

Evaluating the use of Shapelets in Traveling Wave based Fault Detection and Classification in Distribution Systems

Shubhasmita Pati, *Student Member, IEEE*, Milan Biswal, *Member, IEEE*, Satish J. Ranade, *Fellow, IEEE*, Olga Lavrova, *Member, IEEE*, and Mathew J. Reno, *Senior Member, IEEE*

Abstract—The application of traveling wave (TW) 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. In [1], we proposed the use of Shapelets to characterize fault signatures as we move towards a more realistic data-driven solution. The Shapelet can be correlated with the physical characteristics of TW generation process, in contrast to choosing a predefined basis such as Wavelets. This paper is an extension of this work to comprehensively evaluate Shapelets for identifying the fault types, in addition to their locations in single-phase, three-phase, and IEEE 13-bus distribution systems. We also propose using the voltage waveforms instead of the forward waves, which eliminates the need for a CT. The proposed method can identify fault locations within a local region with $\approx 99\%$ accuracy and classify the fault types with an accuracy of over 90%.

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 $\approx 1.2 pu$ [2]. In addition, the faster control mechanisms of the DERs and low inertia systems may require ultra-fast tripping protection schemes (FTPS). Promising approaches include Traveling Wave (TW) based protection.

In this extended version of [1], we present a comprehensive evaluation of the use of Shapelets as a tool for extracting traveling wave features for protective relaying in *distribution systems*. The scope selected for the investigation assumes a communication-free, single-ended, scheme for a main feeder section with laterals. The present paper extends the work in [1] and presents comprehensive studies and evaluations using a simple 5-bus three-phase system and the IEEE 13 bus distribution test system. The development of an actual relay paradigm will be reported in the future.

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 [3]. Observation of wave arrival times of TW 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 a small time window of a few microseconds or milliseconds duration. Further research [4, 5], and advancements in microprocessors, led to practical applications for EHV Transmission [5]. The application of TW relays to a long, untapped, section of EHV, or HVDC lines is a mature area [6]. In [7], a dynamic state estimation technique based on the TW principles is proposed for double-ended transmission lines. It requires GPS-synchronized measurements at both ends of the line and a high-fidelity model of the protected line which is not suitable for the distribution systems and a communication-free approach is desired. Therefore, TW protection is a subject of active research for the distribution systems [8, 9, 10, 11, 12, 13].

There are two primary issues that need attention in TW-based protection for distribution systems. First, both the detection and location of the fault are important. We suspect that a lower resolution in location may be acceptable in relaying applications; the high resolution is desired, post-fault, for needed repair and maintenance operations. Second, the distribution feeders represent a more challenging application because of short travel times, tapped laterals, and complex terminations. This means that TW reflected from a number of laterals results in an overlap or ‘clutter’. Fig. 1 shows the TWs extracted from the measured phase-A voltages in the case of a 3-phase to ground fault. The system used to generate it is a 3-phase 5-bus test system with the laterals shown in Fig. 2. For the case without laterals, the upstream and downstream laterals were simply removed. In the absence of lateral TWs can be observed to have a nearly periodic pattern making it easier to detect wave arrival times. For the case of laterals, it can be observed that the TW waveform becomes more complex and cluttered. Thus, it becomes difficult to isolate waves reflected from different nodes, which in turn makes detection as well as location determination more difficult.

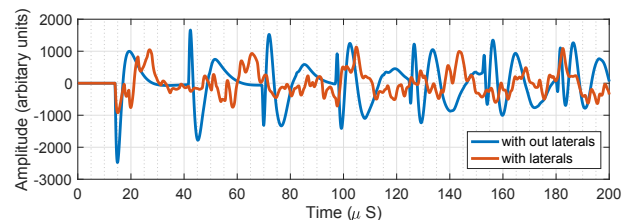


Fig. 1: *Clutter* due to the presence of laterals shown for a TW waveform in case of a 3-phase to ground fault.

Section II reviews the recent advances in TW-based relaying

S. Pati, M. Biswal, S. Ranade, and O. Lavrova are with ¹ New Mexico State University, Las Cruces NM 88003 USA (e-mail: shubha, milanb, srnade, olavrova}@nmsu.edu). M. J. Reno is with the ²Sandia National Laboratories, Albuquerque, NM, 87185, USA (e-mail: mjreno@sandia.gov).

principles applied for distribution systems protection. The gaps in the literature are identified and the use of Shapelets as a tool for TW feature extraction is summarized. Section III describes the proposed approach to building an algorithm for fault type and location classification. The simulation and results are discussed in Section IV. The comprehensive performance evaluation of the proposed approach is discussed in Section V, followed by the conclusion in Section VI.

II. PROBLEM FORMULATION

Faults generate TW which travel through the distribution line from the location of the fault to the opposite ends, at a speed near the velocity of light. The TW relay is placed at one end of the distribution line. The occurrence, location, and type of fault are estimated by observing these wave arrival times. However, due to the presence of multiple lateral lines in the distribution lines, these waves are also reflected at the junctions of each lateral before reaching the TW relay. The superposition of these TW reflections results in clutter, which in turn makes it extremely difficult to estimate the wave arrival times. TW-based protection in tapped (multiterminal) lines has been addressed in some recent literature e.g., [5, 6, 14]. 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 [9, 11]. Reference [15] provides a rather comprehensive survey of recent research in this area with an emphasis on the increasing penetration of distributed resources.

