Electric Power Systems Research 236 (2024) 110940

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Grid-aware learning of characterized waveform measurements for power quality and transient events situational awareness

Mohammad MansourLakouraj [a,*], Hadis Hosseinpour [b], Hanif Livani [a], Mohammed Benidris [b]

[a] Department of Electrical and Biomedical Engineering, University of Nevada, Reno, Reno, NV 89557, USA
[b] Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Keywords:
Event characterization and classification
Grid-aware learning
Waveform measurement units
Power distribution grids
Power quality and transient events
Time-frequency decomposition

ABSTRACT

The emerging class of waveform measurement units (WMUs) can enhance event situational awareness in distribution grids by better characterizing the signatures of events and reporting higher sampling rates compared to traditional measurement units. This study aims to improve the classification and localization of power quality events and transient disturbances in active distribution grids by integrating time–frequency characterization of WMU data and a grid-aware learning algorithm. A time–frequency decomposition technique with a moving window is proposed to extract complex frequencies and residues from a limited set of available WMU data. This endeavor seeks to characterize the unique signatures of events occurring in various locations. Extracted features are then used as nodal graph signals in a grid-aware autoregressive moving average (ARMA) graph convolution to classify event type and location. Despite the scarcity of waveform measurements, the proposed grid-aware model captures the spatial relationship between derived signatures of measured signals at different nodes. Through our numerical studies, the proposed strategy improves the classification and localization of events such as capacitor bank switching, abrupt load changes, generation outages, and various types of faults (e.g., arcing high impedance faults), when compared to state-of-the-art models. The proposed method is validated under various scenarios, including noisy data, different measurement configurations and reporting rates, diverse operational conditions, and unbalanced feeders.

1. Introduction

1.1. Motivations and challenges

Power distribution grids are susceptible to disturbances and events, such as sudden load changes, faults, sudden variability and outages of distributed generations, and malfunctioning of utility equipment and assets, including transformer tap changers and capacitor banks (CBs) [1,2]. Waveform measurement units (WMUs), as a class of high-fidelity sensors, are being adapted into power grids to improve monitoring capabilities for various applications such as power quality and event monitoring and analytics [3]. Analyzing the captured waveform data from these new sensors leads to a better understanding of the roots of events as well as enhanced grid asset health situational awareness [4].

In recent years, phasor measurement units (PMUs) have been introduced at the distribution level, typically with a sampling rate ranging from 30 to 120 samples-per-second (SPS) [5]. PMUs with high sampling rates (120 SPS) can receive at most two phasor samples-per-cycle, therefore making them suitable for analyzing grid conditions during steady-state as well as events with slow dynamics. However, they may not be able to capture fast grid response during non-nominal frequencies and fast transient events. Therefore, there has been rising interest in the deployment of WMUs with higher sampling rates for fast transient events detection, classification, and localization [6,7]. Sampling rates of WMUs can exceed 256 samples-per-cycle, enabling us to better monitor fluctuations and transients in voltage and current waveforms caused by particular events and incipient faults [4]. Furthermore, some power quality events can be observed better at the waveform level with efficient waveform analytics tools [8]. In fact, waveform analysis is the key step to successfully deploying WMUs and achieving a higher level of event situational awareness in electric networks. In this regard, several studies have focused on analyzing the features of measured signals using signal processing techniques, such as wavelets, for power system studies [9,10].

Although the deployment of WMUs in different areas of distribution grids is currently limited, their use enhances event analysis by enabling the collection of more data, which assists in identifying both the

* Corresponding author.
E-mail addresses: mmansourlakouraj@unr.edu (M. MansourLakouraj), hossei13@msu.edu (H. Hosseinpour), hlivani@unr.edu (H. Livani), benidris@msu.edu (M. Benidris).

https://doi.org/10.1016/j.epsr.2024.110940
Received 1 April 2024; Received in revised form 22 June 2024; Accepted 27 July 2024
Available online 9 August 2024

type and location of events. Learning the spatial relationships among voltage and current captured by WMUs, which are physically related to each other based on electric circuit laws, can also advance the event situational awareness [6]. In this regard, three main challenges exist: (1) The widely used data-driven models often do not leverage the valuable physical relationships among the measurements. In addition, implementing model-based techniques is also challenging due to the complexity of grid conditions and occurring events; (2) The high sampling rate of WMUs makes the post-processing of the captured signals more challenging for event classification. Therefore, waveform analytics is required to derive the most important and informative features from waveform data. (3) Some event-related distortions on AC waveform signals are small and/or transient. Therefore, the characterization of events' signatures using the less frequently examined time–frequency features in the moving small steps could lead to enhanced event identification. To address these challenges, our contributions are discussed in the following section.

1.2. Summary of technical contributions

This paper leverages the topological characteristics of distribution grids and time-frequency signatures of events for enhanced event situational awareness. The contributions are listed as follows:

- The proposed method enhances situational awareness by characterizing, classifying, and locating a wide variety of events, particularly those causing minor and/or transient distortion in recorded waveform measurements. These events include capacitor bank (CB) switching, sudden load changes, regular and arcing faults, as well as distributed energy resource (DER) outages. The method enables system operators to respond quickly to identified disturbances, preventing consequences and ensuring operational reliability.
- A time-frequency waveform analytics algorithm is employed to characterize event signatures from waveform measurement units (WMUs), representing a new class of sensors deployed within distribution networks. This method decomposes the waveform measurement samples with the matrix pencil concept in a sliding window into a less explored set of three dominant distinctive features, including complex frequencies (damping factor and angular frequency) and residues. These features are presented as distinctive time-varying series using a moving window to better characterize the event signatures.
- A grid-aware learning model equipped with Auto-Regressive Moving Average (ARMA) graph convolution is proposed to seamlessly integrate the topological properties of distribution grids as structured graph data consisting of characterized nodal waveform data. The proposed method aggregates and learns the spatially correlated nodal features from characterized voltage and current at a few locations equipped with WMUs for enhanced event classification and localization. This contrasts sharply with most learning-based approaches, which ignore the informative topological context of the problem. Incorporating such context into graph representation learning can enhance the learning of relevant features across the distribution system, thereby enhancing the performance.
- The practicality of the approach is validated on an unbalanced distribution grid and a three-phase microgrid with different distributed generators. The model is extensively examined under different conditions, such as different waveform features, configurations of WMUs, signal-to-noise ratios, and reporting sampling rates of WMUs.

1.3. Relevant studies and research gaps

Event analysis in distribution grids is mainly performed using three approaches such as data-driven, model-based, and grid-informed data-driven models. As for data-driven analysis, a PMU-based data-driven model with autoencoders is proposed in [5] to classify disruptive and normal events from each other. A multi-class support vector machine (SVM) classifies the CB switching, oscillations, and other events in distribution grids using PMU data, as discussed in [11]. However, waveform analytics and incorporating the grid information have not been studied in [5,11]. In [12], authors propose a human-level learning approach to extract general shapes and residuals of waveform measurements in distribution grids. Then, the hierarchical Bayesian method uses the extracted features to classify the incipient faults versus other events without locating them. However, the topological correlation of measured data cannot be incorporated into the proposed learning model [12]. In [13], voltage sags are categorized using the Clarke components ellipse within distribution grids; however, the scope of their research is only focused on faults.

Data-driven methods usually do not incorporate the topological features of the grid, but the model-based approach uses circuit theories for event detection and location. The study in [14] leverages an impedance-based technique using the fundamental phasor components of signals to find faulty nodes. In [15], an efficient fault location technique is proposed using the voltage sag in distribution systems, but the localization of various events is not considered. In [4], the use of the Prony technique to find the modes of waveform data is proposed. The circuit model is then created to locate transient faults and events. However, the model-based technique focuses on circuit modeling, which can be challenging work for unbalanced distribution feeders with mutual impedances and varying generations and loads. Moreover, the randomness of the events adds more complexity to the model-based studies.

The stated challenges and advantages of both model-based and purely data-driven techniques have motivated us to close the gap between these methods by using the grid-aware data-driven model. This method is developed using the concept of graph neural network (GNN) [16]. In [17], GNN is used for event classification and localization with PMU data and distance-based physical features of the distribution systems. Ref. [18] implements GNN for event clustering with fundamental and harmonic PMUs, while the lines' physical features and event localization are not considered. Importantly, neither of the studies in [17] nor in [18] investigates the use of waveform measurements with grid-aware learning. In [19], we present our preliminary and brief analysis focusing on waveform analytics for event type classification using GNN with polynomial filters. However, we did not explore more recent graph filters and the task of locating the events.

