# Exploring the use of Shapelets in Traveling Wave based Fault Detection in Distribution Systems

Milan Biswal[1], Shubhasmita Pati[2], Satish J. Ranade[2], Olga Lavrova[2], Matthew J. Reno[3]
[1]*Computer Science Department*
[2]*Klipsch School of Electrical & Computer Engineering*
*New Mexico State University, Las Cruces, NM 88003, USA*
[3]*Sandia National Laboratories, Albuquerque, NM, 87185, USA*
{*milanb, shubha, sranade, olavrova*}*@nmsu.edu, mjreno@sandia.gov*

***Abstract*—The application of traveling wave principles for fault detection in distribution systems is challenging because of multiple reflections from the laterals and other lumped elements, particularly when we consider communication-free applications. We propose and explore the use of Shapelets to characterize fault signatures and a data-driven machine learning model to accurately classify the faults based on their distance. Studies of a simple 5-bus system suggest that the use of Shapelets for detecting faults is promising. The application to practical three-phase distribution feeders is the subject of continuing research.**

***Index Terms*—traveling Waves, Distribution Systems, Shapelets, Fault Classification, Machine learning**

## I. Introduction

The protection of distribution systems with high penetrations of inverter-based distributed energy resources (DERs) is a major challenge for utilities. The traditional over current relays may not detect faults due to the current limiting circuitry in the DERs that could limit the fault current to $\approx 1.2\ pu$ [1]. In addition, the faster control mechanisms of the DERs and low inertia systems may require ultra-fast tripping protection schemes (FTPS).

Faults generate traveling waves (TW) which travel through the transmission line from the fault to the opposite ends. These waves travel at a speed near to the velocity of light and are reflected at junctions of the distribution lines. Observation of wave arrival times at a specific location can be used to infer the occurrence, location and type of fault. This phenomenon is transitory and requires the analysis to be performed on small time window of a few microseconds or milliseconds duration. The use of fault-generated TW as a means for ultra-high-speed fault detection and location has been a subject of extensive research, with implementations reported in the early seventies [2]. Further research [3, 4], and advancements in microprocessors, led to practical applications to EHV Transmission [4]. The application of TW relays to a long, untapped, section of EHV, or HVDC lines are a mature area [5]. Both detection and location are important. We suspect that a lower resolution in location may be acceptable in relaying applications; high resolution is desired, post-fault, for needed repair and maintenance operations. Distribution feeders represent a more challenging application because of short travel times, tapped laterals and complex terminations. This means that waves reflected from a number of laterals result in an overlap or *'clutter'*. Thus, it becomes difficult to isolate waves reflected from different nodes, which in turn makes detection as well as location more difficult. traveling wave based protection in with tapped (multiterminal) lines has been addressed in literature e.g.,[4, 5, 6]. The basic ideas rely on a knowledge of the sequence of wave arrivals followed by some form of pattern matching with the actual waves during a fault. Our exploration draws upon these ideas. One of the ways to address this challenge is to use sophisticated signal processing and machine learning (ML) algorithms to extract relevant information from the TW. Reference [7] provides a rather comprehensive survey of recent research in this area with an emphasis on the increasing penetration of distributed resources.

## II. Problem Formulation

### A. Traveling waves in distribution systems

We consider the feeder shown in Fig. 1, taken to be a single phase feeder for our initial study. The relay R at bus 1 is to detect faults on the primary section $M2$, between buses 2 and 5, using TW principles. The desired scheme is single-ended (communication free) The primary feeder has two laterals separated by 2 km. The upstream and downstream laterals have lengths of 1 km and 2 km, respectively. All lines have identical characteristics with a travel time of $5.5$ µS/km Suppose a fault occurs at $X$ on $M2$ a distance of 1 km from the downstream lateral. The Lattice diagram [8] in Fig. 1 shows the pattern of TW due to the fault. The solid and dotted lines in the lattice diagram represent the forward and the backward waves, respectively.