A. Traveling waves in distribution systems

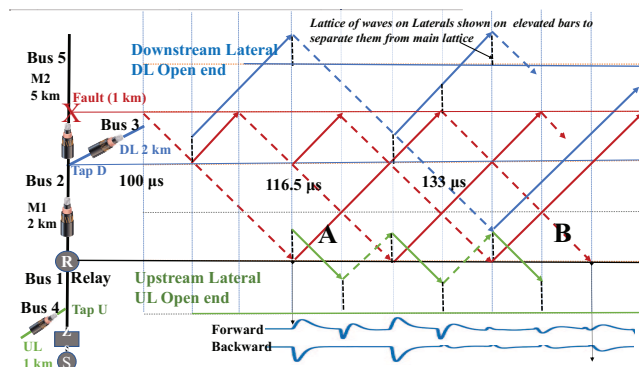


Fig. 2: Lattice diagram illustrating the TW phenomena in a distribution feeder with laterals.

We consider the feeder shown in Fig. 2, taken to be a single-phase feeder for our initial study. The relay R at bus 1 is to detect faults on the primary feeder sections $M1$ and $M2$, 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 $5.5 \mu\text{s}/\text{km}$. Suppose a fault occurs at X on $M2$

a distance of 1 km from the downstream lateral. The Lattice diagram [16] in Fig. 2 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 upstream 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. In the absence of 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, 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. In fact, the lateral return precedes the one from the fault. 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 [6, 14, 15, 17] to detect and locate the fault by observing the arrival times with respect to the reference.

The system in Fig. 2 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. 2, and are keyed to the lattice diagram. The TW relay must detect the periodic arrival of reflected waveforms. It should be noted that the periodicity will be different for the main feeder and each of the laterals (with the exception of two or more laterals of the 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 clutter and introduce ambiguity. Other sources of ambiguity include transformers (loads), capacitor banks, DERS, impedance and arcing in the fault, and 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.

Wavelet analysis is one of the prominent approaches for extracting information from TW clutters [18, 19, 20]. Despite their success in some scenarios, there is no consensus on what the best mother wavelet is [21, 22]. By using Shapelets, we can select the best basis that is not constrained to the existing wavelet basis functions and is governed by the physical phenomenon for TW generation.

B. Hypothesis

Shapelets are discriminative sub-sequences of time series that best characterize inter-class differences and have been successful in several classification tasks [23, 24, 25]. Motivated by this we formulate the following hypotheses:

- 1) *Shapelets as discriminants*: If we discover a set of discriminative sub-sequence/Shapelets and treat them as

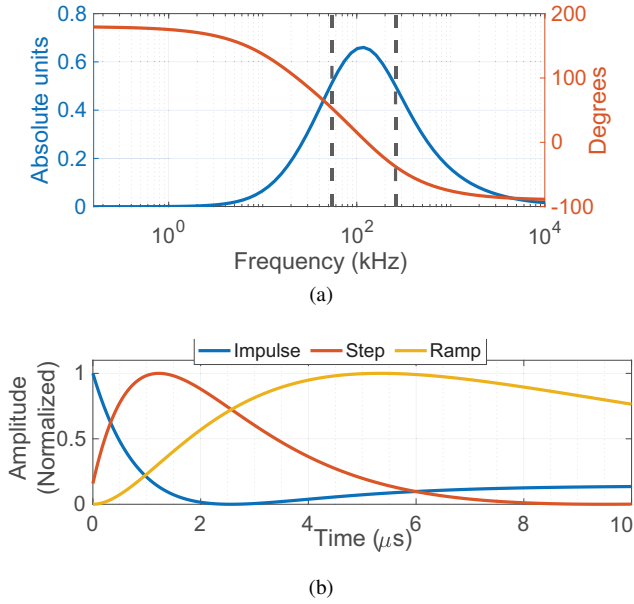


Fig. 3: (a) Magnitude and Phase response, (b) impulse, step, and ramp response of the bandpass filter.

basis functions for extracting information/features from the TW signals, can these features be used to classify different types of faults and find their locations?

- 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?
- 3) *Applicability*: Is it reliable and computationally feasible to develop a fault detection and classification technique based on features extracted from Shapelet analysis?

We propose and explore the use of Shapelets to implicitly ‘isolate’ wave arrivals and then 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) and the type 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 TW in distribution systems are discussed in [26]. The report emphasizes the need for additional modeling, simulation, and experiments to fully understand TW phenomena in distribution systems and to use them for fault detection and identification. Our study uses a simplified 5-bus distribution system simulation setup

as shown in Fig. 2. The system under study has *two* different configurations and associated sets of parameters.

In [1] we 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 lines without parameter frequency dependence. We explored the detection of faults at various distances from the relay with varying fault resistances and inception angles. This paper extends the studies using a three-phase 5-bus distribution system as well as the IEEE 13 bus Distribution Test System [27]. For all 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 can provide sub-100 meters resolution of the fault location. Additionally, at the proof of concept stage, we have started with the higher sampling rate as a baseline from which sensitivity studies can be performed. The available TW relays (e.g., SEL T401L) have a sampling rate of 1 *MHz* [28], we propose that the sampling rate of 10 *MHz* for applications in distribution systems. For a separate project, we have installed 10 *MHz* sensors in the field to capture the TW, so we have shown that this level of sampling rate is possible.

