Title

Data-Driven Analysis of Power Distribution Synchrophasors with Applications to Situational Awareness, Load Modeling, and Reliability

Permalink

https://escholarship.org/uc/item/1h2921m2

Author

Shahsavari, Alireza

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Data-Driven Analysis of Power Distribution Synchrophasors with Applications to Situational Awareness, Load Modeling, and Reliability

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Alireza Shahsavari

September 2019

Dissertation Committee:

Dr. Hamed Mohsenian-Rad, Chairperson
Dr. Nael Abu-Ghazaleh
Dr. Konstantinos Karydis

The Dissertation of Alireza Shahsavari is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

First and foremost I want to thank my advisor Dr. Hamed Mohsenian-Rad [1], whose collective support, guidance, and expertise have made this dissertation a reality. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for his research was contagious and motivational for me. Many thanks to Dr. Nael Abu-Ghazaleh [2] and Dr. Konstantinos Karydis [3] for their great feedback and being in my thesis committee. Thank you all for sharing your knowledge and passion with me during my Ph.D. program.

I would like to thank Dr. Alexandra von Meier [4], Dr. Emma M Stewart [5], and Ciaran Roberts [6] for their collaborations and contributions in various projects related to this dissertation. I am also grateful to my industrial collaborators in Riverside Public Utilities. More specifically, I would like to thank Ed Cortez [7], for his continuous support and for providing me the great opportunity to collaborate with Riverside Public Utilities.

The members of the Smart Grid lab have contributed immensely to my personal and professional time. The group has been a source of friendships as well as good advice and collaboration. To my labmate Mohammad Farajollahi, thank you for your wonderful patience, continual support and warm humour; I am lucky to have made such great friend.

I also extend my gratitude to my family. I specifically, thank my beloved parents for their unconditional kindness, love, patience, and support throughout my life. They have

---

[1] Associate Professor, Department of ECE, University of California Riverside

[2] Professor, Department of ECE, University of California Riverside

[3] Assistant Professor, Department of ECE, University of California Riverside

[4] Adjunct Professor in University of California Berkeley

[5] Lawrence Livermore National Laboratory

[6] Lawrence Berkeley National Laboratory

[7] Principal Electrical Engineer, Riverside Public Utilities

always been and continue to be my shelter and comfort in this life and especially during my academic career. I also thank my wonderful sister and brother for their soothing words and endless encouragement which comforted me during many hard moments along the way.

This work was supported in part by DoE grant EE 0008001, UCOP grant LFR-18-548175, NASA MIRO grant NNX15AP99ANSF, NSF grants 1462530, 1253516, 1405330, and 1307756, RPU grant EIG 14-2853, and a University of California CITRIS seed grant.

The content of this thesis is a reprint of the material that are appeared in the following publications:

- **A. Shahsavari**, M. Farajollahi, H. Mohsenian Rad, "Individual Load Model Parameter Estimation in Power Distribution Network using Load Switching Events", *IEEE Trans. on Power System*, Accepted for Publication, 2019.

- **A. Shahsavari**, M. Farajollahi, E. Stewart, E. Cortez, H. Mohsenian Rad, "Situational Awareness in Distribution Grid Using Micro-PMU Data: A Machine Learning Approach", *IEEE Trans. on Smart Grid*, Accepted for Publication, 2019.

- **A. Shahsavari**, A. Sadeghi-Mobarakeh, E. Stewart, E. Cortez, L. Alvarez, F. Megala, H. Mohsenian Rad, "Distribution Grid Reliability versus Regulation Market Efficiency: An Analysis based on Micro-PMU Data," *IEEE Trans. on Smart Grid*, vol. 8, pp. 2916 - 2925, Jun. 2017.

- M. Farajollahi, **A. Shahsavari**, E. Stewart, H. Mohsenian Rad, "Locating the Source of Events in Power Distribution Systems Using Micro-PMU Data", *IEEE Trans. on Power Systems*, vol.33, pp. 6343 - 6354, May. 2018.

- **A. Shahsavari**, M. Farajollahi, E. Stewart, H. Mohsenian Rad, "A Machine Learning Approach to Event Analysis in Distribution Feeders Using Distribution Synchrophasors," in *Proc. of IEEE Smart Grid Synchronized Measurements and Analytics*, College Station, TX, May. 2019.

- M. Farajollahi, **A. Shahsavari**, H. Mohsenian Rad, "Linear Distribution System State Estimation Using Synchrophasor Data and Pseudo-Measurement," in *Proc. of IEEE Smart Grid Synchronized Measurements and Analytics*, College Station, TX, May. 2019.

- M. Farajollahi, **A. Shahsavari**, H. Mohsenian Rad, "Tracking State Estimation in Distribution Network Using Synchrophasor Data," in *Proc. of IEEE PES General Meeting*, Portland, OR, 2018.

- **A. Shahsavari**, M. Farajollahi, E. Stewart, C. Roberts, F. Megala, L. Alvarez, E. Cortez, H. Mohsenian Rad, "Autopsy on Active Distribution Networks: A Data-Driven Fault Analysis Using Micro-PMU Data," in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

- **A. Shahsavari**, M. Farajollahi, E. Stewart, C. Roberts, H. Mohsenian Rad, "A Data-Driven Analysis of Lightning-Initiated Contingencies at a Distribution Grid with a PV Farm Using Micro-PMU Data," in *Proc. of IEEE PES Society NAPS*, Morgantown, WV, Sep. 2017.

- M. Farajollahi, **A. Shahsavari**, H. Mohsenian Rad, "Location Identification of Distribution Network Events Using Synchrophasor Data," in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

- **A. Shahsavari**, M. Farajollahi, E. Stewart, A. Meier, L. Alvarez, E. Cortez, H. Mohsenian Rad, “A Data-Driven Analysis of Capacitor Bank Operation at a Distribution Feeder Using Micro-PMU Data,” in *Proc. of IEEE Power & Energy Society Conference on Innovative Smart Grid Technologies (ISGT)*, Washington, DC, Apr. 2017.
- M. Farajollahi, **A. Shahsavari**, H. Mohsenian Rad, “High Impedance Faults Location Using Synchronized Harmonic Phasors,” *in proc. of IEEE Power & Energy Society Conference on Innovative Smart Grid Technologies (ISGT)*, Washington, DC, Apr. 2017.
- **A. Shahsavari**, A. Sadeghi, E. Stewart, H. Mohsenian Rad, “Distribution Grid Reliability Analysis Considering Regulation Down Load Resources via Micro-PMU Data,” *in proc. of IEEE International Conference on Smart Grid Communications*, Sydney, Australia, Nov. 2016.

This dissertation is dedicated to my beloved parents Aghdas and Siroos for all of their endless love, support, and encouragement.

ABSTRACT OF THE DISSERTATION

Data-Driven Analysis of Power Distribution Synchrophasors with
Applications to Situational Awareness, Load Modeling, and Reliability

by

Alireza Shahsavari

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, September 2019
Dr. Hamed Mohsenian-Rad, Chairperson

The recent development of distribution-level phasor measurement units, a.k.a. micro-PMUs, has been an important step towards achieving situational awareness in power distribution networks. The challenge however is to transform the large amount of data that is generated by micro-PMUs to actionable information and then match the information to use-cases with practical value to system operators. This open problem is addressed in this thesis. First, we introduce novel *data-driven event detection* techniques to pick out valuable portion of data from extremely large raw micro-PMU data. Subsequently, *data-driven event classifier* are developed to effectively classify power quality events. Importantly, we use field expert knowledge and utility records to conduct an extensive *data-driven event labeling*. Moreover, certain aspects from event detection analysis are adopted as additional features to be fed into the classifier model. In this regard, a multi-class support vector machine (multi-SVM) classifier is trained and tested over 15 days of real-world data from two micro-PMUs on a distribution feeder in Riverside, CA. In total, we analyze 1.2 billion measurement points, and 10,700 events. The effectiveness of the developed event classifier

is compared with prevalent multi-class classification methods, including k-nearest neighbor method as well as decision-tree method. Importantly, five real-world use-cases are presented for the proposed data analytics tools, including:

- Transient Load Modeling for Application in Frequency Regulation Market;
- Static Load Modeling;
- Remote Asset Monitoring;
- Protection System Diagnosis;
- Lightning Initiated Contingency Analysis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The proliferation in distributed energy resources (DERs), electric vehicles, and controllable loads has introduced new and unpredictable sources of disturbance in distribution networks. This calls for developing new monitoring systems that can support achieving situational awareness at distribution-level; thus, allowing the distribution system operator to make the best operational decisions in response to such disturbances.

Traditionally, there have been three major challenges in achieving situational awareness in power distribution systems. First is the lack of high resolution measurements. Metering in distribution systems is often limited to supervisory control and data acquisition (SCADA) at substations with minutely reporting intervals. As for smart meters, their report measurements once every 15 minutes or hourly. Second is the lack of accurate and up-to-date models for most practical distribution circuits. Third, due to the lower voltage and the larger number and variety of utility and customer equipment, distribution systems are subject to a huge number of events on a daily basis.

The first challenge above has recently been resolved by the advent of micro-PMUs [1, 2]. A typical micro-PMU is connected to single- or three-phase distribution circuits to measure GPS time-referenced magnitudes and phase angles of voltage and current phasors at 120 readings per second. This is 108,000 times faster than the reporting rate of a typical smart meter. Since 2015, several micro-PMUs have been installed at pilot test sites in the state of California, including some in the city of Riverside [1]. The data generated by micro-PMUs are a prominent example of big data in power systems. Each micro-PMUs generates 124,416,600 readings per day. Therefore, micro-PMUs installed on a handful of utility distribution feeders can generate terabytes of data on daily basis. Because micro-PMUs stream their measurements continuously, the data must be collected, cleansed, and processed, all in real time [1]. The early studies on micro-PMUs focused on innovative case studies, e.g., in [1, 3–13].

This thesis aims to develop a data-driven package for real-world micro-PMU data to transfer the terabytes of micro-PMU data to actionable use-cases for electric utilities. To such aim, we propose a novel *model-free* situational awareness framework for power distribution systems to turn micro-PMU data in to actionable information for tangible use-cases. The proposed distribution-level situational awareness is built on novel *data-driven event detection* techniques as well as novel *data-driven event classification* techniques.

1.1 Related Works

This section presents the related work on event detection and event classification using micro-PMU data.

1.1.1 Literature Review on Event Detection

In [14], a model-based event detection method is proposed to detect changes in the admittance matrix of the distribution grid using micro-PMUs data. In [15], a model-based anomaly detection method is developed to detect permanent faults in distribution lines, which requires full observability in all buses. However, model-based techniques are often prone to failure due to lack of model accuracy, particularly in case of detecting power quality events. Detecting partially-labeled events in micro-PMU data is proposed in [16–19]. However, given the complexity of distribution systems, it is difficult for utilities to predetermine the variety of distribution-level events [20]. Thus, relying on expert knowledge and labeling events at event detection phase may inevitably result in overlooking some events. To resolve this issue, we propose a model-free event detection approach to capture *unlabeled* data.

1.1.2 Literature Review on Event Classification

In the context of event classification in power distribution systems, prior studies have classified the various causes of fault events [21–24], and power quality events [25–28]. However, broadly speaking, the current literature is still limited when it comes to studying large-scale real-world micro-PMU data sets; therefore, the challenges that may arise in practical event classification problems are yet to be understood and addressed. Followed by proliferation of deploying micro-PMUs at distribution-level, a few studies have been conducted to fill the gap in literature. In [16], events are classified for partially labeled events. In Finally, in [29], classifiers are trained to identify malfunctioned capacitor bank switching and malfunctioned on-load tap-changer switching events using data from hardware-in-the-loop simulations. The transient signatures of these malfunctions are derived by simulating different test systems and test scenarios, i.e., not directly from real-world data streams. In contrast, the labeling in this study is done by using real-world data combined with field knowledge from utility staff and utility event logs.

1.2 Motivations and Contributions

This thesis makes use of real-world micro-PMU data from a feeder in Riverside, CA, to address the second and the third challenges toward situational awareness in distribution-level, that are listed in the introduction. Specifically, we propose a novel *model-free* situational awareness framework for power distribution systems to turn micro-PMU data in to actionable information for tangible use cases. The proposed distribution-level situational awareness framework is shown in Fig. 1.1. This figure also depicts the under-study distribution feeder in Riverside, CA. This feeder is monitored by two micro-PMUs. The proposed situational awareness framework includes three functions, i.e., *perception*, *comprehension*, and *projection*. The data perception conducts data collection from two micro-PMUs as well as data cleansing such as synchronization and possible data missing. Perception is the stage of collecting and cleansing the terabytes of data from the two micro-PMUs. Comprehension is done by introducing novel *data-driven event detection* techniques as well as novel *data-driven event classification* techniques. In this thesis, the event detection is applied to non-linearly dependent data streams from micro-PMUs, including voltage magnitude, current magnitude, active power, and reactive power. Event classification is done

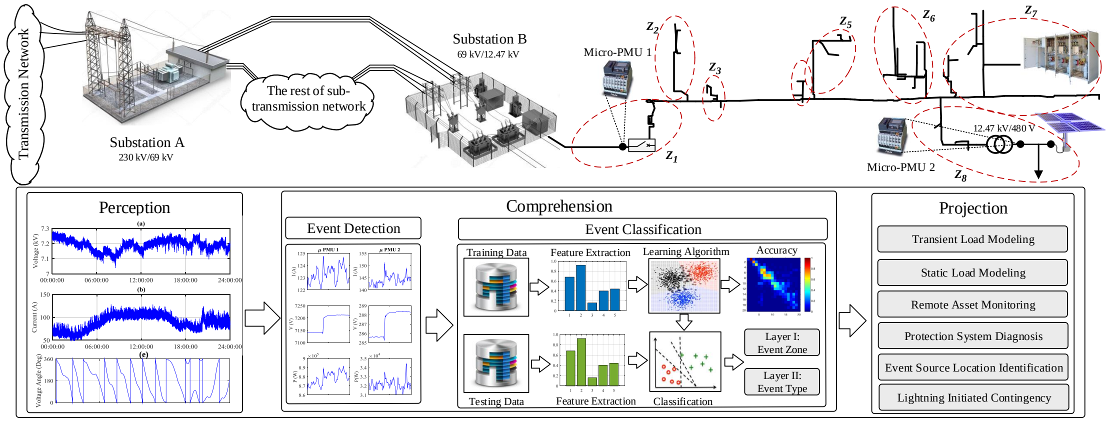

Figure 1.1: The real-world distribution feeder that is studied in this study.

by extracting the inherent features of detected events, and by constructing an algorithm that can learn from and make predictions of various events [30, 31].

The main contributions in this thesis can be summarized as follows:

1. A novel situational awareness framework is introduced for power distribution systems using micro-PMU data, that is model-free; it works by going through a sequence of *event detection*, *event classification*, and *event scrutinization* efforts to transform the large amount of measurement data from micro-PMUs to information that are useful for distribution system operators.

2. The approach in this study makes use of field expert knowledge and utility records in order to conduct an extensive *data-driven event labeling* for micro-PMU data. The detected events are labeled according to *event zone* and *event type*. As for the event detection phase prior to event labeling, our approach is comprehensive; it involves *moving windows* to help compensate the lack of information about the *start time* of

each event. It also involves *dynamic window sizes* to help compensate the lack of information about the *duration* of each event.

3. Different feature selection approaches and different classification methods are examined and compared, including multi-SVM, k-nearest neighbor, and decision-tree, with considering certain aspects of events from micro-PMUs, e.g., *uneven* datasets and features of *multi-stream* signals. It is shown that the use of the proposed detection features, such as detection window and detection indicator, is critical, regardless of the method of classification. It is also observed that multi-SVM is a better classifier compared to k-NN and DT in this particular application domain, whether or not we use the detection features.

4. Six important real-world use-cases are proposed and investigated. The main contributions for the developed use-cases can be summarized as follows:

   - **Transient Load Modeling for Application in Frequency Regulation Market;**

     (a) A novel *data-driven* approach is developed to use experimental micro-PMU data, on three phases, to analyze *transient* behaviors of different regulation down load resources.

     (b) A new method is developed to model the aggregate load transient profile, in form of an aggregate surge current profile, that is induced on a distribution feeder once a group of load resources responds to a regulation down event. This is done by applying unsupervised learning methods, and taking into account factors such as sensing delay, communications delay, and load response delay.

(c) A comprehensive distribution grid reliability analysis is conducted for the under-study distribution system, in presence of regulation down load resources, by taking into account the models and different characteristics of the main feeder's protection system as well as each lateral's protection system. Based on whether or not a recloser device is used in the protection system, both *momentary* and *permanent* reliability indexes are analyzed.

(d) The above reliability analysis is combined with an analysis on performance score calculation in performance-based regulation markets. Accordingly, a methodology is derived to investigate the potential *trade-off* between distribution grid reliability and regulation market efficiency.

(e) It is shown that the surge current induced by regulation down load resources can have severe adverse effect on the protection system, and thus on the reliability of distribution networks. One may attempt to mitigate such adverse effect by adding sufficiently large and randomly selected intentional delays to the response time of the regulation down load resources. However, this has to be done carefully, because while *reshaping* the load resources' aggregate surge current may help to avoid jeopardizing reliability, it should *not* be to the extent that it jeopardizes their performance in offering regulation down service. In practice, there may or may not exist a *safe choice* for the amount of added random delays to break the trade-off. The careful construction and analysis of the reliability-efficiency curves would be critical.

- **Static Load Modeling;**

  (a) Instead of or in addition to using up-stream voltage events as the enabler for load modeling, which is commonly used in the literature, we make use of load switching events across the distribution feeder itself. In principle, once a load is switched, the switching event changes the voltage in the rest of the loads, which causes variation in their active power and reactive power usage; thus allowing us to estimate load parameters of the rest of the individual loads.

  (b) The proposed method can estimate load modeling parameters of individual loads using measurements only at the feeder-head, i.e., at the distribution substation. No measurement is needed at individual loads.

  (c) We provide a theoretical foundation to determine the conditions on the extent of measurements needed to successfully achieve the individual load models.

  (d) The proposed individual load modeling can be done *sequentially* to obtain the load models for a subset of loads as more measurements become gradually available. Moreover, we proposed a variation of our method by using a forgetting factor so as to support estimating the parameters for time-varying individual loads.

  (e) By solving a non-liner least-squares problem, it is shown that the proposed methodology can be extended to utilize different types of *redundancy* in measurements in order to improve load modeling accuracy and robustness.

(f) We develop a residue-based bad data detection and identification method to identify and drop load configuration measurements with erroneous switching status; thus to ensure the accuracy of the load models. This also helps with the cases with imperfect knowledge of line impedances, switch statuses, and other system parameters.

- **Remote Asset Monitoring;**

  (a) We propose a novel data-driven approach to use experimental distribution-level voltage and current synchrophasor data, on three phases, to identify the operational parameters of the switched capacitor bank with no need to install separate asset sensors.

  (b) We conduct the transient and dynamic analysis to identify malfunction in capacitor bank switching events. This is a novel contribution because it studies fast-scale analysis of capacitor bank switching events, for both impact, performance verification and parameterization at distribution level using micro-PMUs.

  (c) We analyze the voltage and current synchrophasor data that is collected from a real distribution system in Riverside, CA during two weeks. The steady-state analysis is conducted to identify unbalanced or underrating operation of the three-phase capacitor bank.

- **Protection System Diagnosis;**

  (a) We proposes a novel data-driven experimental analysis on a single-phase-to-neutral fault to identify the operation of protection devices in fast-time scale. Data from five micro-PMUs on a real-life distribution and sub-transmission network is analyzed with the focus on an animal-caused fault on one feeder.

  (b) We identify the *fault time-line* using feeder-level as well as customer-level micro-PMUs. In this regard, any potential fuse-recloser miscoordination could be detected using experimental data.

  (c) We investigate the response of the PV resources to the fault to explore any miscoordination between feeder protection scheme and inverter protection systems, i.e., its built-in anti-islanding schemes.

  (d) We investigate the effect of the fault on outlying areas covering feeder-level and customer-level impacts.

- **Lightning Initiated Contingency Analysis.**

  (a) We propose a novel data-driven approach to use experimental distribution-level synchrophasor data, i.e., voltage and current measurements, on three phases, to analyze transient behaviors of flashovers caused by lightning strikes and the corresponding responses of the system to this phenomenon.

  (b) The synchrophasor data during three actual lightning strikes in a real-world distribution system in Riverside, CA are studied to identify the transient behavior of short-circuit accompanied lightning-initiated flashover.

  (c) We investigate the response of 7.5 MW PV farm to the lightning-initiated flashover.

Chapter 2

Data Analytic Package for Distribution Synchrophasors

2.1 Introduction

In this chapter, we introduce the data-driven event detection techniques as well as data-driven event classification techniques for real-world micro-PMU data to develop a model-free situational awareness for power distribution systems.

2.2 Data-Driven Event Detection

This section proposes two data-driven event detection techniques to pick out valuable portion of data from extremely large raw micro-PMU data.

2.2.1 Method I: Absolute Deviation Around Median

Let $D_i := [d_1, ..., d_n]^T$ denote a sequence of measurements from a micro-PMU, such as current magnitude on one phase, where n is the number of observation samples in the sequence. Subscript i is the index of the data sequence within the overall micro-PMU data stream. We define MAD_i as the median absolute deviation (MAD) in data sequence D_i as follows:

$$MAD_i = \gamma \cdot M\Big[\, |D_i - M[D_i]| \,\Big], \tag{2.1}$$

where $M[\cdot]$ and $|\cdot|$ denote median and absolute values. A typical value for coefficient γ is 1.4826 [32]. In this study, we detect an event within data sequence D_i if there exists a data point $k = 1, \ldots, n$ for which any of the following holds:

$$\begin{aligned} & d_k \leqslant M[D_i] - \zeta^- MAD_i \\ & M[D_i] + \zeta^+ MAD_i \leqslant d_k, \end{aligned} \tag{2.2}$$

where ζ^- and ζ^+ denote the threshold to detect overshoot and undershoot in the data sequence, respectively. Here, $M[D_i] - \xi^- MAD_i$ and $M[D_i] + \xi^+ MAD_i$ denote the lower-bound and the upper-bound margins for data sequence D_i, respectively. We define an indicator function $\mathbb{I}\{\cdot\}$ such that $\mathbb{I}\{D_i\} = 1$, if the condition in (2.2) holds for data sequence D_i; and $\mathbb{I}\{D_i\} = 0$ otherwise. Note that, the above approach to detect an event in a micro-PMU data sequence is a statistical anomaly detection technique which uses the absolute deviation around median test. Other statistical anomaly detection methods could also be used, such as the extreme studentized deviate test, or the standard deviation around mean

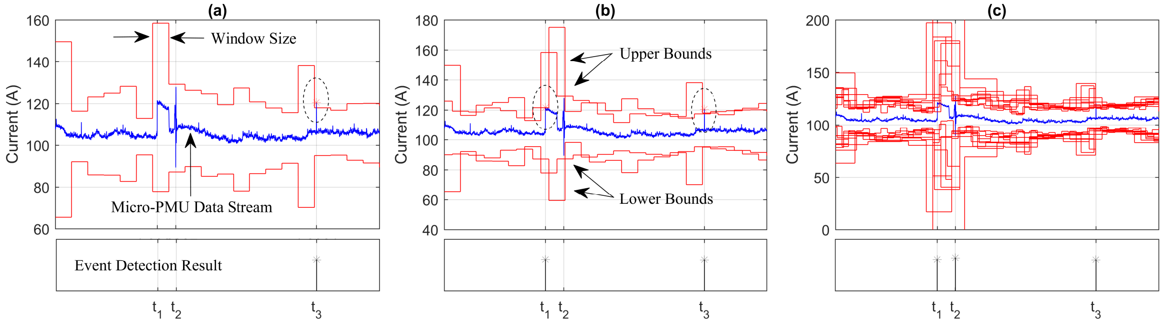

Figure 2.1: Effect of *moving window* and *dynamic window size* on event detection: (a) static window size without moving window: one event is detected at t_3; (b) static window size with moving window: another event is detected at t_1; (c) dynamic window size with moving window: all three events are detected.

value test, c.f. [32, 33].

The choice of parameters ζ^- and ζ^+ and the size of the data sequence window n have impact on the performance of the detection method. While ζ^- and ζ^+ are often selected empirically, choosing the right window size n is very challenging. In fact, we observed that it may *not* be possible to detect all events based on only one value for parameter n.

In order to overcome the above challenges, we propose to use a *dynamic window size* as well as a *moving window* such that we can detect as many events as possible. On one hand, the *dynamic window size* can help to compensate the lack of information about the *duration of each event*. On the other hand, the *moving window* can help to compensate the lack of information about the *start time of each event*.

The impact of applying dynamic window sizes and moving windows is shown in Fig. 2.1. Here, the entire data stream takes 100 seconds. Three major events can be visually detected, with start time stamps t_1, t_2, and t_3. We can see that different events have different natures and different lengths. The first event is long lasting. It can be detected either at its step-up edge or at its step-down edge; or both. The second event includes some transient oscillations. The third event is a momentary spike.

Fig. 2.1(a) shows the case where the window size is fixed at $n = 600$ micro-PMU samples, i.e., five seconds. Therefore, in the 100 seconds of data shown in this figure, there exist $20 = 100/5$ upper-bound and lower-bound margins of the form in (2.2). Only the third event at time t_3 is detected in this case. Fig. 2.1(b) shows the case where there is a second window of the same size, a moving window, that is shifted by 300 samples, i.e., half of the window size. Therefore, besides those 20 upper-bound and lower-bound margins that we saw in Fig. 2.1(a), there are additional 20 upper-bound and lower-bound margins in this figure. Accordingly, for each micro-PMU sample, there exist two upper-bound margins and two lower-bound margins in Fig. 2.1(b). The new upper-bound and lower-bound margins in Fig. 2.1(b) can detect the first event at time t_1. Fig. 2.1(c) shows the windows of $n = 120, 360, 600, 840, 1080$ samples for their sizes, where each window is also moved by half of its own size. The window sizes are 1, 3, 5, 7, and 9 seconds, respectively. Note that, for each micro-PMU sample in this figure, there exist 10 upper-bound margins and 10 lower-bound margins. As it can be seen from this figure, by applying both moving windows and dynamic window sizes, we can detect *all three* events.

Both dynamic window sizes and moving windows are necessary to assure detecting *all* events. This point is illustrated in Fig. 2.2(a). The percentage of correctly identified events versus the window size are shown in this figure; for both static and moving window types. The micro-PMU data stream in this example takes one day and includes 564 events. We can see that the use of moving window is always more effective than the use

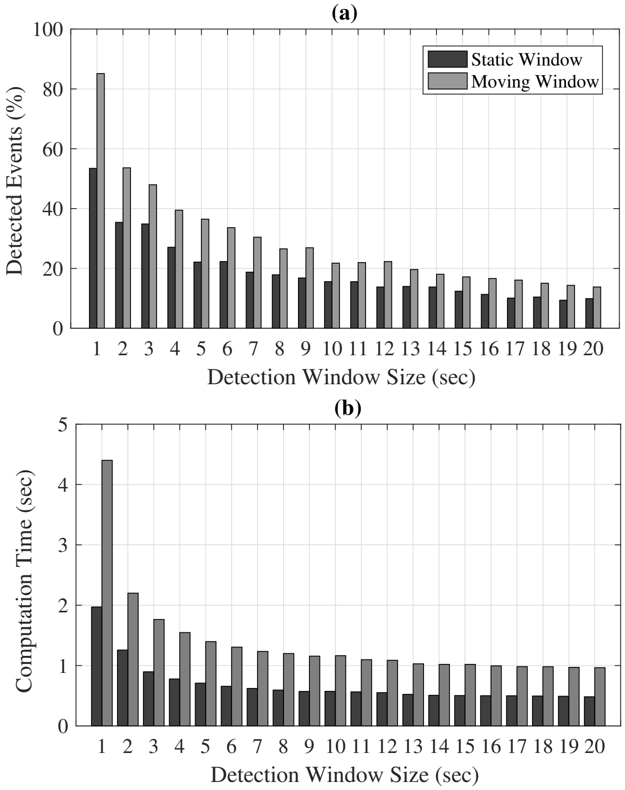

Figure 2.2: (a) The percentage of detected events; (b) Computation time.

of a static window. No single window size can detect all events. However, collectively, a combination of different window sizes and moving windows can detect all 564 events. Also, Fig. 2.2(b) shows the computation time corresponding to static windows and moving windows for different window sizes. As it can be seen in this figure, in each window size, the computation time of the moving windows is twice of the static window. Also, as expected, the computation time decreases as the detection method is applied to wider window sizes.

2.2.2 Method II: Residual Test on Non-Linear Estimation

Let $V_i := [|v_1|, \ldots, |v_n|]^T$ and $I_i := [|i_1|, \ldots, |i_n|]^T$ denote sequences of voltage magnitude and current magnitude from a micro-PMU, respectively. From Circuit Theory, we know that in steady-state the relation between $|v_k|$ and $|i_k|$ at data point $k = 1, \ldots, n$ is as:

$$s_k = |v_k| \cdot |i_k|, \tag{2.3}$$

where s_k denotes apparent power at downstream of micro-PMU. During the transient of events, the (2.3) includes higher order harmonics beside the fundamental harmonic.

Let assume that there exists no major event in a window with n data points. Thus, we can assure that the downstream load of micro-PMU is almost constant in n data points. A non-linear estimator can be designed to estimate I_i by solving the following optimization problem:

$$\underset{B}{\text{minimize}} \quad \left\| \tilde{I}_i - I_i \right\|_2, \tag{2.4}$$

where $\tilde{I}_i := [|\tilde{i_1}|, \ldots, |\tilde{i_n}|]^T$ is the estimated current magnitude in data sequence i, as:

$$\tilde{I}_i = b_1 + \frac{b_2}{V_i}. \tag{2.5}$$

where $B := [b_1, b_2]^T$ is the regression coefficient vector. If there exists an event in data sequence $i = 1, \ldots, n$, e.g., such as load switching event, during transient of the event, (2.3) does not hold. While, it holds for pre-event and post-event. Consequently, the non-linear estimator, that is designed only fundamental frequency, fails to estimate the $|i_k|$ during transient period of event. Thus, the residues corresponding to event data points are larger

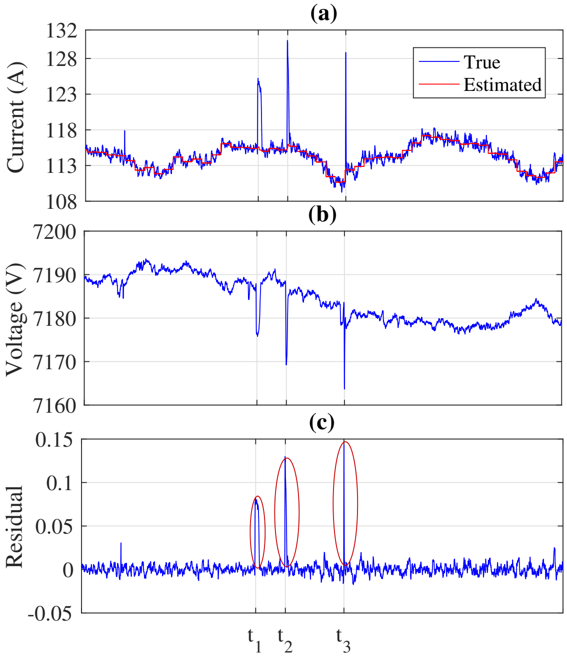

Figure 2.3: Event detection by estimating current magnitude: (a) current magnitude; (b) voltage magnitude; (c) residual.

than those during pre-event and post-event.

Fig. 2.3(a) shows the true current magnitude and estimated current magnitude of micro-PMU 1 during 50 seconds. Here, the non-linear estimation is conducted every second, i.e., $n = 120$ micro-PMU data points. The data sequence includes three major events at t_1, t_2, and t_3. From Fig. 2.3(a) and (b), it seems that all three events are load turning on events. From Fig. 2.3(a), we can see that the non-linear estimator fails in estimating current magnitude during these major events. Fig. 2.3(c) shows the residues in estimating current magnitude. As it can be seen, the residues at t_1, t_2, and t_3 are the largest residuals.

Also, Fig. 2.4(a) shows the true and estimated voltage magnitude of micro-PMU 1 during 50 seconds. The non-linear estimation is conducted every second to estimate the voltage magnitude. The data sequence includes two major events at t_1 and t_2, the former event is initiated from load switching, while latter one is a voltage step-down event which may be initiated from operating a voltage regulator in upstream-level, see Figs. 2.4(a) and (b). From Fig. 2.4(b), we can see that the non-linear estimator fails in estimating voltage magnitude during these events. Fig. 2.3(c) shows the residues in estimating voltage magnitude. As it can be seen, the events at t_1 and t_2 can be detected from residue test.

In Figs. 2.3 and 2.4, we detect events by considering a *fixed* threshold on residues or by applying largest normalized residual test. One may ask, what is the advantage of considering a threshold on residues against applying a threshold on data stream? The answer to this question is that the threshold on residues is *fixed*, while if we want to apply a threshold on data sequence, it should be updated for each window to consider the steady state changes. Also, the residual test based event detection method is not sensitive to the estimation window size. For instance, in Fig. 2.3, we can obtain almost the same residues by considering $n = 1200$ instead of $n = 120$.

2.3 Data-Driven Event Labeling and Feature Selection

Given the events that are captured by using the event detection method in Section 2.2, in this section, we aim to classify the events through conducting a comprehensive event labeling and feature selection approach.

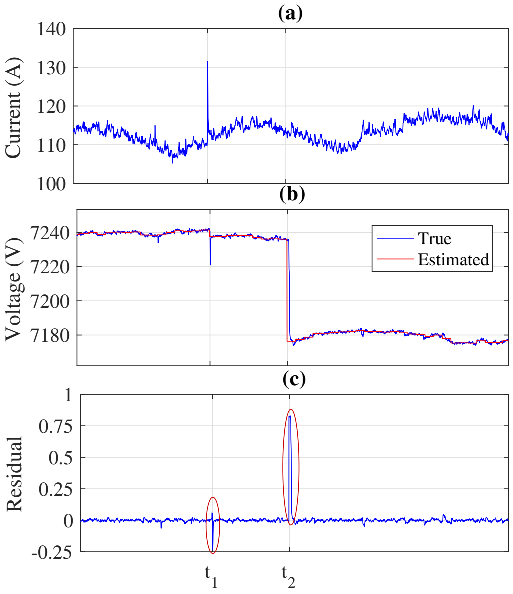

Figure 2.4: Event detection by estimating voltage magnitude: (a) current magnitude; (b) voltage magnitude; (c) residual.

2.3.1 Two-Layered Event Labeling

One can label power system events based on different aspects of their characteristics. Here, we seek to label the events according to *event zone* and *event type*.

Layer 1 Labeling based on Event Zone

In this first layer of classification, each event can take one of the below labels:

- Class I. Events initiated from upstream of micro-PMU 1, i.e., at transmission-level or another distribution feeder;
- Class II. Events initiated from downstream of micro-PMU 2, i.e., at customer location that hosting micro-PMU 2;
- Class III. Events initiated from somewhere between the two micro-PMUs across the

distribution feeder of interest.

An example for a Class I event is shown in Fig. 2.5. Class I events often appear as sustained steps or temporary fluctuations in voltage magnitudes at both feeder-level, seen by micro-PMU 1, and customer-level, seen by micro-PMU 2. However, they do *not* cause any major change in the current magnitudes. Thus, Class I events create signatures on reactive power measurements but not on active power measurements. Class I events could be due to transformer, capacitor bank, generator, or load switching at sub-transmission or transmission networks. They could also be due to momentary faults on another neighbouring distribution feeder, e.g., see [1].

An example for a Class II event is shown in Fig. 2.6. Class II events often appear as sustained steps or temporary fluctuations in voltage magnitude, current magnitude, active power, and reactive power at customer level, seen by micro-PMU 2. Depending on the size of the event, the event signature is noticeable also in the measurements at the feeder level, seen by micro-PMU 1. Class II events could be due to load switching, such as motor and HVAC loads, DER switching, such as PVs and batteries, among other customer-level causes.

An example for a Class III event is shown in Fig. 2.7. Class III events often appear as sustained steps or temporary fluctuations in voltage magnitude, current magnitude, active power, and reactive power at feeder-level, seen by micro-PMU 1. They may also affect the voltage magnitude at customer-level, seen by micro-PMU 2. However, they do *not* have a major impact on current magnitude, active power, and reactive power at

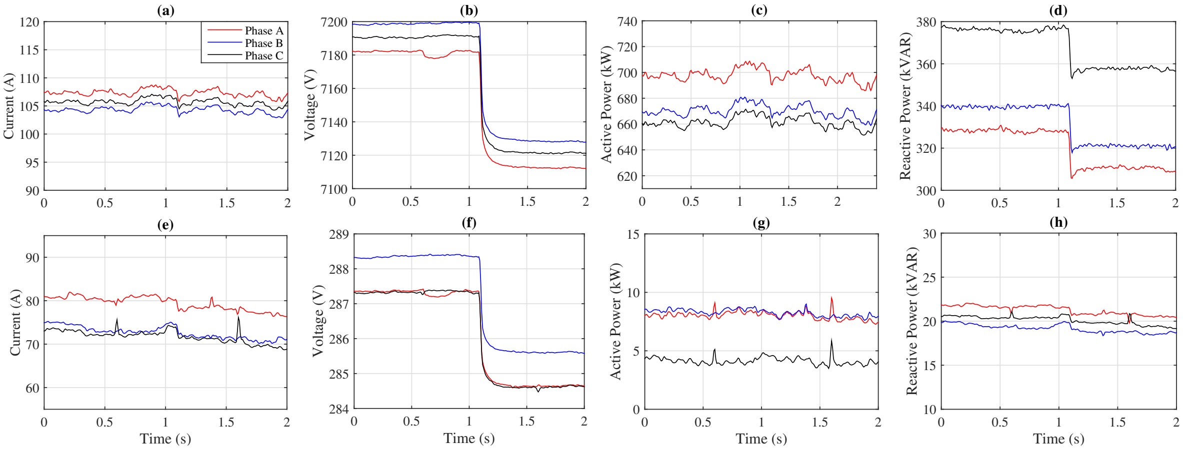

Figure 2.5: An example Class I event: (a) and (e): current; (b) and (f): voltage; (c) and (g): active power; (d) and (h): reactive power. First row corresponds to the measurements from micro-PMU 1. Second row corresponds to the measurements from micro-PMU 2.

customer-level. Class III events can be due to a wide verity of causes, such as distribution-level transformer and capacitor bank switching, lateral fuse blowing, primary protection operation, load switching, DER switching, etc.