In this paper, a time–frequency decomposition technique is used to extract the embedded poles and residues from the signal at each sliding snapshot of time. The obtained complex poles and residues have a time-varying behavior, allowing them to be processed as informative signal features. The concept, called the short-time matrix pencil method (STMPM) with a moving window, has been originally introduced for radio-frequency identification applications [20]. This technique was applied in our fault location study by extracting the complex frequencies from voltage signals [6], while other challenging events and signal features are not explored. In the current work, we expand the use of time–frequency decomposition to extract more time-varying features under a wide range of power quality events and transient disturbances from voltage and current measurements. In the current study, the waveform measurements are thoroughly generated using electromagnetic transient simulation techniques, including automated real-time digital simulation (RTDS), which is highly practical for industrial and real-world applications. Additionally, this paper advances the proposed grid-aware learning algorithm using the recently introduced robust ARMA graph convolutional layers [21], leading to more accurate results in different experiments. The proposed ARMA-based method classifies different types and locations of events, such as arcing faults, switching, and outages of circuit breakers (CB), loads, and distributed energy resources (DER), while [6] uses graph convolution network (GCN) model to locate regular faults.

The rest of the paper is organized as follows. Section 2 discusses the waveform analytics for the characterization of events. Section 3 explains the grid-aware learning method. The case studies and numerical analysis are given in Section 5. Finally, Section 6 presents the conclusions.

2. Characterization of events with waveform analytics

In a distribution grid, WMUs might be located at a few nodes, measuring voltage and current waveforms. Once an event occurs along distribution feeders, the sinusoidal waveform measurements are distorted, creating signatures for that particular event. These changes may not be significant, making them hard to detect. Therefore, an efficient waveform analytics method based on the short-time sliding window is used to extract features of abrupt and/or small changes, enhancing event situational awareness.

2.1. Background in modal analysis

The Matrix Pencil method, originally developed by Hua and Sarkar [22], has been applied to pole estimation and used for extracting poles from the electromagnetic transient responses of antennas. Unlike polynomial methods such as Prony Analysis, which requires a two-step process, the Matrix Pencil approach identifies signal poles directly from the eigenvalues of a single matrix. This method is tailored to estimate the parameters of exponential terms directly by fitting a function to the observed measurements. The authors of [23] comprehensively discuss the advantages of the Matrix Pencil method over the Prony method in terms of accurately extracting modes under noise and reducing computation time.

2.2. Extracting complex frequencies and residues

A measured signal $y(t)$ with a period T can be expressed through M summation of damped sinusoidal terms as follows,

$$y(t) = \sum_{i=1}^{M} r_i \, e^{-\alpha_i t} \, cos\,(\omega_i \, t + \theta_i) \tag{1}$$

where the residue denotes as $\zeta_i = r_i \angle \theta_i$, and the frequency and angular frequency are represented by $f = \frac{1}{T}$ and $\omega_i = 2i\pi f$, respectively. Moreover, the parameter α_i shows the damping factor ($\sec^{-1}$). Finally, t represents the time variable for the signal. Note that parameters α_i and ω_i create the complex frequency.

Using the samples in an N-sample sliding window, the matrix X is built as

$$\mathbf{X} = \begin{bmatrix} y(1) & y(2) & \dots & y(P+1) \\ y(2) & y(3) & \dots & y(P+2) \\ \vdots & \vdots & \vdots & \vdots \\ y(N-P) & y(N-P+1) & \dots & y(N) \end{bmatrix} \tag{2}$$

In (2), $y(.)$ and N indicate the original sample and number of samples in the particular window, respectively. Additionally, the matrix pencil parameter P is defined between N/3 to N/2 [24]. The matrix X has dimension of $(N-P)\times(P+1)$. The singular-value decomposition (SVD) of X gives

$$\mathbf{X} = \mathbf{U}_Z \, \mathbf{S} \, \mathbf{U}_V^H \tag{3}$$

where $\mathbf{U}_Z$ and $\mathbf{U}_V$ denote unitary matrix including eigenvectors of $\mathbf{Z} = \mathbf{X}\mathbf{X}^H$ and $\mathbf{V} = \mathbf{X}^H\mathbf{X}$, respectively. Also, $(.)^H$ represents complex conjugate. The diagonal matrix $\boldsymbol{S}$ is formed through the singular values of $\mathbf{X}$, and a criterion is set as $10^{-k} \le \eta_d/\eta_{max}$ to separate the dominant singular values, reducing the noise impact [25].

In the proposed criterion, k defines the filtering factor. Moreover, η_d and η_{max} represent the dominant and maximum singular amounts, respectively. The columns of $\mathbf{S}$ associated with the dominant singular values are stored, and the remaining columns are removed. Then, the reduced matrix is obtained as $\mathbf{S}'$. Matrix $\mathbf{U}_V$ is diminished to $\mathbf{U}'_V$ by saving its columns associated with the dominant singular values and deleting the remained columns [24]. Then, we can define two new matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ as

$$\mathbf{X}_1 = \mathbf{U}_Z \, \mathbf{S}' \, \mathbf{U}'^H_{V,1} \;\; , \;\; \mathbf{X}_2 = \mathbf{U}_Z \, \mathbf{S}' \, \mathbf{U}'^H_{V,2} \tag{4}$$

where $\mathbf{U}_{V,1}$ and $\mathbf{U}_{V,2}$ are calculated by omitting the last row and first row of $\mathbf{U}'_V$, respectively. Moreover, $z_i = e^{s_i T_s}$ denotes the generalized eigenvalues of the matrix pair $\{\mathbf{X}_1, \mathbf{X}_2\}$ with the sampling period of T_s. The problem of solving for z_i can be determined by the ordinary eigenvalue problem [26] as

$$|\mathbf{X}_1^+\mathbf{X}_2 - \lambda \, \mathbf{I}| = 0 \tag{5}$$

where λ shows the non-zero eigenvalue and $(.)^+$ denotes Moore–Penrose pseudoinverse.

Through the calculation of eigenvalues and reformulations, the complex frequencies s_i in each moving window are obtained as

$$\alpha_i \mp j\omega_i = -\, ln(\lambda_i)/\, T_s \;, \quad \forall i = 1, 2, \dots, M \tag{6}$$

Since the eigenvalues are obtained by this step, the residues ζ_i are calculated from Eq. (7). These residues represent the magnitude and phase information of each pole in the measurement.

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \dots & z_M^{N-1} \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_M \end{bmatrix} \tag{7}$$

The procedure described in (2)–(7) is executed repeatedly within each moving window, introducing the STMPM concept by giving time-varying event signatures.

2.3. Proposed moving short-time waveform analytics for feature selection

The algorithmic feature selection is developed to detect signal distortions during events. Three unique features, such as damping factor, frequency, and residue's magnitude, are obtained from pre-processed waveform data using the following algorithm:

Algorithm 1 Waveform Analytics Algorithm

Data: Step size, length of the sliding window, pre-processed waveform voltage and current data, event monitoring period.
Result: Time-varying event signatures as inputs for grid-aware learning
Initialization;
while *sliding window is within the event monitoring period* **do**
- Move the sliding STMPM window along the waveform signal with an adjusted step size;
- Construct a sample matrix as (2) in each moving window;
- Apply singular value decomposition (SVD) to the matrix (2) to compute matrices $\mathbf{U}_z$, $\mathbf{S}$ and $\mathbf{U}_v$;
- Choose dominant singular values from $\mathbf{S}$ using the selected filtering factor k;
- Create matrices $\mathbf{S}'$ and $\mathbf{U}'_v$ by selecting the dominant singular values;
- Construct matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ by formulas in (4);
- Calculate the eigenvalues from (5) for matrix $\mathbf{X}_1^+\mathbf{X}_2$;
- Calculate the dominant complex frequencies by eigenvalues and sampling period as (6);
- Obtain the residues associated with dominant complex frequencies s_i by (7);
- Rank the residues based on their magnitude;
- Store the largest amplitude and its corresponding complex frequency in each sliding window;

end

These derived event signatures are then given to the grid-aware learning model to locate and classify the events, which is discussed in the next section.

3. Grid-aware event type and location classification

A grid-aware learning model is designed for classifying event type and location. To apply this method, the electric grid and its measurements are represented as graph-structured data. A distribution grid is represented by a graph $\mathbf{Gr} = (B, E)$, where B is the set of N nodes, and E shows the set of I edges. The graph signal is the nodal information, defined as $\mathbf{Z} \in \mathbb{R}^{N\times M}$. M is the length of the nodal feature. The adjacency matrix of this graph is defined as $\mathbf{A} \in \mathbb{R}^{N\times N}$. The matrix $\mathbf{A}$ contains the grid's topological information such as the lines' connectivity status between nodes. By including the edges' features with the distance-based metric [17], $\mathbf{A}$ becomes a weighted adjacency matrix incorporating more physical features of the grid. The diagonal matrix $\mathbf{D}$ shows the nodal degree of the graph, indicating the number of neighboring nodes to each particular node. According to the given fundamental definitions, the $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-0.5}\mathbf{A}\mathbf{D}^{-0.5}$ calculates the normalized graph Laplacian matrix, and its spectral decomposition is represented as (8) based on the Fourier basis $\mathbf{U}$.

$$\mathbf{L} = \sum_{n=1}^{N} \lambda_n \, \mathbf{u}_n \, \mathbf{u}_n^T = \mathbf{U} \; diag\{\lambda_1, \ldots, \lambda_N\} \; \mathbf{U}^T \tag{8}$$

Graph filter works as an operator to modify input $\mathbf{Z}$ with the frequency response h for each eigenvalue λ_n. The outcome of this filtering is

$$\overline{\mathbf{Z}} = \sum_{n=1}^{N} h(\lambda_n) \, \mathbf{u}_n \, \mathbf{u}_n^T \, \mathbf{Z} \tag{9}$$

where the filter response $h(\lambda_n)$ has an important impact on the performance of the model [16].