B. Pre-processing

The voltage and current signals were decomposed into the forward and backward waves based on the principles of the Telegraphers equation [16]. These waveforms were processed through an analog *band-pass filter* (BPF) with a center frequency of 118.5 kHz. The magnitude, phase, impulse, step, and ramp response of the filter are shown in Fig. 3(a), and (b). For a sinusoidal 200 kHz filter input, the filter delay was observed to be approximately 0.4 μs compared to the wave travel time of 5.5 $\mu s/km$. As such we anticipate errors and mis-classifications due to this filter. The filter has an almost linear phase response within the pass band (shown as vertical dotted lines). The transformations (e.g. symmetrical components) are also useful in extracting features. However, in this paper, we use the raw filtered *forward waveform* for all the single-phase experiments. For the three-phase system and IEEE 13 bus experiments, instead of using forward waves, we used filtered *voltage waveforms*. Using voltage waveforms instead of forward waves results in a potential reduction in one pre-processing stage (no current sensor and forward-backward decomposition required). The voltage waveforms corresponding to each phase in the 3-phase system were processed individually through the BPF.

C. Shapelet Discovery

We use the algorithm in [24] 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 offline process is computationally intensive. The hyper-parameters for the Shapelet discovery algorithm were chosen as learning rate = 0.1, weight regularizer = 0.1,

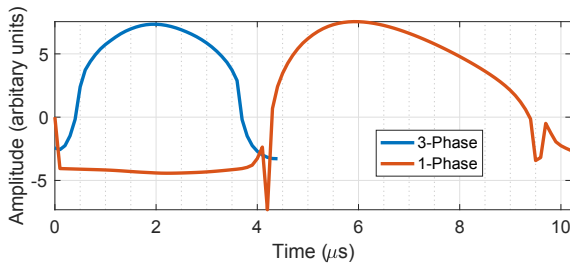


Fig. 4: Mother Shapelets discovered from the single-phase and three-phase datasets with the objective to maximally discriminate faults from steady-state.

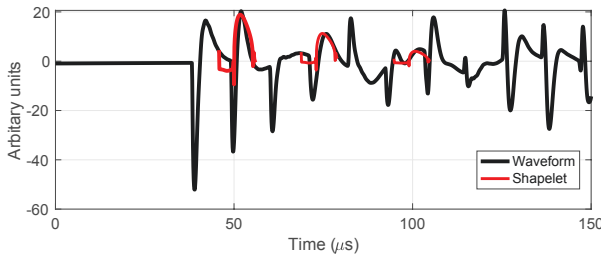


Fig. 5: Mother Shapelet at different translations and scales superimposed over an arbitrary TW waveform.

and the number of iterations = 46. The Shapelet length was chosen as 102 samples based on empirical observations and the algorithm suggested in [24]. 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. 4, we show the best (mother) Shapelets discovered from fault datasets created from the single-phase and three-phase configurations (explained in Section IV-A1). The similarity of this shape to the step and ramp response of the pre-processing band pass filter shown in Fig. 3, 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 5 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. 6. The peaks represent the highest correlation points in one traversal of the TW wave and alternatively, imply the time of arrival. The valley also implies 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 than 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 of $0.5 \mu s$ between any two consecutive peaks. A separation of $0.5 \mu s$ corresponds to approximately 100 meters, which in turn defines the minimum resolution in fault location.

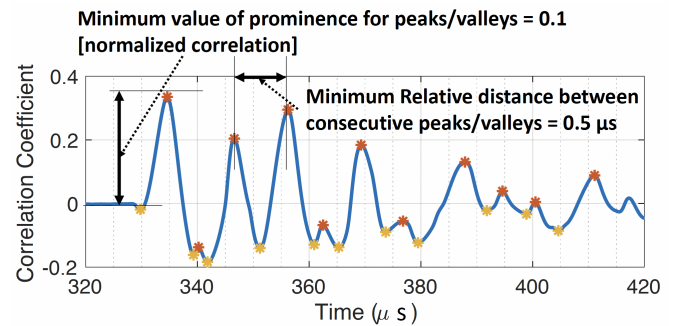


Fig. 6: Constraints to identify peaks and valleys.

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

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

For the three-phase system, we processed the voltage and current waveform from individual phases separately to extract three sets of features. These features were then merged to create a feature set for a pattern.

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 [29]. 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 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.

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

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).

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

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 and 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 is extracted. These features are then fed as input to the trained LDA model to predict the class to which the fault belongs.

We use two different classifiers, (i) fault-location classifier to identify approximate fault locations; and (ii) fault-type classifier to determine the type of the fault. Both the classifiers are fed with the same set of input features. The block diagram of the proposed scheme is shown in Fig. 7.

IV. SIMULATIONS & RESULTS

A. Case 1: Single phase test system

This subsection summarizes results for the single-phase system reported in [1]. Additionally, we discuss a number of parametric studies to explore the robustness of the classification results.

1) *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 setup. For the single-phase systems, as there is only one type of fault (line-to-ground), we only demonstrate the fault-location classification. We created a set of total 3600 patterns and divided them into 9 classes based on the location of the fault. The faults at the upstream and downstream laterals are combined to form one class. The faults on the main feeder were categorized into 8 segments of 500 meters each, corresponding to the other 8 classes, making the distance resolution to be 500 meters. This choice of 500 meters is arbitrary. As the placement of DER evolves, the resolution could be chosen accordingly.

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 Ω with a uniform distribution, to create an exhaustive set of patterns.

2) *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 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. 8(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 misclassified patterns were found to lie near the decision boundary of two neighboring classes. For instance, the 6 misclassified 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. 8(b), to identify the classes that are vulnerable 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., a few patterns at the boundary of Class-5 and 8 are misclassified to their adjoining class.

