The Dissertation Committee for Kyung Woo Min
certifies that this is the approved version of the following dissertation:

Data Analytics Applications to Fault Locations and Overcurrent Protection Devices

Committee:

Surya Santoso, Supervisor

Ross Baldick

Gary Hallock

Aristotle Arapostathis

Unmil Karadkar

Data Analytics Applications to Fault Locations and Overcurrent Protection Devices

by

Kyung Woo Min

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2019

Dedicated to my parents.

Acknowledgments

First and foremost, I thank my advisor, Dr. Surya Santoso, for his support and mindful guidance throughout my Ph.D. studies. He gave me an incredible opportunity to join his group and pursue research in power systems. I am very fortunate and grateful to be his student. Without his help, I would not have been able to finish this dissertation.

I am also grateful to Dr. Ross Baldick, Dr. Gary Hallock, Dr. Aristotle Arapostathis, and Dr. Unmil Karadkar for kindly agreeing to serve on my dissertation committee. Their comments and feedback have been valuable in conducting my research and finishing my dissertation.

Thanks as well to my colleagues in Dr. Santoso's research group: Min Lwin, Tuan Ngo, Pisitpol Chirapongsananurak, Harsha V. Padullaparti, Suma Jothibasu, Swagata Das, Anamika Dubey, Yichuan Niu, Quan Nguyen, Alvaro F. Bastos, Sundaravaradan N. Ananthan, Hunter Estes, David Rosewater, and Tae Hyung Kim. I learned a lot from our discussions. I must thank my Korean friends in Austin, with whom I share great memories. I really enjoyed meeting my friends and spending time together.

Finally, I thank my parents for their unconditional love and support. Their love gave me strength to overcome difficult situations and successfully finish my dissertation.

Data Analytics Applications to Fault Locations and Overcurrent Protection Devices

Publication No. _____

Kyung Woo Min, Ph.D.
The University of Texas at Austin, 2019

Supervisor: Surya Santoso

Power quality (PQ) monitors installed in transmission and distribution systems record disturbance events occurring in the system, such as root mean square (RMS) variations and transients caused by short-circuit faults, transformer energizing, or capacitor switching around the clock, resulting in a large amount of data. Although the collected data contain valuable information about the system, they are often merely stored without any further analysis. Analysis of these data presents opportunities for improving the performance of power systems as well as for monitoring the health of the grid as a whole. The general objective of this proposal is to develop algorithms that make use of three phase voltage and current measurements recorded from the power quality monitors. Specifically, algorithms are developed for the analysis of (1) short circuit faults with their locations (fault analytics) and (2) overcurrent protection devices installed in the system (device analytics). The fault analytics module is used to identify fault events among other power quality events

and estimates the location to the fault occurring in the system. The proposed method uses variable window size in calculating phasors and estimates a single fault location that is more accurate than the multiple locations estimated by the conventional approach using Fourier and cosine filters. The device analytics module aims to evaluate the overcurrent protection devices operating to isolate short-circuit faults from the system. This module identifies recloser and fuse operations and estimates the empirical inverse time-current characteristics of the devices. The results of the device analytics are used to evaluate device opening intervals and coordination and to further narrow down fault location because faults are located downstream from the clearing device. Finally, the dissertation presents a data analytics framework and an open power quality disturbance event schema. The schema is developed to promote the sharing of data recording PQ disturbance events and the metadata providing descriptive and quantitative analysis of the events.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Objective

Power quality (PQ) monitors and intelligent electronic devices such as digital relays and digital fault recorders collect a large amount of data, providing the foundation for detailed analysis of power system disturbances. Knowledge extracted from the data also provides insights to help understand the power system conditions and how to prevent possible disturbance events from recurring. For example, the voltage and current measurement data collected during capacity switching operations are used to estimate system parameters such as damping factors and resonant frequencies. Authors in [1, 2, 3] present signal processing techniques such as Hilbert and wavelet transforms to estimate the magnitude and time constants of the system damping factor.

Similarly, fault events captured by power quality monitors and digital relays have been used in various perspectives for enhancing the reliability of the system operation. Most fault-locating algorithms use voltage and current recorded during short-circuit faults. Since the relationship between the voltage and current during a fault event is defined by the circuit model, advanced algorithms are proposed to minimize the error estimates coming from uncertain

parameters in the system. The sources of errors may include inaccurate phasor estimation, fault resistance, and line impedance [4, 5, 6, 7]. The fault records can be used to estimate circuit parameters such as line impedance [8] and fault resistance [6, 9]. These algorithms use measurements captured from one or both ends of the power line and are used to evaluate predefined parameters and zone settings within digital distance relays.