The waves on up-stream laterals are shown in green and those from the downstream laterals are shown in blue. The reflections from the fault arriving at the relay through the primary (main) feeder are shown in red. The first arrival and reflection (A) from the fault through the main feeder acts as the reference. Absent laterals and load taps, a single-ended scheme can detect and locate the fault by 1) detecting the forward and backward waves labeled 'A' and 'B', respectively

978-1-6654-7902-8/22/$31.00 ©2022 IEEE

2022 IEEE Texas Power and Energy Conference (TPEC) | 978-1-6654-7902-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/TPEC54980.2022.9750728

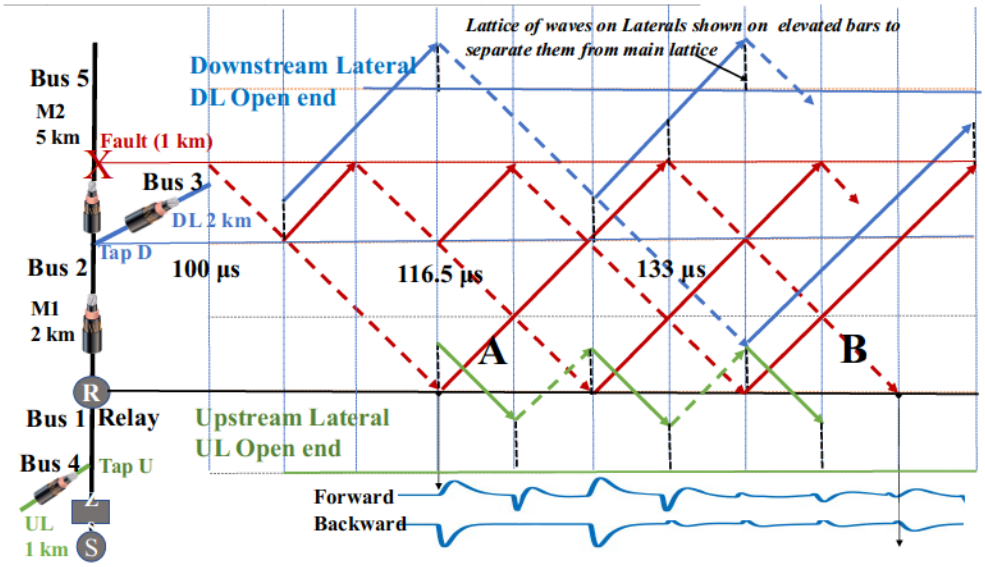



Fig. 1: Lattice diagram illustrating the TW phenomena in the distribution systems.

and 2) estimating the time between the two. It can be observed that before wave B from the main feeder arrives, there are several wave arrivals from the fault as well as the laterals. The shapes of the waves will depend on distortion (attenuation, dispersion), the coefficient of reflection at various junctions and the shape of the incident wave, and will fade with time. For a distribution feeder the Relay R must account for these intermediate waves [5, 6, 7] to detect and locate the fault by observing the arrival times with respect to the reference.

The system in Fig. 1 was simulated using Matlab/Simulink/Simscape. Details are provided in Section IV. The forward(top) and backward (bottom) waves seen by the relay are shown on the bottom right of Fig. 1, and are keyed to the lattice diagram. The TW relay must detect periodic arrival of reflected waveforms. It should be noted that the periodicity will be different for main feeder and each of the laterals (with the exception of two or more laterals of same impedance at a fixed location). If some of these reflections have very close time-of-arrival at the relay then constructive or destructive interference may occur. This will result in a clutter and introduce ambiguity. Other sources of ambiguity include transformers (loads), capacitor banks, DERS, impedance and arcing in the fault, and the measurement noise. TWs similar to those due to a fault can also be produced by switching operations. This 'clutter' makes it extremely challenging to estimate the location of the fault by analyzing, particularly if high location resolution is desired in real-time. The relay must isolate waves or wave arrivals of interest to detect faults.

### B. Hypothesis

*Shapelets* are discriminative sub-sequences of time series that best characterize inter-class differences and have been successful in several classification tasks [9, 10, 11]. Motivated by this we formulate two hypotheses:

1) *Shapelets Discovery:* If we discover a set of discriminative sub-sequence/Shapelets and treat them as basis functions for extracting information/features from the TW signals, can these features be used to classify and find the accurate location of faults?
2) *The mother-Shapelet:* Can we design an analytical *Shapelet* that can be interpreted as a global template (similar to the mother wavelet or basis function)? Do the dilated and translated versions of this template capture the fault signatures?

We propose and explore the use the Shapelets to implicitly 'isolate' wave arrivals and then to extract the most important features from the TW reflection clutter. These features are fed to an ML classifier to estimate the approximate location of the fault (categorized as different classes).