TABLE I: 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

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 a severe impact. To assess the effect of noise we added white Gaussian noise to the filtered waveform before extracting the features. As we do not know the actual measurement noise characteristics for the TW, white Gaussian distribution seemed a reasonable assumption [30]. The mean and standard deviations of 10-fold cross-validation accuracy for different signal-to-noise ratio (SNR) conditions are presented in Table I. 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 to 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

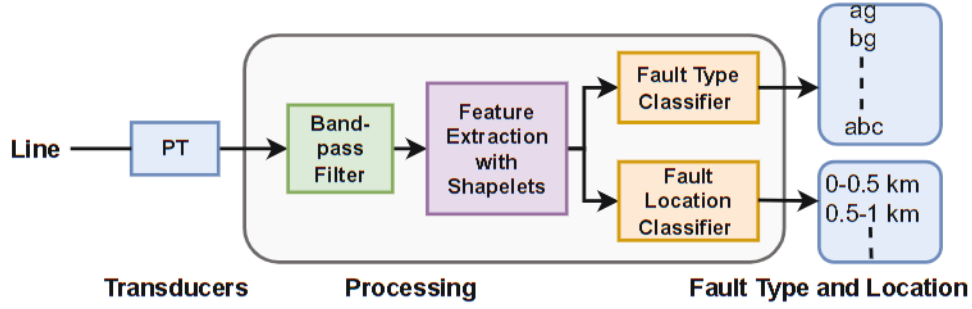


Fig. 7: Block diagram showing the modules of a TW relay based on Shapelets.

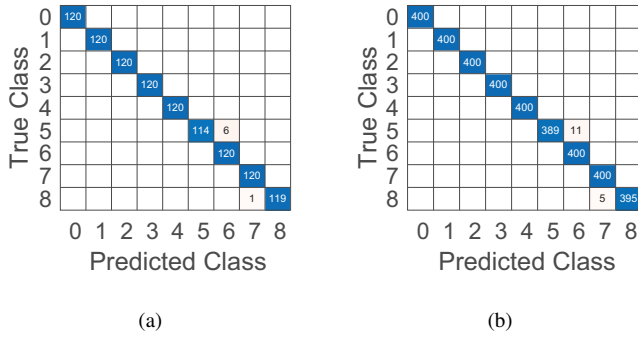


Fig. 8: Confusion matrix with (a) 30% test data and 70% training data, and (b) for the entire data set, without noise.

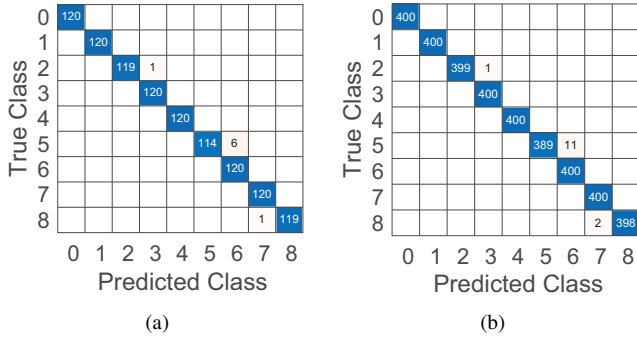


Fig. 9: Confusion matrix with (a) 30% test data and 70% training data, and (b) for the entire data set, at 30 dB SNR.

very extreme test. Therefore, the proposed method can be considered robust against measurement noise.

In Fig. 9(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. 8), only one additional pattern in Class-2 got misclassified into neighboring Class-3. It further demonstrates the robustness to additive noise. The confusion matrix for the entire dataset shown in Fig. 9(b) also reinforces the same conclusion that few patterns at the boundary of Class-2, 5, and 8 are at the risk of being misclassified to their adjoining class, at 30 dB SNR.

B. Case 2: Three-phase test system

The single-phase case evaluated the performance of the Shapelet in localizing faults. We extend the evaluation to 3-phase systems to not only assess the ability to identify fault location but also to diagnose the type of fault.

1) *Dataset*: The dataset for the 3-phase comprised of different types of faults (line-to-ground, line-line-to-ground, line-to-line, three-phase, and three-phase-to-ground) listed in Table II. For each of these fault types, a set of 1350 patterns was created with the parametric variations and class labels as discussed in Section IV-A1, except that the fault resistance took a value between 0.025 – 2 ohm with incremental steps of 0.4. We also varied the inception angle of the fault between 0 – 360° at increments of 70° by varying the source angle. A total of 14,850 (11 × 1350) patterns were generated. In order to mimic the real-world characteristics, we added white Gaussian noise to keep the SNR at 40 dB. The Shapelet discovery for the 3-phase dataset followed the same principle as that of the single-phase. The best Shapelet ($S_1^{3\Phi}$) was used as the mother shapelet (shown in Fig. 4) to extract the features discussed in Section III-D.

2) *Results*: In the 3-phase case, we have two classifiers in action; first, to classify the type of the fault, and second, to identify the approximate location of the fault. Both the classifiers were fed with the same set of features and carried out the inferences in parallel.

(a) *Fault-type classification*: The confusion matrix in Fig. 10 summarizes the classification performance for the identification of fault types with features comprising of first 41 peaks/valleys. The mean classification accuracy with 10-fold cross-validation for all the 11 classes is 86.43% with a standard deviation of 0.74%. The three-phase (*abc* and *abcg*) faults were observed to have a relatively lower classification accuracy.

(b) *Fault-location classification*: Table II summarizes the fault location classification performance for each type of fault with 10-fold cross-validation. The mean fault location classification accuracy for *abc* faults was observed to be the least (98.86%). For all other 10 types of faults, the accuracy was $\geq 99\%$.