3.1. Convolutional ARMA filtering on graph

The robust ARMA filter is able to approximate a large variety of filter responses [21]. Mathematically, the ARMA filter's response with order G reads as follows,

$$h_{A,G}(\lambda) = \left(1 + \sum_{g=1}^{G} q_g \lambda^g\right)^{-1} \left(\sum_{g=0}^{G-1} p_g \lambda^g\right) \tag{10}$$

Consequently, the filtering representation on node space using the ARMA filter is

$$\overline{\mathbf{Z}} = \left(\mathbf{I} + \sum_{g=1}^{G} q_g \mathbf{L}^g\right)^{-1} \left(\sum_{g=0}^{G-1} p_g \mathbf{L}^g\right) \mathbf{Z} \tag{11}$$

If one equates q_g to zero in (11), the polynomial filter is obtained, which has been used in previous works [17,27]. This is one of the advantages of the ARMA filter that can generalize different filters' functions. Also, the added auto-regressive expression makes the model robust against the noises and captures the more global graph structure [21].

Since Eq. (11) requires the slow inversion calculation of the matrix, an iterative-based method is proposed [21] through which the effect of the first-order ARMA filter is approximated as

$$\overline{\mathbf{Z}}^{(t+1)} = a\bar{\mathbf{L}} \, \overline{\mathbf{Z}}^{(t)} + b\mathbf{Z} \tag{12}$$

where $\bar{\mathbf{L}} = 0.5\left(\lambda_{max} - \lambda_{min}\right)\mathbf{I} - \mathbf{L}$, showing a linear transformation of $\mathbf{L}$. The eigenvectors of $\bar{\mathbf{L}}$ is the same as $\mathbf{L}$, but its nth eigenvalue is derived by $\bar{\lambda}_n = 0.5\left(\lambda_{max} - \lambda_{min}\right) - \lambda_n$, where λ_n is the nth eigenvalue of $\mathbf{L}$. Moreover, $\mathbf{Z}^{(t)}$ is the output of filter in tth iteration, $\mathbf{Z}$ contains the inputs, and a and b denote arbitrary coefficients. According to [28], the response of the ARMA filter with the first order can be analyzed by the convergence in

$$\overline{\mathbf{Z}} = \lim_{t\to\infty} \left[(a\bar{\mathbf{L}})^{(t)}\overline{\mathbf{Z}}^{(0)} + b\sum_{i=0}^{t}(a\bar{\mathbf{L}})^i\mathbf{Z}\right] = b(\mathbf{I} - a\bar{\mathbf{L}})^{-1} \tag{13}$$

Consequently, the response of the 1-order ARMA filter is

$$h_{A,1}(\bar{\lambda}_n) = b\,(1 - a\bar{\lambda}_n)^{-1} \tag{14}$$

The summation of G ARMA filters obtains the analytical form of the ARMA filter with order G, resulting in the following filtering operation on the inputs:

$$\overline{\mathbf{Z}} = \sum_{g=1}^{G}\sum_{n=1}^{N} b_g(1 - a_g\bar{\lambda}_n)^{-1} \, \mathbf{u}_n \, \mathbf{u}_n^T \, \mathbf{Z} \tag{15}$$

where the frequency response of ARMA with order G is

$$h_{A,G}(\bar{\lambda}_n) = \sum_{g=1}^{G} b_g(1 - a_g\bar{\lambda}_n)^{-1} \tag{16}$$

The neural network implementation for 1-order ARMA filter uses recursive update in Eq. (12) in a specified T iterations using the following graph convolution skip (GCS) layer,

$$\overline{\mathbf{Z}}^{(t+1)} = \sigma\left(\bar{\mathbf{L}} \, \overline{\mathbf{Z}}^{(t)}\Theta + \mathbf{Z} \, \Omega\right) \tag{17}$$

where the Θ and Ω represent the trainable weights, and $\mathbf{Z}$ contains the initial feature of nodes. The filter is generalized to G-order with $\overline{\mathbf{Z}} = \frac{1}{G}\sum_{g=1}^{G}\overline{\mathbf{Z}}_g^{(T)}$, where $\overline{\mathbf{Z}}_g^{(T)}$ indicates the output of the gth stack in the last GCS layer. The schematic of the ARMA graph convolutional is given in the next section.

4. The proposed multi-stage framework for enhanced situational awareness

The WMUs are assumed to be located on a few critical three-phase nodes of the grid, and the Clarke transformation is employed to convert three-phase waveform data to decoupled modal waveform signals.

The WMUs are assumed to be located at a few critical three-phase nodes of the grid, and the Clarke transformation is employed to convert three-phase waveform data into decoupled modal waveform signals. This transformation yields the space phasor mode of three-phase measurements, which is informative for event analysis and has been used in various studies, such as voltage dip classification [29], fault detection and classification [30], and fault localization [6]. This three-phase transformation is referred to as our pre-processing step, which also facilitates post-processing of the signal in the waveform analytics algorithm, preserving the information from all three phases and reducing the feature dimensions with aerial mode [6]. The waveform analytics Algorithm 1 is depicted in Fig. 1 and represents the feature selection stage, characterizing events' signatures.

The computed features from voltage and current measurements at different locations are stored as a measurement matrix ($\mathbf{Z}$) under a particular labeled event with its location. Specifically, each row of $\mathbf{Z}$ contains the stream of features from one specific measurement location in the network. As an example, $\mathbf{Z}$ has 13 rows when representing the structured graph data of a 13-node system. Moreover, if we label the nodes of the 13-node distribution system (graph) from 0 to 12, and WMUs are located at nodes 1, 6, and 8, $\mathbf{Z}$ contains features for the rows associated with these nodes. Thus, the remaining rows representing nodes without WMUs inevitably have zero values [31]. The length of the features (column dimension of $\mathbf{Z}$) is equal to the total number of STMPM-based features, i.e., damping factor, angular frequency magnitude, and residue's absolute values. Thus, each feature matrix ($\mathbf{Z}$) is a sample associated with a particular label of event and its location, used for training the grid-aware model. Theoretically, the rows of $\mathbf{Z}$ containing characterized voltage and current data in each sample (scenario) are related to each other through the topology of the distribution grid, although they are monitored by different independent WMUs.

The proposed grid-aware learning approach takes into account the spatial relationships by transforming the grid's topology and measurements into structured graph data, which is processed through graph

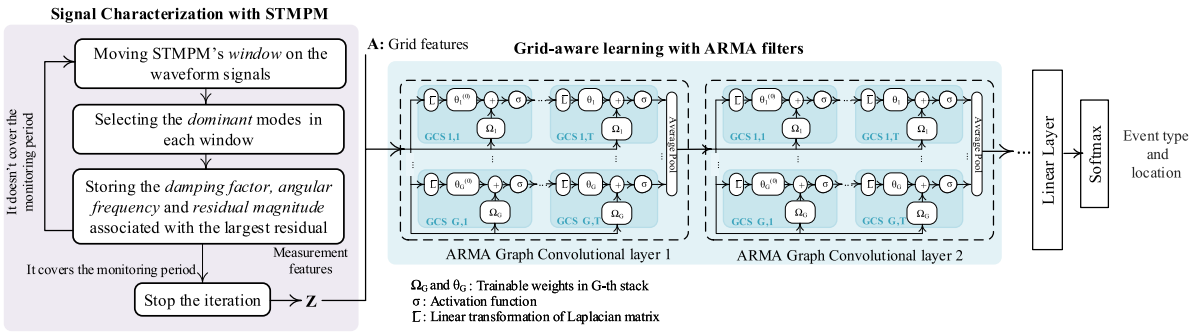

Fig. 1. The proposed framework for event characterization, classification, and localization.

convolution filters. The outputs of graph convolution layers are given to a linear layer, transferring the features into the Softmax classifier as illustrated in Fig. 1. Adam optimizer is used, and the cross-entropy is chosen as the loss function [16]. The loss is minimized by updating the model parameters in the learning process and comparing the predicted and actual classes. After training the model, test samples are provided for model evaluation during the event identification and localization period. In real-world applications, the event identification or monitoring period is updated every few cycles. The grid-aware model consisting of ARMA graph convolution, transformation layers, and softmax is designed by Pytorch Geometric [32].