Layer 2 Labeling based on Event Type

Each event can be further labeled based on its type. This can be done for all events, whether they are in Classes I, II, or III. However, labeling the type of Class I events is not of interest; because distribution-level PMUs are not intended to investigate transmission-level events. Labeling the type of Class II events is not of great interest either; because the customer that hosts micro-PMU 2 is being monitored directly. In fact, it is only Class III events that are of interest to be further classified; because those are the events that occur across the distribution feeder. Thus, in this section, we define a second layer for labeling

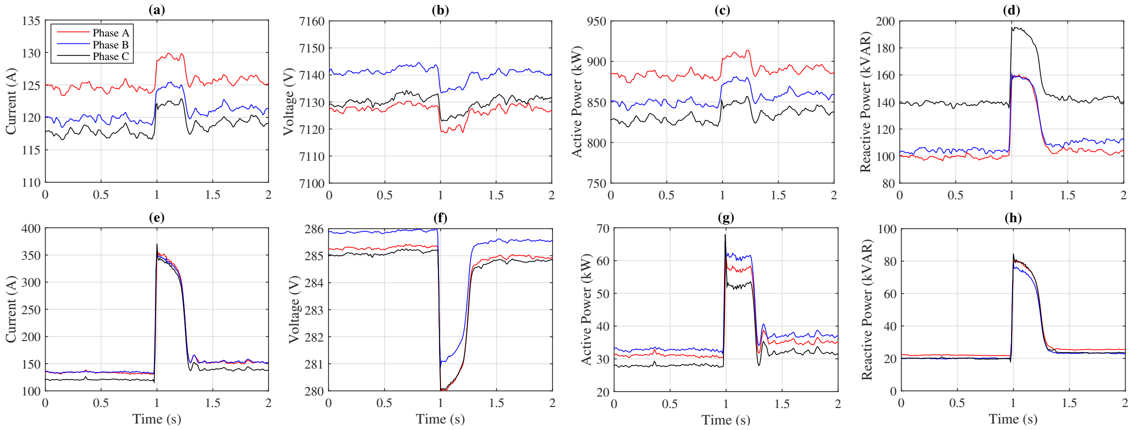

Figure 2.6: An example Class II event: (a) (h) are defined the same way as in Fig. 2.5.

Class III events, as follows:

- Class III.A. Capacitor bank switching; e.g., see Fig. 2.7;
- Class III.B. Distribution-level oscillation, e.g., see Fig. 2.8;
- Class III.C. Other events, e.g., see Fig. 2.9.

The above distinction is based on the fact that capacitor bank switching is an important event in distribution systems and the subject of several studies [1,34]. Oscillation events too are important. There is currently a limited understanding of the oscillation events within distribution systems [2].

Of course, one can define additional labels if one can identify other types of Class III events by observing the available micro-PMU data. For example, the *other event* in Fig. 2.9 is in fact a *load switch on* event that occurred somewhere along the feeder between micro-PMU 1 and micro-PMU 2. The foundation of the event classification methodology that is presented in this study would remain the same.

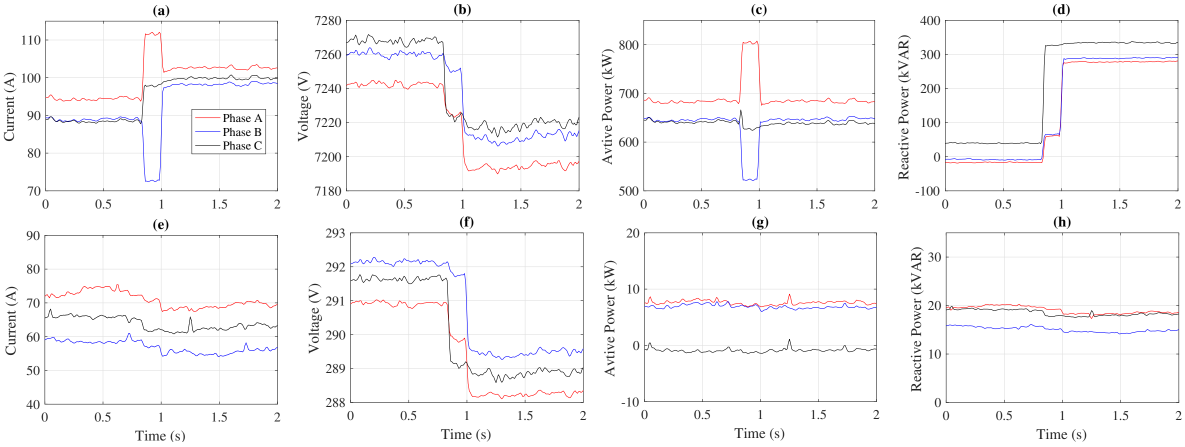

Figure 2.7: An example Class III event, i.e., Class III.A: (a) (h) are defined the same way as in Fig. 2.5

2.3.2 Feature Selection

A crucial task in any classification problem that involves Machine Learning is to choose adequate quantifiable features that can help distinguish classes. In this study, we propose the features in Table 2.1, which consists of three broad categories:

- **Single-Stream Features:** These are quantifiable properties that are derived from single data streams $D_i \in \{I, V, P, Q\}$. They could be obtained by applying the mean, standard deviation, median, difference, or other operators to each of such single data streams within the detected window. In this study, we use standard deviation and absolute difference. Note that, notations d_1 and d_n denote the first and the last data samples in D_i.

- **Multi-Stream Features:** These are quantities that are defined for various combinations of two synchronized data sequences $D_i, D_j \in \{I, V, P, Q\}$, whether from the same micro-PMU or two micro-PMUs. Different operators could be applied to the

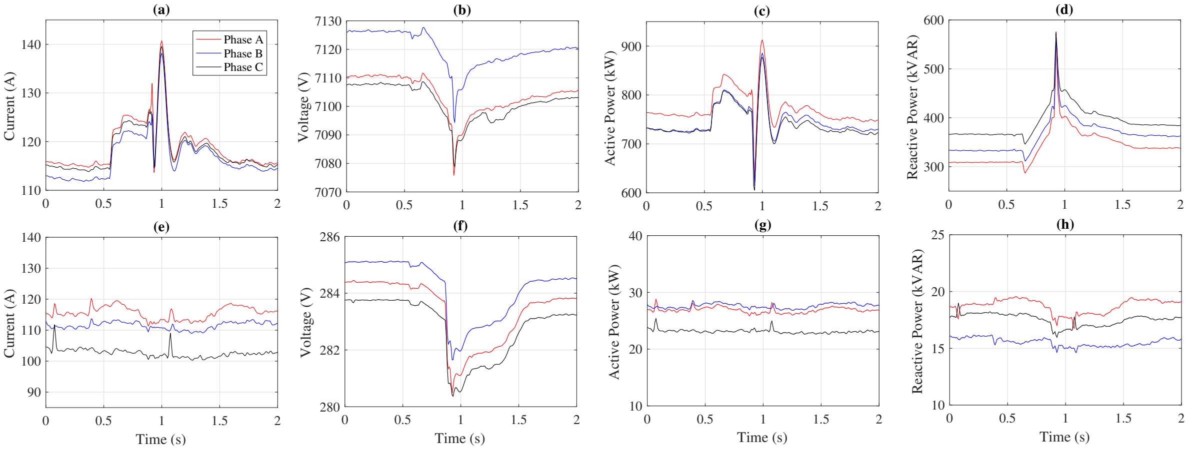

Figure 2.8: An example Class III event, i.e., Class III.B: (a) (h) are defined the same way as in Fig. 2.5

Table 2.1: Proposed features for classification

Feature	Feature Description		Number
Single-stream	Statistics	$\mathrm{std}(D_i)$	8
	Difference	$\lvert d_n - d_1 \rvert$	8
Multi-stream	Correlation	$corr(D_i, D_j)$	28
Detection	Detection Window	ω	1
	Detection Indicator	$\mathbb{I}\{D_i\}$	8

data streams. Here, we use the correlation between any two of the eight available data sequences to construct the multi-stream features.

- **Detection Features:** The way that an event is detected can itself carry useful information to classify the event. We use the following detection features for classification: 1) the smallest window at which the event was detected, denoted by ω; and 2) the binary detection indicators $\mathbb{I}\{D_i\}$ for $D_i \in \{I, V, P, Q\}$, for both micro-PMUs.

Note that, the events that are listed in Table 2.1 have significantly different magnitudes. Therefore, we propose to apply *Z-score normalization*, where for each initial feature x, we replace it with a new feature $z = (x - \mu)/\sigma$, where μ and σ denote mean value and

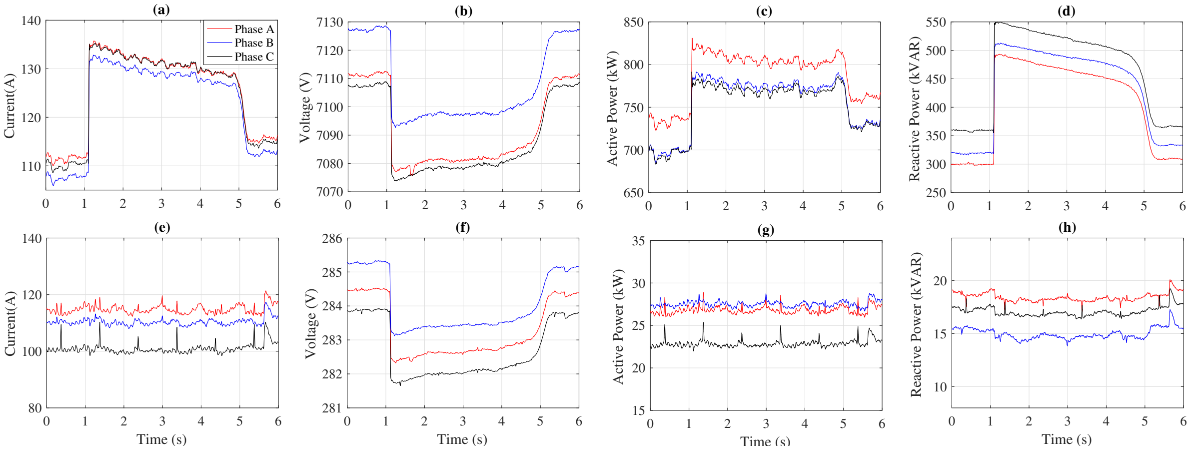

Figure 2.9: An example Class III event, i.e., Class III.C: (a) (h) are defined the same way as in Fig. 2.5

standard deviation, respectively [35].

2.4 Data-Driven Event Classification

We now use the event labeling and feature selection strategies in Section 2.3 to train different types of event classifiers.

2.4.1 Binary-SVM Classifier

Consider m events that are detected by using the method in Section 2.2. We use these events to train an SVM classifier. For each training event $i = 1, \ldots, m$, let X_i denote the 53×1 vector of extracted features, where $53 = 28 + 3 \times 8 + 1$, as in Table 2.1. Also, let y_i denote the assigned label for event i.

When it comes to binary classification, there are only two types of labels. We define $y_i \in \{-1, 1\}$, where $y_i = -1$ is the label for the first class; and $y_i = 1$ is the label for the second class. Let $W^T X + b = 0$ denote a separating hyperplane in the 53×53 feature space that separates the two classes, where W is a 53×1 coefficients vector, and b is the intercept. The SVM training problem seeks to find the optimal hyperplane that has the maximum total distance between the two classes across the training samples. If the training samples are not linearly separable, we should add some slack variables so as to turn the SVM into a *soft margin SVM*, which is formulated as:

$$\underset{W,b,\xi}{\text{minimize}} \quad \frac{1}{2}\|W\|_2^2 + \lambda \sum_{i=1}^{n} \xi_i \tag{2.6a}$$

$$\text{subject to} \quad y_i \left(W^T X_i + b\right) \geq 1 - \xi_i, \quad i = 1, \ldots, m \tag{2.6b}$$

$$\xi_i \geq 0, \; i = 1, \ldots, m, \tag{2.6c}$$

where ξ_i is a slack variable corresponding to training event i. If $y_i = -1$, then constraint (2.6b) requires that $W^T X_i + b \leq -1 + \xi_i$; and if $y_i = 1$, then constraint (2.6b) requires that $W^T X_i + b \geq 1 - \xi_i$; thus, making $W^T X + b = 0$ a separating hyperplane with a soft SVM margin of length $1 - \xi_i$ on both sides. Parameter λ is a tuning parameter. If the extracted features of a training event results in a point that falls on the correct side of the separating hyperplane with respect to the label of the event, then $0 \leq \xi_i < 1$; otherwise $\xi_i \geq 1$, c.f. [36–39].

2.4.2 Multi-SVM Classifier

A multi-class classification problem can be decomposed into several binary classification problems. This can be done by using methods such as one-against-one (OvO), one-against-all (OvA), directed acyclic graph SVM (DAGSVM), and binary tree of SVM

(BTSVM) [36,37]. In this study, we use OvA decomposition. We construct c binary SVM problems, where c is the number of classes. Each binary SVM problem obtains a separating hyperplane to separate one of the c classes from the rest of the $c-1$ classes. This process results in obtaining c different separating hyperplanes, denoted by $W_l^T X + b_l = 0$ where $l = 1, \ldots, c$. We have $c = 3$ for both Layer 1 and Layer 2 classification; see Section 2.3.1. In total, six sets of separating hyperplanes are trained; three sets for each layer. Once the training process is complete, the decision on class prediction for testing event i is made as [36]:

$$y_i = \underset{l=1,\ldots,c}{\arg\max} \left(W_l^T X_i + b_l \right). \tag{2.7}$$

In (2.7), we say that event i is predicted to belong in class l, which has the largest value of the decision function.

An alternative training separating hyperplanes in the form of $W_l^T X + b_l$ is to use non-linear classifiers, such as separating quadratic planes [40]. However, our experimental results based on real-world data have shown that there is no advantage in using nonlinear classifiers which are computationally more complex. What matters the most is to choose the right classification features, i.e., as in Table 2.1, as we will further discuss in the next section.

2.4.3 k-NN and Decision-Tree Classifiers

There are other classifiers that one can consider for this study. One example is the *k-nearest neighbors* (k-NN) classifier [41]. Another example is the *decision-tree* (DT) classifier [42]. The k-NN method classifies an unknown sample based on the known labels of its k-closet, e.g., in the Euclidean sense, neighbors [41]. As for the DT classifier, a decision tree is constructed by creating branches as conjunctions of features as well as leaves as class labels. Then, a test data sample is classified based on branches conjunctions [42]. These additional methods are not discussed here in details due to space limitation. However, detailed performance comparisons are provided across these methods later in the next section.

2.4.4 Metrics to Compare Different Classifiers

First, consider the binary classifiers as in Section 2.4.1. The correctness of each classifier can be evaluated by computing the following four quantities: *True Positive (TP)*, which is the number of events that are correctly classified to be inside of the target class; *True Negative (TN)*, which is the number of events that are correctly classified to be outside of the target class; *False Positive (FP)*, which is the number of events that are incorrectly classified to be inside of the target class; and *False Negative (FN)*, which is the number of events that are incorrectly classified to be outside of the target class. Accordingly, for each binary classifier, we can calculate the following five standard performance evaluation metrics [43]:

$$TPR = \frac{TP}{TP + FN}, \tag{2.8}$$

$$FPR = \frac{FP}{FP + TN}, \tag{2.9}$$

$$PPV = \frac{TP}{TP + FP}, \tag{2.10}$$

$$FOR = \frac{FN}{FN + TN}, \tag{2.11}$$

$$MCC=\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}, \tag{2.12}$$

where TPR, FPR, PPV, FOR, and MCC stand for the true positive rate, false positive rate, positive predictive values, false omission rate, and Matthews correlation coefficient, respectively. It should be noted that, in some machine learning literature, such as in [43], TPR, PPV, and FPR are also known as *recall*, *precision*, and *fall-out* metrices, respectively. The overall accuracy of a binary classifier can be assessed also by using the following metric:

$$F_1\text{Score} = \Big(\frac{TPR^{-1} + PPV^{-1}}{2}\Big)^{-1}. \tag{2.13}$$

Next, consider the multi-class classifier, as in Section 2.4.2. The recall, precision, and F_1*score* for multi-class classifier can be calculated by using either *Macro-averaging* or *Micro-averaging* [43]. Macro-averaging simply normalizes the sum of all metrics. Thus, Macro-averaging does not consider the number of events in each class. Micro-averaging however computes the metrics from sum of *TP*, *TN*, *FP*, and *FN* values of all classes. Thus, Micro-averaging takes the frequency of classes into consideration. Accordingly, one can prove that if Micro-averaging is used, then recall, precision, and F_1Score, all become equal, as follows:

$$\text{Recall} = \frac{\sum_{l=1}^{c} TP_l}{\sum_{l=1}^{c} \Big(TP_l + FN_l\Big)}. \tag{2.14}$$

2.5 Case Studies

The proposed event detection and event classification methods are applied to data from the two micro-PMUs in Fig. 1.1, during 15 days in July 2016. In total, we analyzed

1.2 billion measurement points.

2.5.1 Event Detection in micro-PMU Data Sequence

This section examines the effectiveness of the proposed event-detection methods introduced in Section 2.2, using one day data of current magnitude from micro-PMU 1. Also, we study the effect of the moving window and dynamic window sizes. To such aim, Method I is applied to the following cases:

- Case I. Static window size without moving window;
- Case II. Static window size with moving window;
- Case III. Dynamic window size without moving window;
- Case IV. Dynamic window size with moving window.

In order to compare the above case-studies, the detected events in each case study are compared with those that are detected in Case IV, which includes 564 events. Table 2.2 reports the number of the detected events in Case I and Case II with considering 10 static window sizes. For instance, considering $n = 120$ micro-PMU data points, the moving window results in detecting 179 more events. From this table, we can conclude that most of the events are detected in smaller windows. Also, Table 2.3 shows the number of *additional* events that are detected in each window size in Cases III and IV. From this table we can conclude that considering moving window results in detecting more events in each dynamic windows.

Table 2.2: Number of detected events in Case I and Case II.

n	120	240	360	480	600	720	840	960	1080	1200
Case I	302	200	197	153	125	125	106	101	95	88
Case II	481	303	271	223	206	190	172	151	150	123

Table 2.3: Number of detected events in Case III and Case IV.

n	120	240	360	480	600	720	840	960	1080	1200
Case III	302	+64	+29	+13	+4	+3	0	0	+1	0
Case IV	481	+50	+19	+5	+5	+4	0	0	0	0

We also compare the effectiveness of the Method I with Method II, considering dynamic window sizes and moving windows. The performance indexes for both detection methods are reported in Table 2.4. The obtained results shows Method II outperforms the Method I.

2.5.2 Classifier Design: A Preliminary Illustrative Example

The proposed event detection and event classification methods are applied to data from the two micro-PMUs in Fig. 1.1, during 15 days in July 2016. In total, we analyzed 1.2 billion measurement points, and 10,700 events. Only 1% of the measurements demonstrated any considerable event. Among the events detected, 1802, 2228, and 6670 events are labeled in Class I, Class II, and Class III, respectively. Among the 6670 Class III events, 27 events are labeled as Class A and 43 events are labeled as Class B. The training dataset includes 4.09% and 4.06% of all Layer I and Layer II data, respectively.

In this section, a multi-SVM classifier is designed to separate the events in Layer I.

Table 2.4: Event detection performance indexes for Methods I and II.

Method	TP	FP	FN	PPV	TPR	F_1 Score
Method I	541	23	8	0.96	0.985	0.97
Method II	546	6	3	0.99	0.99	0.99

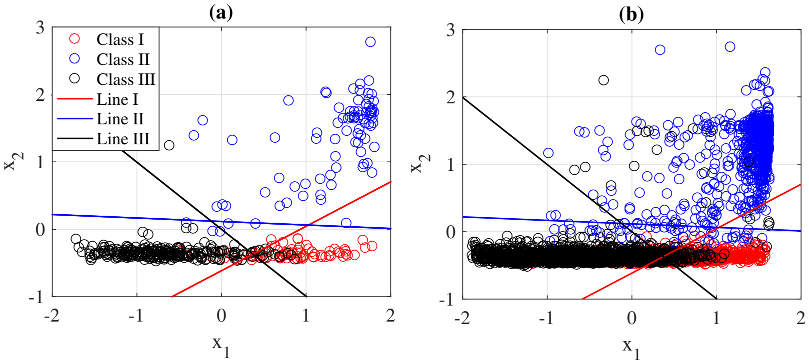

Figure 2.10: Target classes in two dimensional feature space and separating lines: (a) training data points; (b) test data points

Recall from Sections 2.3 and 2.4 that the resulting three separating hyperplanes are in the 53×53 feature space that cannot be visualized. Therefore, in order to develop an example that is easy to illustrate, we use only the two and three most dominant features, out of 53, so as to visualize the separating hyperplanes in 2×2 and 3×3 spaces, as shown in Figs. 2.10 and 2.12, respectively. Fig. 2.10(a) shows the training data points in 2×2 feature space as well as three separating lines obtained from three binary SVMs. Each separating line is somehow designed to separate events of one class from the rest of the events. The confusion matrix corresponding to training dataset is shown in Fig. 2.11(a). Here, the accuracy of training is 90.6%. Next, the separating hyperplanes are applied to the test dataset, and decision on class prediction is made using (2.7). Fig. 2.10(b) shows the test dataset in 2×2 feature space and separating lines. The overall classifier testing accuracy is 74.96%. Fig. 2.11(b) shows the confusion matrix for test dataset.

Fig. 2.12(a) shows the training data points in 3×3 feature space as well as three separating lines obtained from three binary SVMs. In this figure, x_1 is the correlation

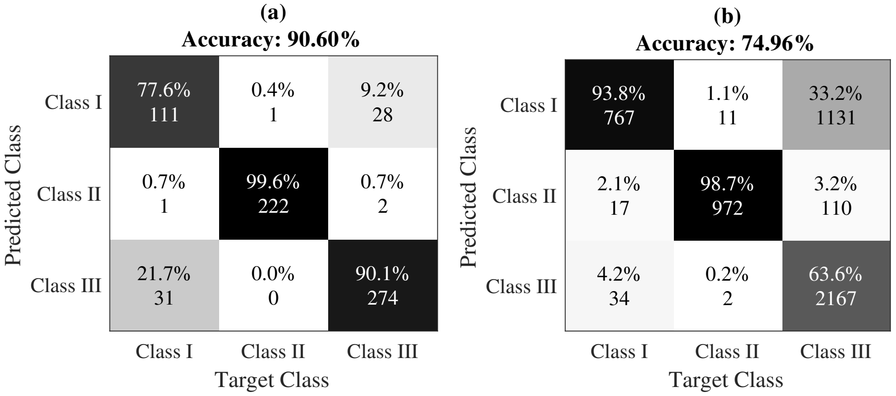

Figure 2.11: Confusion matrices obtained by classifier: (a) training data considering two dominant features; (b) test data considering two dominant features.

coefficients between reactive power of the two micro-PMUs, x_2 is the standard deviation of the current magnitude of micro-PMU 1, and x_3 is the standard deviation of active power at micro-PMU 2. The total number of training events is 438. We can see that the events are properly separated across the three classes. For instance, Hyperplane I separates events corresponding to Class I from the rest of the events. The overall classifier training accuracy is 91%. Next, the above separating hyperplanes are applied to the test dataset, and the decision on class prediction is made using (2.7). Fig. 2.13 shows the predicted classes across 10262 test events. The overall classifier testing accuracy is 89%. Thus, several events in this preliminary example are *not* predicted in the right classes, mainly due to not using all 53 features.

2.5.3 Classification Results and Impact of Detection Features

In order to demonstrate the importance of features, in this section, we separately examine the following two cases:

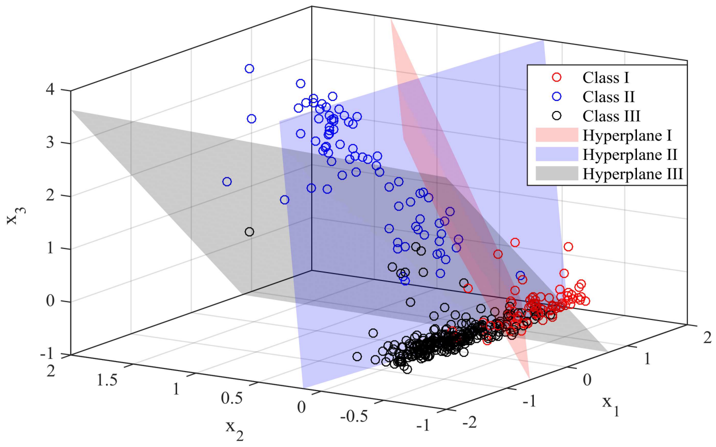

Figure 2.12: Target classes and separating hyperplanes of Layer I in a 3 × 3 feature space. The circles indicate training data points.

- *Case 1:* Classification *without* detection features.
- *Case 2:* Classification *with* detection features.

Interestingly, the overall multi-SVM classifier training accuracy is 100% in both cases; not shown here. However, when it comes to using the classifiers to identify the classes for *test events*, the performance is considerably better for *Case 2* than *Case 1*. The confusion matrices for the multi-SVM, k-NN, and DT classifiers are shown in Fig. 2.14, for both Case 1 and Case 2. Parameter k for the k-NN classifier is set to 3 based on an exhaustive search. All results are based on the data for test events. Each confusion matrix shows the *recall* metric in percentage for binary-classifiers as well as overall classification *recall* using Micro-averaging.

Tables 2.5 to 2.7 show the performance metrics in percentage for binary classifiers in Layer I. We can make two important observations from Fig. 2.14 and Tables 2.5 to 2.7. First, the performance is always better in Case 2 compared to Case 1. In other words,

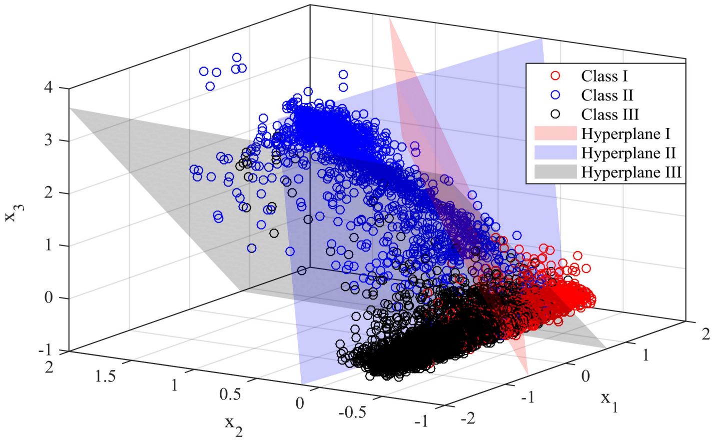

Figure 2.13: Predicted classes of Layer I obtained by the use of the separating hyperplanes in Fig. 2.12. The circles indicate test data points.

Table 2.5: Performance metric in percentage corresponding to the first binary-classifiers in Layer I classification.

Classifier	Case	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	1	95.23	0.66	96.63	0.95	95.12	95.93
	2	100	0.09	99.53	0	99.72	99.76
k-NN	1	97.23	1.31	93.73	0.54	94.59	95.49
	2	99.24	0.50	97.54	0.15	98.06	98.38
DT	1	74.60	0.55	96.46	4.89	52.35	84.14
	2	100	0.02	99.88	0	99.93	99.94

the use of detection features is indeed critical, regardless of the method of classification. Second, the multi-SVM classifier outperforms the k-NN classifier and the DT classifier in both cases. In other words, the multi-SVM classifier is a better choice in this study, whether or not we use the detection features.

Based on the above results, for the rest of this study, we always include the detection features, i.e., we use Case 2. It should be noted that the training accuracy of multi-SVM classifier is 100% in both cases. The results however are not shown here due to

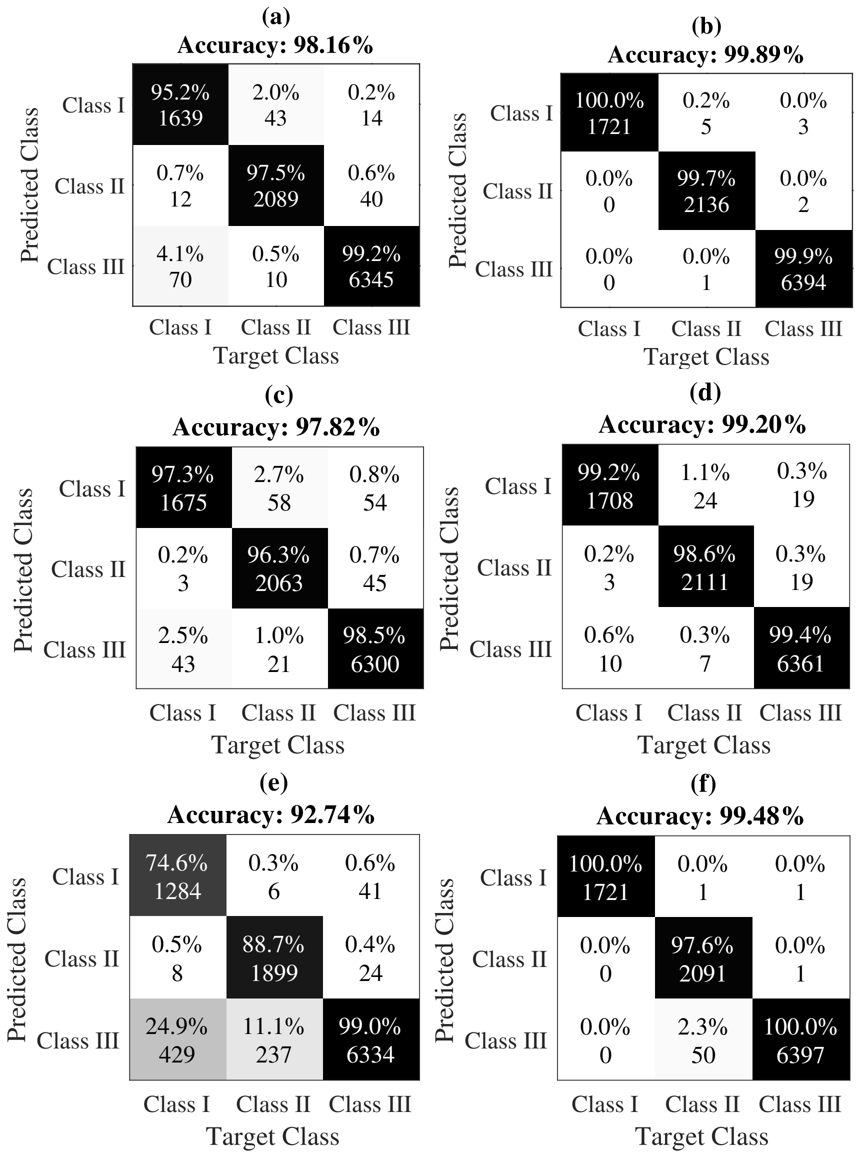

Figure 2.14: Confusion matrix for test data of Layer I classification, i.e., with respect to event zone, obtained by various classifiers: (a) multi-SVM classifier, *Case 1*; (b) multi-SVM classifier, *Case 2*; (c) k-NN classifier, *Case 1*; (d) k-NN classifier, *Case 2*; (e) DT classifier, *Case 1*; (f) DT classifier, *Case 2*.

Table 2.6: Performance metric in percentage corresponding to the second binary-classifiers in Layer I classification.

Classifier	Case	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	1	97.52	0.64	97.57	0.65	96.90	97.54
	2	99.71	0.02	99.90	0.07	99.76	99.81
k-NN	1	96.31	0.59	97.72	0.96	96.23	97.01
	2	98.55	0.27	98.96	0.38	98.43	98.76
DT	1	88.65	0.39	98.34	2.91	91.77	93.24
	2	97.61	0.01	99.95	0.62	98.46	98.77

Table 2.7: Performance metric in percentage corresponding to the third binary-classifiers in Layer I classification.

Classifier	Case	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	1	99.15	2.07	98.75	1.40	97.21	98.95
	2	99.92	0.02	99.98	0.12	99.87	99.95
k-NN	1	98.45	1.65	98.99	2.53	96.62	98.72
	2	99.40	0.44	99.73	0.97	98.86	99.56
DT	1	98.98	17.24	90.48	1.99	85.05	94.54
	2	99.96	1.29	99.22	0.05	98.92	99.59

space limitation.

2.5.4 Classification Results for Second Layer

Unlike in Sections 2.5.2 and 2.5.3, where our focus was on Layer I classification, i.e., with respect to event zone, in this section, we examine the performance for Layer II classification, i.e., with respect to event type. The Layer II events that we observed in the real-world micro-PMU data demonstrated a very *uneven* distribution across different classes. Only 2 and 4 events out of 271 Class III training events belong to Class III.A and Class III.B, respectively. Also, only 0.4% and 0.6% of events in the test dataset are in Class III.A and Class III.B, respectively.

Figs. 2.15(a) and (b) show the confusion matrices corresponding to the training data and the test data for Layer II, respectively. The *recall* metric for the binary-classifiers

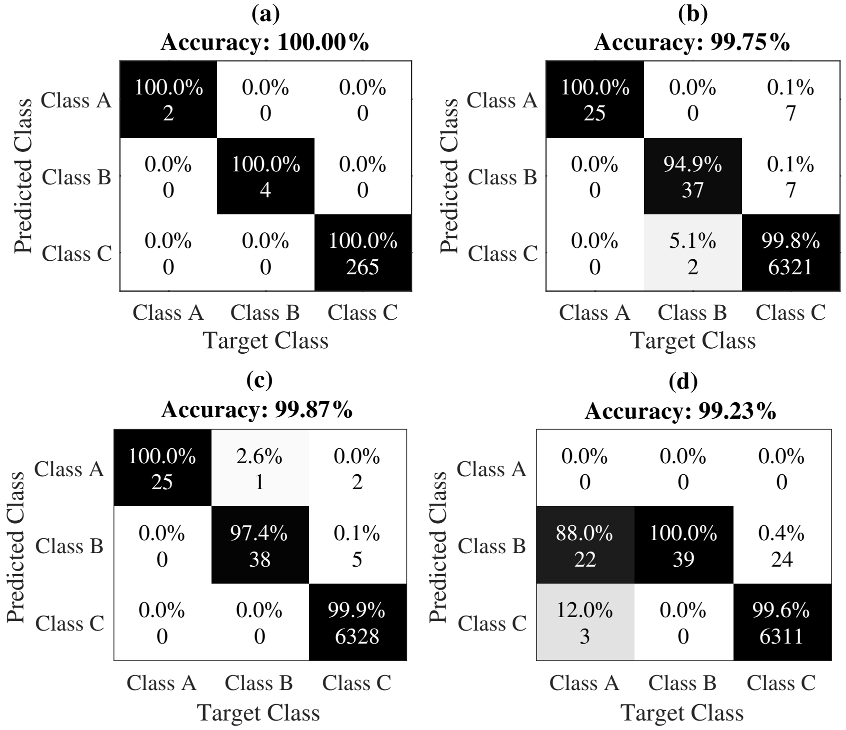

Figure 2.15: Confusion matrix for Layer II classification, i.e., with respect to classifying event type: (a) multi-SVM classifier, training data; (b) multi-SVM classifier, test data; (c) k-NN classifier, test data; (d) DT classifier, test data.

and multi-class classifier are presented in confusion matrices. Similarly, Figs. 2.15(c) and (d) show the confusion matrices of the k-NN classifier and DT classifier for Layer II events. Due to space limitation, the confusion matrices are shown only for the test data. Parameter k for the k-NN classifier is set to 3.

The above results verify the performance of the proposed classifiers in separating uneven datasets. Also, Tables 2.8 to 2.10 report the performance metrics corresponding to binary-classifiers of Layer II. From these results, we can conclude that the performance of the k-NN classifier is slightly better than the multi-SVM classifier in this particular case; although, there is a caveat about the k-NN classifier, which we will explain in the next

Table 2.8: Performance metric in percentage corresponding to the first binary-classifiers in Layer II classification.

Classifier	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	100	0.11	78.12	0	88.33	87.71
k-NN	100	0.04	89.28	0	94.46	94.33
DT	0	0	N/A	0.39	N/A	N/A

Table 2.9: Performance metric in percentage corresponding to the second binary-classifiers in Layer II classification.

Classifier	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	94.87	0.11	84.09	0.03	89.25	89.15
k-NN	97.43	0.07	88.37	0.01	92.74	92.68
DT	100	72.32	45.88	0	67.49	62.90

paragraph. As for the DT classifier, we can see in the last rows of Tables 2.8 to 2.10 that it fails to classify the events in Class III.A. Such events are incorrectly classified as Class III-B events.

We saw in the previous paragraph that the k-NN method can perform slightly better than the multi-SVM method. However, there is a catch, such performance is highly sensitive to the choice of parameter k. What was shown earlier was in fact the best possible result for the k-NN method. To see this, the impact of parameter k on the overall classification performance as well as the performance in classifying Class III.B are shown in Fig. 2.16(a) and (b), respectively. As it can be seen for both cases, the minimum error is achieved by setting $k = 3$. The results are poor in other choices of k. In particular, the

Table 2.10: Performance metric in percentage corresponding to the third binary-classifiers in Layer II classification.

Classifier	TPR	FPR	PPV	FOR	MCC	F_1Score
SVM	99.77	3.12	99.96	18.42	88.77	99.87
k-NN	99.88	0	100	9.85	94.89	99.94
DT	99.62	4.68	99.95	28.23	82.51	99.78

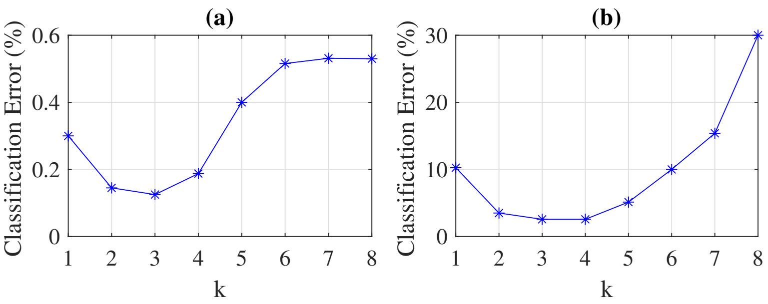

Figure 2.16: Error for the k-NN versus parameter k, in classifying Layer II events: (a) overall classification, (b) classifying Class III.B.

error in classifying Class III.B can be very high if parameter k is not carefully selected; thus, the results for the k-NN method are not as robust as those for the multi-SVM method.

2.6 Conclusions

This chapter develops new methods to resolve Big-Data challenges in transforming unprocessed data of distribution-level PMUs to bytes of informative data through data-driven event detection techniques as well as data-driven event classification techniques. Two model-free event detection methods have been proposed to pick out events from an ongoing stream of micro-PMU data. We investigate the importance of moving windows and dynamic window sizes in the event detection. Also, the effectiveness of the proposed event detection methods are compared on real-world data from two micro-PMUs on a distribution feeder in Riverside, CA. Subsequently, a novel data-driven event labeling technique was combined with different methods of classification to classify the detected events at two layers. Interestingly, we concluded that adopting classification features from the detection process can considerably improve the overall classification accuracy.

Chapter 3

Transient Load Modeling for Application in Frequency Regulation Market

3.1 Introduction

Frequency regulation is a required ancillary service in power system operation in order to maintain the grid nominal frequency, e.g., at 60 Hz in North America, by balancing power generation and power consumption on a second-by-second basis. Regulation up service is concerned with increasing generation or decreasing consumption when frequency drops below a certain threshold. Regulation down service is concerned with decreasing generation or increasing consumption when frequency exceeds above a certain threshold, c.f. [44].

System frequency in power systems is maintained with a careful balance of load and generation, mostly by adjusting the output level of generation resources. However, there is now a growing trend in practice to encourage offering frequency regulation also by controllable loads [45–47].

3.2 Related Works

Examples of controllable loads that are considered to offer frequency regulation include air-conditioning units [45, 48–51] and electric vehicles [52–56]. Also, examples of independent system operators (ISOs) that allow load participation in regulation market include the California ISO [57] and PJM [58]. In practice, load resources participate in regulation market through aggregators [45].