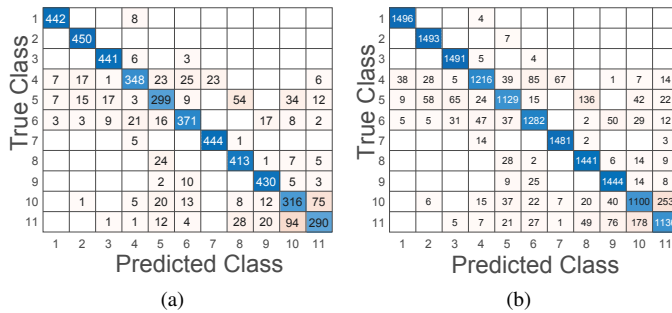


Fig. 10: Confusion matrix for fault type classification with (a) 30% test data and 70% training data, and (b) for the entire 3-phase data set at 40 dB SNR. The class labels are mapped onto fault types as: ag:1, bg:2, cg:3, abg:4, bcg:5, acg:6, ab:7, bc:8, ac:9, abcg:10, abc:11

TABLE II: Fault location classification accuracy with 10-fold cross-validation for 3-phase dataset.

Fault Type		Accuracy (%)	
		Mean (μ)	Std. Dev. (σ)
Single Phase-Ground	ag	99.87	0.45
	bg	99.93	0
	cg	100	0.28
Double Phase-Ground	abg	99.33	1.53
	bcg	99.2	0.7
	acg	99.6	1.17
Phase-Phase	ab	100	0.21
	bc	99.93	0.28
	ac	99.93	0
Three Phase	abcg	99	0.66
	abc	98.86	0.63

C. Case 3: IEEE 13-bus distribution feeder

We considered the 13-bus distribution test system [27] shown in Fig. 11 for a comprehensive evaluation of the proposed approach. The line segment between node 632 and 671 was considered as the observed line for protection and the TW relay was assumed to be placed on node 632. The 632-671 is a 0.61 km long line that was divided into 7 segments of 87 meters to create Classes 1-7, based on the location of the fault. The faults were simulated at three segments

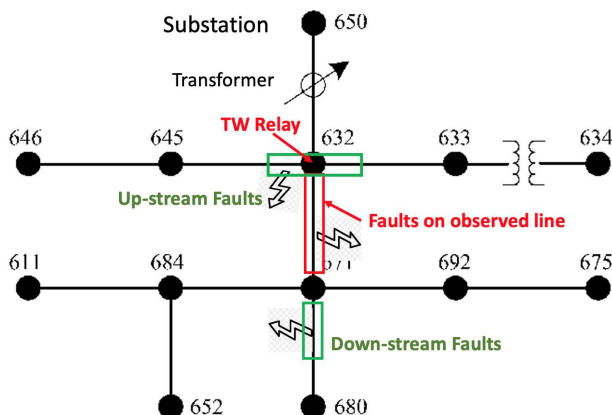


Fig. 11: Experimental setup for 13-bus distribution test feeder.

corresponding upstream (at bus 632), on the observed line (632 – 671), and downstream (671 – 680), respectively. The faults at the upstream and the downstream were labeled as Class-0, resulting in a total of 8 classes for fault location identification. The fault resistance and inception angle were varied between 0.05 – 2 Ω and 0 – 360°, respectively, to create 120 patterns for each class. The noise was added to each TW pattern to keep the SNR at 40 dB. The fault location classification accuracy with 10-fold cross-validation corresponding to each type of fault is summarized in Table III. The mean fault type classification accuracy with 10-fold cross-validation was found to be 92.84% with a standard deviation of 0.81.

TABLE III: Fault-location classification accuracy with 10-fold cross-validation for IEEE 13-bus dataset.

	ag	bg	cg	abg	bcg	acg	ab	bc	ac	abc	abcg
μ	99.6	100	100	99.9	99.9	99.5	100	100	99.8	99.5	99.7
σ	0.4	0	0	0.3	0.3	1.1	0	0	0.4	0.7	0.7

V. PERFORMANCE EVALUATIONS

A. Impact of feature count on classification performance

The feature count is an important factor in the performance of the classifiers. The number of peaks/valleys considered for computing feature values as a function of fault location accuracy for the 3-phase test system dataset is shown in Fig. 12. It was observed that although the mean fault location classification accuracy increases with increasing feature count, the performance with 21, 31 and 41 peaks have relatively similar profiles.

However, for the fault type classification, the feature count has a much higher significance. The fault type classification accuracy for 11, 21, 31, and 41 peaks were 67.04, 83.83, and 86.43, respectively. The set of features from 41-peaks was found to be a reasonable choice for both fault type classification and identification of fault locations.

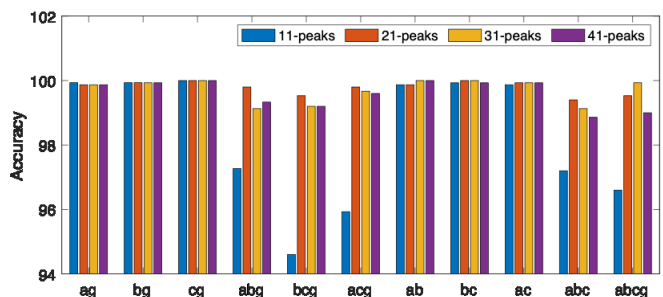


Fig. 12: Impact of feature count on fault location classification accuracy.