5. Case studies and simulation setup

In this section, several events are studied on the modified active unbalanced IEEE 13-bus system using PSCAD automated by its Python API [33]. Different events in various locations, in addition to normal operation scenarios, are simulated under different loading, DER generation, event inception angles, and noises [6]. The schematic of the distribution system is also shown in Fig. 2. Table 1 indicates the event labels, types, locations, and occurrence duration. *P* stands for permanent events in the table. We also generate numerous transient disturbances by simulating events within 1–2 cycles.

The events are broadly categorized as CB switching, sudden load changes, different faults (types 1 and 2), and DER outages in different locations. Fault identified with type 1 refers to faults that are often accompanied by arcing with high impedance contact to the earth. This fault requires accurate circuit modeling, consisting of voltage sources, diodes, and variable resistors, which are simulated in our work as described in [34]. The accuracy of the circuit model is experimentally validated in [35]. Since asymmetrical single-phase high impedance faults (HIFs) are frequent, they are studied as type 1 on different phases and locations of the grid. Furthermore, this study also examines other prevalent asymmetric and symmetric regular faults that have lower contact impedance with the Earth. These fault types are categorized as type 2 faults, specified by low impedance faults (LIFs). It is worth noting that faults may occur in any location, as studied in our previous work [36].

It is assumed that there are three synchronized WMUs (SWMUs) located at buses 1, 6, and 8, recording the voltage and current waveform data with a reporting rate of 256 samples-per-cycle. Lower sampling rates, such as 32, will also be studied in another case study (Section 5.5) to test the robustness of the approach. Gaussian noise with different signal-to-noise ratios (SNR) of 45 dB, 50 dB, 55 dB, and 60 dB is added to the waveform signals to make them more representative of real-world scenarios [18].

In order to train and test the model, 6540 scenarios representing normal conditions and different possible events occurring in various locations are prepared. The datasets are divided into 20% for testing and 80% for validation and training. Two ARMA graph convolution layers are set with the same sizes of 128 as input channels. The ARMA

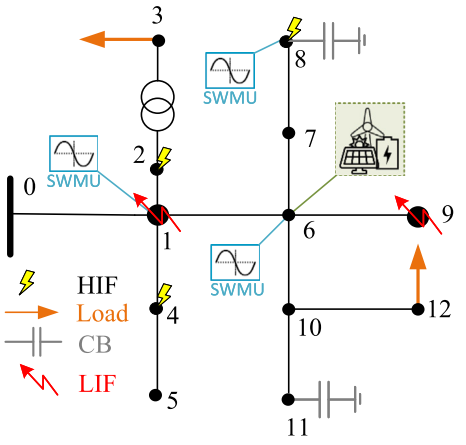

Fig. 2. The schematic of the unbalanced three-phase test system.

Table 1
Events' information.

Labels	Type	Duration	Location
0	3-phase CB switching	1 cycle, *P*	8
1	1-phase CB switching	1 cycle, *P*	11
2	3-phase DER outage	*P*	6
3	3-phase Load switching	1, 2 cycles	3
4	1-phase Load switching	1, 2 cycles	12
5	Fault (type 1)	1 cycle, *P*	2
6	Fault (type 1)	1 cycle, *P*	4
7	Fault (type 1)	1 cycle, *P*	8
8	Fault (type 2)	1, 2 cycles	1
9	Fault (type 2)	1, 2 cycles	9
10	Normal operation	N/A	N/A

layers are configured with $G = 2$ and $T = 3$. The batch size is 8. These values are adjusted with search-based tuning experiments in 100 epochs [6]. The ARMA layers are followed by a linear layer with 256 channels. The learning rate and dropout value are 0.0001 and 0.5, respectively [17]. To better assess the performance of the proposed grid-aware learning method, additional metrics are defined in addition to the accuracy metric, such as macro-average $F1$ ($M-F1$), macro-average recall ($M-Rec$), and macro-average precision ($M-Pre$).

5.1. Time-varying characterization of event signatures

In this section, two transient events, such as CB switching and fault, recorded by 256 samples-per-cycle sensors, are tested using the STMPM method. The length of the sliding window is adjusted to cover 140 samples of the signal, and the step size of the moving window is equal to 40 samples. These parameters are tuned with search-based experiments and optimized for analyzing the simulated events. Another alternative length is 256 samples (equal to one cycle), which gives

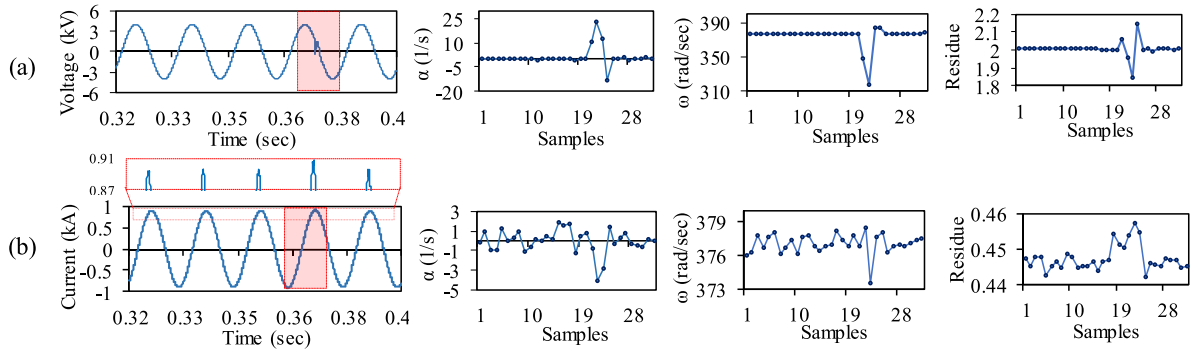

Fig. 3. The waveform signal and extracted time-based features under (a) CB switching, (b) HIF.

distinctive features. The filtering factor k and the pencil parameter $P = N/2$ are set as 3 and 70, respectively [20].

Fig. 3.a demonstrates a transient distortion of alpha Clarke mode voltage on bus 1 when the CB switching occurs at bus 8. The switching of CB creates new oscillation modes in the system, which are characterized by three time-varying parameters: damping factor, angular frequency, and residue magnitude. The dominant parameters associated with the largest residues are obtained by sliding the STMPM window over the waveform signal, shown on the right-hand side of the sinusoidal measurement for 32 consecutive steps. This sliding process covers the entire 5-cycle monitoring period, which is then used for event classification and localization. This duration can be adjusted without loss of generality, but we found it effective in immediate and accurate identification. For instance, the angular frequency drops to 348.29 rad/sec and then to the minimum of 316.95 rad/sec when the signal distortion is observed in the sliding window in the 21st and 22nd steps, respectively. However, the angular frequency is around the fundamental frequency of $377 \approx 2\pi 60$ rad/sec before the event occurrence. Other obtained parameters also show abrupt changes once the event is detected in the moving window. Note that different events may either create new dominant modes, like in the CB switching, or change the existing fundamental modes. Usually, new modes are not created during HIFs, but the existing fundamental modes are changed slightly.

Now, we consider a second scenario where an HIF occurs at bus 2, and the WMU records the noisy current at bus 1. The transformed current is shown in Fig. 3.b. Since the impact of HIF is hard to observe on the recorded waveform data, the zoomed version of the current is shown on the top of the raw waveform signal. As shown in Fig. 3.b, the current magnitude increases slightly under the occurrence of this fault, making it hard to detect by the protection devices in real-world situations. However, the STMPM helps capture sudden insignificant variations, although considerable noise is added to the signal. Three computed parameters pinpoint the instantaneous changes caused by this transient event. Interestingly, the damping factor shows more sensitivity against the hard-to-detect HIF compared to other computed features. This informative feature is obtained thanks to the proposed STMPM algorithm.

In this study, the distinctive time-varying features are collected for both current and voltage at different locations, which includes 96 features for each alpha mode voltage and current signal. Features are normalized with a z-score for use in the grid-aware event type and location classification algorithm.

5.2. Classification performance under different waveform features

In this section, we evaluate the impact of the extracted features from synchronized measurements on the performance of event type and location classification. We intend to show that the waveform measurements provide richer information and that relying solely on signal frequency or magnitude is inadequate for improved situational awareness.

Table 2
Grid-aware classification performance with different computed features.