## III. Methods

TW relays are intended to detect high-frequency electromagnetic transient signatures corresponding to a fault. A TW relay uses a current transformer (CT) and/or a potential transformer (PT) to measure voltage ($v(t)$) and current ($i(t)$) waveforms on a line and digitizes them with a high sampling rate.

### A. Test System

Modeling requirements to recreate and study traveling waves in distribution systems are discussed in [12]. The report emphasizes the need for additional modeling, simulation, and experiments to fully understand traveling wave phenomena in distribution systems and to use it for fault detection and identification. Our initial study uses a simplified simulation setup is shown in Fig. 6. We have considered a single-phase distribution system with an ideal ac source connected to a feeder with two laterals (same as described in Section II). The feeder and laterals are modeled as distributed parameter

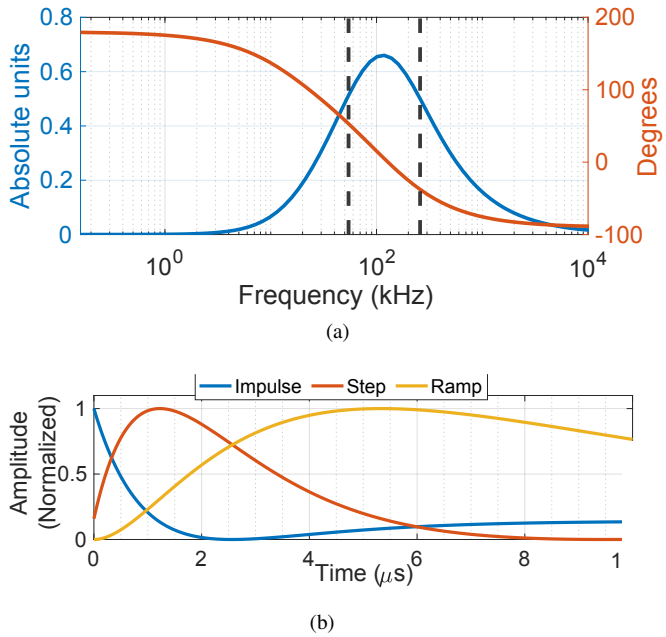



Fig. 2: (a) Magnitude and Phase response, (b) impulse, step, and ramp response of the band pass filter.

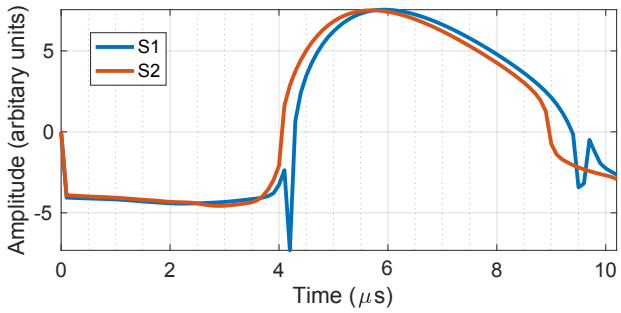



Fig. 3: Shapelets discovered from the dataset with the objective to maximally discriminate faults from steady-state.

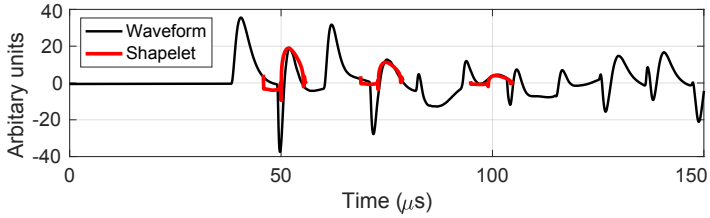



Fig. 4: Mother Shapelet at different translations and scales superimposed over an arbitrary TW waveform.

lines without parameter frequency dependence. We created line-to-ground faults at various distances from the relay with different fault resistances. For the experiments in this paper, we consider a sampling rate of 10 MHz which results in a temporal resolution of 0.1 µs. The very high sampling rate is chosen to satisfy Nyquist criterion.

### *B. Pre-processing*

The voltage and current signals were decomposed into the forward and backward waves based on the principles of Telegraphers equation [8]. These waveforms were processed through an analog band pass filter with a center frequency of 118.5 kHz. The magnitude, phase, impulse, step, and ramp response of the filter are shown in Fig. 2(a), and (b). For a sinusoidal 200 kHz the filter input the filter delay is approximately $0.4\ \mu S$ which as compared to the wave travel time of $5.5\ \mu S/km$. As such we anticipate errors and misclassifications due to this filter. The filter has an almost linear phase response within the pass band (shown as vertical dotted lines). For the experiments in this paper, we utilized the filtered forward waves.