B. Shapelet resemblance to band-pass filter response

In order to evaluate the contribution of the BPF to the shape of the Shapelets, we considered the normalized step response of the BPF (shown in Fig. 3b) as the mother Shapelet. We

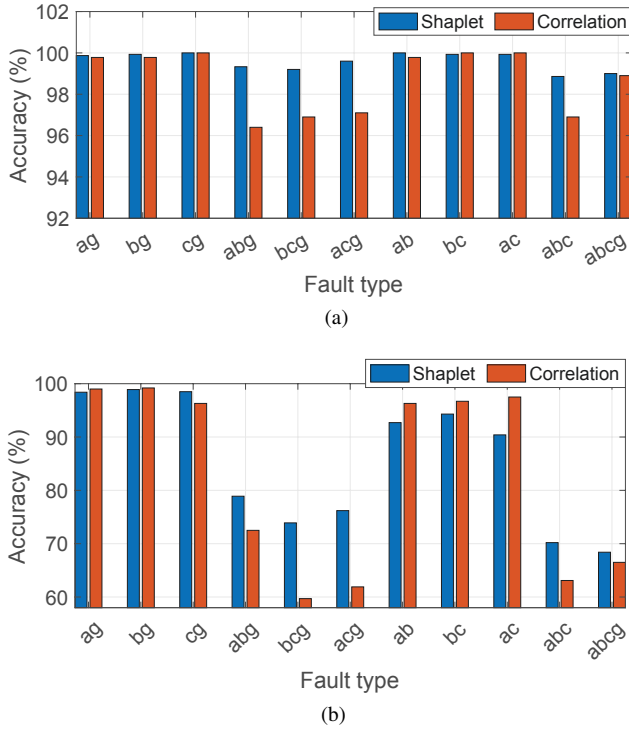


Fig. 13: Comparison of (a) fault-location classification and (b) fault-type classification accuracy, with auto-correlation and Shapelets.

extracted the features (with 41 peaks preserved) from the 3-phase test system dataset and evaluated the performance of the classifiers. The mean accuracy for the fault type classification with 10 fold cross validation was observed to be 85.55% with a standard deviation of 1.11. This is close to the accuracy obtained with the discovered Shapelet, which was 86.43%. For fault location classification, the mean of the average accuracy corresponding to each type of fault was found to be $> 95\%$. This implies that the BPF response has a significant contribution to the shape of the mother Shapelet. The significance of this observation is that, if indeed the BPF response can serve as a good wavelet, the Shapelet discovery process could be significantly simplified.

C. Effect of clutter complexity

The clutter complexity for the TW signature could largely vary based on the network configurations. In order to evaluate the generalization of Shapelets, we conducted another experiment with highly complex TW clutters. As the clutter complexity depends on the number of reflections that get superimposed when the TW signal is observed, the short lateral segments will consequently make the observed TW clutter more complex. Therefore, we reduced the length of upstream and downstream laterals in the 3-phase test system to 0.75 and 0.5 km, respectively. We also added noise to the measured voltage waveforms to keep the SNR at 40 dB. The fault-location accuracy with 10-fold cross-validation is summarized in Table IV. The fault-type classification accuracy has a mean of 81.4502% with a standard deviation of 0.87092. Comparing

these with that of without clutter complexity, although the classification accuracy decreases by a small amount they are still $> 99\%$ for fault location and $> 80\%$ for fault type.

TABLE IV: Fault-location classification accuracy with 10-fold cross-validation for the 3-phase dataset with short segments.

	ag	bg	cg	abg	bcb	acg	ab	bc	ac	abc	abcb
μ	100	99.2	100	98.3	98	98.5	99.9	99.6	99.5	98.4	98.4
σ	0	0.73	0	0.72	0.88	0.98	0.2	0.6	0.5	0.9	0.6

D. Computational complexity

The Shapelet discovery is an offline process that needs to be performed during the design phase to identify the mother Shapelet. As the step response of the BPF resembles the shape of the mother Shapelet, it can be treated as the mother Shapelet without significantly affecting the classification performance. The computational complexity for Shapelet matching is $O(N \times M)$ for a Shapelet of length M over N samples, which is minimal as $M \ll N$. The peak/valley picking can be efficiently implemented without the need for any high-performance processors. These factors make the feature extraction lightweight. Finally, we use a simple LDA-based classifier that is easy to train and extremely fast in making inferences/predictions. For a MATLAB R2022a implementation of the 3-phase fault type and location classification system in the Apple M2 processor with 8 GB RAM, the entire process of feature computation for a signal takes a mean time of 1.5 ms, and prediction takes a mean time of 0.0065 ms. These computational times will further go down for an optimized embedded implementation.

E. Comparison with Auto-correlation

We analyzed an auto-correlation-based technique in order to evaluate the relative importance of the Shapelets in characterizing faults. The auto-correlation coefficients for individual waveforms were computed over a fixed time window for the 3-phase test dataset. The auto-correlation peaks and valleys were identified and the same set of features was extracted as described in Section III-D. The comparison of mean accuracy for fault-location classification corresponding to each type of fault and fault-type classification are shown in Fig. 13. We observed that the line-line-ground (abg, bcb, acg) and 3-phase (abc, abcb) faults are relatively difficult to classify and have lower accuracies for both locations as well as types. With auto-correlation, patterns belonging to these classes are even harder to distinguish. Whereas, the Shapelet approach consistently boosts classification performance for these fault classes. For the fault locations classification, the mean accuracy with auto-correlation and Shapelets are 98.68% and 99.61%, respectively, with 10-fold cross-validation. For fault types, the overall accuracy with auto-correlation is 82.6% which gets boosted to 85.53% with Shapelet. Another advantage of Shapelets is the reduction in computational complexity. Auto-correlation has a computational complexity of $O(N^2)$ for a sequence of length N samples and this reduces to $O(N \times M)$ for a Shapelet of length $M \ll N$.