Features	A-Acc %	M-F1%	M-Pre %	M-Rec %
ω	95.90	95.93	96.12	95.84
α	96.36	96.45	96.71	96.39
r	97.99	98.00	98.09	97.95
α, ω	97.64	97.67	97.76	97.63
α, r	98.51	98.48	98.66	98.35
ω, r	98.73	98.74	98.86	98.66
α, ω, r	99.12	99.11	99.14	99.09

In Table 2, the performance of the proposed grid-aware approach is shown considering different computed features. All scores are obtained by taking the average of five independent experiments. In the case of the individually selected feature, the residue's magnitude contains a wealth of distinct information about the event, resulting in an impressive 97.99% average accuracy (A-Acc). Considering two sets of features, the combination of angular frequency with residue enhances A-Acc to 98.73%, while the other two features show slightly lower scores. This indicates that the residue contains more distinct event-related features, which is intuitive as it captures the features associated with the magnitude of the measured signal. However, when all three extracted features are included in the model, accuracy reaches the maximum level of 99.12%, and there is also an enhancement in other classification metrics. This indicates the fact that although a feature like residue's magnitude helps more in the identification of the event, including other features (sensitive damping factors) enhances the performance by creating more unique characteristics for events with minor and transient signal distortions. Thus, using all three features is suggested to classify and localize the events in the following analysis.

5.3. Configuration of WMUs

In practice, the captured measurements from various WMUs positioned along the feeders are becoming accessible, motivating us to design the grid-informed method to take advantage of the spatial correlation of waveform measurements at different locations to improve event classification and localization performance. However, the scarcity of buses equipped with WMUs along the feeders persists due to factors such as installation and purchase costs. Thus, we are testing the robustness of the proposed grid-aware method for as few as two WMUs with different configurations, addressing the challenges associated with locational scarcity in sensors' information.

Three configurations are defined for two WMUs in two different buses as *A*: [1, 6], *B*: [1, 8], and *C*: [6, 8]. The locations of the WMUs are heuristically selected on three-phase nodes to improve the observability region of the feeder. Optimal placement of sensors is not among the scope of this work. The classification accuracy decreases

Table 3
The impact of WMUs' configuration on the model performance.

Configuration	A-Acc %	M-F1%	M-Pre %	M-Rec %
A	98.96	98.95	98.99	98.91
B	95.15	95.41	95.50	95.44
C	95.27	95.41	95.46	95.41

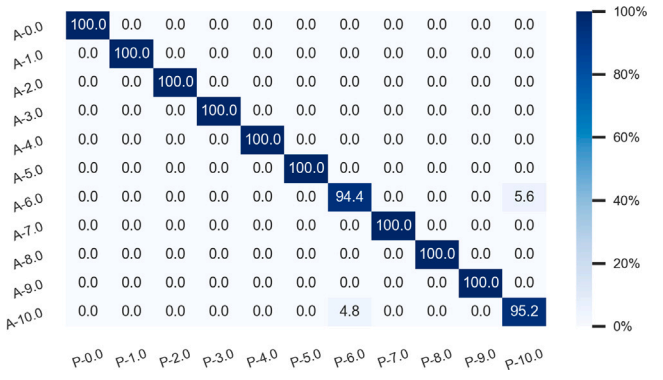

Fig. 4. Confusion matrix for 13-bus system (A: actual label P: predicted label).

slightly to 98.96% for configuration *A* compared to the base case with three WMUs. This accuracy is higher than the classification accuracy for configurations *B* and *C*. Other classification scores are also higher with this configuration, as shown in Table 3. This is because the placement of WMUs at nodes 1 and 6, as configured in *A*, enables the GNN layers to learn the situational changes better in neighboring buses with this configuration. This assessment indicates the robustness of the model for as few as two WMUs in particular locations.

5.4. Comparison with state-of-the-art models

To highlight the superior performance of the grid-aware learning models (GCN and ARMA), a comprehensive comparison is made with well-established classification models decision tree (DT), k-nearest neighbors (kNN), random forest (RF), Gaussian Naive Bayes (GNB), and Bernoulli Naive Bayes (BNB). These implemented state-of-the-art models have shown promising performance in recent power system studies. kNN and DT are proposed as efficient classifiers for event and fault identification in distribution feeders using real-world data and comprehensive simulations [11,37,38]. In the event classification problem with imperfect real-world data, the RF model outperforms all classifiers analyzed in [39]. RF shows effective performance in classifying power quality events in [40]. In [41], Naive Bayes (probabilistic-based) classifier improves the fault identification in power grids.

5.4.1. Tuning parameters in baseline models

To ensure a reliable comparison, the parameters of the state-of-the-art models, such as RF, DT, and kNN, are optimized using a random search method to enhance their performance as much as possible [42]. Random search is an efficient method that can find the optimal model in most cases while incurring a low computational burden [43]. The hyperparameters are optimized through a 4-fold cross-validation process [44], with data split into 20% for testing and 80% for training and validation. The characterized waveform data are employed across all folds. As for GCN, parameters and the optimizer are selected based on the previous studies and experiments [16,17], leading to good performance for the current problem. The same learning rate and dropout values are used, as discussed earlier, for the ARMA model.

Table 4
The comparison between the proposed method and state-of-the-arts.

Approach	A-Acc %	M-F1%	M-Pre %	M-Rec %
BNB	83.32	84.51	85.17	84.45
GNB	83.79	84.78	85.76	84.60
kNN	89.95	90.55	93.43	91.55
DT	93.32	93.61	93.68	93.63
RF	97.31	97.37	97.78	97.20
GCN	98.28	98.31	98.40	98.25
$ARMA$	99.12	99.11	99.14	99.09

Table 5
The computation time comparison of different classifiers.

Classifiers	BNB	GNB	kNN	DT	RF	GCN	ARMA
Computation time (s)	0.016	0.015	0.717	0.005	0.024	0.565	1.068

5.4.2. Performance comparisons

After adjusting the parameters for the baseline models, a detailed comparison is conducted, and the results are tabulated in Table 4. Table 4 reports that the grid-aware ARMA model outperforms GCN and widely used machine learning (ML) models that do not incorporate topological information. This implies that the inclusion of topological features and the use of more robust graph filters, such as ARMA, lead to improved classification accuracy. As shown in Table 4, after the ARMA convolution, grid-aware GCN showcases better classification capabilities compared to topology-agnostic ML algorithms due to including topological features. It is important to highlight that the superior performance of the ARMA model over the GCN is attributed to ARMA's more generalized filter, facilitating enhanced learning of spatial features, as detailed in Section 3. Additionally, it should be noted that both grid-aware methods surpass grid-agnostic machine learning approaches in performance.

The performance of the grid-aware ARMA model is shown by the confusion matrix in Fig. 4 for the discussed experiment in detail. This matrix indicates the classification accuracy for each event and their location. It displays the accurate classification percentages on the diagonal elements, while the off-diagonal elements represent wrong classification percentages. Small misclassifications exist between arcing HIFs in three locations, indicated by labels 5, 6, and 7, and normal operations, represented by label 10 in the confusion matrix, supporting our discussion and numerical analysis. Other events are accurately classified and located in this experiment. This distinction enables accurate crew dispatch and immediate, as well as appropriate, responses to events. This is particularly more important for arcing HIFs, which could be life-threatening and initiate wildfires [45].

Furthermore, we provide details on the computation time of each method for the test dataset, which is 20% of our large dataset in Table 5. It is shown that the decision tree has the minimum computation time of 0.005 s due to the simplicity of its algorithm. However, the computation times of the graph-based solutions are 0.565 s and 1.068 s for GCN and ARMA, respectively, which are higher compared to other conventional ML models. This is because graph learning-based has a more sophisticated structure for learning the features, leading to better results. Please note that their computation time still makes them suitable for practical implementation, as it is only around one second or less, depending on the type of graph filters used.

5.5. Evaluation in a microgrid with slower sampling rate of WMUs

In this section, we test our approach on another system using the real-time simulation setup to classify and locate events. This setup is illustrated in Fig. 5 and has five main components:

(1) The real-time digital simulation (RTDS) is used for electromagnetic transient (EMT) analysis of an existing 8-bus three-phase

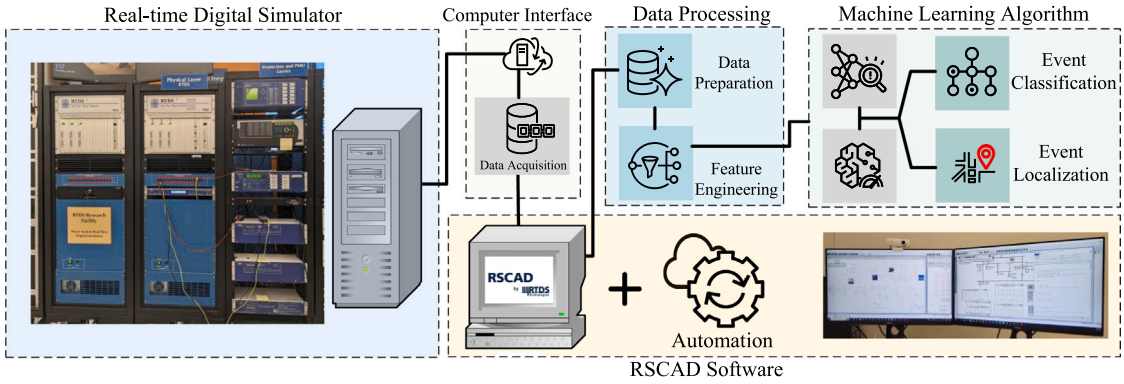

Fig. 5. The framework of the real-time simulator setup.