### *C. Shapelet Discovery*

We use the algorithm in [10] to discover a set of 10 Shapelets, with the objective of distinguishing faults from steady-state waveforms. The algorithm uses signal subsequences to propose test Shapelets (similar to wavelets but with no admissibility criteria) and ranks them based on the Shapelet signal 'distance' (correlation) and the ability to classify. This off-line process is computationally intensive. The hyper parameters for Shapelet discovery algorithm were chosen as: learning rate = 0.1, weight regularizer = 0.1, number of iterations = 46. The Shapelet length was chosen as 102 samples based on empirical observations and the algorithm suggested in [10]. The discovered Shapelets were ranked based on their discriminating ability. This essentially means that the Shapelet having the highest rank possibly captures the most frequently occurring transient portion of the fault waveform. In Fig 3, we show the two best Shapelets discovered from a fault dataset (explained in Section IV-A). The Shapelet (S1) (shown in blue), was considered as the mother Shapelet for the simulations in this paper. The similarity of this shape to the step and ramp response of the pre-processing band pass filter shown in Fig. 2, or a combination of these, suggests that the Shapelet in a TW could be tied to the pre-processing filter response. The Shapelet also reflects partial contributions from reflections at the reactive impedance of the source. Fig 4 shows the magnitude scaled mother Shapelet superimposed over an arbitrarily chosen TW waveform, at multiple locations. The similarity of the scaled and translated mother Shapelet to a segment of the TW waveform implies that it can act as a good basis for teasing out constituent waves.

### *D. Feature Extraction*

We used the Shapelet with the highest rank as the mother Shapelet and computed its correlation with the composite TW waveform observed at the relay. The mother Shapelet was essentially translated in time and its Euclidian distance from the composite waveform is referred to as the correlation coefficient. The correlation coefficients for an arbitrarily chosen composite waveform are shown in Fig. 5. The peaks represent the highest correlation points in one traversal of the TW

wave and alternatively imply the time of arrival. The valley also imply the time of arrival, however the negative values indicate phase reversal upon reflection. The local peaks can be identified automatically by finding points that are greater in magnitude from their adjoining (immediate before and after) points. However, such an approach is sensitive to noise and based on the TW clutter there could be several local peaks. To solve this problem, we specify two peak selection criteria. The first criterion introduces a constraint on the relative peak prominence to be less than $0.1$. Peak prominence is essentially the height of the peak from the mean value of the signal. As the correlation coefficients are normalized in the range of $\{-1, +1\}$, the choice of peak prominence of $0.1$ implies that every identified peak must correspond to at least $10\%$ matching with the chosen Shapelet basis. The second criterion for peak identification is based on the relative location of consecutive peaks. We specify this as a minimum separation

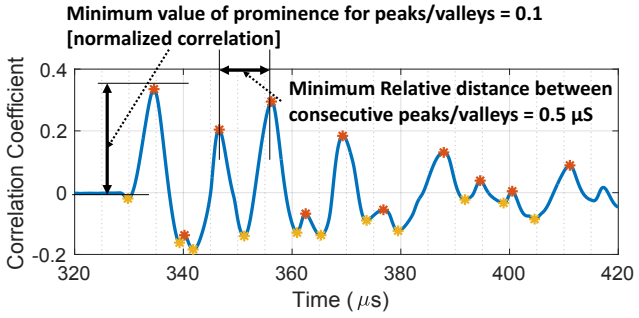



Fig. 5: Constraints to identify peaks and valleys.

We extracted the following set of features based on the location and magnitude of the peaks and valleys.

- **Feature 1**: Magnitude of first 10 correlation peaks excluding the first peak
- **Feature 2** : Distance between first 10 pairs of consecutive correlation peak locations/ time lags
- **Feature 3** : Magnitude of first 10 correlation valleys excluding the first valleys
- **Feature 4** : Distance between first 10 pairs of consecutive correlation valley locations/ time lags

### *E. Building a classifier*

Instead of choosing a non-linear classifier such as a neural network, we first evaluated a linear classifier. The linear classifier works when the decision boundary separating different classes can be defined by lines, planes or hyper-planes [13]. If the extracted features have sufficient discriminating ability then a linear combination of the features may transform the data onto a linearly separable plane. Motivated by the intuition that the Shapelet features can infer about the fault location, we used the multi-class linear discriminant analysis (LDA) classifier. We assume that each class has the same covariance but with different means.