F. Comparison with other techniques

In this sub-section, we present an objective comparison of the performance of the proposed approach with some recently proposed techniques, as there is significant variability in the dataset size, fault types considered, fault location tasks, fault parameters, measurement noise, performance metric calculation, and training & validation of the ML models. The proposed approach is demonstrated for a feeder in the distribution systems. For larger systems, it can simply be replicated individually for each feeder without scalability issues. In [19] a continuous wavelet transform (CWT) based multi-resolution analysis, along with boosted trees/random forest classifier is proposed and results in an accuracy of 93.97% for fault type classification in the IEEE 34 bus test system. An improvement in this accuracy to 96% was reported in [20] with CWT features and a deep convolutional neural network (CNN). In both these approaches, the fault types are categorized into either of three classes, single-line-to-ground fault, line-to-line fault, or 3-Phase fault. Compared to these, the proposed approach distinguishes among individual sub-types of faults resulting in 11 classes. To make the comparison more relevant, we merged the 11 classes to form the three classes for the IEEE 13 bus test system and observed a mean accuracy of 96.7% with 10-fold cross-validation. This fault-type classification accuracy is slightly higher compared to both the previous approaches. Although this does not appear to be a significant improvement, the proposed technique alleviates the need to use computationally demanding CWT or deep CNN. In addition, the deep Learning classifiers may also suffer from over-fitting if not trained on a sufficiently large dataset. This implies that the models could have high validation accuracy but may have poor generalization. In our evaluations, we used a simple discriminant analysis classifier that emphasizes the discrimination ability of the extracted features and performed 10-fold cross-validation. Another difference is that for simulations in [19, 20] faults are created only at the system nodes, whereas we simulate faults at various locations of the lines. We also performed the evaluation under the measurement noise with an SNR of 40 dB.

In [31] a comprehensive performance comparison of various fault location algorithms is presented. In addition, the authors propose a Mathematical Morphology (MM) and Stationary Wavelet Transform (SWT) based feature extraction algorithm. The features are then fed to the RF classifier and regressor to predict the fault locations. For an RF model trained with 70% training data and 30% validation data at an SNR 45 dB, they report an average error of 1367.2 ft (416.72 meters). Compared to this, the proposed fault location classification algorithm for the IEEE 13 bus distribution system, has a resolution of 87 meters and achieves an overall accuracy of $> 99\%$, for all the fault types. Faults can be located and classified with less than 0.5 ms of measured data after the TW arrival.

VI. DISCUSSIONS AND CONCLUSION

This paper proposes a relatively simple Shapelet-based approach to classify faults in the distribution systems through TW principles. In contrast to choosing a pre-defined basis function such as wavelets, 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. The use of Shapelets was comprehensively tested using 5-bus single- and three-phase systems, as well as the more extensive IEEE 13-bus distribution test system. For the 3-phase systems, we used the voltage waveforms instead of forward waves, eliminating the need for current transformers. We observed that the extracted features result in a linear decision boundary, and with discriminant analysis, the fault locations 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 $> 90\%$. The Shapelet features were also used to classify different types of faults in the distribution system. For the IEEE 13 bus test system, the fault type classification accuracy was found to be 92.84% for 11 classes and 96.7% for 3 classes (LG, LL, and 3-phase). These results indicate the promising ability of Shapelets in classifying distribution system faults.

Future work will address hardware implementation, the use of advanced ML methods and sensitivity analysis, and critical comparisons with existing relay solutions, as well as techniques such as the ones described in [7, 8, 9, 11, 12].

VII. ACKNOWLEDGEMENTS

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan <https://www.energy.gov/downloads/doe-public-access-plan>. This work was also partially supported by the NSF Grants OIA-1757207 (NM EPSCoR), HRD-1345232 and HRD-1914635.

REFERENCES

- [1] M. Biswal, S. Pati, S. J. Ranade, O. Lavrova, and M. J. Reno, "Exploring the use of shapelets in traveling wave based fault detection in distribution systems," in *2022 IEEE Texas Power and Energy Conference (TPEC)*, 2022, pp. 1–6.
- [2] IEEE PES Industry Technical Support Leadership Committee, "Impact of IEEE 1547 standard on smart inverters and the applications in power systems," 2020.
- [3] 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.