Besides the literature on load-assisted frequency regulation, this study is also related to the broad literature on distribution system reliability. While the majority of studies in this field addressed contingencies that are triggered by *non-electric* causes, such as a downed power line [59–63, 63, 64] our analysis is more comparable to the smaller group of papers, e.g., in [65, pp. 40-45], that addressed contingencies that are triggered by electric causes. In our case, the contingency is due to the aggregate surge current induced by a regulation down event.

3.3 Motivations and Contributions

While the system-wide benefits of using load resources for frequency regulation are studied well, e.g., see [45–47], the current literature often overlooks the potential impact on power distribution feeders, due to the lack of available monitoring and accurate distribution system models. Note that, based on the hierarchical structure of the power grid, any system-wide service that is offered by load resources is physically mediated by distribution feeders. Therefore, it is of critical to examine whether and to what extent the use of load resources could have an adverse effect on the operation of distribution systems. To the best of our knowledge this problem is not addressed yet.

Addressing the above open problem is the focus of this chapter. Specifically, we consider the regulation *down* service, which requires decreasing generation or increasing consumption when frequency exceeds a threshold [66]. We seek to understand the impact of offering regulation down service by load resources on distribution grid *reliability*. To obtain realistic results, a real-world distribution feeder is considered in Riverside, CA, where distribution-level phasor measurements units, i.e., micro-PMUs, are used to collect *high resolution, time synchronized* voltage and current data at 120 readings per second. The contributions in this study can be summarized as follows:

1. A novel *data-driven* approach is developed to use experimental micro-PMU data, on three phases, to analyze *transient* behaviors of different regulation down load resources.

2. A new method is developed to model the aggregate load transient profile, in form of an aggregate surge current profile, that is induced on a distribution feeder once a group of load resources responds to a regulation down event. This is done by applying pattern recognition methods, and taking into account factors such as sensing delay, communications delay, and load response delay,

3. A comprehensive distribution grid reliability analysis is conducted for the under-study distribution system, in presence of regulation down load resources, by taking into account the models and different characteristics of the main feeder's protection system as well as each lateral's protection system. Based on whether or not a recloser device is used in the protection system, both *momentary* and *permanent* reliability indexes are analyzed.

4. The above reliability analysis is combined with an analysis on performance score calculation in performance-based regulation markets. Accordingly, a methodology is derived to investigate the potential *trade-off* between distribution grid reliability and regulation market efficiency.

5. It is shown that the surge current induced by regulation down load resources can have severe adverse effect on the protection system, and thus on the reliability of distribution networks. One may attempt to mitigate such adverse effect by adding sufficiently large and randomly selected intentional delays to the response time of the regulation down load resources. However, this has to be done carefully, because while *reshaping* the load resources' aggregate surge current may help to avoid jeopardizing reliability, it should *not* be to the extent that it jeopardizes their performance in offering regulation down service. In practice, there may or may not exist a *safe choice* for the amount of added random delays to break the trade-off. The careful construction and analysis of the reliability-efficiency curves would be critical.

3.4 Regulation down Service and Load Resources

Frequency regulation is the mechanism of balancing power generation and power consumption in real time to maintain the stability and reliability of the power system. It can be provided by on-line generation, storage, or load resources. Frequency regulation services are either *regulation up* or *regulation down.* If generation is less than consumption, then frequency drops and *regulation up* service is needed. If generation is greater than demand, then frequency increases and *regulation down* service is needed.

The focus in this study is on regulation down service, which is needed during regulation down events. A *regulation down event* occurs when the frequency exceeds its nominal value, e.g., 60 Hz in the U.S. Regulation down service is provided by a generator, when it decreases its generation, or by a load resource, often through an aggregator, when it increases its consumption during the requested time frame [51].

It is worth pointing out that the impact of regulation up service is not considered in this study. Offering regulation up service at distribution level may potentially cause issues with power quality. However, it often does not involve surge currents and it is unlikely to raise any major reliability issue.

Many independent system operators (ISOs) have recently adopted mechanisms to allow load resources to offer regulation services. For example, the California ISO (CAISO) has a program for Non-Generator Resources (NGRs) with Regulation Energy Management (REM) to enable resources with limited energy capacities to competitively bid in the regulation market [51]. The PJM inter-connection has also introduced the RegD and RegA regulation signals to encourage fast responding loads, generators, and storage units to provide regulation services [51, 53].

Both CAISO and PJM run *performance-based* regulation markets, where the payments to regulation resources are calculated based on how fast and accurately they respond to regulation signals. Specifically, the payments to regulation resources are adjusted based on their *performance accuracy score*, a.k.a., energy precision score, which is a number between 0 and 1. A higher score indicates better regulation performance and leads to a higher payment. If the performance accuracy score of a regulation resource drops below a certain threshold, e.g., 0.5 in the CAISO regulation market, then the resource is ultimately *disqualified* to provide regulation services [66].

We denote the performance score as PS. It is calculated once for each market interval. Mathematically, we can write [66]:

$$\mathrm{PS} = \left[1 - \frac{\sum_{\sigma=1}^{T} \mid s[\sigma] - y[\sigma] \mid}{\sum_{\sigma=1}^{T} s[\sigma]}\right]^{+}, \tag{3.1}$$

where $s[\sigma]$ denotes the regulation set point at each performance accuracy evaluation time slot σ and $y[\sigma]$ denotes the mechanical output of the regulation resource at that time slot. The length of the time slots may vary depending on the ISO market. For example, CAISO examines performance accuracy once every four seconds [66]. PJM examines performance accuracy once every two seconds for fast resources and once every ten seconds for slow resources [67]. Note that, $[x]^{+} = \max\{0, x\}$ and T is the total number of time slots within the market interval. The fraction in (3.1) is a normalized measure of performance *inaccuracy* in following the regulation set points, where $\mid s[\sigma] - y[\sigma] \mid$ is the *error* in following the regulation set point.

Note that, the focus in this study is *not* on determining or coordinating the frequency regulation threshold parameters for load resources. Instead, the focus is on the moment (and a few seconds after) when a frequency event is triggered, with a preset threshold parameter. In other words, what happens *after* such event is triggered, with the load resources and consequently also with the distribution feeders, is of our concern in this chapter.

3.5 Data-Driven Load Transient Models

The impact of flexible load resources providing regulation down service on distribution system reliability can be studied under both *steady-state* and *transient* frameworks. In the steady-sate, the load resources must satisfy *load flow* constraints. This is addressed, e.g., in [68]. What is less understood is the potential adverse impact on power distribution transients, and consequently network reliability. To address this open problem, one shall investigate the transient behavior of load resources at the moment that they are called upon, i.e., when a regulation down event occurs. One option is to derive a mathematical dynamic model for each load. This would require accurate knowledge about all loads, which is difficult because these resources are owned and operated by customers. With hundreds of feeders serving hundreds of thousands of customers, the human and computational resources that are needed to build the modeling framework to represent every load would significantly outwith that of the average utility. Alternatively, in this section, we develop a *data-driven* model for load transients using data from distribution-level phasor measurement units.

3.5.1 Individual Load Transient Signatures

Distribution-level phasor measurement units, a.k.a., micro-PMUs, are new power system sensors that are deployed on distribution feeders to provide precise GPS-synchronized reading of voltage and current phasors. In this section, we use the experimental data from two micro-PMUs that are installed at the secondary side of two pad-mounted 12.47 kV to 480 V transformers in Riverside, CA. The sampling rate is 120 Hz, i.e., one sample every 8.333 msec [1, 3]. This high sampling rate, and the fact that we have access to voltage and current data, allow us capture the load transient within a data-driven framework.

Four fundamental measurements on three phases are derived from each micro-PMU: voltage magnitude, voltage phase angle, current magnitude, and current phase angle. In this study, we use the *current magnitude* to understand the dynamics of the current surge when a major load device, such as an air-conditioner, is hypothetically called upon at a regulation down event. This is because, as we will explain in Section 3.6, most practical distribution protection relays and fuses are sensitive to spikes in the current magnitude.

The central idea in our data-driven load transient modeling approach is to examine the current magnitude data during one week, July, 1 to July 7, 2016, to identify and analyze all major *current surge signatures.* To such aim, we examined the maximum magnitude of the current synchrophasors during each spike event, and compared it with the average magnitude across a time window of raw synchrophasor data, before and after the spike. If the ratio was above a certain threshold, across all three phases, then the spike signature would be captured.

In total, we analyzed $120 \times 60 \times 60 \times 24 \times 3 \times 7 = 217,728,000$ current magnitude readings per micro-PMU. Accordingly, we identified 1,803 current surge signatures with at least 40% momentary spikes in the current magnitude across all three phases. Each current spike takes from only a few milliseconds to a several hundred milliseconds. A closer look at the collected data revealed that these signatures can be clustered into a few groups with similar-shaped signatures in each group. Clustering was done using the *fuzzy C-means clustering method* from pattern recognition, c.f. [69], where C indicates the number of clusters, fixed apriori based on the following features:

- Peak to mean magnitude value;
- Start to peak time, i.e., rise time;
- Start to end time, i.e., Settling time.

The clustering problem is solved by considering five clusters, which resulted in 99.13% average dependency to centres, which indicates that the signatures in each cluster are indeed repetitions or slight variations of each other. Thus, we identified *five load clusters*, that are responsible for the majority of the current surges seen by the micro-PMUs. They are shown in Fig. 3.1. All five signatures are believed to belong to building Heating Ventilation and Air-Conditioning (HVAC) loads. Accordingly, their corresponding five load units are proper candidates to be recruited as frequency regulation load resources, see [51].

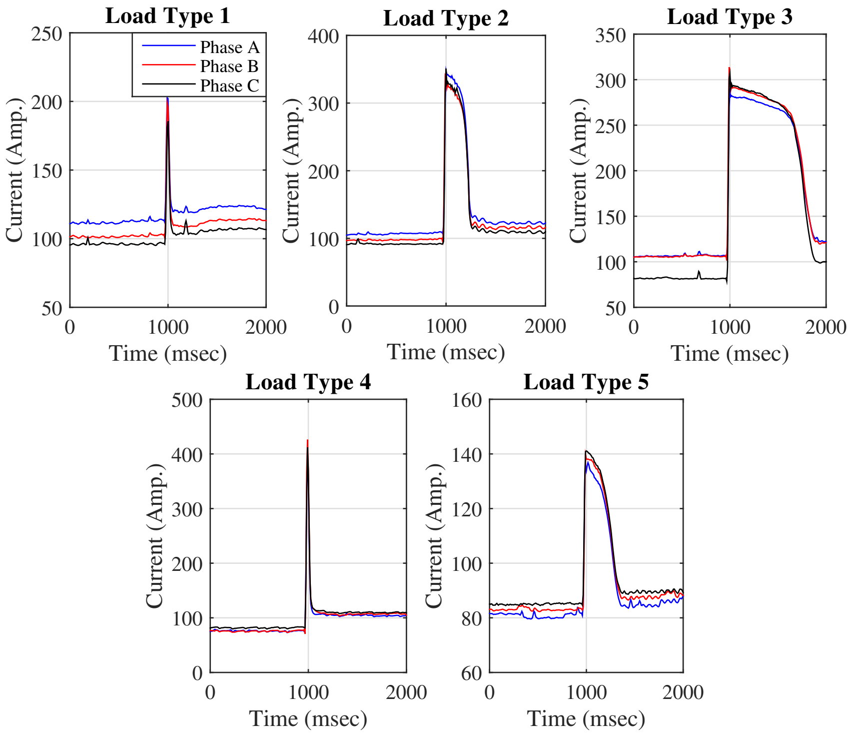

Figure 3.1: The five *current surge signatures* that represent the five load types with significant current surges that are identified through analysis of micro-PMUs data.

3.5.2 Aggregate Load Transient Profiles

Suppose there are N regulation down load resources of M different load types across a distribution feeder. In a typical distribution feeder in today's power systems, where the penetration of DERs is still relatively small, the instantaneous current that goes through the protection relay at the feeder head is the summation of the instantaneous current that is drawn by each load. Therefore, we can use the setup in Fig. 3.2 to calculate the transient current that goes through the feeder's protection system at and a few moments after a regulation down event occurs. In this figure, without loss of generality, we assume that $N > 5$ and the load resources are of the $M = 5$ types in Fig. 3.1. The transformers are considered to be ideal and linear with turns-ratio $TR = n_P/n_S$.

For the model in Fig. 3.2, the aggregate current load profile is decomposed into two components. The first component is the *feeder background load* that is the summation of the current that is drawn by all loads that do *not* offer regulation down service. At the millisecond time resolution in our analysis, the feeder background load is considered as a constant at and around the moment that a regulation down event occurs. The second component is an *aggregate current profile* of the N loads that *do* offer regulation down service. This second component is the one that generates the transient response at and around the moment that a regulation down event occurs. Here, we are essentially modelling the loads at and around the moment that a regulation down event occurs as constant current, c.f. [70].

In practice, there is often a slight *lagging*, i.e., a small delay (typically bounded by 2 or 4 seconds due to regulation market requirements), between the moment that the regulation down event occurs and the moment that the current surge signature appears for each regulation down frequency responsive load. For each load unit $i = 1, \ldots, N$ such delay is modelled as:

$$\begin{aligned}\text{Delay } \tau_i = {} & \text{Sensing Delay} + \\ & \text{Communications Delay} + \\ & \text{Load Response Delay.}\end{aligned} \tag{3.2}$$

If frequency is sensed locally by each load, then communications delay is not a factor, but the sensing delay is a factor as it is often different at different loads. If regulation down

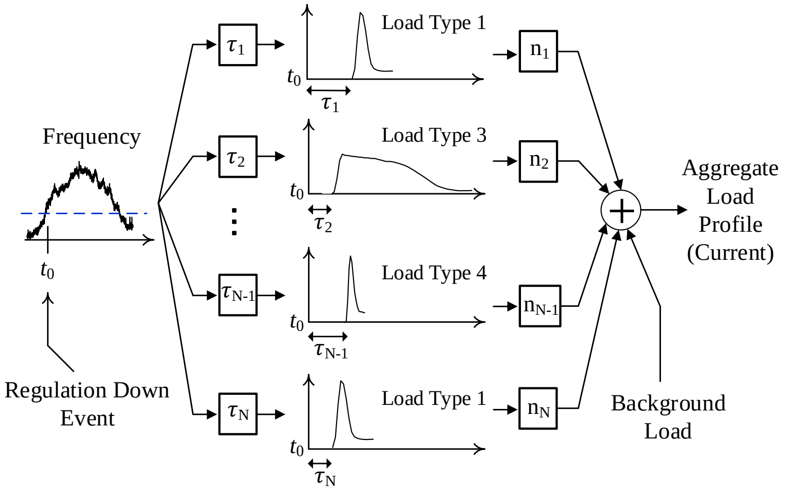

Figure 3.2: The method of calculating the aggregate load current profile at a regulation down event under different load types and different delay values.

commands are dispatched by a central entity, then communications delay has significant impact. Regardless of the method of sensing/communicating, load response delay could be different for different loads based on their internal control mechanisms. Note that, the load response delay may include an *intentional* delay component, as we will discuss below and in Section 3.7.

To gain insight on the role of delay, suppose 30% of the loads at downstream of the under-study 12.47 kV to 480 V transformer, are a mix of $M = 5$ load types that offer regulation service. If $\tau_1 = \ldots = \tau_N = 0$, i.e., there is absolutely no delay, then the single phase aggregate load transient profile is obtained with a large spike, as marked in Fig. 3.3. Of course, in practice, there are always *some* delays, as noted in (3.2). Therefore, this figure also shows two different realizations where the delays are random with a discrete uniform distribution up to 400 msec. We can see that random delays can result in different

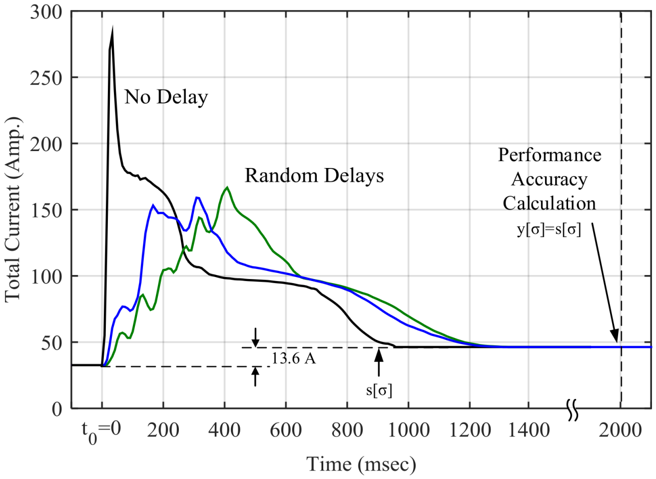

Figure 3.3: Examples of the aggregate load transient profile under different delay scenarios. The regulation down event is assumed to occur at time zero.

aggregate load transient profiles. In particular, they naturally help reducing the size of the spike in the aggregate current signature. However, if the delay is, possibly intentionally, too large, then it can affect performance accuracy, creating a *trade-off* between distribution grid reliability and regulation market efficiency. We will further evaluate and characterize such trade-off in the rest of this study.

3.6 Analysis of Distribution Grid Reliability

3.6.1 Feeder Main Protection System Model

The most common relays at distribution level are electromechanical over-current, equipped *with* or *without* reclosing capability [71]. The over-current relay operates, i.e., picks up, when the feeder current exceeds a threshold, a.k.a., the *pick-up current*. The over-current relays are categorized based upon their time-current characteristics (TCC).

The relays in the under-study feeder are the inverse definite minimum time (IDMT) relays, which are commonly used by most utilities [72].

In IDMT relays, which are of interest in this study, the TCC inversely depends on the current. Therefore, these relays are sub-categorized based on the inverseness, including short time inverse, moderately inverse, long time inverse, very inverse, and extremely inverse. The TCC of relays in pick-up/reset modes can be derived from the dynamic equation of the relay's induction disk rotation with respect to contingency current [72]:

$$K_I I^2 = m\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + K_d\frac{\mathrm{d}\theta}{\mathrm{d}t} + \frac{\tau_F - \tau_s}{\theta_{max}}\theta + \tau_s, \tag{3.3}$$

where, K_I denotes constant that relates torque to current. Also, θ and θ_{max} denote disk travel and maximum disk travel, respectively. In (3.3), m and K_d are moment inertia of disk and drag magnet damping factor, respectively. Also, τ_F and τ_s denote spring torque at maximum travel and inertia spring torque, respectively. The above equation is often approximated by disregarding the second derivation and the linear terms, as follows:

$$K_I I^2 \approx K_d\frac{\mathrm{d}\theta}{\mathrm{d}t} + \tau_s. \tag{3.4}$$

Using (3.4) to model the induction disk rotation around the pick-up current, we can now define the maximum disk rotation with respect to operation time of the relay in tripping mode as

$$\theta_{max} = \int_0^{t_{op}} \frac{\tau_s}{K_d}(I^2/I_P^2 - 1)dt. \tag{3.5}$$

where, I and I_P denote input current and peak-up current, respectively. The tripping time for each disk rotation is calculated as:

$$\begin{aligned} t_{trip} &= K_d \theta_{max} / \tau_s \left(I^2 / I_P^2 - 1 \right) \\ &= K_{trip} / \left(I^2 / I_P^2 - 1 \right). \end{aligned} \tag{3.6}$$

Finally, the tripping criteria for each disk rotation becomes:

$$\int_0^{t_{op}} (1/t_{trip}) dt = 1. \tag{3.7}$$

Note that, the operating time of relay t_{op} can be moved up/down by considering time multiplier setting (TMS) as well as by imposing intentional delay time to provide different TCC for different relay technologies. Consequently, the following equation provides a modification of (3.6) with respect to the desired exponent of plug setting multiplier, TMS, and the intentional delay time:

$$t_{op} = TMS \left(\frac{K_{op}}{(I/I_P)^{\alpha_{op}} - 1} + L \right). \tag{3.8}$$

The above procedure can be repeated to obtain the reset time for $I \leq I_P$, while reset time multiplier setting (RTMS) is imposed. The final result is as follows:

$$t_{reset} = RTMS \left(\frac{K_{reset}}{1 - (I/I_P)^{\alpha_{reset}}} \right). \tag{3.9}$$

The constant coefficients K and α depend on the relay type and the standard being used. The TCC coefficients of tripping and resetting for the common relays are given in [73]. In our analysis, we also consider the case where the feeder protection system is reinforced by an instantaneous over-current element, which operates with no intentional delay, when the current exceeds its pick-up setting.

3.6.2 Lateral Protection System Model

The lateral branches in this study are equipped with cut-out fuses, where a fusible element made of tin or silver melts under current surges and overloads. Thus, the melting period lasts from sensing an over-current to when the fuse link melts. The melting dead-time of each fuse depends on the magnitude and the duration of surge current, which is presented in TCC as minimum melt curve and the maximum total clear curve.

The melting time of a fuse-link can be calculated using either joule-integral equations [74] or heat transfer equations [75]. However, most existing models need inaccessible data of fuses in practice and rarely provide a straightforward dynamic model for intermittent heating and cooling periods. Therefore, in this study, we reformulate the thermal heat transfer model and address to include the dynamic equation for heating period, when I exceeds the minimum melting current corresponding to TCC, with respect to the contingency current:

$$\dot{T} = \frac{1}{m \cdot c}\left(R_f\left(1+\alpha_f T\right) I^2 - K_f\left(T - T_a\right)\right), \tag{3.10}$$

where $T(0) = T_a$, which denotes that equilibrium temperature of fuse at $t = 0$ is equal to ambient temperature. Also, m and c denote equilibrium mass and thermal capacity of fuse, respectively. In (3.10), the R_f and α_f denote resistance of fuse at 25 °C and resistance temperature coefficient. Also, K_f denotes the thermal conductivity coefficient of fuse. Suppose the solution of the above differential equation is $T(t) = f(I, \alpha_f, T_a, R_f, m, c, K_f)$.

In theory, I is the only input into this solution; the rest of parameters are supposed to be known. However, in practice, m, c, and K are not provided by the manufacturer. Therefore, in this study, we consider n arbitrary points I_i and t_i from the TCC of each fuse and set $T(t_i) = T_m$. We then solve a system of n non-linear equations to obtain all unknown parameters. This is done using exhaustive search. The known parameters are set based on the S&C 100 A QR speed fuse: $T_m = 800\,^{\circ}\mathrm{C}$, $R_f = 1\Omega$, and $\alpha_f = 0.0001\Omega/^{\circ}\mathrm{C}$ [75], which results $mc = 162.91$ and $K_f = 47.18$. The impact of pre-loading and ambient temperature on melting time are considered based on [76]. The dynamic model during cooling period is derived from [77].

3.6.3 Reliability Evaluation Under Regulation Down Service

We are now ready to evaluate how the distribution grid reliability is affected due to recruiting and integrating load resources into the regulation down market. We use the following three models that we developed earlier: the aggregate current profiles under regulation down events from Section 3.5.2; the relay response models from Section 3.6.1; and the fuse response models from Section 3.6.2. We will investigate whether and how the feeder's main and lateral protection systems, and consequently the distribution grid reliability, are affected in presence of distributed regulation down load resources.

The impacts on network reliability depend on the protective device settings and the magnitude and surviving time of surge current. For instance, the customers located at the downstream of a lateral experience an interruption, if the aggregated surge current at lateral exceeds the minimum melting current of fuse as well as survives enough in heating period to melt the fuse link. Also, it is possible that the aggregated surge current exceeds its predetermined pick-up current of the relay at the main feeder and stay long enough to trip the relay. Accordingly, *full* or *zonal* interruptions may occur across the feeder.

The impact of such service interruptions can be analyzed using the prevalent reliability indexes: System Average Interruption Duration Index (SAIDI), System Average Interruption Frequency Index (SAIFI), and Average Energy Not Supplied (AENS) [59–63,63]. For the purpose of the study in this study, the standard calculation of these indexes are adjusted as follows:

$$\text{SAIFI} = \frac{1}{N} \times \left(\sum_{i \in \{\Phi \cup \Gamma\}} \lambda_i^p n_i + \sum_{i \in \Psi} \rho_i n_i \right), \tag{3.11}$$

$$\text{SAIDI} = \frac{1}{N} \times \left(\sum_{i \in \{\Phi \cup \Gamma\}} \lambda_i^p r_i n_i + \sum_{i \in \Psi} \rho_i r_i n_i \right), \tag{3.12}$$

$$\text{AENS} = \frac{1}{N} \times \left(\sum_{i \in \{\Phi \cup \Gamma\}} \lambda_i^p r_i P_i + \sum_{i \in \Psi} \rho_i r_i P_i \right). \tag{3.13}$$

where N and n are the total number of customers and number of interrupted customers, respectively. Also, Φ, Γ, and Ψ denote set of annual line contingencies, set of annual transformers contingencies, and set of annual regulation down contingencies, respectively. In the above equations, λ^p denotes permanent interruption frequency. Also, r is interruption duration in hour. Here, the p denotes the probability of regulation service contingency. The first term inside the parenthesis in each case indicates the interruption frequency/duration due to typical fault occurrences in the main feeder and in the laterals, that are *not* related to the market participation of load resources. The second term, however, indicates the imposed interruption frequency/duration due to the surge current caused by the switch on events of the regulation down market participating load resources.

If the main protection system is reinforced by a recloser, see Section 3.7.2, then we shall investigate the *momentary* reliability indexes [78], such as the Momentary Average Interruption Frequency Index (MAIFI), where the recloser and lateral fuses are coordinated, see [79, 80]. Again, for the purpose of this study, the standard calculation of MAIFI is adjusted as follows:

$$\text{MAIFI} = \frac{1}{N} \times \left(\sum_{i \in \{\Phi \cup \Gamma\}} \lambda_i^m n_i + \sum_{i \in \Psi} \rho_i n_i \right). \tag{3.14}$$

where λ^m denotes momentary interruption frequency. While the first terms in (3.11)-(3.14) can be set based on the utility's reliability documents and historical data, the second terms in (3.11)-(3.14) are currently unknown in the literature and the power engineering community. However, in this study, and as we will see in the next section, we use experimental data and the methodologies described in Sections 3.7.1 and 3.7.2 to calculate these additional reliability terms.

3.7 Case Studies

The single-line diagram of the under-study 12.47 kV three-phase feeder in Riverside, CA is shown in Fig. 3.4. It is assumed to serve 10 MVA load, mostly commercial. The main feeder is 4.3 miles long, that is carved up into 10 zones based on its available protective and control devices as listed in Table 3.1. The main feeder is protected by over-current electromechanical phase relays with instantaneous unit connected to a three-phase circuit breaker. The phase over-current relays work in extremely inverse mode with operation parameters $K_{op} = 28.2$, $\alpha_{op} = 2$, $L = 0.1217$ and reseting parameters $K_r = 29.1$ and

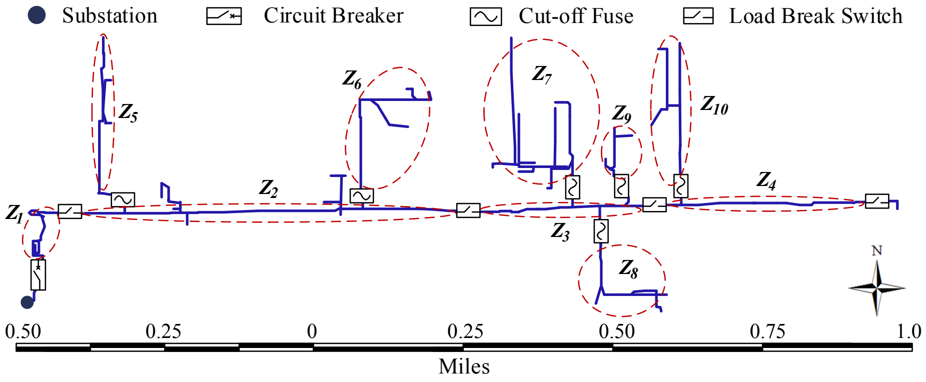

Figure 3.4: The single-line diagram of the under-study feeder in Riverside, CA.

Table 3.1: The characteristics and parameters of each zone.

Zone	Number of Customers	Length (Mile)	Number of Transformers	Load (kVA)
Z_1	0	0.12	0	0
Z_2	30	0.62	3	750
Z_3	20	0.30	2	500
Z_4	20	0.32	1	500
Z_5	30	0.35	3	750
Z_6	60	0.33	5	1,500
Z_7	90	1.05	5	2,250
Z_8	60	0.45	2	1,500
Z_9	40	0.12	3	1,000
Z_{10}	50	0.64	3	1,250
Total	400	4.30	27	10,000

$\alpha_r = 2$ [73]. The time dial setting tap and the pick-up current setting tap are set to one and eight, respectively. The instantaneous pick-up current setting tap is set to 25. Most laterals are protected by fuses as shown in Fig. 3.4. For instance, the lateral marked as Z_8, which serves two 750 kVA transformers, is protected by an S&C 100 A QR speed fuse. The real-life under-study feeder is not reinforced by a recloser; nevertheless, we do study the impact of a recloser on the overall reliability indexes.

Unless stated otherwise, we assume that the regulation down load resources are called once a day. The default probability of annual permanent and annual momentary failures on lines are set to 0.065 and 0.06 faults per km, respectively [81]. The annual permanent and annual momentary failure rates of transformers are set to 0.015 and 0.050 [81]. Based on the length of the under-study feeder, the restoration time for both tripped relay and blown fuse will be 30 minutes, while the average switching period is one hour. Line repair time is assumed to be three hours [81]. Given the often long repair time for a faulted transformer, it is typical to set the repairing time to be equal to the replacement time, i.e., five hours.

The distribution of the load types across the transformers was random, but it followed the same ratio at which the transient signatures appeared in the signature database. Those ratios are 6.99%, 7.41%, 42.98%, 20.03%, and 22.38%, for load type 1 to 5, respectively. For example, we assumed that each transformer may on average experience 6.99% of its surge current during a regulation down event from load resources of type one.

A uniform delay distribution is considered for each load resource, which represents the sensing delay, the communication delay, and the response delay, see (3.2). Even though the response delays are similar within each load type, the two former delays are different for each individual resource.

3.7.1 Impact of Regulation Down Service on Lateral Protection

In this section, we study the dynamic response of the lateral fuses to the surge current caused by regulation down load resources. Load point reliability analysis is done for two cases:

- *Case I:* with no delay in load responses,
- *Case II:* with natural delay in load responses.

For *Case II*, the added natural delay has a uniform distribution between 0 to 1 second, see (3.1). The number of each signature in aggregated surge current is considered as random to meet about 33% load participation in regulation down service.

Without loss of generality, we focus on zone Z_8, which serves two 750 kVA transformers and is protected by an S&C 100 A QR speed fuse. The analysis is done for 10,000 random scenarios to obtain the probability of fuse blowing and its impact on load point reliability indexes. The results on load point reliability evaluation for the loads located in zone Z_8 under *natural* permanent contingencies, i.e., all permanent contingencies *other than* those caused by regulation down service, are shown in Table 3.2. Note that, a fault occurred in zones Z_5, Z_6, Z_7, Z_9, and Z_{10} has no impact on the load point reliability in zone Z_8. Therefore, all entries in the rows corresponding to these zones are zero. The momentary interruption frequency due to natural momentary faults is 1.7652 f/yr in this case. The probability of fuse melting for *Case I* and *Case II* is 0.0040 and 0.0005, respectively. Accordingly, in *Case I*, customers located in zone Z_8 experience 1.46 permanent contingencies per year due to the presence of regulation down load resources, which last for 0.73 hours per year. While, in *Case II*, the frequency and duration of interruptions that are caused due to regulation down resources will be 0.1825 f/yr and 0.0912 hr/yr, respectively.

The overall momentary and permanent load point reliability indexes for zone Z_8 are shown in Table 3.3. The interruption frequency, i.e., λ, and the interruption duration, i.e., U, increase more in *Case I* than *Case II*, because the natural delay leads to damping

Table 3.2: Load point reliability evaluation under natural permanent contingencies for the loads that are located in zone Z_8.

Faulted Zone	λ^p (f/hr)		r (hr/f)	U (hr/yr)	ENS (kWh/yr)
	Line	Trans.			
Z_1	0.0125	0	t_s	0.0125	18.8292
Z_2	0.0648	0.045	t_s	0.1098	164.784
Z_3	0.0313	0.030	t_r	0.2441	366.219
Z_4	0.0334	0.015	t_s	0.0484	72.7114
Z_5	0	0	0	0	0
Z_6	0	0	0	0	0
Z_7	0	0	0	0	0
Z_8	0.0470	0.030	t_r	0.2912	436.829
Z_9	0	0	0	0	0
Z_{10}	0	0	0	0	0
Total	0.3093		2.2830	0.7062	1059.37

Table 3.3: Load point reliability indexes under natural faults and regulation down contingencies

Case	λ^m	λ^p	r (hr/f)	U (hr/yr)	ENS (kWh/yr)
I	3.2252	1.7693	0.8117	1.4362	2154.37
II	1.9477	0.4918	1.6214	0.7974	1196.24

surge current over time. Note that, the interruption duration per fault, i.e., r, is less than the base cases for both *Case I* and *Case II*, see the last row in Table 3.2. Therefore, regulation down service with and without delay increases frequency and duration indexes, while the interruption duration per fault reduces due to the imposed short-time interruptions.

Fig. 3.5 shows the dynamic responses of fuse to two sample three-phase surge currents induced by regulation down load resources for *Case I* and *Case II*, respectively. In *Case I*, the transient of surge current is 915 msec, while the heating period is from $t = 0$ to $t = 833$ msec. This results in melting the fusible element in *Phase C*. However, in *Case II*, even though the surge current survives more than 1,800 msec and the heating period lasts

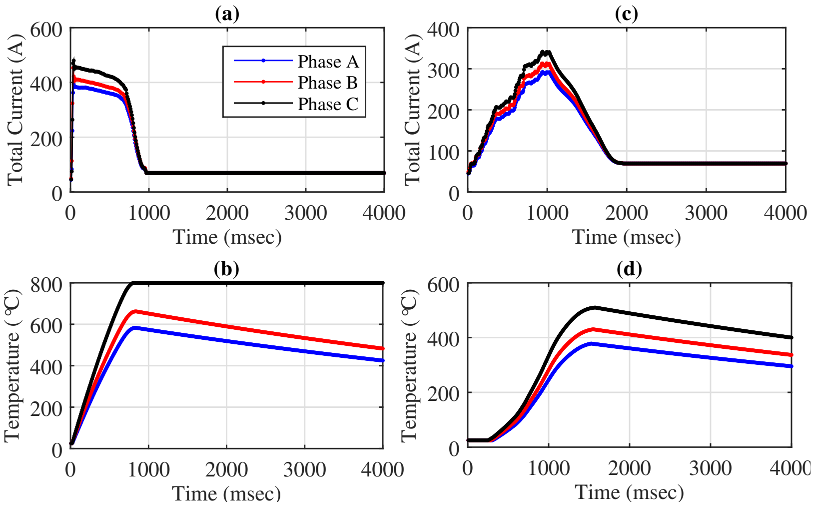

Figure 3.5: Dynamic response of fuse, in terms of total current and temperature, to a regulation down surge current: (a) and (b) are for the case *without* added random delay; (c) and (d) are for the case *with* added random delay.

for 1,250 msec, the magnitude of the surge current is not enough to melt the fuse. Thus, while customers experience outage due to regulation down service contingency in *Case I*, there is no service interruption in *Case II*.

So far, and based on the obtained results, we can conclude that the adverse effect of regulation down surge currents can be mitigated by adding sufficiently large and randomly chosen delays to the response time of regulation down load resources. However, large intentional delays can in turn have adverse effect on the performance of the regulation down service. Therefore, next, we investigate the sensitivity of fuse melting probability to intentional delay. The results are plotted in Fig. 3.6. As expected, the added intentional delay mitigates the adverse impact on lateral protection by decreasing the probability of fuse melting. Note that, the exact decaying rate in the curve in Fig. 3.6 depends on the features of the understudy protection system as well as the type, size, and number of load

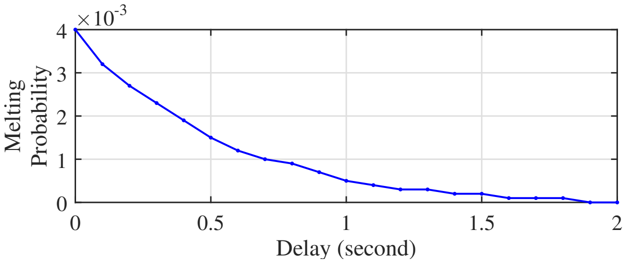

Figure 3.6: Fuse melting probability versus the delay in response time of load resources. The fuses are located on laterals.

resources.

3.7.2 Impact of Regulation Down Service on Main Protection

In this section, we study the dynamic response of the main protection relay to the surge current caused by regulation down load resources. Again, we compare *Case I* and *Case II* as defined in Section 3.7.1. It is assumed that 20-25% of loads participate in regulation down service, while the number of each load type in aggregated surge current is considered as random variable. The problem is solved 10,000 times for each case study to obtain probability of relay tripping, followed by reliability evaluation with respect to natural permanent and momentary contingencies. Network reliability indexes due to natural faults are: SAIFI = 0.3277, SAIDI = 0.8219, CAIDI = 2.5081, AENS = 20.5484, ASAI = 0.999906, and MAIFI = 1.7652. For each random scenario, the analysis is done for four different setups: *with* and *without* recloser as well as *with* and *without* instantaneous relay element. Recall that, if the feeder is reinforced by recloser, then the current surge that is induced by regulation down load resources may affect only the momentary reliability indexes, not the permanent reliability indexes.

The obtained results are presented in Tables 3.4 to 3.6, including the tripping probabilities of over-current relay, denoted by 51P, instantaneous element, denoted by 50P, as well as both over-current relay and instantaneous element, denoted by 50P & 51P. Although the tripping probabilities under contingencies are calculated per phase, the reported values are probability of circuit-breaker tripping since as mentioned before, the feeder is reinforced by a three-phase breaker. We can see that, for *Case I*, where there is no delay, the instantaneous element is more sensitive to the surge current than the over-current relay. This means that the main challenge with regulation down surge current is the magnitude rather than the transient period. In *Case II*, where there is some natural delay, the surge currents of different load resources are spread out in time; hence, the aggregate surge current often, i.e., in 97% of the random scenarios, does *not* exceed the setting of the instantaneous element. Of course, it takes longer for the aggregate surge current to settle down in this case.

We can see in Tables 3.4 to 3.6 that adding delays results in significant reduction in permanent and momentary reliability indexes, with respect to frequency, duration, and energy.

Table 3.4: Different Relay Cases and Their Corresponding Tripping Probabilities Under Regulation Down contingencies

Case	Tripping Probability		
	51P	50P	50P & 51P
I	0.1795	0.7835	0.1720
II	0.0300	0	0

Table 3.5: Different Relay Cases and Their Corresponding Reliability Indexes without Instantaneous Over-Current element Under Regulation Down contingencies

Case	Without 50P			
	Without Recloser			With Recloser
	SAIFI	SAIDI	AENS	MAIFI
I	65.8452	33.5806	839.51	67.2827
II	11.2777	6.29693	157.42	12.7152

Table 3.6: Different Relay Cases and Their Corresponding Reliability Indexes with Instantaneous Over-Current element Under Regulation Down contingencies

Case	With 50P			
	Without Recloser			With Recloser
	SAIFI	SAIDI	AENS	MAIFI
I	286.305	143.810	3595.2	287.742
II	11.2777	6.29693	157.42	12.7152

Figs. 3.7 shows the dynamic responses of the over-current relay to two sample three-phase surge currents induced by regulation down load resources for *Case I* and *Case II*, respectively. Similar results could be obtained for the instantaneous element relay. From Fig. 3.7(a), the transient of surge current is almost 1000 msec, while the tripping period last for about 700 msec. Accordingly, the surge current corresponding to *phase C* satisfies (3.7), while in the rest of phases the disk rotates in the resetting direction before travelling one cycle. However, in Fig. 3.7(c), though the surge current survives more than 1800 msec and the tripping period lasts for almost 1100 msec, the magnitude of surge current is not large enough to satisfy (3.7), see the slope of the tripping period in Fig. 3.7(b) and (d).

