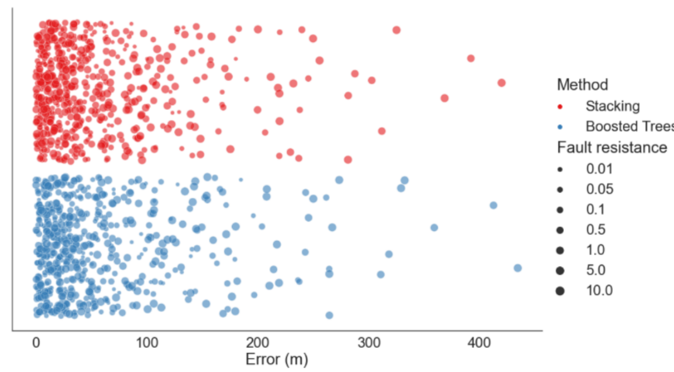

Figure 11. Distribution of location errors per fault resistance values

V. Discussion

The method presented in this paper has shown that a faster fault location and classification is achievable using much lower post-fault data. The method works on top of several assumptions, such as that the fault has already been detected (so the time needed for this purpose is not taken into account) and that the sampling rate is 10 MHz (otherwise, time windows and DWT parameters would need to be adjusted, leading to probably larger delays). Provided that, this method is an excellent trade-off between speed and accuracy. Faults can be located and classified with just 50 µs of measured data after the TW arrival.

As a comparison with other references for fault classification, such as [10], [11], [12], and [14], although they use different test systems, they need time windows that range between 2 ms up to 100 ms. The accuracies reported in those works are similar to ours. Regarding fault location, methods on other references do have lower error average and less maximum error on their respective systems, but they typically use several cycles for estimation, such as in [8], [9], and [13]. Only [16] reports a good performance going down to the limit of 200 µs of measured data for a DC microgrid. As far as the authors know, this is the first proposed method that goes that low in amount of needed measured data.

In this study, fault classification and location under noisy conditions has not been addressed and is left for future work. Due to the decomposition of the signals in several frequency bands, it is expected that moderate noise will not have a significant impact on all levels, allowing a fairly accurate prediction. Also, the next step would be to check the performance of this method on a larger system.

VI. Conclusion

This paper presents a new addition to Discrete Wavelet Transform-based fault classification and location approaches. First, the signal is decomposed into several frequency bands, then it is reconstructed before its energy is calculated using Parseval's Theorem. Second, each energy level is summarized into one value taken after 50 μs from the fault detection. This leads to a huge reduction of data for the Machine Learning prediction stage. A few other considerations have been explained in this paper, such as the convenience of high-sampling rates for achieving lower time windows and the accuracy of ensemble techniques (usually forgotten on power systems). Finally, it has been discussed how this method gives a great performance given the considered time window, making fault location and classification in AC distribution systems possible using just 50 μs of measured data after the traveling wave arrival.

Acknowledgment

This material is based upon work supported by the Laboratory Directed Research and Development program at Sandia National Laboratories and the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number 36533. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

[1] F. Wilches-Bernal. A. Bidram, M. J. Reno, J. Hernandez-Alvidrez, P. Barba, B. Reimer, C. Carr, O. Lavrova and S. Ranade "A Survey of Signal Processing Methods to Enable Protection Schemes in Electric Power Systems" IEEE Access, 2021.

[2] F. Wilches-Bernal, M. J. Reno, J. Hernandez-Alvidrez, "A Dynamic Mode Decomposition Scheme to Analyze Power Quality Events," IEEE Access, 2021.

[3] E. O. Schweitzer, B. Kasztenny, A. Guzmán, V. Skendzic and M. V. Mynam, "Speed of line protection - can we break free of phasor limitations?," 2015 68th Annual Conference for Protective Relay Engineers, 2015, pp. 448-461.

[4] M. S. Mamiş and M. Arkan, "FFT based fault location algorithm for transmission lines," 2011 7th International Conference on Electrical and Electronics Engineering (ELECO), 2011, pp. I-71-I-75.

[5] K. Lout and R. K. Aggarwal, "A feedforward Artificial Neural Network approach to fault classification and location on a 132kV transmission line using current signals only," 2012 47th International Universities Power Engineering Conference (UPEC), 2012, pp. 1-6.

[6] M. Jiménez Aparicio, S. Grijalva, and M. J. Reno, "Fast Fault Location Method for a Distribution System with High Penetration of PV," Hawaii International Conference on System Sciences (HICSS), 2021.

[7] R. Polikar, Multiresolution Analysis: The Discrete Wavelet Transform. Available at: https://ccrma.stanford.edu/~unjung/mylec/WTpart4.html

[8] P. Chiradeja and C. Pothisarn, "Identification of the fault location for three-terminal transmission lines using discrete wavelet transforms," 2009 Transmission & Distribution Conference & Exposition: Asia and Pacific, 2009, pp. 1-4.

[9] V. Vanitha, D. Kavitha, R. Resmi and M. S. Pranav, "Fault classification and location in Three Phase Transmission Lines using Discrete Wavelet Transform," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019, pp. 1-4.

[10] X. G. Magagula, Y. Hamam, J. A. Jordaan and A. A. Yusuff, "Fault detection and classification method using DWT and SVM in a power distribution network," 2017 IEEE PES PowerAfrica, 2017, pp. 1-6.

[11] A. Ngaopitakkul, C. Pothisarn, S. Bunjongjit and B. Suechoey, "DWT and RBF neural networks algorithm for identifying the fault types in underground cable," TENCON 2011 - 2011 IEEE Region 10 Conference, 2011, pp. 1379-1382.

[12] A. Yadav and A. Swetapadma, "Combined DWT and Naive Bayes based fault classifier for protection of double circuit transmission line," International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), 2014, pp. 1-6.

[13] M. J. B. Reddy, D. V. Rajesh, P. Gopakumar and D. K. Mohanta, "Smart Fault Location for Smart Grid Operation Using RTUs and Computational Intelligence Techniques," in IEEE Systems Journal, vol. 8, no. 4, pp. 1260-1271, Dec. 2014.

[14] A. Abdullah, "Ultrafast Transmission Line Fault Detection Using a DWT-Based ANN," in IEEE Transactions on Industry Applications, vol. 54, no. 2, pp. 1182-1193, March-April 2018.

[15] D. K. J. S. Jayamaha, N. W. A. Lidula and A. D. Rajapakse, "Wavelet Based Artificial Neural Networks for Detection and Classification of DC Microgrid Faults," 2019 IEEE Power & Energy Society General Meeting (PESGM), 2019, pp. 1-5.

[16] R. Montoya, B. Poudel, A. Bidram, M. J. Reno, "DC Microgrid Fault Detection Using Multiresolution Analysis of Traveling Waves", Submitted to International Journal of Electrical Power and Energy Systems.

[17] H. Jia, "An Improved Traveling-Wave-Based Fault Location Method with Compensating the Dispersion Effect of Traveling Wave in Wavelet Domain," Mathematical Problems in Engineering, 2017.

[18] I. Baqui, I. Zamora, J. Mazn and G. Buigues, "High impedance fault detection methodology using wavelet transform and artificial neural networks", Electric Power Systems Research, vol. 81, no. 7, pp. 1325-1333, July 2011.

[19] Z. Wang, S. McConnell, R. S. Balog and J. Johnson, "Arc fault signal detection - Fourier transformation vs. wavelet decomposition techniques using synthesized data," 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), 2014, pp. 3239-3244.