- [4] 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.
- [5] 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.
- [6] R. Khuzyashev, I. Kuzmin, S. Tukaev, L. Tikhvatullin, 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.
- [7] D. Celeita, A. S. Meliopoulos, G. Ramos, and L. Romero, "Dynamic state estimation for double-end traveling wave arrival identification in transmission lines," *Electric Power Systems Research*, vol. 170, pp. 138–149, 2019.
- [8] T. E. McDermott, N. Shepard, S. Meliopoulos, M. Ramesh, J. D. Doty, and J. T. Kolln, "Protection of distribution circuits with high penetration of solar pv: Distance, learning, and estimation-based methods," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), Tech. Rep., 2021.
- [9] F. Wilches-Bernal, M. Jiménez-Aparicio, and M. J. Reno, "A machine learning-based method using the dynamic mode decomposition for fault location and classification," in *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2022, pp. 1–5.
- [10] —, "An algorithm for fast fault location and classification based on mathematical morphology and machine learning," in *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2022, pp. 1–5.
- [11] M. Jiménez-Aparicio, M. J. Reno, and F. Wilches-Bernal, "Asynchronous traveling wave-based distribution system protection with graph neural networks," in *2022 IEEE Kansas Power and Energy Conference (KPEC)*. IEEE, 2022, pp. 1–6.
- [12] M. Atrigna, A. Buonanno, R. Carli, G. Cavone, P. Scarabaggio, M. Valenti, G. Graditi, and M. Dotoli, "A machine learning approach to fault prediction of power distribution grids under heatwaves," *IEEE Transactions on Industry Applications*, 2023.
- [13] H. Mirshekali, A. Keshavarz, R. Dashti, S. Hafezi, and H. R. Shaker, "Deep learning-based fault location framework in power distribution grids employing convolutional neural network based on capsule network," *Electric Power Systems Research*, vol. 223, p. 109529, 2023.
- [14] 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.
- [15] C. Y. Evrenosoglu and A. Abur, "Fault location in distribution systems with distributed generation," in *Proc. Power Systems Computation Conference*, 2005.
- [16] J. D. Glover, M. S. Sarma, and T. Overbye, *Power system analysis & design, SI version*. Cengage Learning, 2012.
- [17] F. Miyagishima, O. Lavrova, S. Augustine, S. Ranade, M. J. Reno, and J. Hernandez-Alvidrez, "Numerical analysis of traveling waves in power systems with grid forming inverters," in *2022 North American Power Symposium (NAPS)*. IEEE, 2022, pp. 1–5.
- [18] F. Wilches-Bernal, A. Bidram, M. J. Reno, J. Hernandez-Alvidrez, P. Barba, B. Reimer, R. Montoya, C. Carr, and O. Lavrova, "A survey of traveling wave protection schemes in electric power systems," *IEEE Access*, vol. 9, pp. 72 949–72 969, 2021.
- [19] M. J. Aparicio, M. J. Reno, P. Barba, and A. Bidram, "Multi-resolution analysis algorithm for fast fault classification and location in distribution systems," in *2021 IEEE 9th International Conference on Smart Energy Grid Engineering (SEGE)*. IEEE, 2021, pp. 134–140.
- [20] S. Paul, S. Grijalva, M. J. Aparicio, and M. J. Reno, "Knowledge-based fault diagnosis for a distribution system with high pv penetration," in *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2022, pp. 1–5.
- [21] M. Jiménez-Aparicio, M. J. Reno, and F. Wilches-Bernal, "Traveling wave energy analysis of faults on power distribution systems," *Energies*, vol. 15, no. 8, p. 2741, 2022.
- [22] M. Jimenez-Aparicio, M. J. Reno, and J. Hernandez-Alvidrez, "Fast traveling wave detection and identification method for power distribution systems using the discrete wavelet transform," *Proceedings of the Innovative Smart Grid Technologies, North America (ISGT NA), Washington, DC, USA*, pp. 16–19, 2023.
- [23] 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.
- [24] 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.
- [25] 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 *Power Systems Computation Conference (PSCC)*, 2016, pp. 1–7.
- [26] 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.
- [27] W. H. Kersting, "Radial distribution test feeders," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975–985, 1991.
- [28] (2023) Schweitzer engineering laboratories, "sel-t4011 relay". [Online]. Available: <https://selinc.com/>
- [29] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [30] 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.

- [31] M. Jimenez-Aparicio, F. Wilches-Bernal, and M. J. Reno, “Local, single-ended, traveling-wave fault location on distribution systems using frequency and time-domain data,” *IEEE Access*, 2023.



Shubhasmita Pati (S'16) received the M.S. and Ph.D. degrees in Electrical Engineering from New Mexico State University, Las Cruces, NM, in 2018 and 2023, respectively. She is currently working as a post-doctoral researcher at the National Renewable Energy Laboratory (NREL). Her current research interests include power system modeling, resiliency enhancement, and microgrid protection.



Mathew J. Reno (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Georgia Institute of Technology. He is currently a Principal Member of Technical Staff with the Electric Power Systems Research Department, Sandia National Laboratories. His research interests include distribution system modeling and analysis with high-penetration PV, including advanced software tools for automated analysis of hosting capacity, PV interconnection studies, and rapid quasi-static time series simulations. He is with the IEEE Power System Relaying Committee for developing guides and standards for the protection of microgrids and systems with high penetrations of inverter-based resources.



Milan Biswal (M'09) received the Ph.D. degree from Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India, in 2014. From 2014 to 2016, he was a Postdoctoral Researcher in the design of intelligent technologies for smart grids with the CREST Interdisciplinary Center of Research Excellence, New Mexico State University, Las Cruces, NM, USA, where he was also a Research Scientist, Research Assistant Professor, and Senior Research Scientist. His research interests include signal processing and machine learning solutions for applications in the smart grid and other domains, wide area monitoring and protection systems, PMU data analytics, integration of distributed energy resources, and power systems instrumentation.

tions in the smart grid and other domains, wide area monitoring and protection systems, PMU data analytics, integration of distributed energy resources, and power systems instrumentation.



Satish J. Ranade (F'22) is an Emeritus Professor in the Klipsch School of Electrical Engineering, New Mexico State University (NMSU). He joined NMSU in 1981. He has served as a Distinguished Professor (2013-), Klipsch Professor (2000-2002), PNM Chair in Utility Management (2003-2020), Director of the Electric Utility Management Program (EUMP) and Department head (2013-2019). His Teaching and Research interests are in the Energy Systems area.



Olga Lavrova (S'00, M'11) is an Associate Professor at the Electrical and Computer Engineering Department at the Klipsch School of Electrical and Computer Engineering at the New Mexico State University. Prior to that, Dr. Lavrova was a Principal Member of Technical Staff at Sandia National Labs in the Photovoltaics and Distributed Systems Integration Department. Prior to that, she held the position of Assistant Professor at the Electrical and Computer Engineering Department at the University of New Mexico. She received her B.Sc. degree in Physics

and M.Sc. degree in EE from the St. Petersburg State Electrical Engineering University, and her Ph.D. degree from UCSB in 2001. Her current work and areas of interest include photovoltaics and nano-scale semiconductor structures for photovoltaic applications, Smart grids, Renewable Energy, Controls of the electric grids, smart grids and renewable energy integration, electric storage systems, and power electronics.