3.7.3 Reliability-Performance Tradeoff

So far, and based on the results that we obtained in Sections 3.7.1 and 3.7.2, we can make two main conclusions:

1. The surge current induced by regulation down load resources can have severe adverse effect on the protection system, and thus the reliability of distribution networks.
2. One can mitigate such adverse effect by adding sufficiently large and randomly chosen intentional delays to the response time of the regulation down load resources.

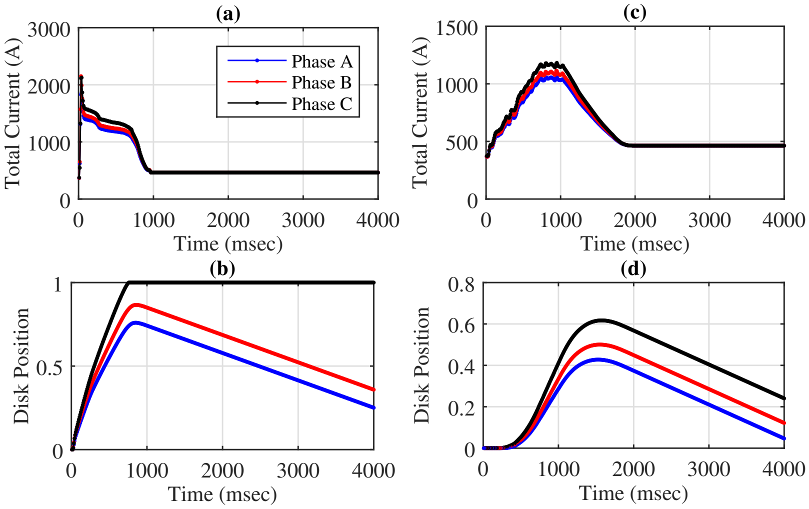

Figure 3.7: Dynamic response of relay, in terms of total current and disk position, to a regulation down surge current: (a) and (b) are for the case *without* added random delay; (c) and (d) are for the case *with* added random delay.

While the second item above is good news, it immediately raises the concern on whether adding intentional delays can have adverse effect on the performance of the regulation down service. Note that, adding intentional delays has been used previously used in other contexts in power systems, e.g., in smart meters at the time of restoring power after a regional service disconnection or blackout. But what is different here is the potential adverse impact that intentional delays may cause on the performance of regulation down service. Note that, if this “solution” ends up jeopardizing regulation down performance then it *defeats the purpose* of offering regulation down service by load resources.

The key to understand and characterize the above trade-off is to reexamine the results in Fig. 3.3 in Section 3.5.2. Accordingly, one may now ask the following two fundamental questions:

1. How much random delay shall we add to the response time of load resources such

that, while we *reshape* their aggregate surge current to avoid jeopardizing reliability, we also do *not* postpone their settling down time to the extent that it jeopardizes their regulation down service?

2. Does there always exist a *safe choice* for the amount of the added random intentional delays in order to satisfy the requirements in the first question above?

One can answer the above questions using the curves in Fig. 3.8, where we plot the tripping probability and the error in regulation service. The latter is defined in Section 3.4. Both curves are plotted versus the delay in the response time of load resources. First, consider Fig. 3.8(a), where performance accuracy is calculated once every two seconds, as in the case of fast resources in PJM. Here, any choice of delay would inevitably degrade either reliability or efficiency. Next, consider Fig. 3.8(b), where performance accuracy is calculated once every four seconds, as in CAISO. Here, there is a *safe region* for the choice of delays, without degrading reliability or efficiency.

The curves in Fig. 3.8 depend on the features of the under-study feeder as well as the type, size, and number of load resources. For example, while the current 20-25% load participation rate in Fig. 3.8(b) is manageable as long as the delays are set properly, increasing the participation rate can reduce or even eliminate the *safe region*, as shown in Fig. 3.8(c).

Ultimately, one needs to obtain the curves such as those in Fig. 3.8(c) for every feeder with large amount of load resources before trying to integrate those load resources into regulation market. One will have to properly limit the participation rate in each feeder, because above a certain level even a carefully selected delay mechanism cannot break the

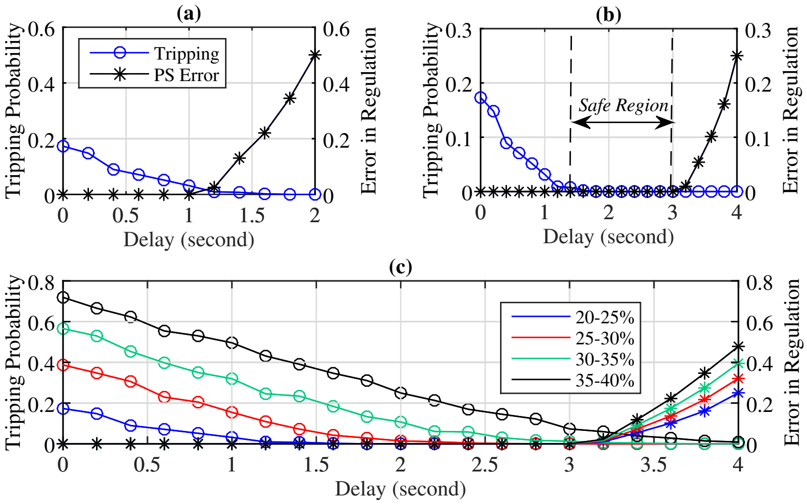

Figure 3.8: Tripping probability versus error in regulation service, both in terms of increasing the delay in response time of load resources: (a) Regulation performance is calculated at 2 sec and 20-25% of loads offer regulation down service; (b) Regulation performance is calculated at 4 sec and 20-25% of loads offer regulation down service; (c) Regulation performance is calculated at 2 see and different percentages of loads offer regulation down service.

tradeoff between distribution grid reliability and regulation market efficiency.

3.8 Conclusions

This study takes the first steps in analyzing the reliability of power distribution systems in presence of regulation down load resources. Using the parameters of a distribution feeder in Riverside, CA, together with experimental micro-PMU data from the same distribution feeder, and also by taking into account the current surge signatures of practical regulation-eligible load types, the characteristics of practical distribution-level protection devices, both on the main feeder and its laterals, and the impact of delay, e.g., due to sensing, communications, and load response, we showed that it is possible to jeopardize distribution grid reliability if several regulation down load resources are on the same feeder. Moreover, we developed a data-driven method to evaluate the trade-off between distribution grid reliability and regulation market efficiency. We showed under what conditions one may or may not break such trade-off by adding properly setup intentional random delays to the response time of the regulation down load resources. The results in this study could be of value to utilities, aggregators, and system operators.

Chapter 4

Static Load Modeling

4.1 Introduction

Accurate load modeling is necessary for power system operation, control, and planning [82,83]. Load modeling has applications in demand response [84], DERs management [85], Volt-VAR control [86], voltage stability [87], and optimal power flow [88]. The existing load modeling methods in the power systems literature can be classified into two main categories [83]: *component-based* methods and *measurement-based* methods. Component-based methods make use of a-priori information or statistical assumptions on customer appliances and other load devices. In contrast, measurement-based methods use field measurements to learn and update the load model parameters in real-time. A recent CIGRE report in [83] has found that the majority of the utilities that were surveyed across 50 countries use measurement-based methods to estimate the parameters of their load models.

Measurement-based load modeling methods can be further classified into three broad categories: static load modeling, c.f., [88–93], composite load modeling, which is a combination of static load modeling and dynamic load modeling, c.f., [94–103], and component-based load modeling, c.f., [104–107]. Our focus in this study is on the first group; which is commonly used to estimate the parameters of the ZIP load model.

When it comes to modeling loads at distribution-level and based on measurements at feeder-head, the common approach in the literature is to obtain a ZIP model for the *aggregate* load of the distribution feeder. A feeder-aggregated load model provides sufficient information to conduct most practical upper-level analysis at sub-transmission and transmission systems. However, there is still a gap in the literature to model the individual loads at each medium to low voltage load transformer. We seek to address this open problem in this study.

Our focus here is *not* on the trivial case where a measurement network such as a network of smart meters is available across the distribution feeder; because in that case the individual loads are monitored rather directly. We are instead interested in achieving individual load models by making use of limited measurements, mainly those at the feeder-head at the substation; thus to support the many utilities that are not yet equipped with a complete network of smart meters.

4.2 Related Works

There exists a rich literature on measurement-based load modeling methods that, similar to the study in chapter, make use of the measurements at feeder-head to estimate the parameters of the load model. In [88], an approximate representation of the ZIP model is proposed using semidefinite programming (SDP) relaxation of the optimal power flow (OPF) problem. In [89], a multi-state load model is developed for distribution system analysis. In [90], a measurement-based method is developed to estimate polynomial as well as exponential load models. In [91, 92], a measurement-based load model is developed for voltage stability analysis. In [93–95], a measurement-based feeder-aggregated load model is developed by using data from phasor measurements units (PMUs). In [96], a measurement-based load model is obtained using a multi-curve identification technique. In [97], measurement-based load modeling is done based on sensitivity analysis. In [98], a measurement-based per-phase load model is developed under unbalanced disturbances. Also, different methods are proposed in [99–103] to estimate composite load model parameters, while the static terms are mostly considered as feeder-head aggregate load model.

Despite the differences in methodologies and applications, the above studies on feeder-aggregated load models make use of a fundamental but similar Circuit Theory concept to analyze the feeder-head measurements during the *voltage events* that occur at the *up-stream* of the understudy distribution feeder at the sub-transmission or transmission networks, c.f. [88–103].

4.3 Motivations and Contributions

There are three limitations in measurements-based methods that rely on up-stream voltage events as the main enabler for load modeling. First, major up-stream voltage events may not occur frequently; therefore, online load modeling may not be possible for several hours until one such event occurs. Second, at every occurrence of an up-stream voltage event, there is practically a different combination of the individual loads across the distribution feeder that are switched on; therefore, the measurements obtained from different up-stream voltage events represent different snapshots of the understudy distribution feeder; thus, such measurements may not be directly comparable even for the purpose of modeling the aggregate feeder load. Third, the aggregate load models are often not useful to conduct distribution-level analysis, such as to study the impact of DERs or issues related to power quality across distribution feeders. An individual load model is needed in these cases.

This study proposes a novel method for individual load modeling in power distribution systems. The main technical contributions in this use-case can be summarized as follows:

1. Instead of or in addition to using up-stream voltage events as the enabler for load modeling, which is commonly used in the literature such as in [88–93], we make use of load switching events across the distribution feeder itself. In principle, once a load is switched, the switching event changes the voltage in the rest of the loads, which causes variation in their active power and reactive power usage; thus allowing us to estimate load parameters of the rest of the individual loads.

2. The proposed method can estimate load modeling parameters of individual loads using measurements only at the feeder-head, i.e., at the distribution substation. No measurement is needed at individual loads.

3. We provide a theoretical foundation to determine the conditions on the extent of measurements needed to successfully achieve the individual load models.

4. The proposed individual load modeling can be done *sequentially* to obtain the load

models for a subset of loads as more measurements become gradually available. Moreover, we proposed a variation of our method by using a forgetting factor so as to support estimating the parameters for time-varying individual loads.

5. By solving a non-liner least-squares problem, it is shown that the proposed methodology can be extended to utilize different types of *redundancy* in measurements in order to improve load modeling accuracy and robustness.
6. We develop a residue-based bad data detection and identification method to identify and drop load configuration measurements with erroneous switching status; thus to ensure the accuracy of the load models. This also helps with the cases with imperfect knowledge of line impedances, switch statuses, and other system parameters.

4.4 Problem Statement

Consider a distribution feeder with $n \geq 2$ buses. For now, and for the simplicity of discussion, suppose the distribution network does not have any lateral, as shown in Fig. 4.1(a). The case for distribution feeders with laterals is discussed in Section 4.6.3. Depending on which individual loads are turned on and which individual loads are turned off, there can be a total of $2^n - 1$ possible load configurations, excluding the no load situation. As time goes by, a variety of load configurations occur, changing the voltage and power that are measured at the feeder-head. The measuring can be done, for example, by using micro-PMU. The measurements corresponding to a total of 10 load configurations are shown in Fig. 4.1(b), indexed as $k = 1, \ldots, 10$. The corresponding load configurations are denoted by $m_1, \ldots, m_{10}$, which take numbers between 1 to $2^6 - 1 = 63$; because $n = 6$ for the network in

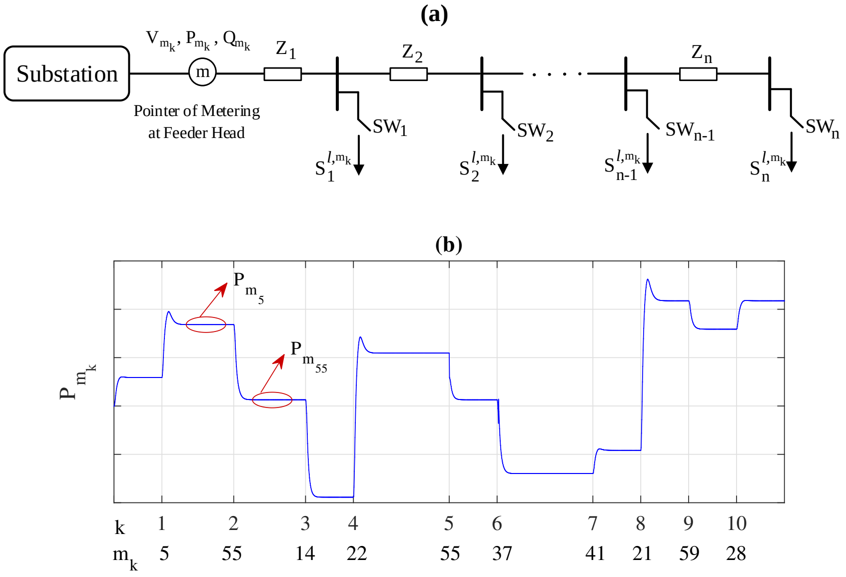

Figure 4.1: An example distribution feeder: (a) the single line diagram of the feeder; (b) the total load that is measured at the feeder head.

this example. Each individual load can be an arbitrary combination of constant impedance, constant current, and constant power load components. The complex power that is drawn by the load at bus i under load configuration m_k is denoted by S_i^{l,m_k}. In this chapter, we seek to answer the following question: *Can we model each of the n individual loads in Fig. 4.1(a) by studying a sequence of measurements at the feeder-head in Fig. 4.1(b)*?

Of course, there exist some special cases for which the above problem is somewhat trivial. For example, if at a load configuration, there is only a single load that is turned on and all other loads are turned off, then modeling that single load is relatively easy, because what is being monitored at the feeder head is the single load itself plus the power loss on distribution lines. However, beyond such relatively trivial special cases, answering the above question can be challenging.

Unlike in the literature, e.g., in [93–98], where load modeling is done only at the very few instances where there is a major step change in the *upstream voltage*, load modeling in this study is done much more frequently, once every time that there is a load switching anywhere across the feeder itself. Moreover, while the studies in [93–98] are concerned with modeling the entire feeder as an *aggregate load*, this study is concerned with modeling the *individual loads* on the feeder.

4.5 Load Modeling Method

4.5.1 The First Set of Equations: Circuit Model

For a given k, consider the measurements that are obtained at the feeder-head during load configuration m_k. According to the law of complex power conservation, we have:

$$\begin{aligned} S_{m_k} = &\sum_{i=1}^{n} \left(S_i^{l,m_k} SW_i^{m_k} \right) \\ &+ \sum_{j=1}^{n} Z_j \left| \sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^{*} SW_d^{m_k} \right|^2, \end{aligned} \tag{4.1}$$

where the first term is the total load, and the second term is the total loss. Binary variable $SW_i^{m_k}$ is one if the individual load i is turned on during load configuration m_k; and zero otherwise. Note that, we could define a notation for current and replace (4.1) with an equation based on the Kirchhoff's Current Law (KCL). However, in this chapter, we present the circuit model only in terms of complex power and voltage phasor.

Next, we write the Kirchhoff's Voltage Law (KVL) for the circuit in Fig. 4.1(a). Any *loop* can be used for this purpose. In fact, there are $n(n + 1)/2$ loops in this circuit that can be used to write the KVL equations. However, as proved by using Graph Theory in [108, 109], one can write only n *independent* KVL equations in this circuit. This can be achieved by writing the KVL equations for the n loops between the substation and every node $i = 1, \cdots, n$, as shown below:

$$V_i^{m_k} = V_{m_k} - \sum_{j=1}^{i} Z_j \Bigg(\sum_{d=j}^{n} \Big(\frac{S_d^{l,m_k}}{V_d^{m_k}} \Big)^* \, SW_d^{m_k} \Bigg). \tag{4.2}$$

Together, the $n+1$ independent complex nonlinear equations in (4.1) and (4.2) provide the model for the circuit. The complex power S_{m_k} and the voltage phasor V_{m_k}, which are measured at the feeder-head, as well as the lines impedances Z_j for $j = 1, \ldots, i$, are the *known* parameters in these equations. In contrast, the individual load complex power S_i^{l,m_k} and the individual node voltage phasor $V_i^{m_k}$ are the *unknowns* to be determined for all $i = 1, \ldots n$. For any load configuration m_k, the number of unknowns can be counted as:

$$n + \sum_{i=1}^{n} SW_i^{m_k}, \tag{4.3}$$

where the first term counts $V_i^{m_k}$ for every node $i = 1, \ldots, n$; and the second term counts S_i^{l,m_k} for every load i that is turned on under load configuration m_k, for which $SW_i^{m_k} = 1$.

For every load configuration m_k, the system of non-linear equations in (4.1) and (4.2) can have a solution only if one single load is switched on, in which case the summation in the second term in (4.3) is one. In all other cases, the system of nonlinear equations in (4.1) and (4.2) is *under-determined*.

Suppose more load configurations occur as time goes by. This will provide new equations and also introduce new unknowns. Suppose the measurements are available for c distinct load configurations. The following upper bound always holds:

$$c \leq \min\{\xi, 2^n - 1\}. \tag{4.4}$$

In Fig. 4.1(b), we have $\xi = 10$ but $c = 9$; because load configurations m_2 and m_5 are the same.

Given the measurements at the feeder-head for c distinct load configurations, the number of unknowns becomes:

$$\text{Number of Unknowns} = c \times n + \sum_{k=1}^{c} \sum_{i=1}^{n} SW_i^{m_k}, \tag{4.5}$$

and the number of independent equations increases to:

$$\text{Number of Equations} = c \times (n+1). \tag{4.6}$$

The first term in (4.5) counts voltage phasor $V_i^{m_k}$ in every node $i = 1, \ldots, n$ and each of the c distinct load configurations $k = 1, \ldots, c$. The second term in (4.5) counts complex power S_i^{l,m_k} for every load $i = 1, \ldots, n$ that is turned on under load configuration m_k for each of the c distinct load configurations.

The system of non-linear equations in (4.1) and (4.2) for c distinct load configurations can have a solution *only if* a single load is switched on in every load configuration

k. This occurs *only if* the following inequality holds:

$$\sum_{k=1}^{c} \sum_{i=1}^{n} SW_i^{m_k} \leq c. \tag{4.7}$$

In that case, the total number of unknowns in (4.1) and (4.2) would be less than or equal to the total number of independent equations in (4.1) and (4.2). In all other cases, we lack sufficient independent equations. This can potentially be resolved by deriving new equations from load models, as we discuss next.

4.5.2 The Second Set of Equations: Load Model

Any load with any combination of active and reactive power consumption can be expressed in the generic form of a ZIP load [83]. Specifically, for any two distinct load configurations m_k and m_h, the complex power consumption at the individual load at bus i can be modeled as:

$$S_i^{l,m_k} = P_i^{l,m_h} \left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|} \right)^{n_{p_i}} + jQ_i^{l,m_h} \left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|} \right)^{n_{q_i}}. \tag{4.8}$$

If $n_p = 0, 1, 2$, then the active power component of the load is constant power, constant current, and constant impedance, respectively. For all other values of n_p, the active power component is a combination of these three load elements. The reactive power component can be defined similarly using n_q.

Note that, for each individual load i, one can construct the load model in (4.8) only if there *do* exist at least two load configurations m_k and m_h in which load i is turned on. In other words, we cannot introduce the load model in (4.8) for individual load i unless the following inequality holds:

$$\sum_{k=1}^{c} SW_i^{m_k} \geq 2. \tag{4.9}$$

Thus, unless we state otherwise, we assume that the above condition holds for all loads $i = 1, \ldots, n$. The special case when (4.9) does *not* hold will be discussed in Section 4.6.2.

The load models of the form in (4.8) can provide additional equations that can be combined with the equations in (4.1) and (4.2) to obtain the unknowns that we identified in (4.5). However, the load equations in (4.8) also introduce new unknowns, because n_{p_i} and n_{q_i} for buses $i = 1, \ldots, n$, are not known. Therefore, the number of new unknowns becomes:

$$\text{Number of Unknowns} = n. \tag{4.10}$$

Here we count $n_p + jn_q$ as one unknown complex number.

Next, we need to identify how many of the new equations in the form of (4.8) are *independent*; and thus useful for identifying the unknown individual loads in our system. This matter is addressed in a Theorem as presented below.

Theorem 1. *Given c distinct load configurations, the number of independent complex equations in the form of (4.8) is:*

$$\text{Number of Equations} = \sum_{i=1}^{n} \sum_{k=1}^{c} SW_i^{m_k} - n. \tag{4.11}$$

Proof: For each individual load i, let us define:

$$c_i = \sum_{k=1}^{c} SW_i^{m_k}. \tag{4.12}$$

One can write a load model equation of the form in (4.8) for *any two distinct* load configurations m_k and m_h. Therefore, the total number of load model equations of the form in (4.8) that we can write for each individual load i is:

$$\binom{c_i}{2} = \frac{1}{2}\, c_i \times (c_i - 1). \tag{4.13}$$

However, these many equations are *not* independent; because the expression in (4.8) preserves *transitive relation* [110]. For instance, for a given individual load i, suppose we have $c_i = 3$. That is, suppose load i is turned on in three load configurations m_k, m_h, and m_g. From (4.8), we have:

$$P_i^{l,m_k} + jQ_i^{l,m_k} = P_i^{l,m_h}\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|}\right)^{n_{p_i}} + jQ_i^{l,m_h}\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|}\right)^{n_{q_i}}, \tag{4.14}$$

$$P_i^{l,m_k} + jQ_i^{l,m_k} = P_i^{l,m_g}\left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|}\right)^{n_{p_i}} + jQ_i^{l,m_g}\left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|}\right)^{n_{q_i}}, \tag{4.15}$$

$$P_i^{l,m_h} + jQ_i^{l,m_h} = P_i^{l,m_g}\left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|}\right)^{n_{p_i}} + jQ_i^{l,m_g}\left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|}\right)^{n_{q_i}}, \tag{4.16}$$

However, the above equations are *not* independent due to their transitive relationship. Any one of the three equations in (4.14)-(4.16) can be obtained by applying proper non-linear operators to the other two equations. For instance, we can obtain (4.16) by *dividing* the real and the imaginary parts of (4.14) by the real and the imaginary parts of (4.15), respectively, as shown below:

$$\begin{aligned} \frac{P_i^{l,m_k}}{P_i^{l,m_k}} &= \frac{P_i^{l,m_h}}{P_i^{l,m_g}} \times \left[\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|} \right)^{n_{p_i}} \Big/ \left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|} \right)^{n_{p_i}} \right] \\ \Rightarrow \quad P_i^{l,m_h} &= P_i^{l,m_g} \left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|} \right)^{n_{p_i}}. \end{aligned} \tag{4.17}$$

The calculation for the imaginary part is similar and omitted for brevity. The above transitive relationship for the system of nonlinear equations in (4.8) is the generalization of the concept of linear dependency in system of linear equations.

In order to obtain the largest subset of independent load equations from the total number of equations in (4.13), first, for each individual load i, we consider one of the load configurations as the *reference configuration*. Next, we write $c_i - 1$ different load equations of the form in (4.8) based on the reference configuration together with each of the $c_i - 1$ remaining load configurations. All such $c_i - 1$ load configurations are, by-construction, independent. Once we repeat this procedure for all n loads, we can obtain several independent load equations, at a total number equal to the one in (4.11). ■

4.5.3 Combining Circuit Model and Load Model

Suppose we have collected the measurements at the feeder head under c distinct load configurations. Also, suppose we combine the Circuit Model in Section 4.5.1 and the Load Model in Section 4.5.2. We propose Algorithm 1 to obtain the individual loads model. This algorithm can be used as long as enough load configuration measurements are available. This raises the following question: What is the smallest c, denoted by $c_{\min}$, in order to solve the system of nonlinear equations in (4.6) and (4.11) to obtain the set of unknowns in (4.5) and (4.10)? This question is answered in a Theorem as presented below.

Theorem 2. *Suppose the inequality in (4.9) holds for all loads* $i = 1, \ldots, n$, *where* $n > 2$. *The minimum number of distinct load configurations in order to solve the system of equations in (4.6) and (4.11) to obtain the unknowns in (4.5) and (4.10) is:*

$$c_{\min} = 2\, n. \tag{4.18}$$

Proof: We need the number of unknowns to be less than or equal to the number of independent equations. From (4.6), (4.11), (4.5), and (4.10), this can be expressed as the following inequality:

$$\begin{aligned} & n \times c + \sum_{i=1}^{n} \sum_{k=1}^{c} SW_i^{m_k} + n \\ & \leq c \times \big(n+1\big) + \sum_{i=1}^{n} \sum_{k=1}^{c} SW_i^{m_k} - n. \end{aligned} \tag{4.19}$$

Once we cancel out the common terms on both sides and reorder the rest of the terms, we can express the above inequality as $c \geq 2n$. Therefore, $c_{\min} = 2n$. ■

Recall from Section 4.4 that our goal is to model each of the n individual loads in Fig. 4.1(a) by studying the sequence of measurements in Fig. 4.1(b). Theorem 2 indicates how far in the sequence of measurements we must go before we can obtain the load models. From Theorem 2, any arbitrary but distinct $2n$ load configurations that satisfy (4.9) for all n loads can be used in Algorithm 1 to obtain the unknowns and model the loads.

There are several algorithms available in the literature to solve a system of nonlinear equations [111,112]. In this study, we use the Levenberg-Marquardt algorithm, which is iterative and commonly used in curve-fitting problems [112]. The initial guess for all unknown parameters in both circuit model and load model are set to 1 per-unit.

Note that, the unknowns that are of interest in this study are n_{p_i} and n_{q_i}. The other unknowns, i.e., $V_i^{m_k}$ and S_i^{l,m_k}, act as auxiliary variables to help us identify n_{p_i} and n_{q_i}.

Algorithm 1 Individual Load Modeling

```
Step 1: Obtain the Circuit Model
for each load configuration $m_k$ do
    Apply law of complex power conservation $\rightarrow$ (1)
    Apply KVL between substation and every node $i$ $\rightarrow$ (2)
end for
Step 2: Obtain the Load Model
for every node $i$ do
    Calculate complex power consumption $\rightarrow$ (8)
end for
Step 3: Solve the system of equations in (1), (2), and (8)
return Solutions
```

4.6 Remarks and Extensions

In this section, we discuss some additional aspects of the proposed load modeling method, including some extensions.

4.6.1 Redundant Load Configurations

Recall from Section 4.4 that there exist $2^n - 1$ possible distinct load configurations for an n-bus system. From Theorem 2, as few as $2n$ of them is sufficient to solve the load modeling problem. But what if we continue collecting new distinct load configurations beyond $c_{\min}$? What can we do with the remaining $2^n - 1 - 2n$ distinct load configurations? Furthermore, is there any benefit to also look into the duplicate load configuration events?

For example, recall that load configuration m_5 was the duplicate of load configuration m_2 in Fig. 4.1(b). Finally, can we also make use of the upstream voltage events? Note that, so far, we did not use such measurements.

To address the above questions, we collectively refer to the following items as redundant load configurations:

- Any additional distinct load configuration for $c > c_{\min}$;
- Any duplicate of an existing load configuration;
- Any major upstream voltage event.

We are interested in using the above redundant configurations to enhance load modeling accuracy in presence of errors in measurements. Note that, the above redundant additional and duplicate load configurations both introduce new unknowns in the Circuit Model, but they do *not* change the number of unknowns in the Load Model.

It is not difficult to construct the new equations and identify the new unknowns similar to (4.6)(4.11) and (4.5)(4.10). The details are omitted due to space limitation. Let $F(V_{m_k}, S_{m_k}, V_i^{m_k}, S_i^{l,m_k}, n_s) = 0$ denote the resulting system of equations. We can obtain the individual load parameters by solving the following non-linear least-squares problem:

$$\underset{V_i^{m_k}, S_i^{l,m_k}, n_s}{\text{minimize}} \left\| F(V_{m_k}, S_{m_k}, V_i^{m_k}, S_i^{l,m_k}, n_s) \right\|_2 . \tag{4.20}$$

Again we can use a variation of the Levenberg-Marquardt algorithm, called the damped least-squares algorithm, to solve the problem in (4.20). As in Section 4.5.3, our ultimate goal here is to obtain the values of n_{p_i} and n_{q_i} for all n loads. The difference compared to

Section 4.5.3 is that we now have more equations and more auxiliary variables. Of course, if the redundant equations are removed and if there is no error in measurements, then solving problem (4.20) reduces to solving a system of nonlinear equations, just like in Section 4.5.3.

4.6.2 Sequential Load Modeling

What if the inequality in (4.9) does *not* hold for all n individual loads? Let us define:

$$\hat{n} = \sum_{i=1}^{n} \mathbb{I}\left(\sum_{k=1}^{c} SW_i^{m_k} \geq 2\right), \tag{4.21}$$

where $\mathbb{I}(\cdot)$ is a 0-1 indicator function. Here, $\hat{n} \leq n$ denotes the number of individual loads that *do* satisfy the inequality in (4.9). Let us also define $\hat{c}$ as a subset of c distinct load configurations in which all individual loads that are turned on do satisfy the inequality in (4.9). Of course, we have $\hat{c} \leq c$. We can now apply Theorem 2 to $\hat{n}$ and $\hat{c}$ and similarly achieve $\hat{c}_{\min} = 2\hat{n}$.

The above analysis can lead to developing a sequential load modeling approach, as shown in Algorithm 2. As time goes by, a variety of load configurations occur. Our algorithm must wait until such time that there exists an $\hat{n} > 2$ and $\hat{c} = 2\hat{n}$. It can then obtain the model for the $\hat{n}$ loads. As measurements from more load configurations become available, Algorithm 2 gradually and sequentially models all individual loads across the feeder.

Algorithm 2 Sequential Load Modeling

```
while there is a new load configuration measurement do
    Consider $\hat{c}$ load configurations measured so far.
    Set $\hat{n}$ according to (21).
    if Theorem 2 is satisfied for $\hat{c}$ and $\hat{n}$ then
        Run Algorithm 1 for $\hat{c}$ and $\hat{n}$
        return Solutions
    end if
end while
```

4.6.3 Distribution Feeder with Laterals

The feeder in Fig. 4.1(a) does not have any lateral. However, in practice, most distribution feeders do have laterals. In order to incorporate laterals in our model, we need to revise and replace the Circuit Model in (4.1) and (4.2) with

$$\begin{aligned} S_{m_k} &= \sum_{i=1}^{n} \left(S_i^{l,m_k} SW_i^{m_k} \right) \\ &\quad + \sum_{j=1}^{n} Z_j \left| \sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} e_j^T B e_d \right|^2, \end{aligned} \tag{4.22}$$

$$\begin{aligned} V_i^{m_k} &= V_{m_k} \\ &\quad - \sum_{j=1}^{i} Z_j e_j^T B e_i \Bigg(\sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} e_j^T B e_d \Bigg), \end{aligned} \tag{4.23}$$

where B is the *bus-injection to branch-flow matrix* [113], e_i is the canonical basis of $\mathbb{R}^n$, and $e_i^T B e_j$ is 1 if the power injection of bus j flows on branch i, and zero otherwise. An example to construct matrix B is given in Section 4.7-A.

4.6.4 Imperfect Knowledge of Load Switching Status

So far, we have assumed that the switching status $SW_i^{m_k}$ of the individual loads is already known in each measured load configuration m_k. This can be achieved by using the existing methods that are designed to identify load switching on distribution networks. For example, in [5], the location of each load switching event along the distribution feeder is identified by applying the Compensation Theorem from Circuit Theory to the measurements that are available at the feeder-head and at the feeder-tail. No measurement is needed at individual loads.

We *do* recognize that, in practice, the methods such as the one in [5] are *not* entirely precise. As a result, our knowledge of the load switching status is imperfect and may carry errors. However, this issue can be resolved by the use of the redundant load configurations that we proposed in Section 4.6.1. Specifically, the use of redundant load configurations allows us to conduct *bad data detection* in order to identify and drop load configurations with erroneous switching status.

Let r denote the *residue vector* for function $F(\cdot)$ in the least-squares problem in (4.20). Using the Chi-squares test [114], we detect erroneous load configurations if:

$$\mathbb{I}\left(r^T r > \chi^2_{v,\alpha}\right) = 1, \tag{4.24}$$

where $\chi^2_{v,\alpha}$ is the Chi-square distribution function corresponding to a detection confidence level with probability α, and with $v = \xi - 2n$ degrees of freedom in solving problem (4.20) due to having $\xi - 2n$ redundant equations. Once the existence of an erroneous switch status is detected using the Chi-square test, we next apply the largest normalized residual (LNR) test to identify such erroneous load configurations and drop them from the analysis to assure accurate and robust load modeling.

4.6.5 Time-Varying Load Modeling

In practice, the parameters of the load models may change over time. In order to track time-varying load modeling parameters, we propose to apply *forgetting factors* to load configurations as well as using the Chi-squares test and the LNR test as in Section 4.6.4. Specifically, suppose the measurements corresponding to c load configurations are available, where $c > c_{\min}$ and $v = \xi - 2n$ is the number of redundant configurations. We define the forgetting factors as follows:

$$\beta(k) = \begin{cases} \lambda^{v-k+1} & k = 1, \ldots, v \\ 1 & k = v+1, \ldots, 2n \end{cases} \tag{4.25}$$

where $0 < \lambda \leq 1$. We use $\beta(k)$ as the *weight* corresponding to equation number k in the non-linear least-squares problem in (20). This allows us to estimate the load model parameters based on the most updated load configurations. Similar to the bad data detection method that we used in Section 4.6.4 to detect error in load switching status, here, we use the Chi-squares test and the LNR test in order to detect, identify, and drop load configurations that impose error to the aforementioned weighted non-linear least-square problem. As a result, whenever a load changes, the first few subsequent load configuration measurements are dropped until the available load configuration measurements are sufficient to achieve an accurate load modeling result with low and consistent residues.

4.6.6 The Use of Multiple Sensors

The methodology that is proposed in this study is intended to allow individual load modeling by making use of measurements from as few as only a single sensor that is installed at the feeder head. Furthermore, in principle, the proposed methodology is applicable to any distribution feeder with an arbitrary number of nodes and laterals. Nevertheless, as the number of nodes increases, one may start facing numerical issues in order to solve the formulated nonlinear system of equations due to its size. Therefore, when it comes to long distribution feeders with a large number of nodes, such as the case of the IEEE 123-bus test system in Section 4.7.9, one needs to start installing additional sensors. This can be done by simply installing the additional sensors at the head of the long laterals. Importantly, the methodology that is needed to solve the load modeling problem remains the same. On one hand, each lateral whose aggregate load is measured at its head, can be seen as an independent load modeling problem, whose individual loads can be modeled using Algorithms 1 and 2. On the other hand, the measured aggregate load of such laterals can be used as known parameters when it comes to solving the combined circuit models and load models for the *rest* of the feeder, again using Algorithms 1 and 2. This essentially breaks down the load modeling problem into smaller problems of the same type.

4.7 Case Study and Results

All the case studies in this section are done in PSCAD [115] to construct the distribution feeder and all load configurations.

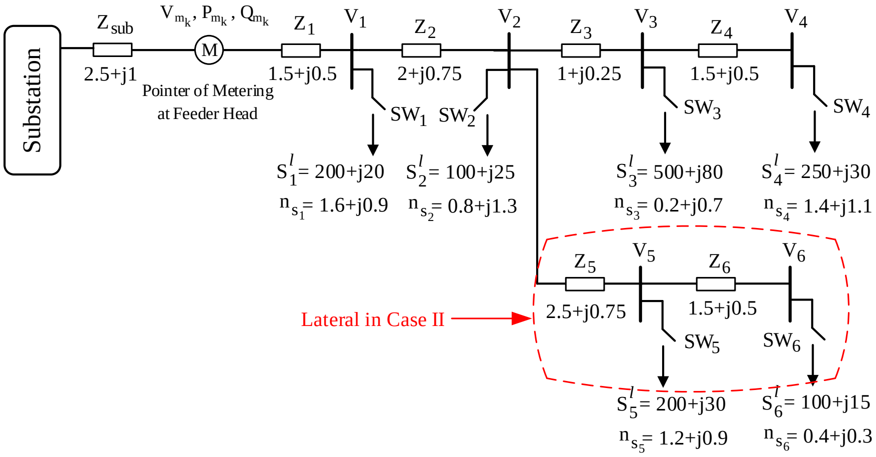

Figure 4.2: Single-line diagram of the feeder in the two basic test cases in Section 4.7.1. Case I does not include the lateral. Case II does include the lateral.

4.7.1 Basic Results With and Without Laterals

This section investigates the effectiveness of the proposed method based on using the sufficient number of load configurations as presented in Theorem 2. To do so, we apply the proposed method to two cases as defined in Fig. 4.2:

- Case I: Feeder *without* no lateral;
- Case II: Feeder *with* lateral.

Both cases are shown in Fig. 4.2. The load model parameters are considered based on the experimental studies in [116, 117]. For example, the load in bus 1 is a combination of residential appliances, including coffee maker, LED and tungsten lights, and LCD television, as described in [116]. Recall from Section 4.4 that there are 2^4-1 and 2^6-1 possible distinct load configurations in Case I and Case II, respectively. From Theorem 2, the individual load

Table 4.1: Load configurations for Case I in Section 4.2.

Configuration	SW_1	SW_2	SW_3	SW_4	Time
m_1	1	0	0	0	$[0\ ,\ t_1]$
m_2	0	1	1	0	$[t_1\ ,\ t_2]$
m_3	0	1	0	1	$[t_2\ ,\ t_3]$
m_4	0	0	1	1	$[t_3\ ,\ t_4]$
m_5	1	1	1	0	$[t_4\ ,\ t_5]$
m_6	1	1	0	1	$[t_5\ ,\ t_6]$
m_7	1	0	1	1	$[t_6\ ,\ t_7]$
m_8	1	1	1	1	$[t_7\ ,\ t_8]$

Table 4.2: Load configurations for Case II in Section 4.2.