The objective of this dissertation is to develop data analytics tools that provide analysis of the short-circuit faults and the overcurrent protection devices clearing the fault. The raw input datasets used for these applications are three phase instantaneous voltage and current waveforms captured by power quality monitors, containing a wide variety of power quality disturbance events. Data analytics tools presented in this dissertation examine the raw datasets collected from power quality monitors, extract knowledge, and provide actionable insights. Utility operators can make use of the fault analytics tool to accurately detect and locate the fault in the system to expedite service restoration and improve reliability. Also, analytics of the overcurrent devices can be used to identify the device clearing the fault, which also helps narrow down the fault locations; evaluate protection coordination and breaker opening intervals; and monitor the health conditions of the devices installed in the system.

The dissertation is organized as follows. Chapter 2 presents fault analytics: detection and location of short-circuit faults. Chapter 3 provides device analytics: evaluation and identification of overcurrent protection devices. Fi-

nally, Chapter 4 proposes a power quality data analytics framework and a database schema, power quality linked data (PQLD).

1.2 Contribution

1.2.1 DC Offset Removal Algorithm for Locating Momentary Faults

The fault analytics module presents a DC offset removal algorithm to improve the fault location estimates of momentary faults. For a momentary fault, the data are limited because of its short duration, and the phasor estimation is complicated by the exponentially decaying DC offsets. The fault analytics module accurately detects fault inception and clearing times, estimates the voltage and current phasors during short-circuit conditions, then provides the input voltage and current phasors to existing impedance-based fault location algorithms, such as the Takagi method [5, 10]. The algorithm uses the RMS-wavelet method for fault detection and estimates voltage and current phasors using the nonlinear least squares algorithm. The proposed method uses a variable window size in calculating phasors and estimates a single fault location that is more accurate than the multiple locations estimated by the Fourier and cosine filters. The method is validated using both simulated and field data.

For the bolted fault, the test results showed that the proposed method provides a vast improvement over the Fourier filter. The proposed method determined a location that was more than 10% closer to the actual location than the worst-case estimate given by the Fourier filter. When there is a

fault resistance, the fault location accuracy of the proposed method does not significantly improve because of other sources of error, such as load current and system nonhomogeneity. Nonetheless, the proposed method can avoid scenarios where phasor estimations contribute additional errors. In the test case, the worst error using the Fourier and cosine filters was -16.72% and -4.76%, respectively, and the worst error using the proposed method was 2.23%. The test also showed that the proposed method can be applied in fault events where the duration is less than a cycle. The cosine filter requires the greatest number of data points, one and a quarter cycles, and therefore resulted in a large error of 60%. The proposed method improved the location accuracy by more than 58%, and the error was only -1.67%. The algorithm is described in Chapter 2 and was published in [7].

1.2.2 Identification and Evaluation of Overcurrent Protection Devices

The device analytics module presents algorithms to identify and evaluate overcurrent protection devices (recloser and fuse) clearing the fault. The contribution consists of two parts. First, the rule-based method is presented to identify the type of device clearing the fault (recloser or fuse). Necessary features that characterize the fault-clearing devices are calculated from the input voltage and current waveforms. The rule-based expert system then uses these features to determine whether the fault was cleared by a recloser or a fuse.

Second, the methodology for estimating empirical inverse time-current characteristics (TCC) curve of the recloser is proposed. Data preprocessing techniques are presented to estimate the current flowing through the recloser and to detect the fault inception and clearing times related to the breaker operating time and delay. Then, a nonlinear least squares algorithm is formulated to estimate the TCC curve parameters of the reclosers clearing the fault. The estimated TCC parameters are used to construct an empirical TCC curve. The empirical TCC curve is then used to evaluate opening intervals of breakers and identify the device (TCC curve) clearing the fault.

The proposed algorithms require only the voltage and current measurements from the substation. No other information, such as the circuit model or load allocation, is required. As opposed to the method presented in [11], where prior knowledge of the TCC curves is assumed, the approach does not require the TCC curve information because the curves are automatically constructed from the field data. The efficacy of the algorithms is validated using simulated data, event reports generated from a digital relay test bench, and field events collected from 24.9 kilovolt (kV) distribution circuits. The algorithm for identifying the type of device is published in [12]. The algorithm developed for estimating the empirical TCC curve is in preparation for submission.