In LDA, each class $(Y)$ generates data $(X)$ using a multivariate normal distribution. In our case, the data $X$ essentially is the extracted set of features. The model assumes $X$ has a Gaussian mixture distribution. The classifier predicts such that the expected classification cost in (1) is minimized.

$$\hat{y} = \arg \min_{y=1,2--K} \sum_{k=1}^{K} P(\hat{k/x}) C(y/k) \tag{1}$$

Where, $\hat{y}$ is the predicted classification. $K$ is the number of classes. $P(\hat{k/x})$ is the posterior probability of class $k$ for observation $x$. $C(y/k)$ is the cost of classifying an observation as $y$ when its true class is $k$, defined in (2).

$$C(y/k) = \begin{cases} 1, & \text{if } y \ = k \\ 0, & \text{if} y = k \end{cases} \tag{2}$$

### *F. Classifying a pattern*

The chosen mother Shapelet and pre-trained LDA classification model are stored/deployed at the TW relay module. Classification requires a computation of the correlation between the Shapelet and the incoming signal, at each successive lag of the Shapelet with respect to the signal. This step would have to be performed in real time, is not very computationally intensive since the length (span) of the Shapelet is short (102 samples in our experiments). The peaks and valleys are detected from the correlation coefficients and the chosen set of features are extracted. These features are then fed as input to the trained LDA model to predict the class to which the fault belongs.

## IV. Simulations & Results

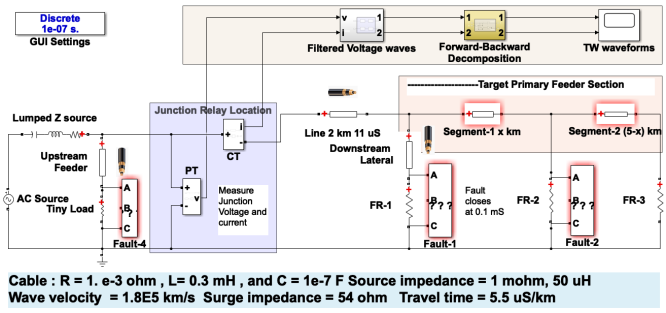



Fig. 6: Experimental setup.

### *A. Dataset*

We created a database with a comprehensive set of fault waveforms by varying the fault resistance and fault location parameters in the simple simulation set up. We created a set of total $3600$ patterns and divided them into $9$ classes based on the location of fault, summarized in Table I. The faults on the main feeder were categorized into $8$ segments of $500$ meters each, corresponding to $8$ classes, making the distance resolution to be $500$ meters. The fault locations were uniformly distributed within each segment/class. At each fault location,

the fault resistance was varied randomly between $0.1$ to $2\ \Omega$ with an uniform distribution, to create an exhaustive set of patterns.

TABLE I: Parametric variation of the data set.

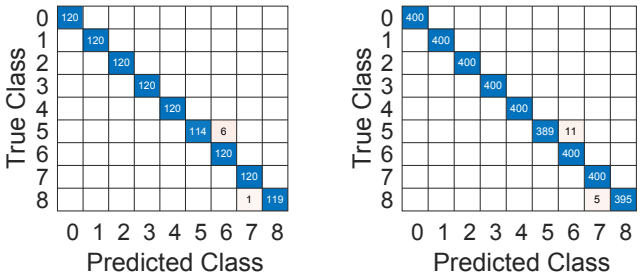
| **Class Labels** | **Fault Location ($x$ km)** | **Fault Resistance ($R\ \Omega$)** | **Number of patterns** |
|---|---|---|---|
| Class-0 | Fault at up-stream and down-stream laterals | $0.1 \leq R < 2$ | 400 |
| Class-1 | $1 \leq x < 1.5$ | $0.1 \leq R < 2$ | 400 |
| Class-2 | $1.5 \leq x < 2$ | $0.1 \leq R < 2$ | 400 |
| Class-3 | $2 \leq x < 2.5$ | $0.1 \leq R < 2$ | 400 |
| Class-4 | $2.5 \leq x < 3$ | $0.1 \leq R < 2$ | 400 |
| Class-5 | $3 \leq x < 3.5$ | $0.1 \leq R < 2$ | 400 |
| Class-6 | $3.5 \leq x < 4$ | $0.1 \leq R < 2$ | 400 |
| Class-7 | $4 \leq x < 4.5$ | $0.1 \leq R < 2$ | 400 |
| Class-8 | $4.5 \leq x < 5$ | $0.1 \leq R < 2$ | 400 |