Configuration	SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_1	1	0	0	1	0	0	$[0\ ,\ t_1]$
m_2	1	0	0	1	0	1	$[t_1\ ,\ t_2]$
m_3	1	1	0	0	1	0	$[t_2\ ,\ t_3]$
m_4	0	0	1	1	0	1	$[t_3\ ,\ t_4]$
m_5	0	1	1	1	0	1	$[t_4\ ,\ t_5]$
m_6	1	1	0	0	1	1	$[t_5\ ,\ t_6]$
m_7	0	1	1	0	1	1	$[t_6\ ,\ t_7]$
m_8	1	0	1	1	1	0	$[t_7\ ,\ t_8]$
m_9	0	1	1	1	1	1	$[t_8\ ,\ t_9]$
m_{10}	1	1	1	0	1	1	$[t_9\ ,\ t_{10}]$
m_{11}	1	1	1	1	1	0	$[t_{10}\ ,\ t_{11}]$
m_{12}	1	1	1	1	1	1	$[t_{11}\ ,\ t_{12}]$

models for Case I and Case II can be obtained from any 8 and 12 distinct load configurations, that hold (4.9), respectively. The examples are given in Tables 4.1 and 4.2, respectively. The resulting feeder-head measurements are shown in Figs. 4.3 and 4.4 for Case I and Case II, respectively.

For Case I, the Circuit Model includes 40 equations and 52 unknowns, see (4.5) and (4.6). The Load Model adds 16 additional equations and 4 new unknowns, see (4.10) and (4.11). Thus, the combined system of nonlinear equations has 56 equations and 56

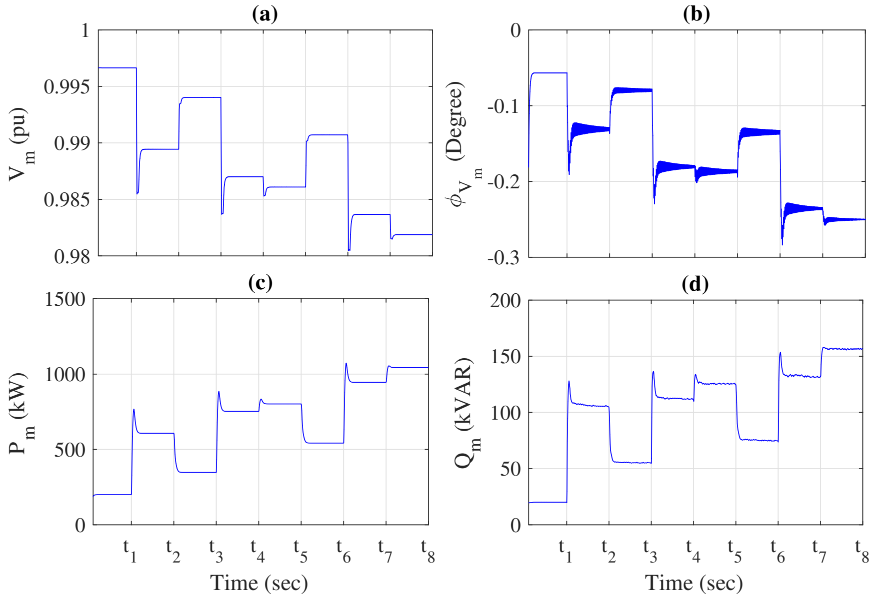

Figure 4.3: Feeder-head measurement parameters for Case I: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

unknowns. For Case II, the total number of equations is 126 and the total number of unknowns is also 126. Matrix B for Case II in Fig. 4.2 is obtained as:

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.26}$$

Figs. 4.5(a) and (b) shows the true and the estimated individual load models for Case I and Case II, respectively. We can see that the proposed method works well on both

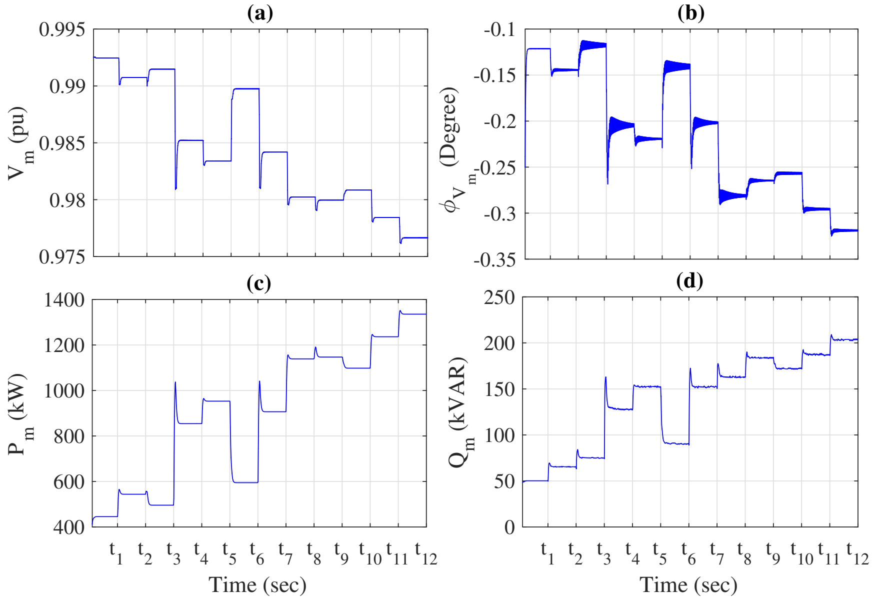

Figure 4.4: Feeder-head measurement parameters for Case II: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

distribution feeders with or without laterals. The error in individual load modeling is less than 0.01% for both Case I and Case II.

Figs. 4.6(a) and (b) show the true and estimated voltage phasors for Case I and Case II, respectively. The proposed method can also estimate auxiliary variables by 0.01% error. Similar results are obtained in estimating active and reactive power consumption, that are omitted due to brevity.

4.7.2 Performance Comparison

Before we get into the details in this section, we must emphasize that, to the best of our knowledge, this use-case is the first study that provides *individual* load model parameters using feeder-head measurements. Therefore, we *cannot* compare the individual

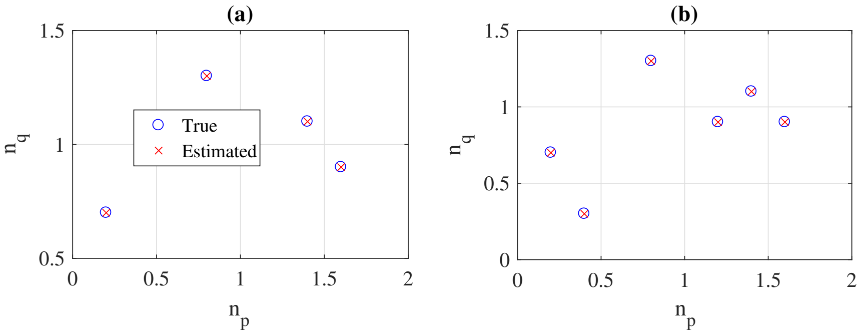

Figure 4.5: True and estimated load model for: (a) Case I; (b) Case II.

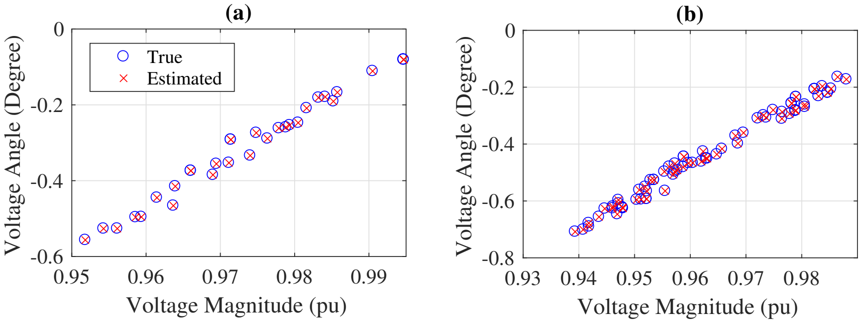

Figure 4.6: True and estimated voltage phasor for: (a) Case I; (b) Case II.

load modeling method in this study with another individual load modeling method. But what we *can* do, is to aggregate the individual load models that we obtain here, and then compare the result with the result of applying the feeder-aggregated load modeling methods in the literature, such as in [88–93]. The test setup to conduct such performance comparison is similar to Case II in Section 4.7.1. A total of 15 load configurations are considered, as given in Table 4.3, out of which three cases are caused by upstream voltage events. They are denoted by m_3, m_7, and m_{12}. The remaining 12 cases are distinct load configurations.

The results are shown in Figs. 4.7(a) and (b) with respect to the feeder-aggregated load parameters n_p and n_q, respectively. Here, comparison is done with the method in [90].

Table 4.3: Load configurations for Case II in Section 4.7.2.

Configuration	V(pu)	SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_1	1.00	1	0	0	1	0	0	$[0\ ,\ t_1]$
m_2	1.00	1	0	0	1	0	1	$[t_1\ ,\ t_2]$
m_3	0.95	1	0	0	1	0	1	$[t_2\ ,\ t_3]$
m_4	0.95	1	1	0	0	1	0	$[t_3\ ,\ t_4]$
m_5	0.95	0	0	1	1	0	1	$[t_4\ ,\ t_5]$
m_6	0.95	0	1	1	1	0	1	$[t_5\ ,\ t_6]$
m_7	1.00	0	1	1	1	0	1	$[t_6\ ,\ t_7]$
m_8	1.00	1	1	0	0	1	1	$[t_7\ ,\ t_8]$
m_9	1.00	0	1	1	0	1	1	$[t_8\ ,\ t_9]$
m_{10}	1.00	1	0	1	1	1	0	$[t_9\ ,\ t_{10}]$
m_{11}	1.00	0	1	1	1	1	1	$[t_{10}\ ,\ t_{11}]$
m_{12}	1.05	0	1	1	1	1	1	$[t_{11}\ ,\ t_{12}]$
m_{13}	1.05	1	1	1	0	1	1	$[t_{12}\ ,\ t_{13}]$
m_{14}	1.05	1	1	1	1	1	0	$[t_{13}\ ,\ t_{14}]$
m_{15}	1.05	1	1	1	1	1	1	$[t_{14}\ ,\ t_{15}]$

We can make three key observations. First, one can use the feeder-aggregated load modeling method in [90] no sooner than time t_2, because time t_2 is the first time in this case study that an upstream voltage event occurs. In contrast, the method that is proposed in this study can be used as early as time t_1, i.e., even before an upstream voltage event occurs. Second, as time goes by, the feeder-aggregated model using the method in [90] is not updated, except at time t_6 and time t_{11}, i.e., when the second and the third upstream voltage events occur, respectively. In contrast, the method that is proposed in this study is updated much more frequently. Finally, and most importantly, we can see that every time that the feeder-aggregated model is updated using the methods in [90], it matches the aggregated model that is obtained based on the method in this study.

All in all, we can conclude that the method in this chapter is equally good compared to the methods in the literature as far as feeder-aggregated load modeling is concerned whenever an upstream voltage event occurs; with an added advantage that it can update

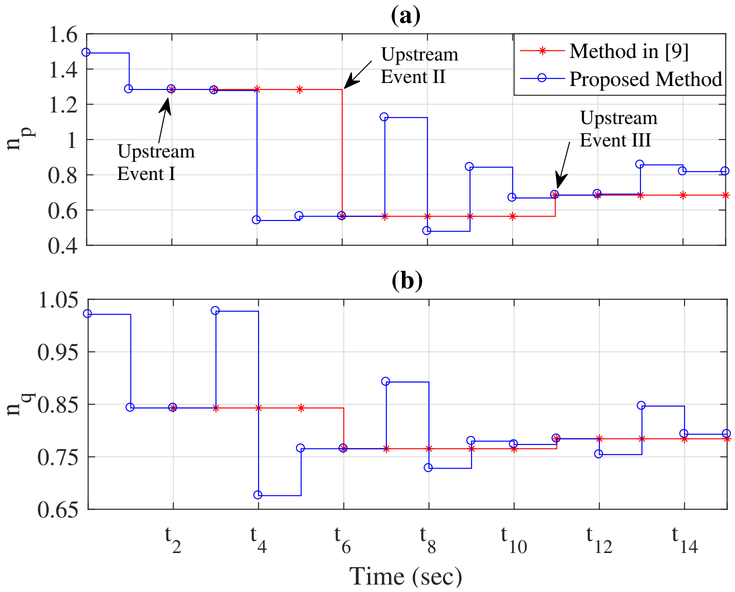

Figure 4.7: Performance comparison with literature in terms of feeder-aggregated load modeling: (a) load model voltage exponents for active power; (b) load model voltage exponents for reactive power.

such feeder-aggregated model more frequently. Of course, our method also provides individual load models, which is the primary goal of this study.

4.7.3 Tackling Errors in Measurements

In practice, sensors, such as micro-PMUs, may *not* be precise, mostly due to errors in CTs and PTs. Several methods have been introduced in the literature to improve data quality in synchrophasors, c.f., [118, 119]. In this section, we assume that the measurments for load modeling applications still include some levels of error. Monte Carlo method is used to generate different scenarios for different levels of error in feeder-head measurements. The test setup is similar to Case II in Section 4.7.1. However, this time, we also consider 12 redundant load configurations of all three types, as in Table 4.4. These scenarios are *in*

Table 4.4: Redundant load configurations for case study in Section 4.7.3.

Configuration		SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_{13}	Additional	1	0	1	0	0	0	$[t_{12}\ ,\ t_{13}]$
m_{14}	Additional	0	0	1	0	0	1	$[t_{13}\ ,\ t_{14}]$
m_{15}	Additional	0	1	1	0	1	0	$[t_{14}\ ,\ t_{15}]$
m_{16}	Additional	1	1	0	1	0	0	$[t_{15}\ ,\ t_{16}]$
m_{17}	Additional	1	1	0	1	1	0	$[t_{16}\ ,\ t_{17}]$
m_{18}	Additional	0	0	1	1	1	1	$[t_{17}\ ,\ t_{18}]$
m_{19}	Additional	1	0	1	1	1	1	$[t_{18}\ ,\ t_{19}]$
m_{20}	Additional	1	1	0	1	1	1	$[t_{19}\ ,\ t_{20}]$
m_{21}	Duplicate	1	0	1	1	1	0	$[t_{20}\ ,\ t_{21}]$
m_{22}	Duplicate	0	1	1	1	1	1	$[t_{21}\ ,\ t_{22}]$
m_{23}	Upstream	1	1	1	0	1	1	$[t_{22}\ ,\ t_{23}]$
m_{24}	Upstream	1	1	1	1	1	1	$[t_{23}\ ,\ t_{24}]$

Table 4.5: Load modeling estimation error in percentage considering error in complex power measurement in feeder-head.

Error	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
n_p	0.2317	0.4532	0.6446	0.8260	1.0638	1.2589
n_q	0.6991	1.3915	2.1482	2.5407	3.4625	4.0320

addition to the 12 load configurations in Table 4.2. The upstream events are 1.05 per unit step-up in feeder-head caused by an upstream voltage regulator. The total number of equations is 258 and the total number of unknowns is 246. Thus, the degrees of freedom in redundant measurements is $v = 258 - 246 = 12$.

Fig. 4.8 shows the feeder-head measurements for all the 24 load configurations in Tables 4.2 and 4.4. Problem (4.20) is solved to obtain individual load parameters. Tables 4.5 and 4.6 show the load modeling error in percentage considering different error levels in the feeder-head complex power measurement and voltage measurement, respectively. We see that, even in the presence of errors in measurements, the results demonstrate an overall satisfactory performance in load model estimation.

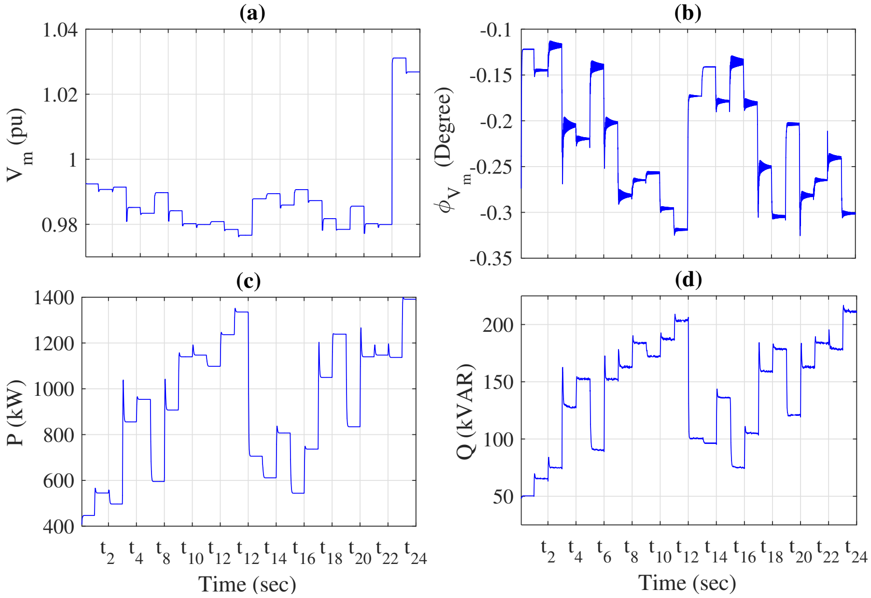

Figure 4.8: Feeder-head measurements for the test case in Section 4.7.3: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

Table 4.6: Load modeling estimation error in percentage considering error in voltage measurement in feeder-head.

Error	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
n_p	0.4157	0.8201	1.2472	1.6862	2.1178	2.5517
n_q	1.3830	2.7107	4.0166	5.4983	6.9226	8.3686

4.7.4 Impact of Error in Line Impedance

In this section, we investigate how the error in our knowledge about line impedances may affect the performance of the load modeling method. We use the Monte Carlo method to generate different scenarios for different levels of error in line impedances. The test setup is similar to Case II in Section 4.7.1. Similar to Section 4.7.3, we consider redundant load configurations, as in Tables 4.2 and 4.4. Again, the total number of equations is 258 and the total number of unknowns is 246. Accordingly, the degrees of freedom in the redundant measurements is $v = 258 - 246 = 12$. The results are shown in Table 4.7. We

Table 4.7: Impact of error in our knowledge about line impedances.

Error in Line Impedance	5%	10%	15%	20%	25%	30%
Error in Estimating n_p	0.09	0.93	1.30	1.98	2.54	3.38
Error in Estimating n_q	0.78	1.95	3.23	5.34	9.16	11.87

see that, even in the presence of some considerable errors in line impedances, the proposed load modeling algorithm can achieve an overall satisfactory performance in estimating the load modeling parameters.

4.7.5 Identifying Erroneous Switch Status

In this section, we examine the effectiveness of the proposed bad data detection and identification method to remove load configuration measurements with erroneous switch status. We use the Monte Carlo method to generate different scenarios with erroneous switch status for Case II, where the redundant load configurations are as in Tables 4.2 and 4.4. The Chi-squares test in (4.24) is performed by assuming v and α to be 12 and 0.005, respectively. The bad data detection method is able to detect erroneous switch status with 100% accuracy.

Next, we apply the LNR test to the residues to identify the load configuration measurements with erroneous switch status. We did so only for the residues of the Load Model equations; because the residues of the Circuit Model equations depend only on one load configuration. The load configuration measurements with erroneous switch status are then dropped. Afterwards, the load modeling is done on the remaining load configurations, and the Chi-squares test is applied again. If the Chi-squares test detects more bad data, then LNR test is applied again. This procedure continues until there is no outlier residue.

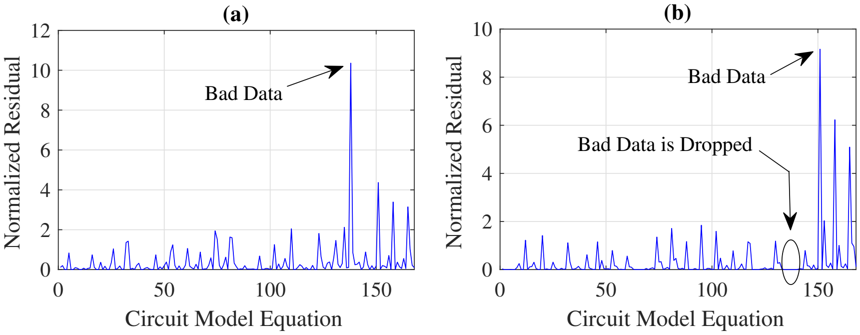

Figure 4.9: Bad data identification based on the LNR test in two load configurations: (a) the first incorrect switch status data; (b) the second incorrect switch status data.

Fig. 4.9(a) shows the normalized residues in the presence of erroneous load configurations 20 and 22. Equation number 138 has the largest residue in this case. This equation corresponds to the Circuit Model in load configuration 20. Once the measurements for this load configuration are dropped, the normalized residues appear as in Fig. 4.9(b). The largest residue corresponds to equation 151, which corresponds to load configuration 22. Once the measurements for this load configuration are also dropped, then the remaining load configuration measurements will result in obtaining accurate load models.

4.7.6 Tracking Time-Varying Load Models

In this section, we examine the effectiveness of the proposed method in Section 4.6.5 to track time-varying load modeling parameters. The test setup is similar to Case II. Load modeling is done based on the measurements from the most recent $c = 18$ load configurations. The degree of freedom in redundant measurements is $v = 6$. In total, suppose $k = 1, \ldots, 42$ load configurations occur. At load configuration number 19, the parameter of the load at bus 2 changes from $n_{p_2} = 0.8$ to $n_{p_2} = 1.8$. Fig. 4.10(a) shows the

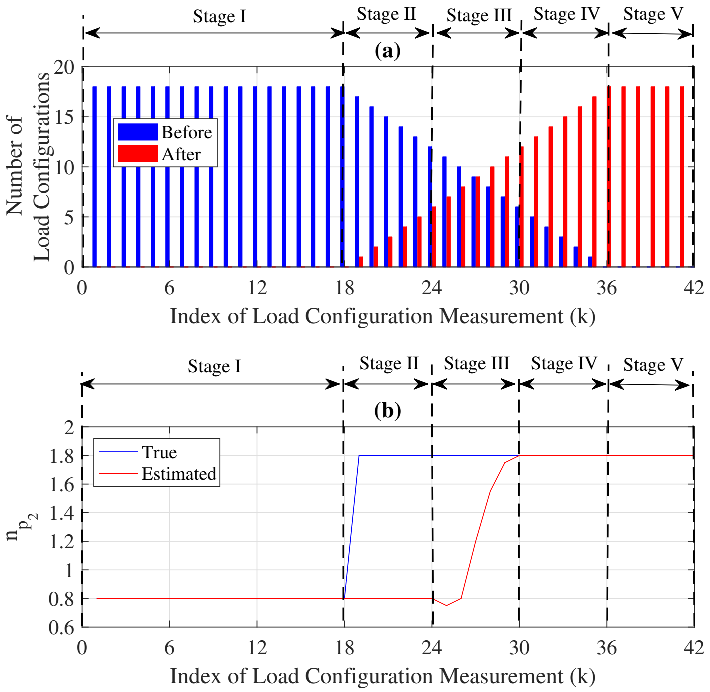

Figure 4.10: An example to track the change in a time-varying load at bus 2 using the proposed method: (a) The number of load configurations corresponding to the cases *before* and *after* the change in the load; (b) the true and estimated load model parameters. Note that, Load modeling at each measurement point k is done over a window of the last 18 load configurations.

number of load configurations at each measurement point. Color black denotes the number of load configurations that occurred *before* the change in the load, while color red denotes the number of load configurations that occured *after* the change in the load. The purpose of such color-coded distinction is to demonstrate the transition in the load modeling process from the measurement points that correspond to the old load model to the measurement points that correspond to the new load model.

Fig. 4.10(b) shows the true and estimated n_{p_2}. In Stage I, the estimated load model is accurate up until the moment when the load changes. In Stage II, the majority of the available measured load configurations correspond to the case *before* the change in the load. Therefore, either they do not trigger an update in the estimated load parameter, or the update in the estimated load parameter is not accepted due to a spike in the residues. In Stage III, the situation is reversed, and the majority of the available measured load configurations correspond to the case *after* the change in the load. Therefore, the estimated load parameter is updated until it gradually approaches the new true load parameter. In Stage IV, the available measured load configurations are sufficient to achieve accurate load modeling. Finally, in Stage V, the load modeling procedure is back to its initial stage where all the available measured load configurations correspond to the same load model.

4.7.7 Load Modeling in Presence of DERs

In this section, we examine the effectiveness of the proposed load modeling method in presence of DERs. The test setup is similar to Case II in Section 4.7.1, where it is assumed that a PV unit is connected to bus number 6 through a line with $1.5 + j0.5$ ohms. The nominal active power and reactive power of the PV unit are 300 kW and -15 kVAR, respectively. Thus, the renewable energy penetration in this test setup is 32%. In this study, we assume that the output power of the PV unit is either directly measured or estimated, while the voltage phasor of the PV unit is estimated as an unknown in the load modeling estimation. Similar to Section 4.7.3, we consider redundant load configurations, as in Tables 4.2 and 4.4. It is assumed that the output power of the PV unit changes during these load configurations, with possible turn off events.

First, the load modeling is conducted under the assumption that the output power of the PV unit is measured accurately. As expected, the error in individual load modeling

Table 4.8: Impact of error in measuring or estimating PV output.

Error in PV Unit	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
Error in Estimating n_p	0.28	0.48	0.69	0.85	1.11	1.29
Error in Estimating n_q	0.75	1.42	2.29	2.73	3.80	4.45

is less than 0.01%. Next, we use the Monte Carlo method to generate different scenarios for different levels of error in measuring or forecasting the output power of the PV unit. The results are shown in Table 4.8. We see that, even in the presence of imperfect knowledge on output power of the PV unit, the results demonstrate an overall satisfactory performance in load modeling.

4.7.8 Sequential Load Modeling on IEEE 33-bus Test System

The detailed data for the IEEE 33-bus test feeder is available in [120]. We modified this feeder by adding one virtual infinite bus with line impedance 0.5+j0.25 Ohms to the first bus of the feeder. We consider the first bus as measuring bus, which has no load. Thus, we aim to find load models for $n = 32$ buses. There are $2^{32} - 1$ possible load configurations for this feeder. From Theorem 2, only 64 distinct load configurations that satisfy (4.9) are needed to obtain the individual load models.

Figs. 4.11(a)-(c) show the measured voltage, active power, and reactive power at the substation. We apply the sequential load modeling method from Section 4.6.2 to $\hat{c}$ load configurations that hold (4.9) for $\hat{n}$ buses, where $\hat{c}_{\min} = 2\hat{n}$, to obtain the load models for $\hat{n}$ buses. Fig. 4.11(d) shows number of load models that are complete. Load modeling starts at time t_1 as soon as sufficient load configuration measurements become available.

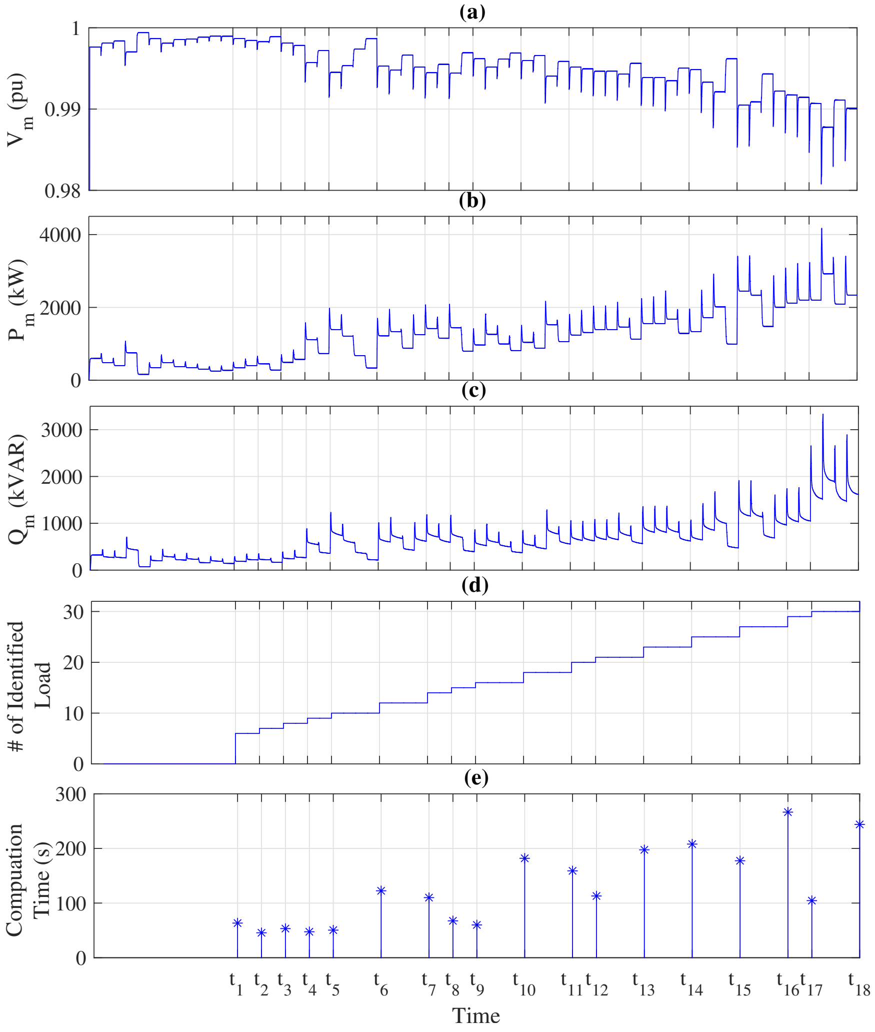

Figure 4.11: Sequential load modeling for the IEEE 33-bus test feeder: (a) feeder-head voltage magnitude; (b) feeder-head active power; (c) feeder-head reactive power; (d) number of modeled load; (e) computation time of each sequence.

Fig. 4.11(e) shows the computation time to run the load modeling algorithm. The computation time at each step depends on the number of equations and the number of unknowns in that step, and these numbers themselves depend on the new load configuration that occurs and measured at each step. For example, in the first execution of the algorithm at time t_1, the number of equations and unknowns are both 116; and the computation time is 62 seconds. As another example, in the second execution of the algorithm at time t_2, the number of equations and unknowns are both 23, and the computation time is 45 seconds. The total computation time across all steps to finish modeling all individual loads in this IEEE 33-bus test case is 38 minutes. This is the summation of the computation times across all 18 steps. That means, on average, each execution takes about two minutes.

4.7.9 Using Multiple Sensors on IEEE 123-bus Test System

In this section, we examine the effectiveness of the proposed load modeling method on the IEEE 123-bus test system, shown in Fig. 4.12. The detailed data for the IEEE 123-bus test feeder is available in [121]. We assume that five micro-PMUs are deployed on the test feeder. This results in breaking down the load modeling problem into five zones, as shown in Fig. 4.12 by using different colors. Zone 1 includes the buses in downstream of sensor M_1 and upstream of sensors M_2, M_4, and M_5. Zone 2 includes the buses between sensors M_2 and M_3. Finally, Zones 3, 4, and 5 include the buses in downstream of sensors M_3, M_4, and M_5, respectively.

Recall from Section 4.6.6 that each zone is essentially a separate load modeling problem. For instance, consider Zone 1 whose lines and buses are shown in black. This zone has 29 buses, out of which 19 buses are single phase. Thus, there exist 51 voltage phasors as unknowns in each load configuration. This zone also includes 21 single-phase loads. Note that, the load for the rest of the network is directly measured using sensors M_2, M_4, and M_5. From Theorem 2, the minimum number of distinct load configurations in order to estimate individual load model parameters for the loads in this zone is 42, i.e., twice the number of the loads in this zone. Here, we consider that 8 out of the 21 loads in

Zone 1 are switched on during every load configuration. Thus, the Circuit Model includes $42 \times (53+1) = 2,268$ equations and $42 \times 53 + 8 \times 42 = 2,562$ unknowns, see (4.5) and (4.6). The Load Model adds 315 additional equations and 21 new unknowns, see (4.10) and (4.11).

The combined system of nonlinear equations for load modeling in Zone 1 has 2,583 equations and unknowns. This number is 1,962 for Zone 2, 2070 for Zone 3, 456 for Zone 4, and 3,870 for Zone 5. To compare these numbers with the case when there is only one sensor installed, i.e., when sensors M_2, M_3, M_4, and M_5 are not available, the combined system of nonlinear equations has 47,405 equations and 47,405 unknowns. Therefore, as expected, installing additional sensors can break down the load modeling problem into smaller problems, which accordingly reduces the computational complexity of the load modeling problem. Importantly, the resulting load modeling accuracy in each zone was 99.9% or more, demonstrating a very accurate load modeling performance. The computation-time of running Algorithm 1 on Zones 1, 2, 3, 4, and 5 was 31, 19, 21, 4, and 42 minutes, respectively.

4.8 Conclusions

A new method was proposed for individual load modeling in power distribution systems. It works by analyzing the measurements at the feeder-head during load switching events as they occur across the feeder. The basic idea is that, once a load is switched, the switching event changes the voltage for the rest of the loads, which causes variation in their active and reactive power consumption; thus allowing us to estimate their load

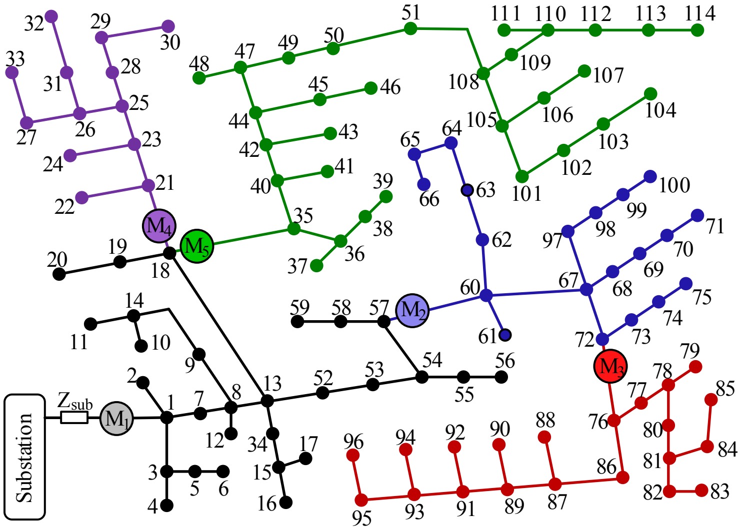

Figure 4.12: Multiple sensor installations on the IEEE 123-bus test system.

parameters. The proposed method can estimate load modeling parameters of individual loads. Therefore, it can support various practical use cases at distribution level, such as distribution-system optimal power flow analysis, DER management, Volt-VAR control, and voltage stability. Of course, if needed, the obtained individual load models can be aggregated to also provide a feeder-aggregated load model. A theoretical foundation was provided to determine the conditions for successful load modeling. The proposed load modeling can be done sequentially to model a subset of loads while more measurements become gradually available. The proposed method is extended to support also estimating the parameters for time-varying individual loads. The proposed method is extended also to utilize different types of redundancy in measurements to improve load modeling accuracy. Further, a residue-based bad data detection and identification method was developed to identify and drop load configuration measurements with erroneous switching status; thus to ensure the accuracy of the load models. Several test cases, including the IEEE 33-bus and

IEEE 123-bus test feeders are studied to assess the performance of the proposed method.

Chapter 5

Remote Asset Monitoring

5.1 Introduction

In the context of distribution system and the load it serves, utilities seek to keep the overall power factor close to 1, or within its 5%, for optimal network efficiency and manageable voltage profiles with respect to the operational constraints. Reactive power regulation methods are employed by utilities to compensate demanded reactive power, and bring the power factor close to optimal values, by using equipment which provides local reactive power (VAR) generation at distribution level. Capacitor bank installation is an efficient approach to achieve Volt-VAR control. Hence, there is an ongoing growth in capacitor bank deployment at distribution level to improve voltage profile, decrease losses and improve power factors.

Fixed capacitor banks are typically used to improve the power factor of individual inductive loads, such as motors, or a group of loads that have a relatively constant demand for reactive power (VAR). In contrast, switched capacitor banks are connected to the grid based on some predetermined operational objectives. Accordingly, switched capacitor banks are often equipped by a controller which sends the closing and tripping commands to a circuit breaker when the predetermined criterion is meet. Even though the automatic capacitor banks typically operate their switched on and off events properly through the Volt-VAR controller, the system operator often does not have any insight regarding the exact status of the capacitor banks due to lack of useful state measurement.

The identification of capacitor banks switching is useful. From operation point of view, some power-factor based identification applications, including network reconfiguration, service restoration, Volt-VAR control, fault detection and location, peak and loss reduction, in distribution level highly depend on the capacitor banks status. Also, switching of capacitor banks leads to high magnitude and high frequency transient disturbances. These disturbances in power systems may damage key equipment and potentially have impact on system reliability, by causing anomalies and unplanned DER trips. Thus, operators desire useful monitoring of capacitor banks for optimal network operation.

5.2 Related Works

Several efforts in the literature investigate the capacitor bank switching events. The effect of switching transmission capacitor banks on phase-to-phase voltages and consumer-end distribution capacitors is studied in [122, 123]. Also in [124], the authors investigate interaction of capacitor bank in the utility system with the failure of a drive system. Real evaluation of impact of capacitor bank switching events, requires a voltage and current view prior to the transient and subsequent switching action. In [125], the transient treatment of capacitor bank voltage and current is during switching action only. Also, the authors in [126, 127] consider key parameters that can determine the transient of inrush current caused by switching of capacitor banks.

Apart from considering the effect of capacitor bank switching on the system, there is lack of investigation regarding useful capacitor bank monitoring in tandem with other analytics. In [128, 129], the authors discuss control schemes using coordinated measurements and communications between field devices. In this regard coordination of the control actions of the voltage regulator and capacitor results in more effective Volt/VAR control. More specifically, [130] reported the usage of micro-PMU to solely detect the capacitor events via considering the instantaneous changes in the reactive power demand. In this regard, reactive power average is calculated every second, so the capacitor bank changes can detected in a short time. In this regard, the analysis in [130] did not move on to use the key synchronization and high sampling rate properties of micro-PMUs to possibly go beyond only detecting a capacitor event.

5.3 Motivations and Contributions

This chapter aims to propose a novel data-driven approach to use experimental distribution-level voltage and current synchrophasor data, on three phases, to identify the operational parameters of the switched capacitor bank with no need to install separate asset sensors. We analyze the voltage and current synchrophasor data that is collected from a real distribution system in Riverside, CA during a 24 hours period. The steady-state analysis is conducted to answer the following question: would the data collected by micro-PMUs allow identifying the operational parameters of the switched capacitor bank, such as the per-phase reactive power support and its comparison with the switched capacitor banks nameplate parameters? Also, the transient and dynamic analysis aims to answer the following question: what are the dynamic effects of capacitor bank switching events on feeder-level and load-level voltage and current synchrophasor data, on three phases? This is a novel contribution because it studies fast-scale analysis of capacitor bank switching events, for both impact, performance verification and parameterization at distribution level using micro-PMUs.