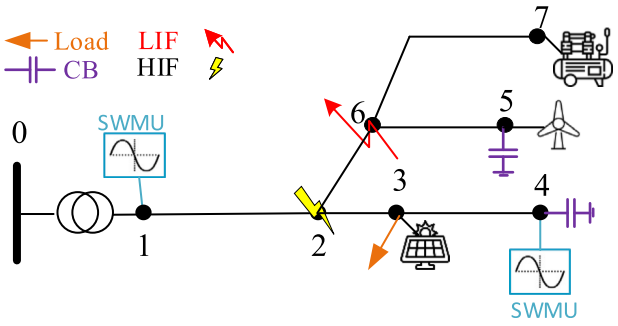

Fig. 6. The schematic of the three-phase microgrid in RTDS.

microgrid with inverter-based PV, wind turbine, and diesel generator [46]. The compact schematic of this system with diverse sources and the base transformer's secondary voltage of 13.2 kV is shown in Fig. 6.

(2) A computer is utilized as an intermediary to collect the measurements under different types of events such as CB switching, sudden load outage, and high and low impedance faults. The events' locations are indicated in Fig. 6.

(3) The RSCAD software is employed to run and control the simulation of events in five distinct locations using our developed automation script in the programming environment of RSCAD software.

(4) In this step, algorithms are developed for preparing the raw waveform data and processing them using the STMPM-based feature extraction model.

(5) Finally, the proposed ARMA graph convolution receives the extracted features to classify the events with their associated locations.

The WMUs are assumed to be located at buses 1 and 4, with a sampling rate of 32 samples per cycle. This lower sampling rate is chosen to evaluate the effectiveness of the proposed method under such conditions. Three measurement noise levels are taken into account: 80 dB (less noisy), 60 dB (noisy), and 40 dB (very noisy). The events are identified by corresponding labels in parentheses, which include CB switching at bus 4 (0), CB switching at bus 5 (1), HIF occurrence at bus 2 (2), LIF occurrence at bus 6 (3), load switching at bus 3 (4), and regular operation (5).

The proposed grid-aware model achieves an average accuracy of 96.97% with an F1 score of 96.06% in five separate experiments. Fig. 7 presents the confusion matrix for one of the experiments. Misclassification between regular operation and HIF also occurs in this case due to minor signal distortion during HIF and significant noise added during regular situations. However, as shown in Fig. 7, most scenarios are accurately predicted across all events, aligning with the calculated high accuracy. Thus, this case further validates the effectiveness of the proposed method under various events, even with low sampling rate WMUs and a more complex grid with diverse generation resources.

5.6. Analysis and discussion about the impact of short-time moving window in STMPM

In this section, we explore two scenarios to highlight the significance of employing STMPM with short-time moving windows in improving event classification in the microgrid. The baseline cases are detailed as follows:

Case 1: Matrix pencil method (MPM) is employed, which was successfully used in [47] for real-world event identification. MPM operates *without* a sliding window, computing event signatures within a window that spans the entire monitoring duration.

Case 2: The STMPM is used with a *larger* sliding window and step size. We have doubled the length of the sliding window and increased the step size to 32.

Table 6 displays the accuracy and F1 scores for both cases, as well as the dropped scores, compared to the results obtained for the microgrid test system in Section 5.5. In case 1, the performance declines significantly when choosing to use MPM without a sliding window. This is because MPM does not provide a detailed time-varying signature that represents the evolving or changing modes' features, such as damping factor, angular frequency, and residues over time. In Case 1, MPM cannot differentiate between normal operating conditions and HIF by only analyzing the modes within a single, large window. This means that the dominant modes usually do not change under HIF compared to normal conditions, so differentiating them may not be possible. We also observe partial misclassification between loads and CB switching in different experiments using MPM. This method also deviates from the localization of events of the same type as CB switching occurring in buses 4 and 5 in different experiments. The maximum misclassification error is 8.2% in the case of locating the CB switching in our various experiments. In Case 2, the performance also declines when a wider window is selected in STMPM, making the characterization of events less specific. However, it still works much better than case 1 in which MPM is used for the characterization of the waveform. The confusion matrix associated with one of the experiments is shown in Fig. 8 for case 2. Misclassification occurs among other types of events, not just HIFs and normal operation, a scenario not observed in the framework of Section 5.5 that employed an optimized smaller window and step length.

The results presented in this section underscore the effectiveness of a precise methodology capable of extracting unique patterns from three time-varying features, thereby enhancing classification accuracy. The significance of incorporating all three time-based elements — damping factor, angular frequency, and residues — is elaborated in Section 5.2. Moreover, this section further demonstrates that identifying time-varying patterns across a moving window yields superior results compared to analyzing these features within a solitary window. The distinctive characterization of waveform measurements becomes more important when studying the same events in different locations, which

Table 6
The impact of moving window on the event classification performance.

Strategy of waveform analytics	A-Acc (Dropped score) %	M-F1 (Dropped score) %
Case 1	80.31 (−16.66)	69.40 (−26.66)
Case 2	91.33 (−5.64)	89.95 (−6.11)

Fig. 7. Confusion matrix for Microgrid (A: actual label P: predicted label).

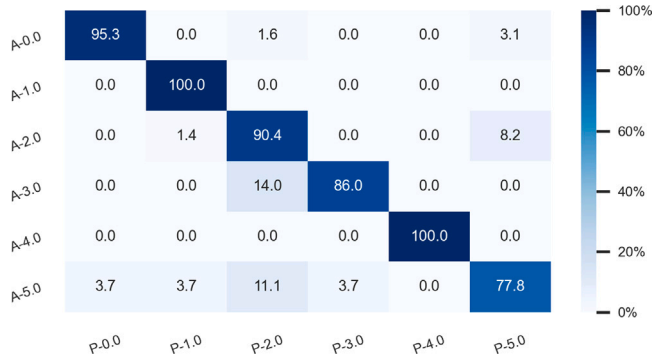

Fig. 8. Confusion matrix when using STMPM with larger moving window and steps (A: actual label P: predicted label).

have closely related signatures, or events that can sometimes have almost similar impacts on the system measurements, e.g., increasing voltage under load reduction and CB connection. Consequently, we propose STMPM-based feature extraction with appropriate adjustment of the window, which moves over small steps and derives *distinctive* shapes over time, associated with the type and location of the events.

5.7. Further discussion about characterizing power quality events

In this section, various scenarios are considered where common patterns, such as swell, sag, and spike, occur, and the waveform analytics algorithm is implemented to characterize these patterns. In Fig. 9.a, a signal swell occurs, and the magnitude of the signal increases. This is nicely characterized by the residue, which steps up, indicating the increased magnitude of the signal. The sharp deterioration seen due to the starting of signal swelling is featured by complex frequency (damping and angular frequency) as spikes. In Fig. 9.b, a small notch is observed and clearly characterized with three time-varying extracted features of the waveform analytics algorithm. Before this distortion, all three features represent fundamental values ($\alpha \approx 0$, $\omega \approx 377$) with a constant magnitude of the signal, as shown in the residue. The damping factor indicates more significant variations, supporting its high and effective sensitivity to transient distortion. In Fig. 9.c, a spike occurs, and this sharp distortion is represented and located by very sharp spikes in the three computed features. In Fig. 9.d, a sag occurs in the signal accompanied by flicker and small oscillations around the peaks. The residue first indicates this event as a step-down drop, followed by a slight increase in magnitude. The complex frequency terms represent this event with fluctuations.

Please note that signals a to d were already created in our RTDS-based simulations and used for the identification and localization of the events in Section 5.5. However, new signals e and f are created, presented in Fig. 9 to further demonstrate the potential of our method. In Fig. 9.e, two notches are created and distinctly characterized over time. The interval between these two distortions, where the sliding window does not observe any distortion, is characterized smoothly, similar to the computed values before the first notch. This indicates the time localization capability of our method. In Fig. 9.f, the frequency of the signal changes from 60 Hz to 40 Hz. The angular frequency clearly shows this reduction as a step change. The damping factor and residue only show sudden changes related to the point where this transition from 60 Hz to 40 Hz occurs in the signal, creating small changes in the data.