### *B. Results*

We performed 10-fold cross-validation to evaluate the performance of the LDA classifier. By performing cross-validation, we use all our patterns in the dataset, both for training and for testing while evaluating the classifier on examples it has *never* seen before. For the dataset without noise, the mean classification accuracy of all the 10-folds was $99.56\%$ with a standard deviation of $0.42\%$. The small value of the standard deviation indicates that the classifier is consistent. This implies that by training the classifier on patterns of the data set chosen randomly, and deploying it will lead to similar performance.

To further analyze the classification performance in terms of the confusion matrix, we divided the data set randomly into $70\%$ training data (280 patterns in each class) and $30\%$ (120 patterns in each class) testing data, chosen randomly. The classifier was trained with $70\%$ training data and the resulting confusion matrix with $30\%$ test data is shown in Fig. 7(a). In this case, the classification accuracy was found to be $99.35\%$. It can be observed that for Class-5, six patterns are wrongly classified as the neighboring Class-6. Similarly, 1 patterns from Class-8 are wrongly classified as Class-7. The mis-classified patterns were found to lie near the decision boundary of two neighboring classes. For instance, the 6 mis-classified patterns in Class-5 correspond to the faults occurring at locations between $3.45$ to $3.5\ km$ (within $50$ meters from the starting class boundary for Class-6). The confusion matrix for the entire data set is shown in Fig. 7(b), to identify the classes that are venerable when the ML model is deployed. In our case, it leads to the similar conclusion we derived from the 70-30 train-test data, i.e, few patterns at the boundary of Class-5 and 8 are missclassified to their adjoining class.

The classifier performs quite well for the simulated data. However, the field data will have measurement noise and for low intensity-high frequency TW waves, this could have severe impact. To assess the effect of noise we added white Gaussian noise to the filtered waveform before extracting the features.




(a) (b)

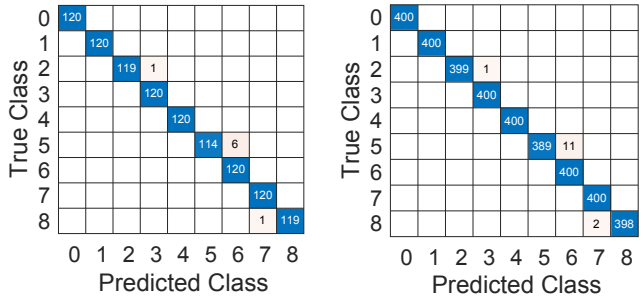
Fig. 7: Confusion Matrix with (a) 30% test data and 70% training data, and (b) for the entire data set, without noise.

TABLE II: Classification accuracy with different SNRs.

| **SNR (dB)** | 45 | 40 | 35 | 30 | 25 | 20 |
|---|---|---|---|---|---|---|
| **Accuracy (%) Mean** | 99.55 | 99.55 | 99.53 | 99.55 | 99.33 | 92.47 |
| **Std. Dev.** | 0.47 | 0.47 | 0.49 | 0.46 | 0.49 | 1.22 |

As we do not know the actual measurement noise characteristics for the TW, white Gaussian distribution seemed a reasonable assumption [14]. The mean and standard deviations of 10-fold cross-validation accuracies for different signal-to-noise ratio (SNR) conditions are presented in Table II. The classification performances were minimally affected (stays at a mean of $\approx 99.5\%$ with $\approx \pm 0.5\%$ standard deviation) for SNR of up $25\ dB$. However, for $20\ dB$ SNR, the accuracy dropped to $92.47\%$ with an increase in standard deviation to $\pm 1.22\%$. Considering that, we are introducing noise after pre-processing with the band pass filter, $20\ dB$ SNR is a very extreme test. Therefore, the proposed method can be considered as robust against measurement noise.

In Fig. 8(a), the confusion matrix computed for the 30% test data at $30\ dB$ SNR is shown. The model was trained with the rest $70\%$ training data. Comparing it with the case without noise (Fig. 7), only one additional pattern in Class-2 got mis-




(a) (b)

Fig. 8: Confusion Matrix with (a) 30% test data and 70% training data, and (b) for the entire data set, at 30 dB SNR.

classified into neighboring Class-3. It further demonstrates the robustness to additive noise. The confusion matrix for the entire dataset shown in Fig. 8(a) also reinforces the same conclusion that few patterns at the boundary of Class-2, 5, and 8 are at the risk of being missclassified to their adjoining class, at 30 dB SNR.