5.4 Under-Study Real-Life Test System

In order to identify the operational parameters of the capacitor bank, a real-life distribution feeder is considered in Riverside, CA. The single line diagram of the under-study feeder is shown in Fig. 5.1. The point of common coupling between the sub-transmission system and the under-study feeder is marked as *Substation*. The under-study feeder includes one three-phase fixed capacitor bank rated at 600 kVAR, which is always connected, as well as a three-phase switched capacitor bank rated at 900 kVAR. The switched capacitor bank is switched by a vacuum circuit breaker and controlled by a *Volt-VAR controller*, i.e., the capacitor bank is *switched-on*, when any phase of the bus voltage is below the low-voltage override threshold. Conversely, the controller *trips* any energized capacitor banks when any phase of the bus voltage is above the high-voltage override threshold. The rated Volt-VAR controller time delay is set to be 300 seconds. As it can be seen from Fig. 5.1, transient limiting inductors (TLIs) are installed in series with each phase of the switched capacitor bank. The TLIs limit transient currents during switching events or faults.

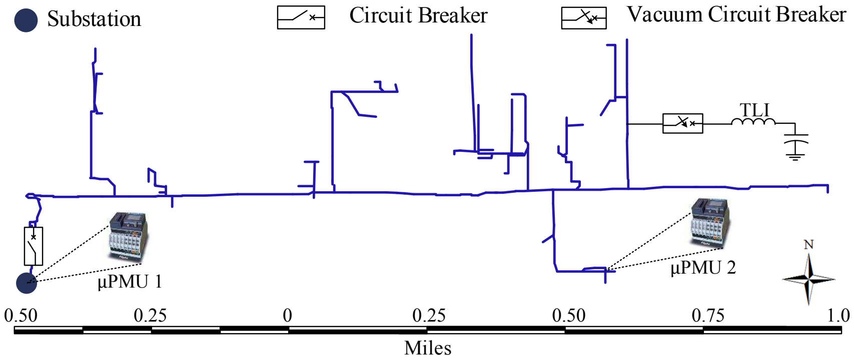

Figure 5.1: Test system is a real-life feeder in Riverside, CA, which includes a switched capacitor bank and is equipped with two micro-PMUs.

Neither the fixed nor the switched capacitor bank are monitored, i.e., they are not equipped with any sensors. Accordingly, in principle, RPU is not aware of the following: 1) how these capacitor banks operate on a daily basis; 2) how the operational parameters might be different from their rated values, possibly due to wear and tear over time. However, the under-study feeder where these two capacitor banks are located is equipped by two distribution-level PMUs, a.k.a, micro-PMUs. One micro-PMU is located at the substation and another micro-PMU is located at a secondary side of a three-phase load transformer that is located about 400 yards from the switched capacitor bank. The micro-PMUs were not installed with the intention of monitoring the operation of the capacitor banks, but more in general for monitoring the feeder performance overall, and as such are not optimally placed. The key question we seek to answer with this paper is: Would the data collected by micro-PMUs allow identifying the operational parameters of the switched capacitor bank, and thus eliminating the need to install separate sensors for the switched capacitor bank?

The installed micro-PMUs report four fundamental measurements on three phases, i.e., in total there are 12 measurement channels: voltage magnitude, voltage phase angle, current magnitude, and current phase angle, with the sampling rate of 120 Hz, i.e., one sample every 8.333 msec. We hypothesize that this high sampling rate, together with direct access to voltage and current data at different yet time synchronized locations, could allow us identify the operation of the switched capacitor bank within a data-driven framework. Our study involves both stead-state, i.e., slow time-scale, and dynamic, i.e., fast time-scale, analysis.

5.5 Steady-State Analysis

The focus in this use-case is on the micro-PMUs data on one day. Note that, the two micro-PMUs generate 248,832,000 data points every day. The recorded single-phase voltage magnitudes and current magnitudes are shown in Fig. 5.2. Generally speaking, the fluctuations in this figure are statistically similar during period of a day, with no significant observation on the potential effect of capacitor switching events. A similar conclusion can be made by plotting voltage phase angles and current phase angles as well as by plotting voltage phase differences. Note that, the latter two cases are not shown here due to space limit.

5.5.1 Analysis of Data from micro-PMU 1

Fig. 5.3 shows the recorded three-phase power factor and reactive power at *Substation*, measured by micro-PMU 1. The immediate observation from this figure is that there exist drastic instantaneous changes in power factor and reactive power. Based on Fig. 5.3(a), at 08:36:45 the power factor increases by 10% in all phases and lasts for almost

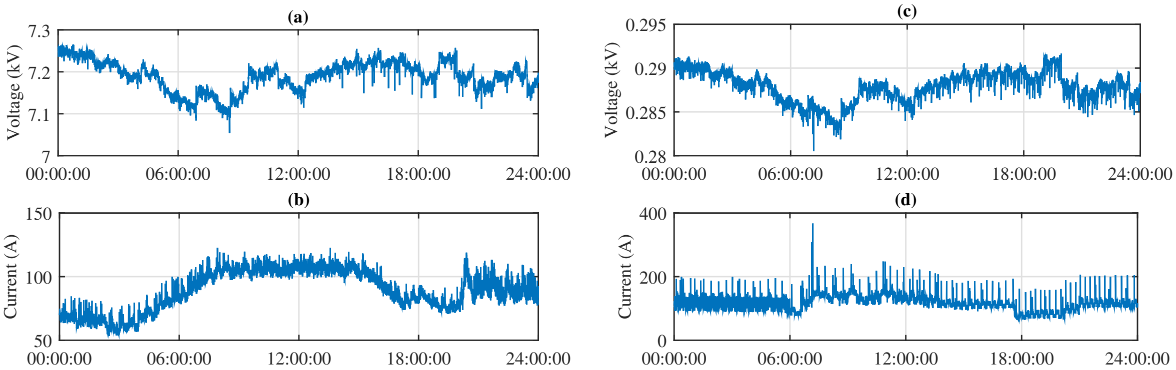

Figure 5.2: Single phase voltage magnitude and current magnitude on one day measured by: (a) and (b) micro-PMU 1, (c) and (d) micro-PMU 2.

11 hours, till 19:55:54. Possible scenarios for power factor fluctuation at 08:36:45 are: 1) switching off a large inductive load, or 2) switching on a large capacitor bank. Likewise, the fluctuation at 19:55:54 could be the effect of: 1) switching on a large inductive load, or 2) switching off a large capacitor bank. To the extent of our knowledge, there is such large inductive load in the under-study network. Therefore, both events should be due to the operation of the switchable capacitor bank.

We can confirm the above conclusion by examining the results on reactive power in Fig. 5.3(b). From this figure, the *per-phase* reactive power change at both events is about 300 kVAR, which is equal to the installed 900 / 3 = 300 kVAR per-phase switched capacitor bank. From now on, we refer to 08:36:45 as *switched-on time* and 19:55:54 as *switched-off* time.

According to the steady-state analysis on per-phase reactive power, the exact injected reactive power by capacitors at phase A, B, and C are 291, 294, and 288, respectively. Therefore, the operational parameters of the capacitor banks are slightly different from their

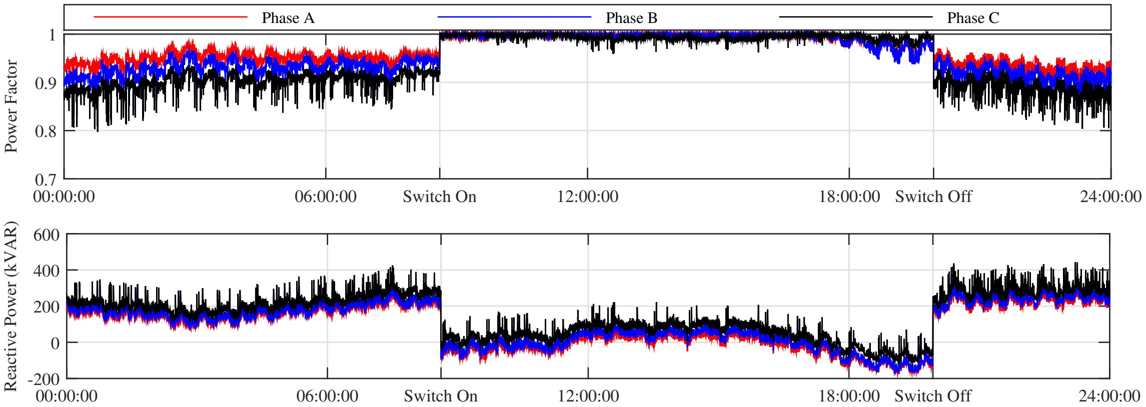

Figure 5.3: Power factor and reactive power at *substation*: (a) Power factor, (b) Reactive power.

rated values, possibly due to error in voltage and current measurements, 0.3% and 1.2% respectively, or due to wear and tear over time. Importantly, the impact of degradation appears to be somewhat imbalanced across the three phases.

Another observation in Fig. 5.3(b) is that quite often there exist *reverse* reactive power flows from the under-study feeder toward the sub-transmission network. This may not be desirable and may suggest the need to possibly update the operational parameters of the switched capacitor bank, possibly in coordination with adjusting the tap changers inside the substation.

5.5.2 Analysis of Data from micro-PMU 2

As for the measurements at micro-PMU 2, the operation of the capacitor bank did not demonstrate any visible impact on the power factor and reactive power measurements as far as steady-state analysis is concerned. This is not surprising because micro-PMU 2 is dedicated to make voltage and current phasor measurements at a particular load point, even though such load is in a close-by location to the capacitor bank. Note that, the related figures are not shown here due to the space limitation.

5.5.3 Possible Applications of Results

The above analysis may find applications in operating the distribution system, e.g., with respect to the following aspects:

- Network reconfiguration;
- Service restoration;
- Volt/VAR control;
- Fault detection and location;
- Peak and loss reduction;
- Load balancing;
- Energy management.

For instance, consider the first item, i.e., suppose *service restoration through reconfiguration* is called upon due to fault occurrence at a feeder that is coupled with the under-study feeder. In this scenario, the under-study network should energize transferred loads from the faulted feeder. The decision maker requires the status of the unmeasured capacitor bank to check the feasibility of the post-reconfiguration operating point, i.e., feeder constraints such as voltage threshold. In addition, if force-switching of capacitor bank is not possible, the power factor estimation based on previous data provides information on feasibility or infeasibility of reconfiguration in next hours.

5.6 Transient and Dynamic Analysis

In this section, the effect of capacitor bank switching events on the under-study feeder are studied in fast time-scale.

5.6.1 Analysis based on an RLC Model

To facilitate the analysis in this section, we consider an *RLC* circuit. Suppose the capacitor bank has initial voltage $v_c(0^-)$ and then it is connected to the circuit at time $t = 0$, which is the switch on time in our context. The differential equation of the *RLC* circuit is obtained as:

$$\frac{d^2i(t)}{dt^2} + \frac{R}{L}\frac{di(t)}{dt} + \frac{1}{LC}i(t) = \frac{\omega}{L}cos(\omega t + \phi), \tag{5.1}$$

The solution of above differential equation with respect to the initial condition is obtained as:

$$i(t) = \gamma_1 e^{-t/\tau_1} + \gamma_2 e^{-t/\tau_2} + \gamma_3 cos(\omega t + \varphi). \tag{5.2}$$

The first two terms indicate the transient current due to the capacitor switching event and the last term is the steady-state current, where the coefficients depend on the initial condition, source phase angle at switching event, and circuit parameters, i.e., $\gamma_i := \Gamma_i\left(v_c(0^-), \phi, \omega, R, L, C\right)$, where Γ_i is a function.

We can conclude that the transient current of the under-study feeder in response to a *switch on* event of the switched capacitor bank depends on both the *initial condition* and the *phase angle* at the moment of switching. Typically, in capacitor bank switch on event, the circuit breaker is closed when the source voltage is nearly equal to zero. However, the zero crossing of phases A, B, and C occur at 0°, 60°, and 120°, relative to the three-phase voltage with no phase shifting. Therefore, when the circuit breaker is *not* equipped with pre-insertion resistors, there will be current transients in at least two phases.

Similarly, we can also conclude that the transient current of the under-study feeder in response to a *switched off* event of the switched capacitor bank depends only on *phase angle* at the moment of switching. Unlike in switch on switching, in switch off event, the circuit breaker is closed in zero-current. Accordingly, due to 120° phase shift between phases, there will be current transients in at least two phases, when the circuit breaker is *not* equipped with pre-insertion resistors.

5.6.2 Analysis of Data from micro-PMU 1

Fig. 5.4 shows the transient response of the under-study feeder during the capacitor bank's *switch on* event. As mentioned in Section 5.5 and shown Fig. 5.4(a), the feeder power factor at *switch on* event increases by 10%. Also, Fig. 5.4(b) represents pre-switching and post-switching reactive power flow on the under-study feeder. Voltage transient of the under-study feeder is shown in Fig. 5.4(c). As expected, the voltage magnitude is increased due to reactive power compensation by the switched capacitor bank. However, there is no significant changes in voltage transient, possibly due to available TLIs or insufficient sampling rate. The three-phase transient current is shown in Fig. 5.4(d). While there is no significant transient change in phase A, the current of phases B and C include an *overshoot* and an *undershoot*, respectively. Therefore, we can conclude that the capacitor bank is

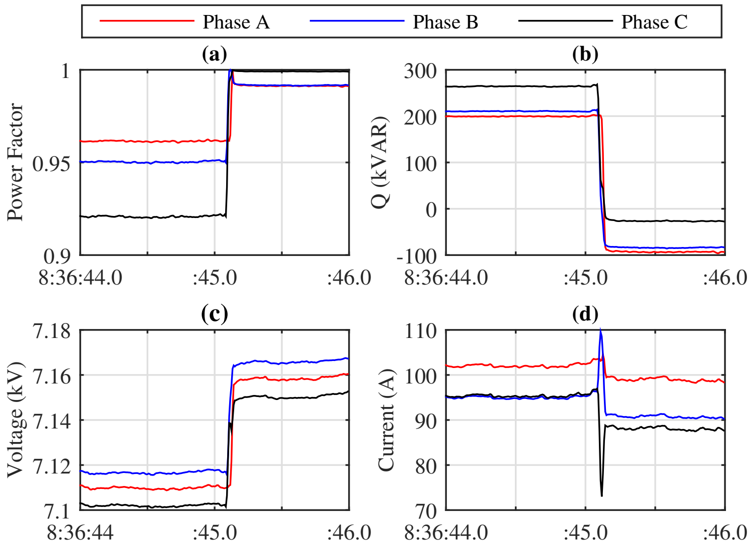

Figure 5.4: Under-study feeder dynamic response to capacitor *switch on* event: (a) Power factor, (b) Reactive power, (c) Voltage magnitude, (d) Current magnitude.

switched on at zero crossing of phase A.

Fig. 5.5 shows the transient response of the under-study feeder during the capacitor bank's *switch off* event. The changes in power factor, reactive power, and voltage magnitude are as expected, see Fig. 5.5(a), (b), and (c). However, in this event, there exists a transient current in phase A and B, there is no significant transient change in phase C. Accordingly, the capacitor bank is switched off at zero crossing of phase C.

5.6.3 Analysis of Data from micro-PMU 2

Fig. 5.5 shows the transient response of the load where micro-PMU 2 is located during the capacitor bank's switching events, both on and off. As it can be seen from Fig. 5.6(a) and (c), voltage magnitude of the load follows the changes in voltage magnitude in substation. Current magnitude depends on changes in voltage when there is no changes in

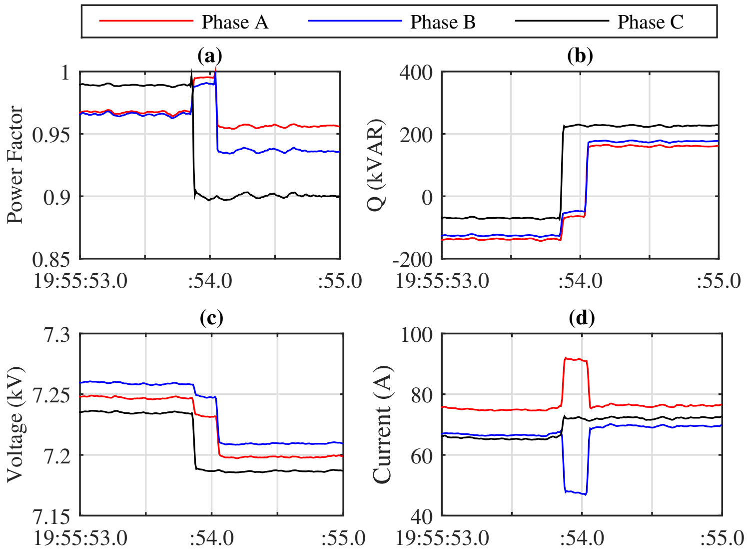

Figure 5.5: Under-study feeder dynamic response to capacitor *switch off* event: (a) Power factor, (b) Reactive power, (c) Voltage magnitude, (d) Current magnitude.

the overall active and reactive power at downstream customers, see Fig. 5.6 (b) and (d).

5.7 Analysis of Data from micro-PMUs for two Weeks

Typically, for a wye-floating capacitor bank, the strategy for switching-off is to open contacts in two steps; first, opening one phase at zero-crossing of its current; second, opening the two other phases at quarter of a cycle later, at 90° relative to zero-crossing of the first phase [34].

Fig. 5.5(d) shows that there is an *overshoot* and an *undershoot* in current magnitude of phase A and phase B, respectively, while phase C experiences an ideal switching with no transient in current magnitude. Also, in Fig. 5.5(b), reactive power of phase C increases due to capacitor bank switching-off on phase C, while after about 200 msec

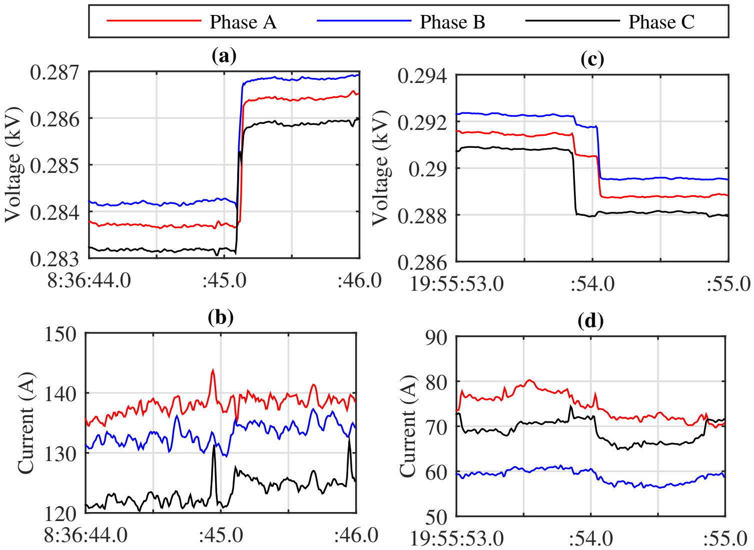

Figure 5.6: Load level dynamic response to capacitor switching event: (a) Voltage transient at *switch on* event, (b) Current transient at *switch on* event, (c) Voltage magnitude at *switch off* event, (d) Current magnitude at *switch off* event.

dead-time, reactive power increases on phases A and B. Thanks to the event detection approach, this issue is further studied across 25 switching-off events that occurred across two weeks. The results are shown in Fig. 5.7. We can confidently conclude that the capacitor bank switching is ideal during its first step, but there is always about 20% overshoot and undershoot transients in current magnitude lasting for 100-200 msec during its second step.

The unbalanced or underrating operation of the capacitor bank was also investigated during the two weeks of this study. Fig. 5.8 shows the change in reactive power compensation across three phases due to switch-on events. Reactive power compensation is 288 kVAR on Phase A, 291 kVAR on Phase B, and 286 kVAR on Phase C. The results are similar during the switching on events are omitted due to space limitation.

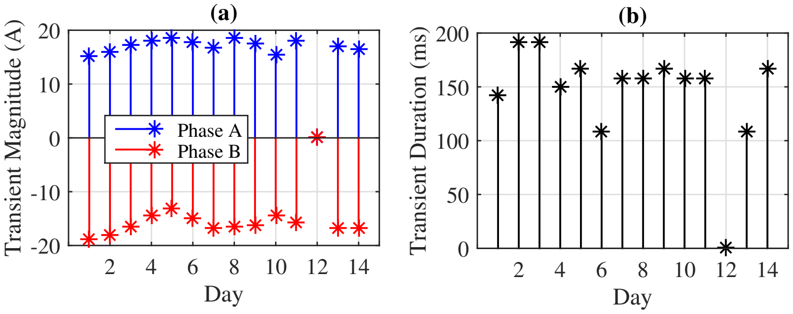

Figure 5.7: Capacitor bank switch-off events: (a) transient current magnitude, (b) duration of transition.

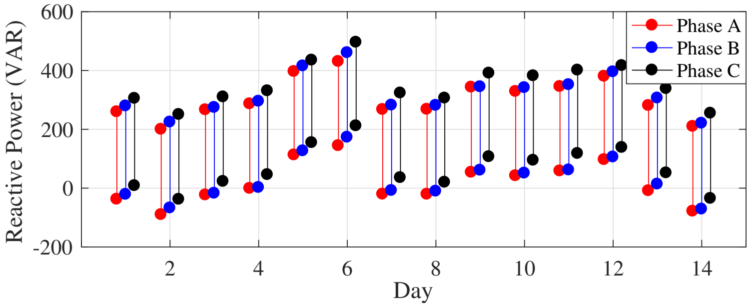

Figure 5.8: Change in reactive power during capacitor bank switching.

5.8 Conclusions

A data-driven experimental analysis is provided to investigate the operational parameters of capacitor banks switching events at a real-life distribution system based on data from two micro-PMUs. According to feeder level micro-PMU, *capacitor bank switching events*, *operational parameters of capacitor banks*, as well as *direction of reactive power flow* are observed. In the fast-time scale analysis, the effects of capacitor bank switching events on the under-study real-life feeder are studied. The results show that the transient current of under-study feeder in response to *switch on* event of the switched capacitor bank depends on both *initial condition* and *phase angle* at the moment of switching. Accordingly, in three-phase capacitor bank *switch on* event both *overshoot* and *undershoot* in phase currents are possible and observed in actual real-life data. While, in three-phase capacitor bank *switch off* event both *overshoot* and *undershoot* in phase currents are possible and observed in actual real-life data.

Chapter 6

Protection System Diagnosis

6.1 Introduction

Most service interruptions and power quality issues are nowadays initiated due to some sort of fault at distribution grid. Of course, it is impossible in practice to entirely prevent such faults, in particular, those that occur by natural causes or physical accidents. Therefore, utilities must be prepared to rather mitigate the fault impact. This can be done by first crystallizing weakness points; and then by modifying the control and protection parameters to reduce the impact.

In the context of distribution systems, the faults are generally categorized based on the interruption duration, including permanent faults and momentary faults. Statistically speaking, about 80% of all faults that occur at distribution level are temporary in nature [131]. Recloser deployment is an efficient protection scheme against temporary faults [78]. Commonly, reclosers are expected to operate in coordination with lateral fuses to minimize the number of affected customers.

For example, consider a fault occurrence in downstream of a lateral fuse. On one hand, the recloser is set to trip before the minimum melting time of the fuse. On the other hand, the maximum clearing time of the fuse must be smaller than the last delayed trip of recloser, i.e., the lock-out status. Therefore, if the fault is momentary, then all customers experience a momentary interruption with no blown fuse. If the fault is permanent, then the customers in downstream of the faulted lateral fuse experience permanent interruption while the rest of the customers experience only a momentary interruption.

6.2 Related Works

Several efforts have been conducted in the literature to detect fuse-recloser miscoordination as well as to improve their coordination [132–134]. Apart from the level of effectiveness of these various methods, there is still a lack of analysis based on experimental data from real-life fault events in order to check and confirm the optimal operation of protection devices during fault time-line. Such experimental analysis has been traditionally a very difficult task due to the difficulty in cross-examining the extremely fast events that occur at different locations of the distribution system during a fault event.

Examining the effects of fault and protection coordination becomes even a more far-reaching task in *active distribution feeders*, i.e., distribution systems that are connected to DERs and distributed generators (DGs). This is a true real-world challenge. For example, the capacity of the installed solar generation units in the US increased by 28% in one year from 2014 to 2015. In addition to harvesting power from photovoltaic (PV) farms, there is an increase in the penetration of customer-owned behind-the-meter PV panels in recent years [135].

The response of such high penetration level of PV generation units to momentary and permanent faults is yet to be understood by electric utilities. Moreover, the feeder protection settings must now be coordinated also with PV inverter protection system, which adds yet another degree of complexity to an already complex protection system coordination problem. For instance, transient overvoltage occurs when a behind-the-meter PV generation unit continues to feed into an isolated feeder that is affected by a permanent fault [136].

In [137], effective anti-islanding protection schemes are proposed for PV inverters in faulted feeders. However, the optimal operation of inverter protection schemes is yet to be investigated in practice to check for any potential miscoordination between feeder protection devices and inverter protection system during permanent and momentary interruptions.

6.3 Motivations and Contributions

This chapter proposes a novel data-driven experimental analysis on a single-phase-to-neutral fault to identify the operation of protection devices in fast-time scale. Data from five micro-PMUs on a real-life distribution and sub-transmission network is analyzed with the focus on an animal-caused fault on one feeder. We aim to identify the *fault timeline* using feeder-level as well as customer-level micro-PMUs. In this regard, any potential fuse-recloser miscoordination could be detected using experimental data. In addition, the under-study network includes several PV resources. Thus, the response of the PV resources to the fault is also investigated to explore any miscoordination between feeder protection scheme and inverter protection systems, i.e., its built-in anti-islanding schemes. Finally, we investigate the effect of the fault on outlying areas covering feeder-level and customer-level impacts.

6.4 Problem Statement

This study investigates the impact of faults on a power distribution system, in steady state as well as transient conditions, utilizing a real-life distribution network in Riverside, CA. The analysis is based on true measurements and verified field information from utility reports and crew interviews.

The single line diagram of the under-study grid is shown in Fig. 6.1. The network is operated by Riverside Public Utilities (RPU). The point of common coupling between the transmission system and the under-study sub-transmission and distribution network is marked as *Substation A*, see Fig. 6.1. In total, the under-study sub-transmission network includes 15 substations, 69 kV and 33 kV. Of interest in this paper is *Substation B* and to a lesser extend *Substation C*. Two distribution feeders in downstream of Substation B as well as one feeder in downstream of Substation C are included in this study, marked by *Feeders I, II, and III* in Fig. 6.1.

Feeder I and Feeder II are active distribution feeders, each hosting one 100 kW behind-the-meter PV generation unit, with balanced three-phase inverters. Also, Substation C is interconnected with a 7.5 MW investor-owned behind-the-meter PV farm comprising 25,000 solar panels. The solar generated power is fed into the local distribution grid.

The entire sub-transmission grid of the under-study network is monitored by RPU through a supervisory control and data acquisition (SCADA) system. The default minutely

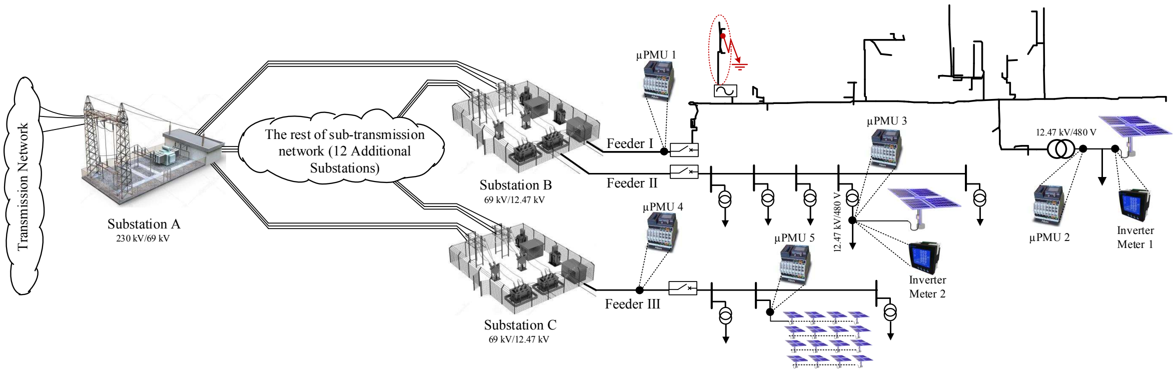

Figure 6.1: Test system is a real-life feeder in Riverside, CA, which is equipped with five micro-PMUs. The fault occurs on Feeder I.

reporting rate of the SCADA system is not sufficient to establish the detailed time-line for the fault event. Fault indicators could be utilized to indicate the presence of a fault, but not the details that are required for evaluating the protection system. The SCADA system could be reconfigured to one reading per-second during a fault, but this additionally would still not be enough to analyze the millisecond nature of the fault events.

Fortunately, RPU is one of few pioneer utilities that has already deployed distribution-level PMUs, a.k.a, micro-PMUs. The RPU network is equipped with five micro-PMUs, one at the secondary of a 69 kV to 12.47 kV transformer at Substation B feeding Feeder I, two at the commercial customer locations with the 100 kW PV units at Feeders I and II, one at the secondary of a 69 kV to 12.47 kV transformer at Substation C feeding Feeder III, and one at the point of common coupling of the 7.5 MW PV farm located on Feeder III, see Fig. 6.1.

Noted that, the last two micro-PMUs, i.e., micro-PMU 4 and micro-PMU 5 are in an outlying area from micro-PMU 1 and 2, which are close to the fault location of interest. Also, despite Feeders I and II being connected to Substation B, they are *electrically isolated* at this substation. Accordingly, to some extent, we can assume that micro-PMU 3 too is in an outlying area from micro-PMU 1 and 2, but it is affected through a different equivalent circuit than micro-PMUs 4 and 5. In addition to the SCADA system and the installed micro-PMUs, the two customer-owned behind-the-meter PV panels are monitored by meters with one sample every minute.

The installed micro-PMUs report four fundamental measurements on three phases, i.e., in total, there are 13 measurement channels: voltage magnitude, voltage phase angle, current magnitude, and current phase angle, with the sampling rate of 120 Hz, i.e., one sample every 8.333 msec. The 13th channel is GPS Lock, utilized to determine if the sensor has established a satellite lock to ensure precise time synchronization. As a novel application of micro-PMU data, we hypothesize that, this high sampling rate, together with direct access to time synchronized voltage and current phasor data at different locations, could allow us study fault event in details within a data-driven framework.

We are interested in answering the following questions with respect to fault events at power distribution systems:

1. What is the time-line of the fault?
2. Is there any miscoordination between lateral fuses and feeder recloser that is revealed during the fault?
3. What are the effects of fault on PVs and other DERs?
4. Is there any miscoordination between the PV and DER inverter protection system and the main feeder recloser?

5. What is the effect of the fault on outlying resources?

The focus in this paper is the fault occurrence in Feeder I, on a day in November 2016. The fault occurred around noon. The location of the fault is highlighted in Fig. 6.1, which is close to (almost 0.3 mile away from) Substation B. Utility crews reported that a blown fuse is founded in phase B of the lateral as well as a diseased bird on pole. With the knowledge on the cause for the fault obtained from the outage report, as well as the data collected by micro-PMUs and inverter meters, the fault event will be studied in details in this paper within a data-driven framework to answer the above questions.

6.5 Fault Time-Line Analysis

According to the utility outage report, the fault was on phase B and caused by a bird crash. Now, let us look at the recorded voltage and current magnitudes on all phases on micro-PMU 1 and micro-PMU 2 that are located at the upstream and the downstream of fault location, as shown in Fig. 6.2. We can make the following immediate observations from this figure:

1. There exist drastic instantaneous changes in voltage magnitude and current magnitude across *all three phases*. This contradicts our expectation based on the utility outage report that the fault was only on Phase B;
2. The transient in voltage magnitude and current magnitude can be divided into *four perceptible periods*;

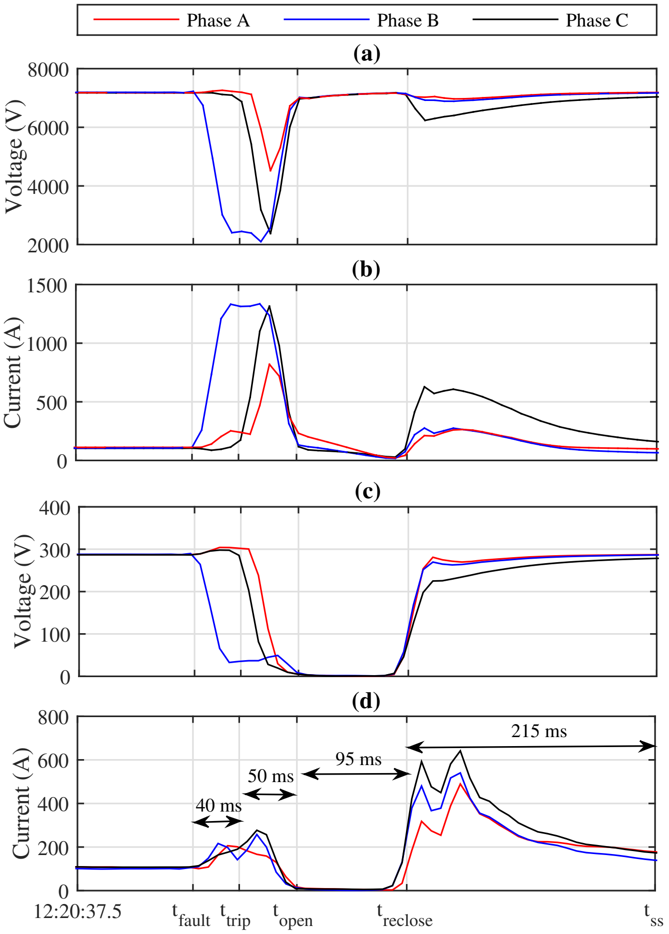

Figure 6.2: Three phase voltage magnitude and current magnitude during fault time-line: (a) and (b) micro-PMU 1, (c) and (d) micro-PMU 2.

3. The changes in voltage and current magnitudes are *not* the same on the upstream versus downstream of the fault.
4. While the voltage magnitude and current magnitude at micro-PMU 2 drop to zero and stay at zero for almost 95 msec, the voltage magnitude at micro-PMU 1 almost remains at is nominal level, i.e., around 7200 volts.

Therefore, it is clear that there is more room to investigate beyond what the utility field crew included in their fault report. Following the second item above, next, we investigate each of the four identified stages that comprise the fault time-line.

6.5.1 Stage I

The first stage starts at time t_{fault} and lasts for almost 40 milliseconds, till t_{trip}. During this period, it is assumed that there exists a single phase to neutral fault on phase B, as highlighted in Fig. 6.1. As it can be seen from Fig. 6.2, voltage on phase B drops at Substation B by 70%. The amount of such drop depends on fault location and fault resistance. We can see that the deviations in customer voltage on phase B more or less follow those of the substation voltage, see Fig. 6.2(c). During this period, substation injects fault current on phase B; consequently substation current increases on phase B to almost 1400 A. The hypothesis of a fault existing in phase B is verified by these observations. Note that, fluctuations in current magnitude of micro-PMU 2 will be studied in the next section with the focus on the effect of fault on PV units.

Another interesting observation during the first stage is the low-magnitude deviation in voltage on phases A and C due to fault occurrence on phase B, see Fig. 6.3. Based on this figure, the magnitude of voltage on phase B decreases as expected. However, there exist deviations in magnitude of phases A and C, due to changing the neutral of the three-phase system. Note that, the phase-to-neutral fault current with specific fault resistance leads to changing the three-phase neutral.

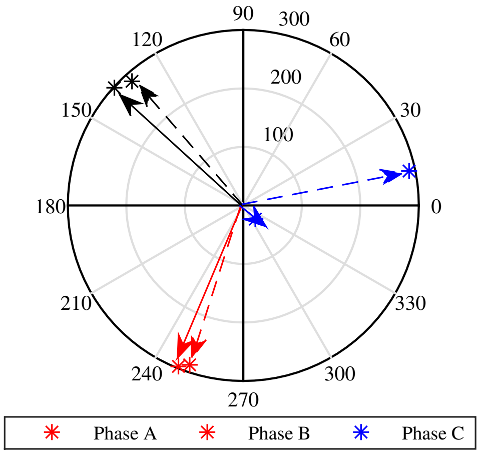

Figure 6.3: Three-phase voltage phase angle and magnitude fluctuation of micro-PMU 2 during fault: dash-line at t_{fault} and solid-line at t_{trip}.

6.5.2 Stage II

At t_{trip}, recloser sends the trip command to the circuit breaker, because the fault current exceeds the recloser instantaneous pick-up current. Thus, it is concluded that during the second stage, the transient is caused by the operation of circuit breaker, which lasts about 50 milliseconds to completely isolate Feeder I from Substation B, at t_{open}. To the extent of our knowledge, the circuit breaker of Feeder I is a three-phase switch, thus all phases are affected during this period.

6.5.3 Stage III

During the third stage, Feeder I is isolated from the rest of network. Thus, micro-PMU 1 reports the grid-side measurements and micro-PMU 2 reports the measurements on the isolated feeder. As expected, on one hand, the three-phase voltage of micro-PMU 1 increased to its nominal voltage during this period. On the other hand, the voltage at customer-side drops to zero in this period, since the feeder is isolated by circuit breaker. Also, the customer-side current is zero during this period. All customers in Feeder I must have similarly experienced a momentary outage during this period.

During the third period, the equivalent circuit from micro-PMU 1 is for the most part the same as an opened RLC circuit. Based on the circuit analysis, when the RLC circuit is opened, its current oscillates around zero with specific damping. As it can be seen from Fig. 6.2(b), the current of micro-PMU 1 gradually decreases to zero. Note that the gradual reduction could also be due to the micro-PMUs internal low pass filters. Also, three-phase current angle in the corresponding period fluctuates about 180 degrees, see Fig. 6.4. These two observations on the substation current magnitude and the fluctuations in current angle explain the presence of current oscillation in form of the one in an open-circuit RLC circuit. A similar conclusion can be made from the active and reactive power oscillations at substation, as shown in Fig. 6.5.

6.5.4 Stage IV

At t_{reclose}, the first shot of recloser is finished and the feeder is reconnected to substation. The fourth stage starts from t_{reclose} to t_{ss}, where t_{ss} is around the beginning time of the post-fault steady-state. During this period, all loads on this feeder, except for the loads on the downstream of the blown fuse, are *reconnected* to the grid. This causes a disturbance in substation voltage and current magnitudes, see Fig. 6.2(a) and (b). For instance, the surge current of turning-on the loads in phase C is about six time greater than the pre-fault current. Also, there exists a three-phase surge current in customer-level, i.e.,

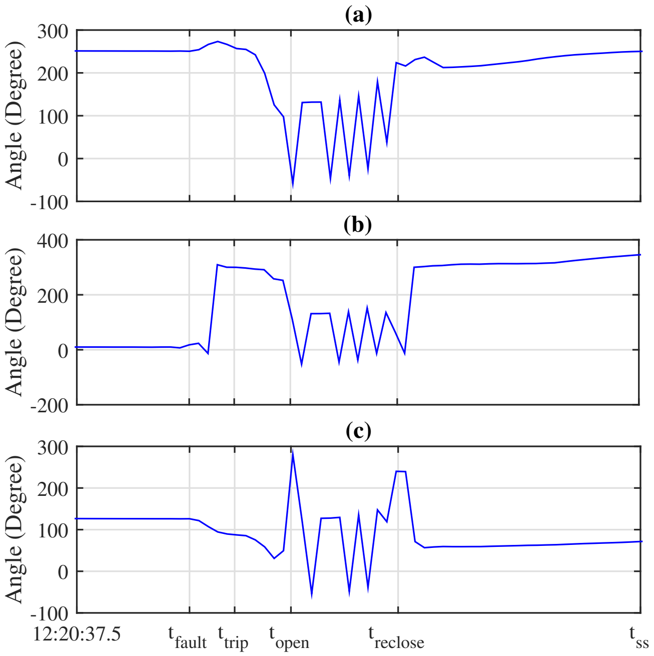

Figure 6.4: Current phase angle of micro-PMU 1 during fault time-line: (a) Phase A (b) Phase B (c) Phase C.

micro-PMU 2, due to reconnecting the load after reclosing.