5.8. Computation time of waveform analytics algorithm

In this section, we explore the impact of parameters on the computation time of the waveform analytics model. A batch of datasets with 1120 streams of data, created by RTDS in Section 5.5, is selected for this evaluation.

In the base case study with $k = 3$, $P = 16 = \frac{N}{2}$, a step size of 8, and a window length of 32, where N indicates the number of samples in the window, computation time for 1120 scenarios is 16.24 s. In practice, we usually analyze one event (one scenario) at a time, so the proposed method is applicable for real-time implementation. However, some studies may not always be concerned with real-time implementation and could focus on post-event analysis of massive, recorded datasets. Therefore, designing computationally efficient and technically solid methods is key for *different* power system studies, and three cases are investigated by adjusting parameters such as k, step size, and window size in this regard:

Case 1: We set k to 8, which increases computation time to 17.37 s. This parameter is intended to select the dominant eigenvalues using the constraint $10^{-k} \leq \eta_d/\eta_{max}$. Thus, increasing this value can lead to choosing more modes and raising the computation time *unnecessarily*.

Case 2: We reduce the matrix pencil parameter P to 5, which changes the dimension of the matrices for the mathematical decomposition and slightly reduces the computation time to 14.68 s. However, it is recommended to set the parameter within the range of $N/2$ and $N/3$ to better *eliminate* the effect of noise [24].

Case 3: The step size is set to 16 samples (doubled compared to the base case) to have the sliding window move faster over the recorded data. As a result, the computation time drops to 8.32 s for the datasets; however, this can *reduce* the event identification accuracy, as discussed in Section 5.6.

6. Conclusion

This study focuses on enhancing situational awareness in active distribution grids using a new class of waveform measurement units (WMUs) for power quality and transient arcing events. The main goal is to improve event characterization, classification, and localization through a hybrid model based on signal processing and topology-aware learning. The time–frequency signal decomposition method is applied to the WMU data using the short-time matrix pencil concept to extract a set of distinctive and less-explored signal features. The

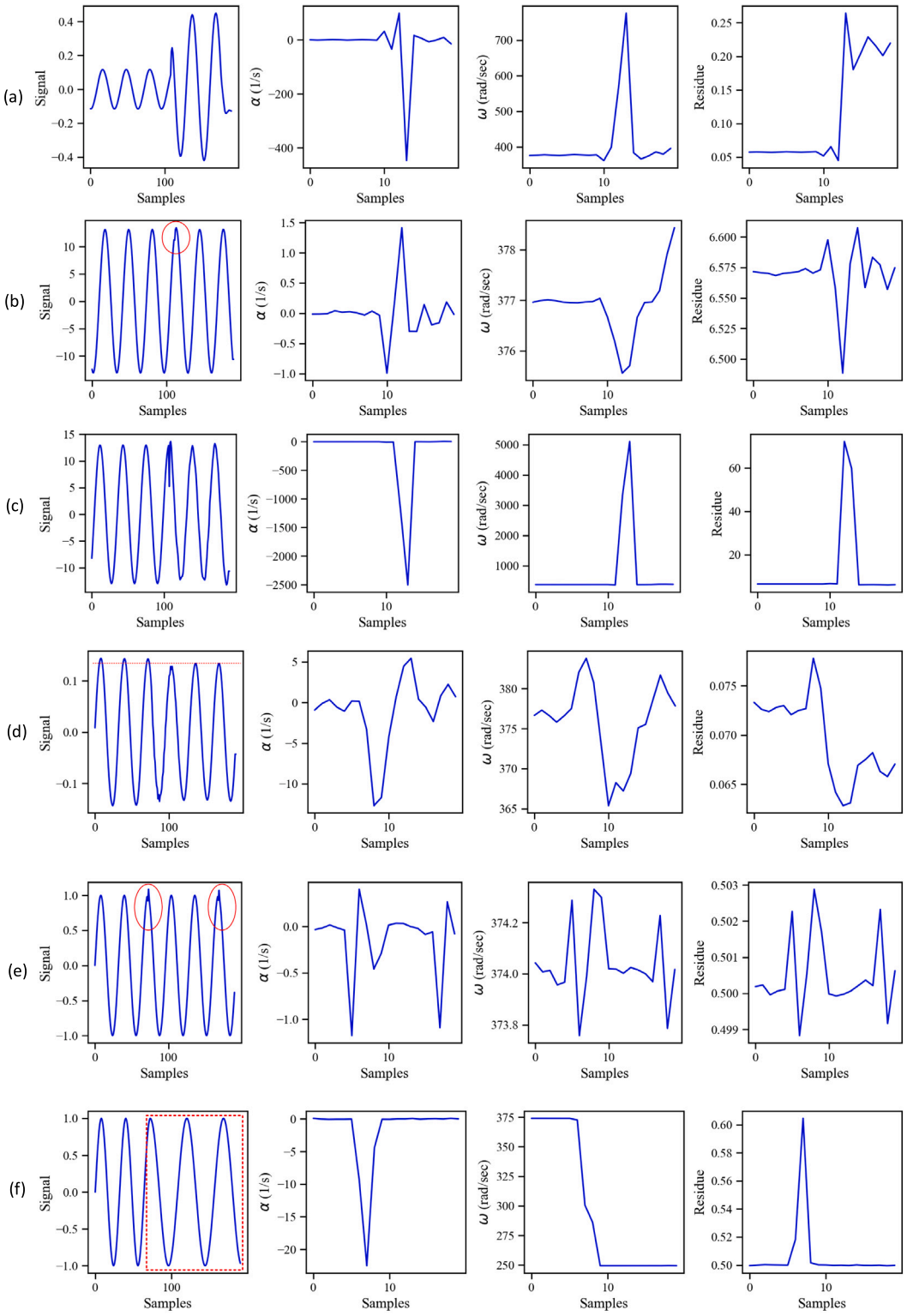

Fig. 9. Different power quality disturbances (a)–(f).

derived features consisting of dominant signal damping factors, angular frequency, and residues' magnitude are then used as distinctive event signatures for training a grid-aware learning model to classify and locate the events. The learning-based method is developed using the ARMA graph convolution filters, which incorporate physical distance information of the lines and nodal measurement data. This model bridges the gap between topology-agnostic machine learning models and hard-to-implement circuit-based models by creating an efficient grid-aware data-driven classifier that utilizes structured data from waveform measurements across the grid. Two test systems, the unbalanced 13-bus IEEE model and three-phase microgrid model, are considered to assess the performance of our model under a variety of events, such as capacitor bank (CB) switching, abrupt load changes, short circuit and unbalanced faults, arcing high impedance faults,

and distributed generation unit outages. Based on our numerical assessments, the proposed time-frequency waveform analytics, equipped with a sliding moving window, demonstrate promising potential for detecting small and transient distortions in high-fidelity waveform data, thereby improving event classification and localization with a grid-aware learning model. The proposed grid-aware model outperforms state-of-the-art classification models under noisy conditions and operational changes. It also demonstrates robust performance even with a limited number of synchronized WMUs and lower sampling rates.

CRediT authorship contribution statement

Mohammad MansourLakouraj: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hadis Hosseinpour:** Writing – review & editing, Visualization, Software, Methodology. **Hanif Livani:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Mohammed Benidris:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This material is based upon work supported by the National Science Foundation, United States under Grant ECCS-2033927.