## V. Discussions and Conclusion

This paper proposes a relatively simple Shapelet based approach to classify faults in the distribution systems through traveling wave principles. The potential for using Shapelets was explored using a simple system. In contrast to choosing a pre-defined basis function, we extracted a set of discriminating sub-sequences (Shapelets) from a dataset with different fault characteristics. One of the extracted Shapelets was considered as the template/basis (mother Shapelet) to extract features from the TW clutter. We observed that the extracted features result in a linear decision boundary, and with discriminant analysis the faults in the distribution system can be classified with reasonable accuracy ($> 99\%$ in the absence of noise). We also found the proposed Shapelet based method to be robust against measurement noise retaining a classification accuracy $> 90\%$. These preliminary results indicates the promising ability of Shapelets in classifying distribution system faults. However, we will perform a detailed comparative evaluation (robustness and computational analysis) of this approach with other established methods [4, 5, 7, 9] in continuing work. We will focus on classifying different types of faults in practical 3-phase distribution systems and on more precise estimation of fault distance, and examine the computational complexity in the context of implementation.

## VI. Acknowledgements

This work was partially supported by the NSF Grants OIA-1757207 (NM EPSCoR), HRD-1345232, HRD-1914635 and funding from Sandia National Laboratories Campus Executive (CE) LDRD Supplemental Project 20-0656, and award number GR0006770 (SHAZAM), DOE NNSA MSIPP STEP2NLs; DOE Financial Assistance Award DE-NA0003983 and funding from the Electric Utility Management Program (EUMP) at the New Mexico State University. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and 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. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

The authors would like to thank Miguel Jimenez Aparicio of Sandia National Laboratories for the feedback on initial version of the manuscript.

## References

[1] IEEE PES Industry Technical Support Leadership Committee, "Impact of IEEE 1547 standard on smart inverters and the applications in power systems," 2020.

[2] M. Chamia and S. Liberman, "Ultra high speed relay for EHV/UHV transmission lines – development, design and application," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 6, pp. 2104–2116, 1978.

[3] P. A. Crossley and P. G. McLaren, "Distance protection based on traveling waves," *IEEE Power Engineering Review*, vol. PER-3, no. 9, pp. 30–31, 1983.

[4] S. Marx, Y. Tong, and M. Mynam, "Traveling-wave fault locating for multiterminal and hybrid transmission lines," in *Proceedings of the 45th Annual Western Protective Relay Conference*, 2018, pp. 16–18.

[5] R. Khuzyashev, I. Kuzmin, S. Tukaev, L. Tukhvatullin, and E. Stepanova, "Traveling-wave fault location algorithms in hybrid multi-terminal networks with a tree-like structure," in *E3S Web of Conferences*, vol. 124. EDP Sciences, 2019, p. 01012.

[6] J. Wang and Y. Zhang, "Traveling wave propagation characteristic-based lcc-mmc hybrid hvdc transmission line fault location method," *IEEE Transactions on Power Delivery*, pp. 1–1, 2021.

[7] C. Y. Evrenosoglu and A. Abur, "Fault location in distribution systems with distributed generation," in *Proc. Power Systems Computation Conference*, 2005.

[8] J. D. Glover, M. S. Sarma, and T. Overbye, *Power system analysis & design, SI version*. Cengage Learning, 2012.

[9] L. Ye and E. Keogh, "Time series shapelets: a new primitive for data mining," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 947–956.

[10] J. Grabocka, N. Schilling, M. Wistuba, and L. Schmidt-Thieme, "Learning time-series shapelets," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 392–401.

[11] M. Biswal, Y. Hao, P. Chen, S. Brahma, H. Cao, and P. De Leon, "Signal features for classification of power system disturbances using PMU data," in *2016 Power Systems Computation Conference (PSCC)*, 2016, pp. 1–7.

[12] K. Prabakar, A. Singh, M. Reynolds, M. Lunacek, L. Monzon, Y. N. Velaga, J. Maack, S. Tiwari, J. Roy, C. Tombari *et al.*, "Use of traveling wave signatures in medium-voltage distribution systems for fault detection and location," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2021.

[13] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[14] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, "Characterizing and quantifying noise in PMU data," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5.