Interestingly, but expectedly, some power quality sensitive loads are reconnected with some delay. This issue is better understood by looking at the per-phase steady-state pre-fault and post-fault active and reactive power of Feeder I, which is read by micro-PMU 1, as in Fig. 6.6. We can see that, after reclosing, some of the sensitive loads are still disconnected. The reduction in active and reactive power of phase B is more than phase A and C, which verifies that the customers fed by phase B in faulted zone are isolated by blown fuse.

We also examined the coordination between lateral fuses and feeder recloser to detect any possible miscoordination. In practice, the recloser in instantaneous mode must

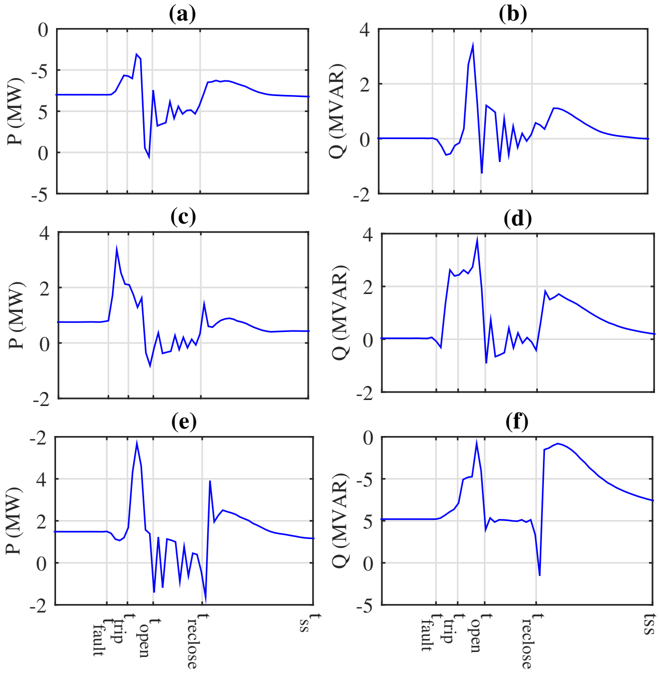

Figure 6.5: Active and reactive power at micro-PMU 1 during fault time-line: (a) and (b) Active and reactive power of phase A (c) and (d) Active and reactive power of phase B (e) and (f) Active and reactive power of phase C.

recognize momentary faults, and operate faster than lateral fuses [134]. This clearly was the case in this real-world fault event. However, as we will discuss in the next section, there might be a fuse-recloser miscoordination caused by the additional fault current supplied by DERs in active networks. In general, the coordination between recloser and lateral fuse in active feeders depends on fault type, fault location, generation level of DER, etc.

6.6 Responses of PV Resources to Fault

This section aims to study the impact of fault occurrence on grid connected PV generation. The faulted feeder, i.e., Feeder I, includes 100 kW behind-the-meter PV generation.

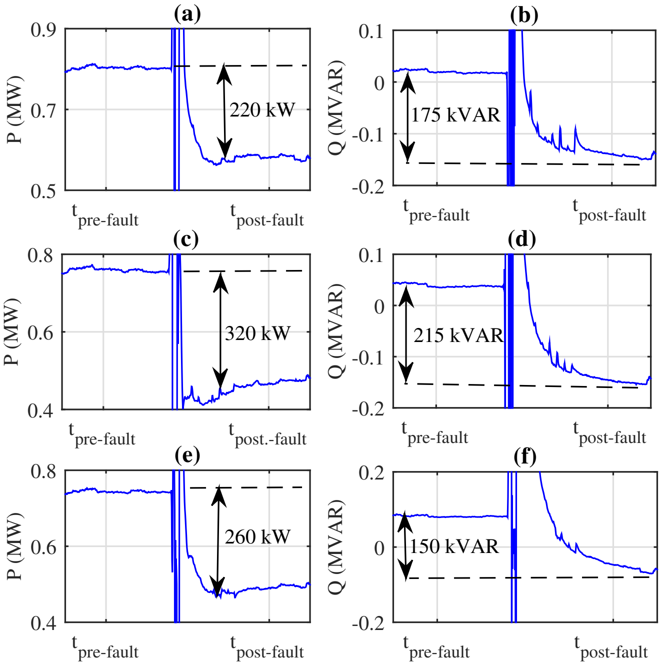

Figure 6.6: Active and reactive power at micro-PMU 1 during fault time-line: (a) and (b) Active and reactive power of phase A (c) and (d) Active and reactive power of phase B (e) and (f) Active and reactive power of phase C.

There was no reverse power flow from PV bus back to the feeder before the fault occurrence. The dash-lines in Fig. 6.7 show the three-phase current magnitude and angle of micro-PMU 2 at t_{fault}. Before the fault occurrence, the PV bus loads are served partially by the PV unit, and the rest by the main grid. After fault occurrence, however, the current of phase B rotates approximately 180 degrees, and increases in magnitude, see the solid-line in Fig. 6.7, which means that during Stage I of the fault, the current phase angle reverses its polarity. Based on the current magnitude and phase angle, we can conclude that after fault occurrence, phase B of the PV generation injects power to the fault location. However, Phase A and C of PV generation still injects power to the nearby loads with low-deviation in magnitude and phase.

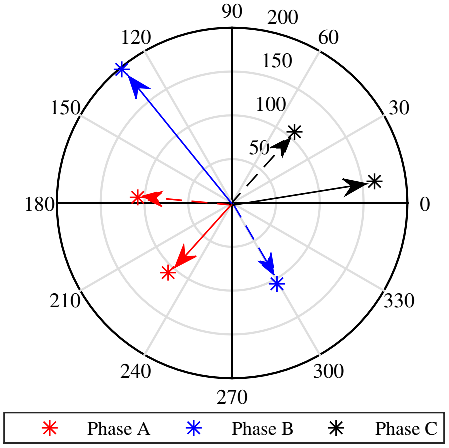

Figure 6.7: Three-phase current phase angle and magnitude fluctuation of micro-PMU 2 during fault: dash-line at t_{fault} and solid-line at t_{trip}.

In practice, almost all grid-tied inverters have a built-in anti-islanding protection scheme to stop feeding power back to the grid during interruption on the side of the feeder. Accordingly, at t_{open} the anti-islanding protection system of PV inverter detects feeder isolation due to recloser tripping, and disconnects itself from the feeder. This hypothesis is justified in Fig. 6.8. Based on this figure, at $t_{\text{pre-fault}}$ the PV derives 8.11 kW, 9.21 kW, and 5.31 kW from the grid and the rest of power is supported by PV panels. Whereas, after fault occurrence and recloser operation, the PV panel is still disconnected and the entire loads should be grid fed. According, the active power measured by micro-PMU 2 increases. After almost 6 minuets, the PV generation is reconnected to the grid and the active power measured by micro-PMU 2 decreases to pre-fault value, see Fig. 6.8. This observation also can be made from Invert Meter 1, where the output power of PV drops to zero at 12:20 PM and lasts for about 6 minutes, as shown and marked by ellipse in Fig. 6.9. Based on this observation, there exists a miscoordination between anti-islanding protection scheme

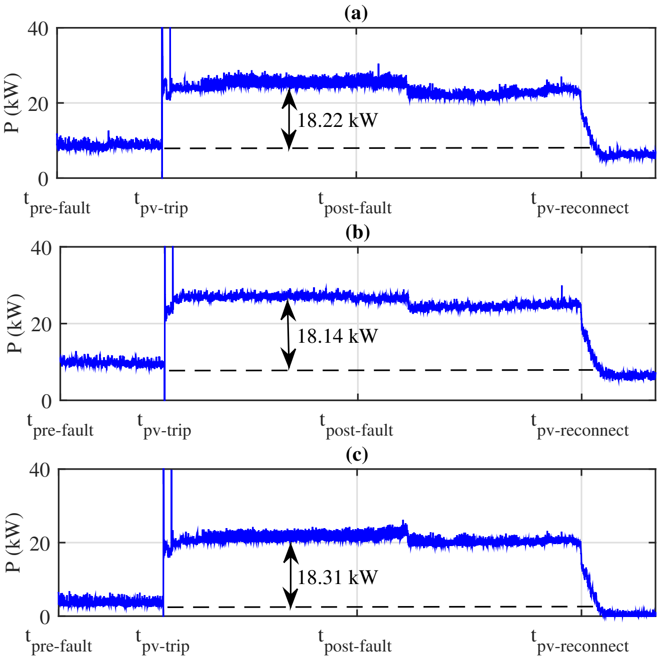

Figure 6.8: Active power at micro-PMU 2 from pre-fault to post-fault: (a) phase A (b) phase B (c) phase C.

of inverter and feeder recloser. Despite the permanent fault, the PV bus experiences a momentary fault and after $t_{reclosing}$, the fault is cleared by the blown lateral fuse.

Also, the fault causes unusual disturbances on outlying PV units, located on Feeder II and III. However, based on the data of micro-PMU 3 and 5 as well as Inverter Meter 2 (not shown here), there was no interruption in PV generation output.

6.7 Fault Effect on Outlying Area

This section briefly investigates the effect of fault on outlying areas, at both customer-level and feeder-level. As mentioned before, though Feeder I and II are fed from the same substation, they are electrically isolated in substation. Therefore, we can treat the

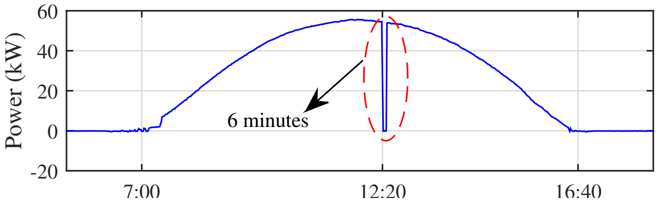

Figure 6.9: Active power of PV panel in Feeeder I measured by Inverter Meter 1 including the fault and reclosing time at 12:20.

data from micro-PMU 3 as customer-level measurements in an outlying area. Furthermore, there are also two micro-PMUs, i.e., micro-PMU 4 and 5, that are physically located tens of miles away from the faulted feeder. They too report the effect of fault in additional outlying areas. The three-phase voltage magnitude of these three outlying micro-PMUs are depicted in Fig. 6.10.

Based on Fig. 6.10(a), the fluctuation in voltage magnitude of micro-PMU 3 during the first stage of the event is similar to the fluctuations observed by micro-PMUs 1 and 2, see Fig. 6.2(a) and (c). The reason behind this observation is that the fault affects the voltage of common point of micro-PMU 2 and 3. Thus micro-PMU 3 follows the same voltage signature from the common point, of course, with lower-deviation from nominal voltage compared to micro-PMUs 1 and 2.

From Figs. 6.10(a) and (b), the voltage magnitude deviations in micro-PMUs 4 and 5 have different signatures compared to those of micro-PMUs 1, 2, and 3. The accurate justification for these differences would require additional detailed information on the rest of the system, including the exact grid topology, the type and locations of phase shifters, transformer windings, and voltage regulators between Feeder I and III.

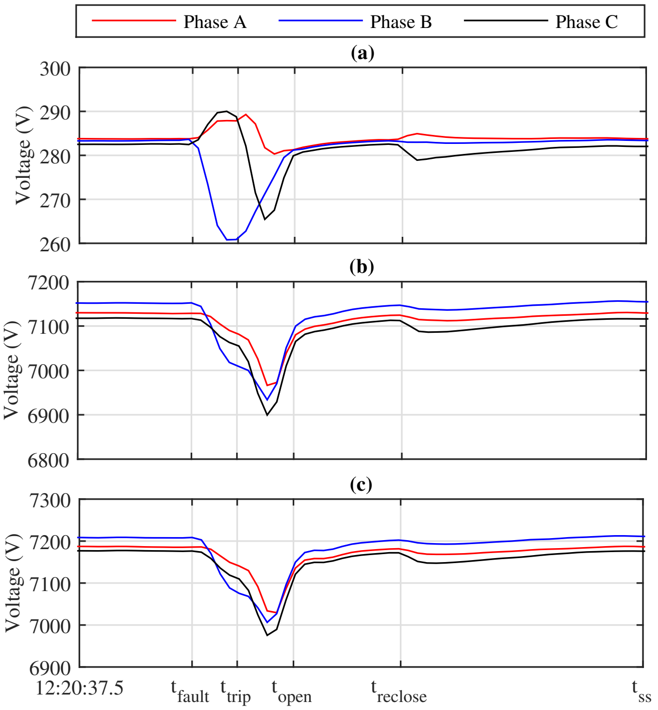

Figure 6.10: Synchronized voltage magnitude measurements during fault in three outlying locations: (a) micro-PMU 3, (b) micro-PMU 4, (c) Micro PMU 5.

6.8 Conclusions

A data-driven experimental fault analysis is presented to identify fault time-line in fast-time scales at a real-life distribution and sub-transmission system in Riverside, CA, based on synchronized measurements from five micro-PMUs. First, the exact time-line of the fault is extracted, followed by discussions on each stage of the time-line. Second, the optimal operation of protective devices during each stage of the time-line is examined. Of particular interest was to investigate the fuse-recloser coordination based on the experimental data from real-life system. Third, the transient response of PV resources to fault, specially during islanding mode, was studied. Our analysis showed a miscoordination between recloser and anti-islanding protection system of the PV unit. Finally, the effect of fault on outlying areas was investigated. It was observed that the outlying areas experienced major power-quality disturbance; the transient of deviations depends on grid parameters.

The results in this study can be extended in several directions. First, the methodology that is used in this analysis can be repeated to other fault events on the understudy distribution and sub-transmission system. This will allow creating a database for similar autopsy results for a variety of fault occurrences. Second, while the methodology that is used in this paper is mostly intuitive and manual; in future, one can extend the analysis to perform similar data extractions, such as in terms of identifying the different stages of the fault, in an automated fashion using proper algorithms.

Chapter 7

Lightning Initiated Contingency Analysis

7.1 Introduction

Lightning strikes may affect quality of power supply as well as service continuity. From power quality point of view, the momentary interruptions made by lightning strikes may cause damages to customers' equipment and revenue loss to power utilities. Lightning strikes may also damage critical elements of the power system such as transformers.

The effect of lightning on power electricity lines can be traced by monitoring the transient impulse voltages, either produced by lightning directly striking a phase conductor or induced by lightning flash striking nearby the power line (indirect lightning). This short tail impulse voltage, with a *microsecond* duration, is large enough to damage the equipment in the system.

In order to protect equipment from lightning impulse voltages, surge arresters are used [138]. In case the overvoltage exceeds the minimum strength voltage of a surge arrester, the flashover will be initiated. This lightning-initiated flashover is a phenomenon which causes an electricity arc due to breaking the minimum threshold of voltage insulator [139]. Such arcing would result in a temporary short-circuit, accompanying with a dip in voltage and increase in current, which might take a few *millisecond*. At distribution level, this kind of temporary short-circuit events can have considerable effects on power quality sensitive loads.

In the context of *active* distribution systems, i.e., distribution systems that are connected to DERs, the effects of lightning-initiated flashover and temporary short-circuit are becoming major challenges for utilities. In the US alone, the solar energy deployment increased from 2014 to 2015 by 28%. In addition to harvesting power from photovoltaic (PV) farms, there has been a significant increase in penetration of customer-owned behind-the-meter PV panels in recent years [135].

7.2 Related Works

Many studies have previously examined the effect of lightning-initiated flashover on *passive* distribution systems. In [140], the performance of a distribution line struck by direct lightning was analyzed, which was done based on the flashover and back-flashover calculation due to lightning. In [141], the authors proposed a statistical evaluation of lightning overvoltage on overhead distribution lines based on neural network method. In [142], the authors studied the cause-and-effect relationships between lightning flashes and the corresponding events on the nearby distribution system, such as permanent and transient system outages. In [143], the authors studied the performance of overvoltage protection via surge arresters against lightning overvoltage, in the presence of DGs and other smart grid resources.

With respect to DERs, several efforts, such as [144,145], reported on experimental and field study regarding the effect of direct or indirect lightning striking in a nearby PV modules. However, they did not examine the effect of lightning-initiated flashover propagated on power grid network on the PV inverters connected to the distribution network. Any such analysis inevitably requires using *synchronized* measurements across the power grid at the moment when the lightning strike occurs.

Because of the random nature of lightning, it is highly insightful to conduct data driven studies of lightening and initiated flashover. In [146], the authors discussed on the use of large data sets obtained from different measurements such as phasor measurement units (PMUs) to observe and trace the effect of lightning events. In [147], the author reported on data obtained from PMUs during a lightning strike, which captured some transient changes on voltage and current phasors caused by lightning. In [148], the authors described a PMU data management system that supports input from multiple PMU data streams for detection of events such as lightning strikes.

7.3 Motivations and Contributions

The studies in [146–148] are all based on PMU data at *transmission level.* In contrast, in this study, we seek to study lightening events using data from distribution-level

PMUs, a.k.a, micro-PMUs. This allows us to focus our analysis on DERs and active distribution networks. Micro-PMUs have a sampling rate that is higher than that of typical commercial transmission PMUs. They also have higher precision compared to their transmission-level counterparts. Some of the benefits of micro-PMUs are described in [19, 149], e.g., with respect to event and fault detection and diagnostics at distribution feeders. Micro-PMUs are gradually becoming commercially available, and several innovative applications of micro-PMU data have recently been proposed in [3, 9].

This study aims to propose a novel data-driven approach to use experimental distribution-level synchrophasor data, i.e., voltage and current measurements, on three phases, to analyze transient behaviors of flashovers caused by lightning strikes and the corresponding responses of the system to this phenomenon. In this regard, the synchrophasor data during three actual lightning strikes in a real-world distribution system in Riverside, CA are considered. Based on the obtained data, the transient behavior of short-circuit accompanied lightning-initiated flashover as well as the response of DERs to this phenomenon is investigated. Of particular interest is to analyze the impact and the response of a 7.5 MW PV farm. To the best of our knowledge, this is the first study that studies lightening-induced events at distribution level using micro-PMUs.

7.4 Under-Study Real-Life Test System

In order to study the effect of lightning and lightning-initiated contingencies, a real-life power network is considered in Riverside, CA. The single line diagram of the under-study network is shown in Fig. 7.1. The point of common coupling between the transmission

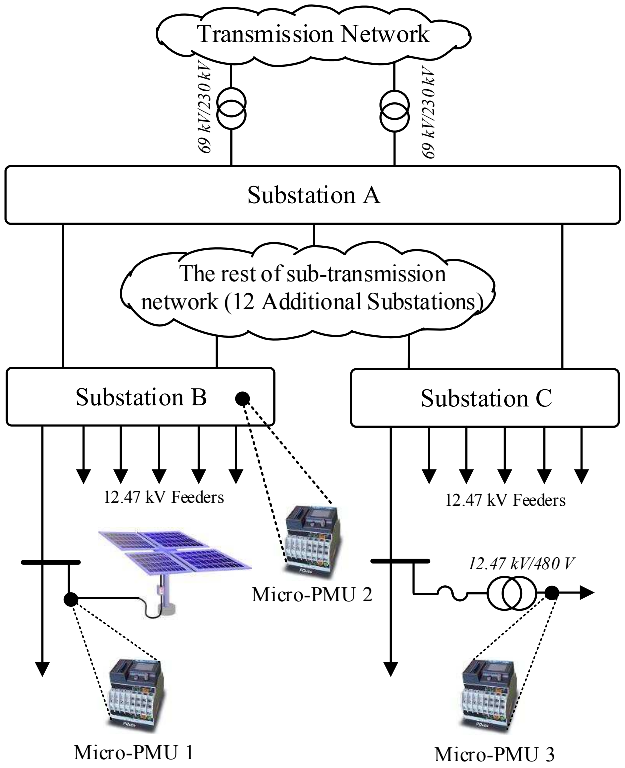

Figure 7.1: The real-world test system that is studied in this chapter is part of a sub-transmission system in Riverside, CA, which includes a PV farm under Substation B. Data from three micro-PMUs are used for this analysis.

system and RPU's sub-transmission network is marked as *Substation A*. In total, the under-study sub-transmission network includes 15 substations, 69 kV and 33 kV. Of interest in this study is *Substation B* and to a lesser extend *Substation C*. *Substation B* is interconnected with a 7.5 MW investor-owned behind-the-meter solar farm comprising 25,000 solar panels. The solar-generated power is fed into the local distribution grid and is enough to power about 1,600 homes.

The network is equipped with several micro-PMUs, while only the data from three micro-PMUs is available on the day of interest in this study. micro-PMU 1 is at grid-connected side of PV farm, which is a local generation node at distribution feeder. micro-PMU 2, is deployed at low voltage side of a 69 kV substation marked as *Substation B*. Finally, micro-PMU 3 is deployed at the secondary side of a 12.47 kV to 480 V transformer at a commercial building located at downstream of *Substation C*. This third micro-PMU is in an outlying area from *Substation B*. It is intended to allow us distinguish system wide events from events that specific to Substation C and the PV farm.

The installed micro-PMUs report four fundamental measurements on three phases, i.e., 13 channels total: voltage magnitude, voltage phase angle, current magnitude, and current phase angle, with the sampling rate of 120 Hz, i.e., one sample every 8.333 msec. The 13th channel is GPS Lock, utilized to determine if the sensor has established a satellite lock to ensure precise time synchronization. This high sampling rate, together with the fact that we have direct access to voltage and current data of different locations, allow us capture the effect of lightning-initiated contingencies within a data-driven framework, and also the response of the PV farm and the overall system to such contingencies. Given the fact that since the micro-PMU 3 is installed at outlying area from micro-PMU 1 and 2, the system-wide and local contingencies can be distinguished.

The focus in this study is on the micro-PMU data that is collected during lightening strikes in Riverside, CA on October 24, 2016 between 11:00:00 AM and 03:00:00 PM local time, which included multiple lightening strikes [150]. The recorded single phase voltage magnitude and current magnitude of micro-PMU 1 is shown in Fig. 7.2. The same data at the same location and during the same time frame, but for a sunny day on October 25, 2016 [150], is shown in Fig. 7.3. Generally speaking, the frequency and magnitude of

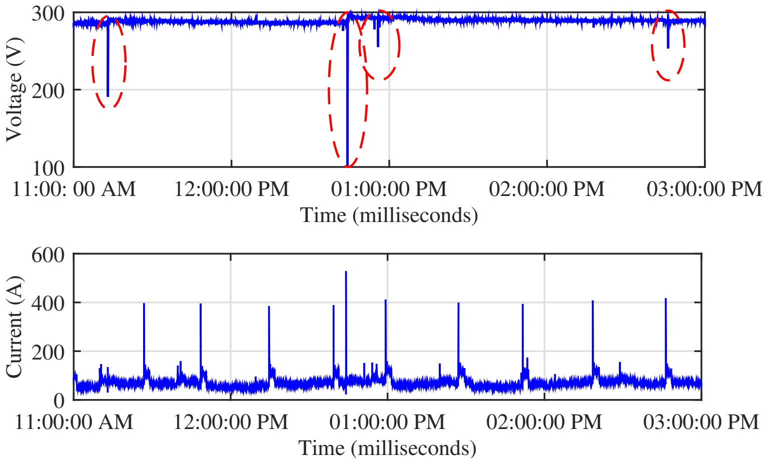

Figure 7.2: Single phase voltage and current of micro-PMU 3 for a four-hours period during a *rainy day*, October 24 2016.

current (as opposed to voltage) fluctuations are similar on both days with no significant observation on the effect of lightning. However, the two figures differ significantly with respect to the voltage measurements. Specifically, during the sunny day, voltage fluctuated between 282.58 to 290.14 volts, i.e., at almost ±2%. In contrast, there are at least four *voltage dips* during the rainy day with 25%-65% voltage drop, marked by dashed-ellipses in Fig. 7.2.

The voltage drops could be initiated by a major disturbance like upstream breaker operation or caused by short-circuit event due to surge arrester operation. To an extent of our knowledge, there is no protection system operation reported by crew. Accordingly, and based on the weather data from [150], we can conclude that the captured transient fluctuations in voltage could have been initiated by lightening strikes.

Generally speaking, the induced voltage of a lightning strike will initiate flashover, arcing, and consequently a transient short-circuit, when the surge arresters damp the spike

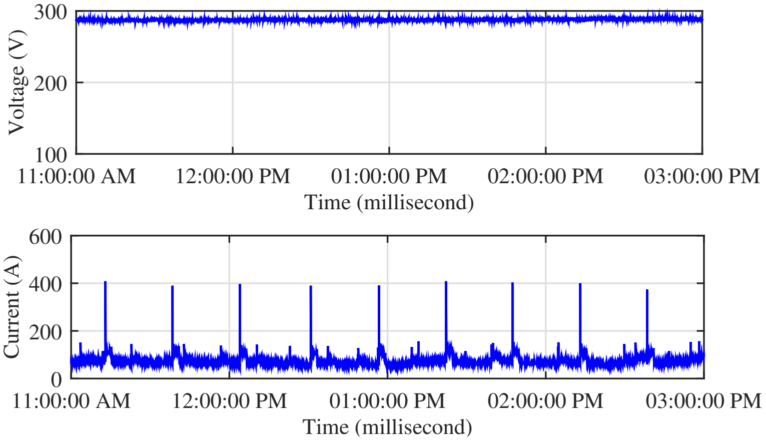

Figure 7.3: Single phase voltage and current of micro-PMU 3 for a four-hours period during a *sunny day*, October 25 2016, one day before the rainy day.

down. The short-circuit current will then flow in the form of an arc, thus, posing a fault in the power system. Next, we provide data driven explanations for the three major voltage dips in Fig. 7.2, at 11:13:01 AM, 12:44:05 PM, and 12:55:46 PM.

7.5 Data-Driven Analysis of Lightning-Initiated Contingencies

In this section, the three lightning-induced transient contingencies that we identified at the end of the previous section are analyzed. Our data-driven approach is conducted for all three installed micro-PMUs so that we can distinguish the local and system-wide events, i.e., in the sense of transmission or sub-transmission levels, followed by per-phase analysis.

7.5.1 First Lightning Event

The first event is occurred at 11:13:01 and lasted for about 200 milliseconds. Fig. 7.4 shows the recorded per-phase voltage and current from the micro-PMUs. The immediate observation is that the under-study transient event is system-wide, i.e., it is not confined to Substation B and its interconnected PV farm, because micro-PMU 3 is also seeing the same event.

According to the captured voltage signatures, the voltage drop at micro-PMUs 1, 2, and 3 are 51%, 56%, and 60%, respectively. As mentioned before, there is no recorded fault in the utility's database during the under-study period of time. Another important observation is that the voltage dip signatures are homologous with lightning-initiated flashover and arcing voltage sags. Consequently, we can conclude that the transient event is indeed system-wide and caused as a result of lightning followed by flashover and arcing on surge arresters, somewhere on the in transmission level or sub-transmission level.

Based on the voltage signatures in Fig. 7.4, there are two voltage stress periods caused by lightning, from :01.1 to :01.17 and from :01.17 to :01.3. The voltage stresses are high enough to exceed the withstand voltage of insulator and consequently cause flashover. Also, it seems that the magnitude of voltage stress is higher on *phase C* is higher than the other two phases.

Since there exists a transient behaviour in the voltage of the entire RPU's network, i.e., the system-wide voltage dip, there must be current responses to the disturbance. However, the current response of each device highly depends on its own dynamic. For instance, the current signature of micro-PMU 1, see Fig. 7.4 (a), shows that there is a three phase

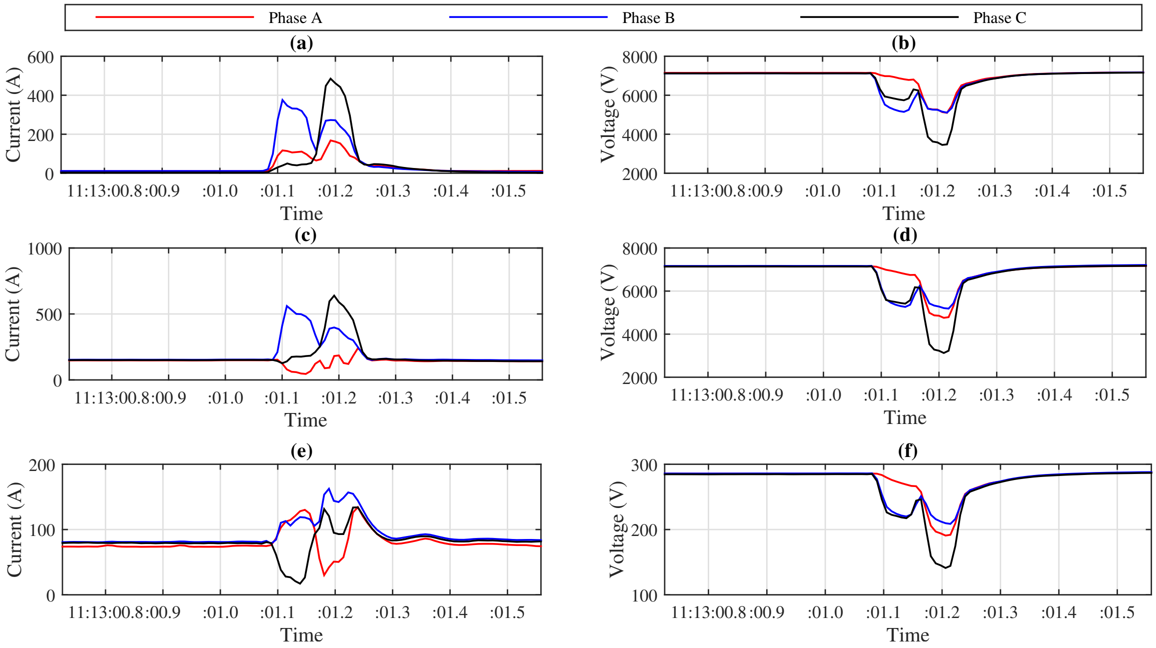

Figure 7.4: Three phase voltage and current transients at the first lightening event: (a) and (b) micro-PMU 1; (c) and (d) micro-PMU 2; (e) and (f) micro-PMU 3.

surge current in the PV farm, which may or may not cause a reverse power flow in transient period, because we know that in normal operation, power flows from *Substation B* towards the *PV Farm*.

Thus, if there is any possible reserve power flow, it should be detected in relative phase angle difference (RPAD) between micro-PMU 1 and 2 or the current angle of micro-PMU 1 should be change 180°. Fig. 7.5 shows pre-event and post-event current of phase C measured by micro-PMU 1. Note that, in the rainy day, the downstream loads of PV farm are greater than the generated power. As it can be seen from this figure, current angle changes about 180° which means there could be a reverse power flow from PV farm toward the short-circuited point. Also, the magnitude of current increases as mentioned in Fig. 7.4. Also, the active power supported by phase C of micro-PMU 1 is shown in Fig. 7.6. As

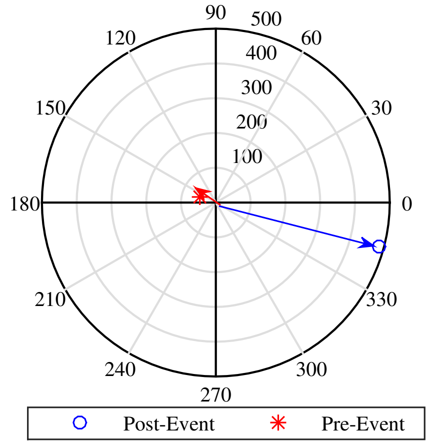

Figure 7.5: Post-event and pre-event current phasors at micro-PMU 1.

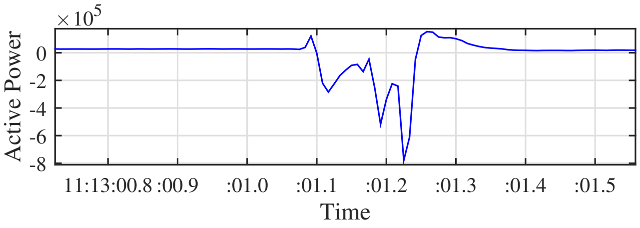

Figure 7.6: Active power of phase C measured by micro-PMU 1.

expected, there exists a reverse power flow in post-event compared to pre-event.

The above observation on momentary reverse power flow is important because most distribution networks are designed to operate on a unidirectional power flow, i.e., power flows from substation to the end of radial feeder. Accordingly, feeder protection system, e.g., overcurrent protection relay, is designed to trigger based on unidirectional power flow and current. While in the case of possible reverse power flow, feeder protection system may not work properly. Accordingly, the protection system should be re-designed based on the possible reverse power flow to ensure the network security and reliability.

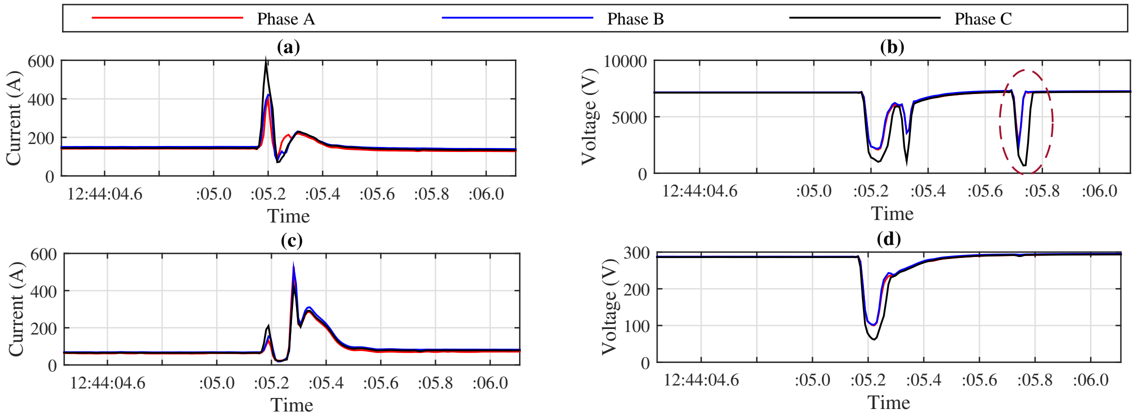

Figure 7.7: Three phase voltage and current transients at the second lightening event: (a) and (b) micro-PMU 2; (c) and (d) micro-PMU 3.

7.5.2 Second Lightning Event

The second event occurred at 12:44:05 and lasted for about 250 milliseconds. Fig. 7.7 shows the recorded per-phase voltage and current from micro-PMUs 2 and 3. Note that, no data was provided for micro-PMU 1 in this figure. The reason is that micro-PMU 1 stopped recording right before this second lightning event. Our understanding is that this happened because the PV farm was temporarily disconnected from the grid due to the drop in solar irradiance beyond a threshold. Although, as we will see later in Section 7.5.3, micro-PMU 1 automatically reset and resumed operation a few minutes later.

Same as in the first event, the under-study transient event is system-wide, because both micro-PMU 2 and 3 detected it. The amount of drop in voltage magnitudes at these two locations is 65% and 71%, respectively. Again, there is no recorded fault in the utility's database for during the under-study period of time. In addition, the voltage dip signatures are homologous with lightning-initiated flashover and arcing voltage sags. Consequently, we can conclude that the transient event is system-wide and caused as a result of lightning followed by flashover and arcing on surge arresters in transmission level.

Unlike in the first event in Section 7.5.1, the magnitude of voltage stress on all phases are roughly the same, see Fig. 7.7. Since there exists a transient response in voltage across RPU's network, i.e., the system-wide voltage dip, there would be current response to the disturbance. However, again, the current responses of various devices highly depend on their own dynamic. For instance, the current fluctuations in customer level, micro-PMU 3, depends on the load type, i.e., capacitive, inductive, or resistive.

Of great interest in the second lightning event is the fact that the micro-PMU 2 shows *another voltage dip* which was *post-event*, marked by a dashed-ellipse in Fig. 7.7 (b). This event was in fact local to Substation B, *because it had no effect on voltage or current at the outlying meter point*, i.e., on the measurements of micro-PMU 3. Our conjecture is that there was an issue at the PV farm, possibly losing some of the inverters at the moment of reconnection, which caused the disturbance.

7.5.3 Third Lightning Event

The last lightning event occurred at 12:55:46 and lasted for about 150 milliseconds. Fig. 7.8 shows the recorded per-phase voltage and current from all three micro-PMUs. Note that, the PV farm has already reconnected and micro-PMU 1 has already started recording by the time that we reach the third lightning event. Again, the event is system-wide because all three micro-PMUs have detected it. There exists voltage drop at all three phases; however, the magnitude of voltage drops are different for each phase, e.g., the voltage stress in phase C is more than phase A and B. Therefore, phase C experienced a

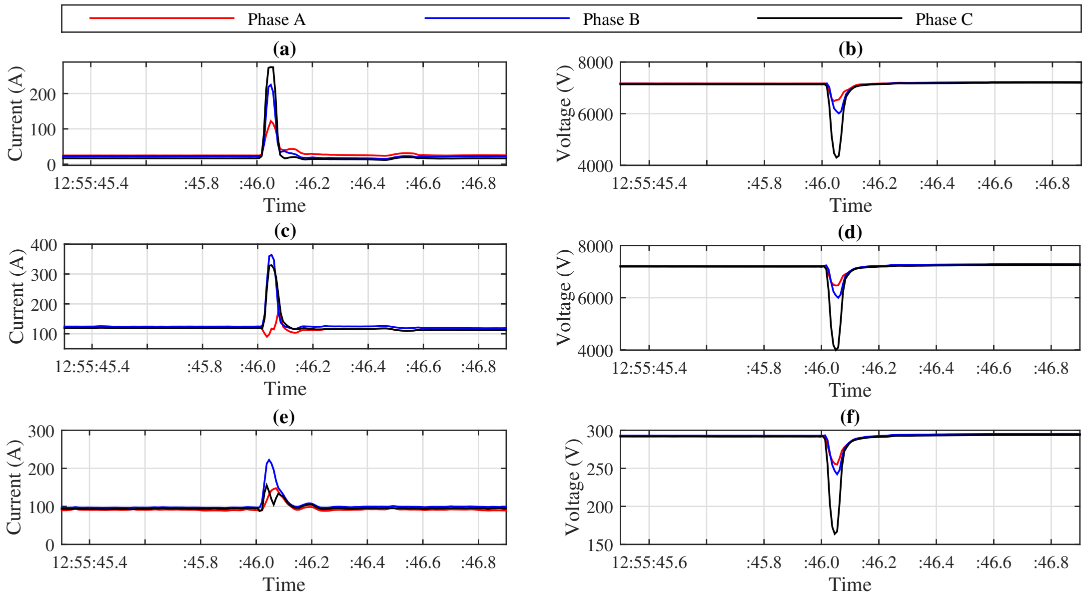

Figure 7.8: Three phase voltage and current transients at the third lightening event: (a) and (b) micro-PMU 1; (c) and (d) micro-PMU 2; (e) and (f) micro-PMU 3.

significant voltage dip. Similar to the first two events, we can conclude that the transient event is caused as a result of lightning, followed by flashover and arcing on surge arresters in transmission level.