References

[1] O. Ozgonenel, T. Yalcin, I. Guney, U. Kurt, A new classification for power quality events in distribution systems, Electr. Power Syst. Res. 95 (2013) 192–199.
[2] R. Pourramezan, H. Karimi, J. Mahseredjian, Synchrophasor network-based detection and classification of power system events: A singular value decomposition approach, Electr. Power Syst. Res. 223 (2023) 109645.
[3] A.F. Bastos, S. Santoso, W. Freitas, W. Xu, Synchrowaveform measurement units and applications, in: 2019 IEEE Power & Energy Society General Meeting, PESGM, IEEE, 2019, pp. 1–5.
[4] M. Izadi, H. Mohsenian-Rad, Synchronous waveform measurements to locate transient events and incipient faults in power distribution networks, IEEE Trans. Smart Grid 12 (5) (2021) 4295–4307.
[5] I. Niazazari, H. Livani, A PMU-data-driven disruptive event classification in distribution systems, Electr. Power Syst. Res. 157 (2018) 251–260.
[6] M. MansourLakouraj, H. Hosseinpour, H. Livani, M. Benidris, Waveform measurement unit-based fault location in distribution feeders via short-time matrix pencil method and graph neural network, IEEE Trans. Ind. Appl. (2022) 1–10, http://dx.doi.org/10.1109/TIA.2022.3231586.
[7] I. Niazazari, et al., Event cause analysis in distribution networks using synchro waveform measurements, in: 2020 52nd North American Power Symposium, NAPS, IEEE, 2021, pp. 1–5.
[8] J.A. Wischkaemper, C.L. Benner, B.D. Russell, K. Manivannan, Application of waveform analytics for improved situational awareness of electric distribution feeders, IEEE Trans. Smart Grid 6 (4) (2015) 2041–2049, http://dx.doi.org/10.1109/TSG.2015.2406757.
[9] O. Ozgonenel, S. Karagol, Transformer differential protection using wavelet transform, Electr. Power Syst. Res. 114 (2014) 60–67.
[10] R.A. Ghunem, R. El-Shatshat, O. Ozgonenel, A novel selection algorithm of a wavelet-based transformer differential current features, IEEE Trans. Power Deliv. 29 (3) (2014) 1120–1126.
[11] A. Shahsavari, M. Farajollahi, E.M. Stewart, E. Cortez, H. mohseni Rad, Situational awareness in distribution grid using micro-PMU data: A machine learning approach, IEEE Trans. Smart Grid 10 (6) (2019) 6167–6177.
[12] S. Xiong, et al., Incipient fault identification in power distribution systems via human-level concept learning, IEEE Trans. Smart Grid 11 (6) (2020) 5239–5248.
[13] M.B. Mohtasham, A. Jalilian, Classification of multi-stage voltage sags and calculation of phase angle jump based on clarke components ellipse, Electr. Power Syst. Res. 205 (2022) 107725.
[14] F.C. Trindade, W. Freitas, J.C. Vieira, Fault location in distribution systems based on smart feeder meters, IEEE Trans. Power Deliv. 29 (1) (2013) 251–260.
[15] S. Lotfifard, M. Kezunovic, M.J. Mousavi, Voltage sag data utilization for distribution fault location, IEEE Trans. Power Deliv. 26 (2) (2011) 1239–1246, http://dx.doi.org/10.1109/TPWRD.2010.2098891.
[16] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, 2016, arXiv preprint arXiv:1609.02907.
[17] M. MansourLakouraj, M. Gautam, H. Livani, M. Benidris, A multi-rate sampling PMU-based event classification in active distribution grids with spectral graph neural network, Electr. Power Syst. Res. 211 (2022) 108145.
[18] A. Aligholian, H. mohseni Rad, GraphPMU: Event clustering via graph representation learning using locationally-scarce distribution-level fundamental and harmonic PMU measurements, IEEE Trans. Smart Grid (2022) 1, http://dx.doi.org/10.1109/TSG.2022.3225373.
[19] M.M. Lakouraj, H. Hosseinpour, H. Livani, M. Benidris, Grid-aware waveform analytics for event classification in distribution grids, in: 2023 IEEE Industry Applications Society Annual Meeting, IAS, 2023, pp. 1–6, http://dx.doi.org/10.1109/IAS54024.2023.10406523.
[20] R. Rezaiesarlak, M. Manteghi, Short-time matrix pencil method for chipless RFID detection applications, IEEE Trans. Antennas and Propagation 61 (5) (2013) 2801–2806.
[21] F.M. Bianchi, D. Grattarola, L. Livi, C. Alippi, Graph neural networks with convolutional ARMA filters, IEEE Trans. Pattern Anal. Mach. Intell. 44 (7) (2022) 3496–3507, http://dx.doi.org/10.1109/TPAMI.2021.3054830.
[22] Y. Hua, T. Sarkar, Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise, IEEE Trans. Acoust. Speech Signal Process. 38 (5) (1990) 814–824, http://dx.doi.org/10.1109/29.56027.
[23] L.L. Grant, M.L. Crow, Comparison of matrix pencil and prony methods for power system modal analysis of noisy signals, in: 2011 North American Power Symposium, 2011, pp. 1–7, http://dx.doi.org/10.1109/NAPS.2011.6024892.
[24] T. Sarkar, O. Pereira, Using the matrix pencil method to estimate the parameters of a sum of complex exponentials, IEEE Antennas Propag. Mag. 37 (1) (1995) 48–55, http://dx.doi.org/10.1109/74.370583.
[25] L.L. Grant, M.L. Crow, Comparison of matrix pencil and prony methods for power system modal analysis of noisy signals, in: 2011 North American Power Symposium, IEEE, 2011, pp. 1–7.
[26] T. Sarkar, O. Pereira, Using the matrix pencil method to estimate the parameters of a sum of complex exponentials, IEEE Antennas Propag. Mag. 37 (1) (1995) 48–55, http://dx.doi.org/10.1109/74.370583.
[27] M. MansourLakouraj, et al., Event classification in active distribution grids using physics-informed graph neural network and PMU measurements, in: 2022 IEEE Industry Applications Society Annual Meeting, IAS, 2022, pp. 1–6, http://dx.doi.org/10.1109/IAS54023.2022.9939922.
[28] E. Isufi, A. Loukas, A. Simonetto, G. Leus, Autoregressive moving average graph filtering, IEEE Trans. Signal Process. 65 (2) (2017) 274–288, http://dx.doi.org/10.1109/TSP.2016.2614793.
[29] A. Bagheri, I.Y. Gu, M.H. Bollen, E. Balouji, A robust transform-domain deep convolutional network for voltage dip classification, IEEE Trans. Power Deliv. 33 (6) (2018) 2794–2802.
[30] O. Gashteroodkhani, M. Majidi, M. Etezadi-Amoli, A combined deep belief network and time-time transform based intelligent protection scheme for microgrids, Electr. Power Syst. Res. 182 (2020) 106239.
[31] K. Chen et al, Fault location in power distribution systems via deep graph convolutional networks, IEEE J. Sel. Areas Commun. 38 (1) (2019) 119–131.
[32] PyTorch Geometric, URL https://pytorch-geometric.readthedocs.io/en/latest/.
[33] PSCAD Software, URL https://www.pscad.com/.
[34] S. Gautam, S.M. Brahma, Detection of high impedance fault in power distribution systems using mathematical morphology, IEEE Trans. Power Syst. 28 (2) (2013) 1226–1234, http://dx.doi.org/10.1109/TPWRS.2012.2215630.
[35] A. Emanuel, D. Cyganski, J. Orr, S. Shiller, E. Gulachenski, High impedance fault arcing on sandy soil in 15 kV distribution feeders: Contributions to the evaluation of the low frequency spectrum, IEEE Trans. Power Deliv. 5 (2) (1990) 676–686, http://dx.doi.org/10.1109/61.53070.
[36] M. MansourLakouraj, R. Hossain, H. Livani, M. Ben-Idris, Application of graph neural network for fault location in PV penetrated distribution grids, in: 2021 North American Power Symposium, NAPS, IEEE, 2021, pp. 01–06.
[37] K. Saleh, A. Ayad, Fault zone identification and phase selection for microgrids using decision trees ensemble, Int. J. Electr. Power Energy Syst. 132 (2021) 107178.
[38] T. Biswal, S. Parida, A novel high impedance fault detection in the micro-grid system by the summation of accumulated difference of residual voltage method and fault event classification using discrete wavelet transforms and a decision tree approach, Electr. Power Syst. Res. 209 (2022) 108042.
[39] Y. Liu, L. Yang, A. Ghasemkhani, H. Livani, V.A. Centeno, P.-Y. Chen, J. Zhang, Robust event classification using imperfect real-world PMU data, IEEE Internet Things J. (2022).

[40] Z. Jiang, Y. Wang, Y. Li, H. Cao, A new method for recognition and classification of power quality disturbances based on IAST and RF, Electr. Power Syst. Res. 226 (2024) 109939.
[41] E. Aker, M.L. Othman, V. Veerasamy, I.b. Aris, N.I.A. Wahab, H. Hizam, Fault detection and classification of shunt compensated transmission line using discrete wavelet transform and naive bayes classifier, Energies 13 (1) (2020) 243.
[42] Y. Yuan, Z. Wang, Y. Wang, Learning latent interactions for event classification via graph neural networks and PMU data, IEEE Trans. Power Syst. 38 (1) (2023) 617–629, http://dx.doi.org/10.1109/TPWRS.2022.3158248.
[43] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. 13 (2) (2012).
[44] K. Chen, J. Hu, Y. Zhang, Z. Yu, J. He, Fault location in power distribution systems via deep graph convolutional networks, IEEE J. Sel. Areas Commun. 38 (1) (2020) 119–131, http://dx.doi.org/10.1109/JSAC.2019.2951964.
[45] O. Gashteroodkhani, M. Majidi, M. Etezadi-Amoli, Fire hazard mitigation in distribution systems through high impedance fault detection, Electr. Power Syst. Res. 192 (2021) 106928.
[46] Real Time Digital Simulator (RTDS), URL http://www.rtds.com.
[47] N. Taghipourbazargani, G. Dasarathy, L. Sankar, O. Kosut, A machine learning framework for event identification via modal analysis of PMU data, IEEE Trans. Power Syst. (2022) 1–12, http://dx.doi.org/10.1109/TPWRS.2022.3212323.