In this case, the current response to the disturbance is resembles the surge current in switching on a load. However, the signatures are not the same for different phases.

Interestingly, similar to the first lightning event, the current of phase A at Substation B decreases at the beginning of the event while the PV farm feeds Phase A, see Fig. 7.4 (c) and Fig. 7.8 (c). The amount of drop in current depends on the level of PV farm generation and the load of the feeder in phase A.

7.6 Conclusions

A data-driven experimental analysis is provided to investigate lightening-induced contingencies at a real-life distribution system based on synchronized measurement data from three micro-PMUs. The core of the analysis is to separately explain the system-wide and local contingencies during three specific lightening events. Of particular interest was to analyze the impact and the response of a 7.5 MW PV farm. Several observations are made and the related causes are discussed. For example, it was shown that during a lightening strike, there can be a reverse power flow from the PV site to the substation due to a transient short-circuit caused by the surge arresters operation. This is despite the fact that during normal operation of the PV farm, i.e., at any time other than during a lightening strike, there is no reverse power flow due to the larger load at the feeder compared to the generation output of the PV. This is important, because most distribution networks have been designed to operate on a unidirectional power flow, and feeder protection system is designed and triggered based on unidirectional power flow. While in the case of reverse power flow caused by lightning event, the protection system should be re-designed to ensure the network security and reliability. These and other similar observations can lead the way to investigate how different DERs could be affected and and/or respond to the lightening-induced contingencies in distribution systems. The results in this study can also ultimately help with improving *resilience* in active distribution systems with PVs and other DERs.

The analysis in this study can be extended in several directions. For example, if a large data set is available for a large number of lightning events, then one can conduct a statistical analysis, for example on whether the observed reverse power flow at the PV farm is rare occurrence or rather a trend; and thus a potential reliability risk. Accordingly, one can also look into finding proper remedy options to the observed potential issues, such as re-calibration of the distribution system protection device in presence of major PV and other DER installations.

Chapter 8

Conclusions

The goal of this thesis was to start from a stream of raw micro-PMU data and turn them into information for tangible use-cases for power distribution systems. The measurement points in this study were obtained from micro-PMUs on a pilot distribution feeder in Riverside, CA. First, novel model-free event detection techniques are proposed to pick out valuable portion of data from micro-PMU data streams from a real-world test site in Riverside, CA. The proposed event detection methods are examined on two weeks real-world data from micro-PMUs. In total, 10,700 events were detected and examined. We investigate the importance of moving windows and dynamic window sizes in the event detection. It is shown that no single window size can detect all events. One should rather use a combination of different window sizes and moving windows. Also, the effectiveness of the proposed event detection methods are compared on real-world data from two micro-PMUs. Subsequently, a novel data-driven event labeling technique was combined with different methods of classification to classify the detected events at two layers. Interestingly, we concluded that adopting classification features from the detection process can considerably improve the overall classification accuracy. The effectiveness of the developed classifier is examined based on three types of classification methods. Finally, five real-world uses-cases were investigated, namely transient load modeling for application in frequency regulation market; static load modeling; remote asset monitoring; protection system diagnosis; lightning initiated contingency analysis. The results in this thesis could be of value to utilities and system operators.

This study can be extended to active distribution networks with higher penetration of DERs. While we expect our approach to perform well with detecting the increased number of events that may occur due to the increased number DERs, it may not be easy to label all such new events due to the limited information about the resources and equipment that are owned and operated by customers.

Bibliography

[1] H. Mohsenian-Rad, E. Stewart, and E. Cortez, “Distribution synchrophasors: Pairing big data with analytics to create actionable information,” *IEEE Power and Energy Magazine*, vol. 16, no. 3, pp. 26 – 34, Apr. 2018.

[2] A. V. Meier, D. Culler, A. McEachern, and R. Arghandeh, “Micro-synchrophasors for distribution systems,” in *Proc. of IEEE ISGT*, Washington, DC, Feb. 2014.

[3] A. Shahsavari, A. Sadeghi-Mobarakeh, E. Stewart, E. Cortez, L. Alvarez, F. Megala, and H. Mohsenian-Rad, “Distribution grid reliability versus regulation market efficiency: An analysis based on micro-PMU data,” *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2916 – 2925, Jun. 2017.

[4] A. Shahsavari, M. Farajollahi, and H. Mohsenian-Rad, “Individual load model parameter estimation in distribution systems using load switching events,” *IEEE Trans. on Power Systems*, to be published, doi: 10.1109/TPWRS.2019.2919901.

[5] M. Farajollahi, A. Shahsavari, E. Stewart, and H. Mohsenian-Rad, “Locating the source of events in power distribution systems using micro-PMU data,” *IEEE Trans. on Power Systems*, vol. 33, no. 6, pp. 6343 – 6354, May 2018.

[6] A. Shahsavari, M. Farajollahi, E. Stewart, A. von Meier, L. Alvarez, E. Cortez, and H. Mohsenian-Rad, “A data-driven analysis of capacitor bank operation at a distribution feeder using micro-PMU data,” in *Proc. of IEEE PES ISGT*, Washington D.C., Apr. 2017.

[7] A. Shahsavari, M. Farajollahi, E. Stewart, C. Roberts, F. Megala, L. Alvarez, E. Cortez, and H. Mohsenian-Rad, “Autopsy on active distribution networks: A data-driven fault analysis using micro-PMU data,” in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

[8] A. Shahsavari, M. Farajollahi, E. Stewart, C. Roberts, and H. Mohsenian-Rad, “A data-driven analysis of lightning-initiated contingencies at a distribution grid with a PV farm using micro-PMU data,” in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

[9] A. Shahsavari, A. Sadeghi-Mobarakeh, E. Stewart, and H. Mohsenian-Rad, “Distribution grid reliability analysis considering regulation down load resources via micro-pmu data,” in *Proc. of the IEEE International Conference on Smart Grid Communications*, Sydney, Australia, Nov. 2016.

[10] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, “Linear distribution system state estimation using synchrophasor data and pseudo-measurement,” in *Proc. of IEEE Smart Grid Synchronized Measurements and Analytics*, May 2019.

[11] ——, “Tracking state estimation in distribution network using synchrophasor data,” in *Proc. of IEEE PES General Meeting*, Aug. 2018.

[12] ——, “Location identification of high impedance faults using synchronized harmonic phasors,” in *Proc. of IEEE PES ISGT*, Washington D.C., Apr. 2017.

[13] ——, “Location identification of distribution network events using synchrophasor data,” in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

[14] O. Ardakanian, Y. Yuan, R. Dobbe, A. V. Meier, S. Low, and C. Tomlin, “Event detection and localization in distribution grids with phasor measurement units,” *arXiv preprint arXiv:1611.04653*, 2016.

[15] M. Pignati, L. Zanni, P. Romano, R. Cherkaoui, and M. Paolone, “Fault detection and faulted line identification in active distribution networks using synchrophasors-based real-time state estimation,” *IEEE Trans. on Power Delivery*, vol. 32, no. 1, pp. 381 – 392, 2017.

[16] Y. Zhou, R. Arghandeh, and C. J. Spanos, “Partial knowledge data-driven event detection for power distribution networks,” *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 5152 – 5162, Mar. 2018.

[17] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, “Anomaly detection using optimally-placed μPMU sensors in distribution grids,” *IEEE Trans. on Power Systems*, vol. 33, no. 4, pp. 3611 – 3623, Oct. 2017.

[18] Y. Zhou, R. Arghandeh, H. Zou, and C. Spanos, “Non-parametric event detection in multiple time series for power distribution networks,” *IEEE Trans. on Industrial Electronics*, vol. 66, no. 2, pp. 1619 – 1628, Jun. 2019.

[19] Y. Zhou, R. Arghandeh, I. Konstantakopoulos, S. Abdullah, A. von Meier, and C. J. Spanos, “Abnormal event detection with high resolution micro-PMU data,” in *Proc. of IEEE Power Systems Computation Conference*, Genoa, Italy, Jun. 2016.

[20] P. K. Ray, N. Kishor, and S. R. Mohanty, “Islanding and power quality disturbance detection in grid-connected hybrid power system using wavelet and s-transform,” *IEEE Trans. on Smart Grid*, vol. 3, no. 3, pp. 1082–1094, Jun. 2012.

[21] L. Xu and M. Chow, "A classification approach for power distribution systems fault cause identification," *IEEE Trans. on Power Systems*, vol. 21, no. 1, pp. 53–60, Jan. 2006.

[22] N. S. Coleman, C. Schegan, and K. N. Miu, "A study of power distribution system fault classification with machine learning techniques," in *Proc. of IEEE PES NAPS*, Charlotte, NC, Nov. 2015.

[23] J. Zhang, Z. He, S. Lin, Y. Zhang, and Q. Qian, "An ANFIS-based fault classification approach in power distribution system," *International Journal of Electrical Power & Energy Systems*, vol. 49, pp. 243–252, Jul. 2013.

[24] L. Xu, M.-Y. Chow, and L. S. Taylor, "Power distribution fault cause identification with imbalanced data using the data mining-based fuzzy classification e-algorithm," *IEEE Trans. on Power Systems*, vol. 22, no. 1, pp. 164–171, 2007.

[25] P. M. Ramos, F. M. Janeiro, A. C. Serra *et al.*, "Pq monitoring system for real-time detection and classification of disturbances in a single-phase power system," *IEEE Trans. on Instrumentation and Measurement*, vol. 57, no. 8, pp. 1725–1733, Jun. 2008.

[26] S. Mishra, C. Bhende, and B. Panigrahi, "Detection and classification of power quality disturbances using s-transform and probabilistic neural network," *IEEE Trans. on Power Delivery*, vol. 23, no. 1, pp. 280–287, Dec. 2008.

[27] E. Styvaktakis, M. H. Bollen, and I. Y. Gu, "Automatic classification of power system events using RMS voltage measurements," in *Proc. of IEEE Power Engineering Society Summer Meeting*, Chicago, IL, Jul. 2002.

[28] P. Dash, B. Panigrahi, D. Sahoo, and G. Panda, "Power quality disturbance data compression, detection, and classification using integrated spline wavelet and s-transform," *IEEE Trans. on Power Delivery*, vol. 18, no. 2, pp. 595–600, 2003.

[29] I. Niazazari and H. Livani, "A PMU data-driven disruptive event classification in distribution systems," *Electric Power Systems Research*, vol. 157, pp. 251–260, Apr. 2018.

[30] A. Shahsavari, M. Farajollahi, E. Stewart, E. Cortez, and H. Mohsenian-Rad, "Situational awareness in distribution grid using micro-PMU data: A machine learning approach," *IEEE Trans. on Power System*, to be published, doi: 10.1109/TSG.2019.2898676.

[31] ——, "A machine learning approach to event analysis in distribution feeders using distribution synchrophasors," in *Proc. of IEEE Smart Grid Synchronized Measurements and Analytics*, May 2019.

[32] P. Huber, "Robust statistics," *Wiley*, pp. 107–109, 2009.

[33] P. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 160–165, 2011.

[34] W. Sima, M. Zou, Q. Yang, M. Yang, and L. Li, "Field experiments on 10 kV switching shunt cap banks using ordinary and phase-controlled vacuum circuit breakers," *Energies*, vol. 9, no. 2, pp. 88–102, Jan. 2016.

[35] L. Shalabi, Z. Shaaban, and B. Kasasbeh, "Data mining: A preprocessing engine," *J of Computer Science*, vol. 2, no. 9, pp. 735–739, 2006.

[36] C. Hsu and C. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. on Neural Networks*, vol. 13, no. 2, pp. 415–425, Aug. 2002.

[37] G. Madzarov, D. Gjorgjevikj, and I. Chorbev, "A multi-class SVM classifier utilizing binary decision tree," *Informatica*, vol. 33, no. 2, 2009.

[38] Y. Liu and Y. F. Zheng, "One-against-all multi-class svm classification using reliability measures," in *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, vol. 2. IEEE, 2005, pp. 849–854.

[39] V. Franc and V. Hlavác, "Multi-class support vector machine," in *Proc. of IEEE International Conference on Pattern Recognition*, Aug. 2002.

[40] S. Suthaharan, "Machine learning models and algorithms for big data classification," *Integrated Series in Information Systems*, vol. 36, 2016.

[41] N. García-Pedrajas, J. A. R. del Castillo, and G. Cerruela-García, "A proposal for local k values for k-nearest neighbor rule," *IEEE Trans. on Neural Networks*, vol. 28, no. 2, pp. 470–475, Dec. 2017.

[42] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and svm-based data analytics for theft detection in smart grid," *IEEE Trans. on Industrial Informatics*, vol. 12, no. 3, pp. 1005–1016, Mar. 2016.

[43] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[44] B. J. Kirby, *Frequency regulation basics and trends.* United States. Department of Energy, 2005.

[45] H. Hao, B. Sanandaji, K. Poolla, and T. Vincent, "Potentials and economics of residential thermal loads providing regulation reserve," *Energy Policy*, vol. 79, pp. 115–126, 2015.

[46] Y. Zhang, M. E. Raoufat, and K. Tomsovic, "Remedial action schemes and defense systems," *Smart Grid Handbook*, 2016.

[47] M. Motalleb, M. Thornton, E. Reihani, and R. Ghorbani, "Providing frequency regulation reserve services using demand response scheduling," *Energy Conversion and Management*, vol. 124, pp. 439–452, Sep. 2016.

[48] O. Erdinç, A. Taşcıkaraoğlu, N. G. Paterakis, Y. Eren, and J. P. Catalão, "End-user comfort oriented day-ahead planning for responsive residential HVAC demand aggregation considering weather forecasts," *IEEE Trans. on Smart Grid*, vol. 8, no. 1, pp. 362–372, Apr. 2017.

[49] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building hvac systems," *IEEE Trans. on Smart Grid*, vol. 6, no. 2, pp. 776–783, Jan. 2015.

[50] J. H. Yoon, R. Baldick, and A. Novoselac, "Dynamic demand response controller based on real-time retail price for residential buildings," *IEEE Trans. on Smart Grid*, vol. 5, no. 1, pp. 121–129, Jan. 2014.

[51] P. Zhao, G. Henze, S. Plamp, and V. J. Cushing, "Evaluation of commercial building HVAC systems as frequency regulation providers," *Energy and Buildings*, vol. 67, pp. 225–235, Dec. 2013.

[52] E. Yao, V. Wong, and R. Schober, "A robust design of electric vehicle frequency regulation service," in *Proc. of IEEE SmartGridComm*, Jan. 2014.

[53] G. He, Q. Chen, C. Kang, P. Pinson, and Q. Xia, "Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life," *IEEE Trans. on Smart Grid*, vol. 7, no. 5, pp. 2359–2367, Sep. 2016.

[54] K. Ko, S. Han, and D. K. Sung, "Performance-based settlement of frequency regulation for electric vehicle aggregators," *IEEE Trans. on Smart Grid*, vol. 9, no. 2, pp. 866 – 875, May 2018.

[55] E. Yao, V. W. Wong, and R. Schober, "Robust frequency regulation capacity scheduling algorithm for electric vehicles," *IEEE Trans. on Smart Grid*, vol. 8, no. 2, pp. 984 – 997, Mar. 2017.

[56] S. Li, M. Brocanelli, W. Zhang, and X. Wang, "Integrated power management of data centers and electric vehicles for energy and regulation market participation," *IEEE Trans. on Smart Grid*, vol. 5, no. 5, pp. 2283–2294, Jun. 2014.

[57] California ISO. [Online]. Available: https://www.caiso.com/Documents/NGR-REMOverview.pdf

[58] PJM. [Online]. Available: http://www.pjm.com/Globals/Training/Courses/~/media/6BE10C20331147CFB740954976EE18D1.ashx

[59] R. N. Allan, *Reliability evaluation of power systems.* Springer Science & Business Media, 2013.

[60] A. Shahsavari, S. M. Mazhari, A. Fereidunian, and H. Lesani, “Fault indicator deployment in distribution systems considering available control and protection devices: a multi-objective formulation approach,” *IEEE Trans. on Power Systems*, vol. 29, no. 5, pp. 2359–2369, Sep. 2014.

[61] A. Shahsavari, A. Fereidunian, and S. M. Mazhari, “A joint automatic and manual switch placement within distribution systems considering operational probabilities of control sequences,” *International Trans. on Electrical Energy Systems*, vol. 25, no. 11, pp. 2745–2768, Sep. 2015.

[62] S. M. Mohammadi-Hosseininejad, A. Fereidunian, A. Shahsavari, and H. Lesani, “A healer reinforcement approach to self-healing in smart grid by PHEVs parking lot allocation,” *IEEE Trans. on Industrial Informatics*, vol. 12, no. 6, pp. 2020–2030, Dec. 2016.

[63] A. Shahsavari, A. Fereidunian, A. Ameli, S. M. Mazhari, and H. Lesani, “A healer reinforcement approach to smart grids by improving fault location function in FLISR,” in *Proc. of IEEE International Conference on Environment and Electrical Engineering*, Nov. 2013.

[64] A. Shahsavari, A. Fereidunian, and H. Lesani, “A healer reinforcement approach to smart grid self-healing by redundancy improvement,” in *Proc. of IEEE Electric Power Distribution Conference*, Apr. 2013.

[65] T. A. Short, *Distribution reliability and power quality.* CRC Press, 2005.

[66] California ISO, “Business Requirements Specification - Pay For Performance Regulation,” May 2013.

[67] PJM. [Online]. Available: https://www.pjm.com/~/media/committees-groups/subcommittees/drs/20150224/20150224-item-03b-draft-manual-12-version-32-regulation-redlines.ashx

[68] E. Mallada, C. Zhao, and S. Low, “Optimal load-side control for frequency regulation in smart grids,” vol. 62, no. 12, Dec. 2017, pp. 6294 – 6309.

[69] J. Bezdek, R. Ehrlich, and W. Full, “FCM: The fuzzy c-means clustering algorithm,” *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, 1984.

[70] W. Kersting, *Distribution system modeling and analysis.* CRC, 2012.

[71] A. Pradhan, A. Routray, and S. M. Gudipalli, “Fault direction estimation in radial distribution system using phase change in sequence current,” *IEEE Trans. on Power System*, vol. 22, no. 4, pp. 2065–2071, Oct. 2007.

[72] T. E. Mcdermott and G. Benmouyal, “Ieee standard inverse-time characteristic equations for overcurrent relays. discussion,” *IEEE Trans. on Power Delivery*, vol. 14, no. 3, pp. 868–872, Jul. 1999.

[73] AREVA. [Online]. Available: http://mt.schneider-electric.be/op_main/micom/p12x_en_da6.pdf

[74] D. Li and L. Qi, "Energy based fuse modeling and simulation," in *Proc. of IEEE Electric Ship Technologies Symposium*, Apr. 2013.

[75] B. Hegde, S. Midlam-Mohler, and P. J. Tulpule, "Thermal model of fuse dynamics for simulation under intermittent dc faults," in *Proc. of ASME Dynamic Systems and Control*, Oct. 2015.

[76] T. A. Short, *Electric power distribution handbook.* CRC press, 2014.

[77] Eaton, "Medium Voltage Expulsion Fuses," May 2002.

[78] M. E. Raoufat, A. Taalimi, K. Tomsovic, and R. Hay, "Event analysis of pulse-reclosers in distribution systems through sparse representation," *Proc. of International Conference on Intelligent System Application to Power Systems*, Sep. 2017.

[79] S. Chaitusaney and A. Yokoyama, "Prevention of reliability degradation from recloser–fuse miscoordination due to distributed generation," *IEEE Trans. on Power Delivery*, vol. 23, no. 4, pp. 2545–2554, 2008.

[80] S. M. Brahma and A. A. Girgis, "Microprocessor-based reclosing to coordinate fuse and recloser in a system with high penetration of distributed generation," in *Proc. of IEEE PES Winter Meeting*, Jan. 2002.

[81] R. N. Allan, R. Billinton, I. Sjarief, L. Goel, and K. So, "A reliability test system for educational purposes-basic distribution system data and results," *IEEE Trans. on Power System*, vol. 6, no. 2, pp. 813–820, May 1991.

[82] A. Arif, Z. Wang, J. Wang, B. Mather, H. Bashualdo, and D. Zhao, "Load modeling - a review," *IEEE Trans. on Smart Grid*, vol. 9, no. 6, pp. 5986 – 5999, Nov. 2018.

[83] J. V. Milanovic, K. Yamashita, S. M. Villanueva, S. Ž. Djokic, and L. M. Korunović, "International industry practice on power system load modeling," *IEEE Trans. on Power Systems*, vol. 28, no. 3, pp. 3038–3046, Aug. 2013.

[84] M. Yao, D. K. Molzahn, and J. L. Mathieu, "The impact of load models in an algorithm for improving voltage stability via demand response," in *Proc. of IEEE Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Jan. 2017.

[85] M. Dabbaghjamanesh, S. Mehraeen, A. Kavousi-Fard, and F. Ferdowsi, "A new efficient stochastic energy management technique for interconnected ac microgrids," in *Proc. of IEEE PES General Meeting*, Portland, OR, Aug. 2018.

[86] K. P. Schneider and T. Weaver, "Volt-var optimization on american electric power feeders in northeast columbus," in *Proc. of IEEE PES T&D*, Orlando, FL, May 2012.

[87] B. Milosevic and M. Begovic, "Voltage-stability protection and control using a wide-area network of phasor measurements," *IEEE Trans. on Power Systems*, vol. 18, no. 1, pp. 121–127, Feb. 2003.

[88] D. K. Molzahn, B. C. Lesieutre, and C. L. DeMarco, "Approximate representation of ZIP loads in a semidefinite relaxation of the OPF problem," *IEEE Trans. on Power Systems*, vol. 29, no. 4, pp. 1864–1865, Jan. 2014.

[89] K. P. Schneider, J. C. Fuller, and D. P. Chassin, "Multi-state load models for distribution system analysis," *IEEE Trans. on Power Systems*, vol. 26, no. 4, pp. 2425–2433, Nov. 2011.

[90] A. S. Carneiro, L. F. Araujo, J. L. R. Pereira, P. A. Garcia, I. D. Melo, and M. B. Amaral, "Static load modeling based on field measurements," in *Proc. of IEEE PowerTech*, Manchester, UK, Jun. 2017.

[91] C. A. Baone, S. Veda, Y. Pan, W. Premerlani, J. Dai, and A. Johnson, "Measurement based static load model identification," in *Proc. of IEEE PES General Meeting*, Denver, CO, Jul. 2015.

[92] J. Zhao, Z. Wang, and J. Wang, "Robust time-varying load modeling for conservation voltage reduction assessment," *IEEE Trans. on Smart Grid*, vol. 9, no. 4, pp. 3304–3312, Nov. 2016.

[93] S. Pandey, A. K. Srivastava, P. Markham, M. Patel *et al.*, "Online estimation of steady-state load models considering data anomalies," *IEEE Trans. on Industry Applications*, vol. 54, no. 1, pp. 712–721, Jan. 2018.

[94] V. Vignesh, S. Chakrabarti, and S. Srivastava, "An experimental study on the load modelling using PMU measurements," in *Proc. of IEEE PES T&D*, Chicago, IL, Apr. 2014.

[95] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "An event-oriented method for online load modeling based on synchrophasor data," *IEEE Trans. on Smart Grid*, vol. 6, no. 4, pp. 2060–2068, Mar. 2015.

[96] H. Renmu, M. Jin, and D. J. Hill, "Composite load modeling via measurement approach," *IEEE Trans. on Power Systems*, vol. 21, no. 2, pp. 663–672, Oct. 2006.

[97] S. Son, S. H. Lee, D. H. Choi, K. B. Song, J. D. Park, Y. H. Kwon, K. Hur, and J. W. Park, "Improvement of composite load modeling based on parameter sensitivity and dependency analyses," *IEEE Trans. on Power Systems*, vol. 29, no. 1, pp. 242–250, Sep. 2014.

[98] V. Vignesh, S. Chakrabarti, and S. C. Srivastava, "Load modeling under unbalanced disturbances," *IEEE Trans. on Power Systems*, vol. 31, no. 2, pp. 1661–1662, Mar. 2016.

[99] P. Regulski, D. Vilchis-Rodriguez, S. Djurovic, and V. Terzija, "Estimation of composite load model parameters using an improved particle swarm optimization method," *IEEE Trans. on Power Delivery*, vol. 30, no. 2, pp. 553–560, Feb. 2015.

[100] M. Jin, H. Dong, H. Ren-Mu, D. Zhao-Yang, and H. David, "Reducing identified parameters of measurement-based composite load model," *IEEE Trans. on Power Systems*, vol. 23, no. 1, pp. 76–83, Feb. 2008.

[101] K. Zhang, H. Zhu, and S. Guo, "Dependency analysis and improved parameter estimation for dynamic composite load modeling," *IEEE Trans. on Power Systems*, vol. 32, no. 4, pp. 3287–3297, Nov. 2017.

[102] D. Han, J. Ma, R.-m. He, and Z.-y. Dong, "A real application of measurement-based load modeling in large-scale power grids and its validation," *IEEE Trans. on Power Systems*, vol. 24, no. 4, pp. 1756–1764, Nov. 2009.

[103] B.-K. Choi and H.-D. Chiang, "Multiple solutions and plateau phenomenon in measurement-based load model development: issues and suggestions," *IEEE Trans. on Power Systems*, vol. 24, no. 2, pp. 824–831, May 2009.

[104] S. Shao, M. Pipattanasomporn, and S. Rahman, "Development of physical-based demand response-enabled residential load models," *IEEE Trans. on power systems*, vol. 28, no. 2, pp. 607–614, May 2013.

[105] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Development of low-voltage load models for the residential load sector," *IEEE Trans. Power System*, vol. 29, no. 5, pp. 2180–2188, Sep. 2014.

[106] R. Torquato, Q. Shi, W. Xu, and W. Freitas, "A monte carlo simulation platform for studying low voltage residential networks," *IEEE Trans. on Smart Grid*, vol. 5, no. 6, pp. 2766–2776, Jul. 2014.

[107] K. McKenna, A. Keane *et al.*, "Residential load modeling of price-based demand response for network impact studies," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2285–2294, Sep. 2016.

[108] P. Feldmann and R. A. Rohrer, "Proof of the number of independent kirchhoff equations in an electrical circuit," *IEEE Trans. on Circuits and Systems*, vol. 38, no. 7, pp. 681–684, Jul. 1991.

[109] M. Parodi and M. Storace, "Linear and nonlinear circuits: Basic & advanced concepts," *Springer*, pp. 230–231, 2018.

[110] D. Smith, M. Eggen, and R. S. Andre, "A transition to advanced mathematics," *Nelson Education*, pp. 147–148, 2014.

[111] T. M. Le, B. Fatahi, H. Khabbaz, and W. Sun, "Numerical optimization applying trust-region reflective least squares algorithm with constraints to optimize the non-linear creep parameters of soft soil," *Applied Mathematical Modelling*, vol. 41, pp. 236–256, Jan. 2017.

[112] M. K. Transtrum and J. P. Sethna, "Improvements to the levenberg-marquardt algorithm for nonlinear least-squares minimization," *arXiv preprint arXiv:1201.5885*, 2012.

[113] J. H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Trans. on Power Delivery*, vol. 18, no. 3, pp. 882–887, Jul. 2003.

[114] E. Caro, A. J. Conejo, R. Mínguez, M. Zima, and G. Andersson, "Multiple bad data identification considering measurement dependencies," *IEEE Trans. on Power Systems*, vol. 26, no. 4, pp. 1953–1961, Nov. 2011.

[115] H. Manitoba, "Research centre," *PSCAD/EMTDC: Electromagnetic transients program including dc systems*, 1994.

[116] A. Bokhari, A. Alkan, R. Dogan, M. Diaz-Aguiló, F. De Leon, D. Czarkowski, Z. Zabar, L. Birenbaum, A. Noel, and R. E. Uosef, "Experimental determination of the ZIP coefficients for modern residential, commercial, and industrial loads," *IEEE Trans. on Power Delivery*, vol. 29, no. 3, pp. 1372–1381, Oct. 2014.

[117] E. Sortomme, A. Negash, S. Venkata, and D. Kirschen, "Voltage dependent load models of charging electric vehicles," in *Proc. of IEEE PES General Meeting*, Vancouver, BC, Canada, Nov. 2013.

[118] K. D. Jones, A. Pal, and J. S. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Trans. on Power Systems*, vol. 30, no. 3, pp. 1121–1130, May 2015.

[119] K. Kirihara, K. E. Reinhard, A. K. Yoon, and P. W. Sauer, "Investigating synchrophasor data quality issues," in *Proc. of IEEE Power and Energy Conference*, Champaign, IL, Feb. 2014.

[120] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.

[121] IEEE, Distribution System Analysis Subcommittee Report. [Online]. Available: http://sites.ieee.org/pes-testfeeders/resources/

[122] M. McGranaghan, W. Reid, S. Law, and D. Gresham, "Overvoltage protection of shunt-capacitor banks using mov arresters," *IEEE Trans. on power apparatus and systems*, no. 8, pp. 2326–2336, 1984.

[123] M. McGranaghan, R. Zavadil, G. Hensley, T. Singh, and M. Samotyj, "Impact of utility switched capacitors on customer systems-magnification at low voltage capacitors," in *Proc. of IEEE Transmission and Distribution Conference*, Sep. 1991.

[124] V. E. Wagner, J. P. Staniak, and T. L. Orloff, "Utility capacitor switching and adjustable-speed drives," *IEEE Trans. on Industry Applications*, vol. 27, no. 4, pp. 645–651, 1991.

[125] G. Olivier, I. Mougharbel, and G. Dobson-Mack, "Minimal transient switching of capacitors," *IEEE Trans. on power delivery*, vol. 8, no. 4, pp. 1988–1994, 1993.

[126] D. V. Coury, C. J. Dos Santos, M. Oleskovicz, and M. C. Tavares, "Transient analysis concerning capacitor bank switching in a distribution system," *Electric Power Systems Research*, vol. 65, no. 1, pp. 13–21, 2003.

[127] R. Aradhya, S. Subash, and K. Meera, "Evaluation of switching concerns related to shunt capacitor bank installations," in *IPST95-International Conference On Power System Transients*, 1995, pp. 3–7.

[128] V. Borozan, M. E. Baran, and D. Novosel, "Integrated volt/var control in distribution systems," in *Proc. of IEEE Power Engineering Society Winter Meeting*, Jan. 2001.

[129] G. Ozdemir, S. Emiroglu, and M. Baran, "Supervisory control for coordinating volt/var control devices on a distribution system," in *Proc. of IEEE Power Energy Society ISGT*, Sep. 2016.

[130] E. C. Kara, C. M. Roberts, M. Tabone, L. Alvarez, D. S. Callaway, and E. M. Stewart, "Towards real-time estimation of solar generation from micro-synchrophasor measurements," *arXiv preprint arXiv:1607.02919*, 2016.

[131] S. Brahma and A. Girgis, "Development of adaptive protection scheme for distribution systems with high penetration of distributed generation," *IEEE Trans. on power delivery*, vol. 19, no. 1, pp. 56–63, 2004.

[132] M.-H. Kim, S.-H. Lim, and J.-C. Kim, "Improvement of recloser-fuse operations and coordination in a power distribution system with SFCL," *IEEE Trans. on Applied Superconductivity*, vol. 21, no. 3, pp. 2209–2212, 2011.

[133] S. Chaitusaney and A. Yokoyama, "Prevention of reliability degradation from recloser–fuse miscoordination due to distributed generation," *IEEE Trans. on Power Delivery*, vol. 23, no. 4, pp. 2545–2554, Oct. 2008.

[134] P. H. Shah and B. R. Bhalja, "New adaptive digital relaying scheme to tackle recloser-fuse miscoordination during distributed generation interconnections," *IET Generation, Transmission & Distribution*, vol. 8, no. 4, pp. 682–688, 2014.

[135] E. Reihani, M. Motalleb, R. Ghorbani, and L. S. Saoud, "Load peak shaving and power smoothing of a distribution grid with high renewable energy penetration," *Renewable Energy*, vol. 86, pp. 1372–1379, 2016.

[136] A. Eshraghi and R. Ghorbani, "Islanding detection and over voltage mitigation using controllable loads," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 125–135, 2016.

[137] M. Liserre, A. Pigazo, A. Dell'Aquila, and V. M. Moreno, "An anti-islanding method for single-phase inverters based on a grid voltage sensorless control," *IEEE Trans. on Industrial Electronics*, vol. 53, no. 5, pp. 1418–1426, 2006.

[138] M. Paolone, C. A. Nucci, E. Petrache, and F. Rachidi, "Mitigation of lightning-induced overvoltages in medium voltage distribution lines by means of periodical grounding of shielding wires and of surge arresters: modeling and experimental validation," *IEEE Trans. on Power Delivery*, vol. 19, no. 1, pp. 423–431, 2004.

[139] F. Mahmood, N. A. Sabiha, and M. Lehtonen, "Probabilistic risk assessment of mv insulator flashover under combined ac and lightning-induced overvoltages," *IEEE Trans. on Power Delivery*, vol. 30, no. 4, pp. 1880–1888, 2015.

[140] R. Cabral, D. Gazzana, R. Leborgne, A. Bretas, G. Dias, and M. Telló, "Analysis of distribution lines performance against lightning using ATP-EMTP," in *Proc. of the IEEE International Symposium on Electromagnetic Compatibility (EMC EUROPE)*, Italy, Dec. 2012.

[141] J. A. Martinez and F. Gonzalez-Molina, "Statistical evaluation of lightning overvoltages on overhead distribution lines using neural networks," *IEEE Trans. on power delivery*, vol. 20, no. 3, pp. 2219–2226, 2005.

[142] M. Lat and J. Kortschinski, "Application guide for surge arresters and field research of lightning effects on distribution systems," in *IET 10th International Conference on Electricity Distribution*, Brighton, UK, May 1989.

[143] M. Araujo, R. Flauzino, O. Batista, L. Moraes, and C. Martins, "Protection of the distribution lines with distributed generation against lightning overvoltages in the context of smart grids," in *Proc. of the IEEE World Congress on Sustainable Technologies (WCST)*, London, UK, Mar. 2013.

[144] Y. M. Hernández, D. Ioannidis, G. Ferlas, E. G. T. Tsovilis, Z. Politis, and K. Samaras, "An experimental approach of the transient effects of lightning currents on the overvoltage protection system in mw-class photovoltaic plants," in *Proc. of the IEEE International Conference on Lightning Protection (ICLP)*, Shanghai, China, Oct. 2014.

[145] Z. Benesova, R. Haller, J. Birkl, and P. Zahlmann, "Overvoltages in photovoltaic systems induced by lightning strikes," in *Proc. of the IEEE International Conference on Lightning Protection (ICLP)*, Vienna, Austria, Sep. 2012.

[146] M. Kezunovic, T. Djokic, P.-C. Chen, and V. Malbasa, "Improved transmission line fault location using automated correlation of big data from lightning strikes and fault-induced traveling waves," in *Proc. of the IEEE 48th Hawaii International Conference on System Sciences (HICSS),*, Kauai, HI, USA, Jan. 2015.

[147] Y. Wan, "Synchronized phasor data for analyzing wind power plant dynamic behavior and model validation," *Contract*, vol. 303, pp. 275–3000, 2013.

[148] B. McCamish, R. Meier, J. Landford, R. Bass, E. Cotilla-Sanchez, and D. Chiu, "A data driven framework for real time power system event detection and visualization," *arXiv preprint arXiv:1501.04038*, 2014.

[149] J. H. Eto, E. Stewart, T. Smith, M. Buckner, H. Kirkham, F. Tuffner, and D. Schoenwald, “Scoping study on research and development priorities for distribution-system phasor measurement units,” 2015.

[150] [Online]. Available: https://www.wunderground.com/history/

[151] M. Motalleb, M. Thornton, E. Reihani, and R. Ghorbani, “A nascent market for contingency reserve services using demand response,” *Applied Energy*, vol. 179, pp. 985–995, 2016.

[152] “IEEE standard for high-accuracy instrument transformers,” *IEEE Std C57.13.6-2005*, 2005.

[153] R. Bayless, J. Selman, D. Truax, and W. Reid, “Capacitor switching and transformer transients,” *IEEE Trans. on power delivery*, vol. 3, no. 1, pp. 349–357, 1988.

[154] E. M. Stewart, S. Kiliccote, C. McParland, C. Roberts, R. Arghandeh, and A. von Meier, “Using micro-synchrophasor data for advanced distribution grid planning and operations analysis,” *Ernest Orlando Lawrence Berkely National Laboratory*, 2014.

[155] S. F. Alwash, V. K. Ramachandaramurthy, and N. Mithulananthan, “Fault-location scheme for power distribution system with distributed generation,” *IEEE Trans. on Power Delivery*, vol. 30, no. 3, pp. 1187–1195, 2015.

[156] S. M. Brahma and A. A. Girgis, “Development of adaptive protection scheme for distribution systems with high penetration of distributed generation,” *IEEE Trans. on power delivery*, vol. 19, no. 1, pp. 56–63, 2004.

[157] V. John, Z. Ye, and A. Kolwalkar, “Investigation of anti-islanding protection of power converter based distributed generators using frequency domain analysis,” *IEEE Trans. on Power Electronics*, vol. 19, no. 5, pp. 1177–1183, 2004.

[158] M. Jamei, E. Stewart, S. Peisert, A. Scaglione, C. McParland, C. Roberts, and A. McEachern, “Micro synchrophasor-based intrusion detection in automated distribution systems: Towards critical infrastructure security,” *IEEE Internet Computing*, vol. 20, no. 5, 2016.

[159] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, “Automated anomaly detection in distribution grids using upmu measurements,” *arXiv preprint arXiv:1610.01107*, 2016.

[160] E. C. Kara, C. M. Roberts, M. Tabone, L. Alvarez, D. S. Callaway, and E. M. Stewart, “Towards real-time estimation of solar generation from micro-synchrophasor measurements,” *arXiv preprint arXiv:1607.02919*, 2016.

[161] M. Jamei, A. Scaglione, C. Roberts, A. McEachern, E. Stewart, S. Peisert, and C. McParland, “Online thevenin parameter tracking using synchrophasor data,” in *Proc. of the IEEE PES General Meeting*, Chicago, IL, Jul. 2017.

[162] Z. Taylor, H. A.-Hejazi, E. Cortez, L. Alvarez, S. Ula, M. Barth, and H. Mohsenian-Rad, "Battery-assisted distribution feeder peak load reduction: Stochastic optimization and utility-scale implementation," in *Proc. of IEEE PES General Meeting*, Boston, MA, Jul. 2016.

[163] C. Shand, A. McMorran, E. Stewart, and G. Taylor, "Exploiting massive pmu data analysis for lv distribution network model validation," in *Proc. of IEEE Power Engineering Conference (UPEC)*, Stoke on Trent, UK, Dec. 2015.

[164] C. Efthymiou and G. Kalogridis, "Applications of the windowed FFT to electric power quality assessment," in *Proc. of the IEEE International Conference on Smart Grid Communications*, Gaithersburg, MD, Oct. 2010.

[165] A. L. Liao, E. M. Stewart, and E. C. Kara, "Micro-synchrophasor data for diagnosis of transmission and distribution level events," in *Proc. of the IEEE Transmission and Distribution Conference and Exposition (T&D)*, Dallas, TX, USA, Jul. 2016.

[166] A. G. Phadke and J. S. Thorp, *Synchronized phasor measurements and their applications.* Springer Science & Business Media, 2008.