1.2.3 Data Analytics Framework for Power Quality Disturbance Events

Although power quality (PQ) disturbance events such as RMS variations and transients occur in transmission and distribution systems, these datasets are mostly managed through proprietary solutions in different data formats. In this dissertation, an open schema, power quality linked data (PQLD), is presented to manage voluminous PQ disturbance events in power systems. The schema promotes sharing of power quality data and combines analyses of data from multiple sources. PQLD takes an incremental approach to data publishing and can be implemented using freeware software such as MongoDB and Python.

The proposed schema defines five classes to store the metadata associated with PQ disturbance events: event, time-series, description, software analysis, and Institute of Electrical and Electronics Engineers (IEEE) disturbance classification. The classes that form the schema are presented and then demonstrated using actual disturbance events captured from a distribution system. This work is published in [13].

1.3 Literature Review

1.3.1 DC Offset Removal Algorithm for Locating Momentary Faults

Several improved methods have been presented to eliminate the effect of the DC offset. The digital mimic filter proposed in [14] is a type of high-pass filter implemented in combination with the conventional full-cycle or half-cycle

discrete Fourier transform (DFT) filter. The low-frequency DC offset terms are suppressed using the predetermined range of the time constants. Other research in [15, 16, 17, 18, 19] derives the phasor at the fundamental frequency by estimating parameters associated with the DC offsets. Authors in [15] investigate the DC offset removal algorithm that requires one cycle or half a cycle plus two sample points. The method takes advantage of the two additional sample points and uses their DFT coefficients to derive the parameters associated with the DC offset. [16] takes a similar approach, where the DC offsets are estimated by subtracting odd-sample-set DFT coefficients from the even-sample-set DFT coefficients. In [18] and [19], one-cycle averages of the fault current waveform are taken to eliminate sinusoidal components and estimate the DC component. Authors in [20, 21, 22] use the least squares approach, where multiple time-domain current waveforms containing DC offsets are used to estimate the DC offset parameters and the phasors at the fundamental frequency. [17] presents three simplified methods that approximate the exponential DC offset to a simpler form to reduce computation burden. [20] linearizes the exponential decaying DC offset by using a Taylor series before implementing the least squares filter in a recursive manner.

The methods described above have been developed primarily for implementation in real-time protective relaying applications. The accuracy of the phasor estimation methods must be balanced with the relay's objective of detecting a fault as quickly as possible. Therefore, the algorithms are limited by the computation time and the speed of switching operation. For the pur-

pose of fault location, there is less restriction on the computation time; it can be done in seconds or minutes [4] as opposed to milliseconds (ms) or cycles for protective relaying applications. Thus, more accurate methods can be employed to reject DC offset and improve estimation results. The method used in the fault analytics module consists of exact detection of the fault inception and clearing times using an RMS-wavelet method and phasor estimations using the nonlinear least squares method. As a result, the algorithm provides more accurate fault location estimates.

1.3.2 Identification and Evaluation of Overcurrent Protection Devices

Very little research has been done on identifying the type of fault clearing devices in distribution systems. In [11], estimated fault current and duration are compared with the fault clearing time in the TCC curves of the protective devices. Although the approach is successful in most cases, the method requires prior knowledge of the TCC curves for all protective devices. In addition, as the number of protective devices increases, having various TCC curves reduces the margin between the decision boundaries. Because [11] makes a few assumptions in estimating fault current and fault duration, the classification accuracy is affected. The approach presented in this dissertation uses features derived from the input measurements and identifies the device by using an algorithm that emulates the decision-making process of those with expertise in power systems [23]. Since the identification is rule-based, it is scalable, debuggable, and highly interpretable compared with other complex classifiers

that use neural networks [24, 25] or support vector machines [26, 27].

In [28], an adaptive neuro fuzzy inference system was used to model the desired TCC curve. This dissertation proposes to estimate the TCC curve parameters of an overcurrent relay/reclosers installed and operating in distribution systems, based on IEEE Standard C37.112-2018 [29]. Since the proposed algorithm is based on the IEEE standard, the applications are easy to interpret and debug. Data preprocessing is also presented, where the fault clearing times and load currents are estimated to formulate the estimation algorithm. The data used are time series voltage and current data captured from a single monitoring location at the substation.

This dissertation presents two possible applications using the empirical TCC curve: evaluating opening intervals of breakers and identifying the device (TCC curve) clearing the fault. Existing algorithms for evaluating and monitoring breaker performances include using current coil data [30, 31] and analyzing event reports generated from digital relays [32]. In [12], rules are developed using extracted features such as real and reactive power differences to identify whether a recloser or fuse has cleared the fault. This approach can be used with existing algorithms to further narrow down the specific device and operating curve that clear the fault.

1.3.3 Data Analytics Framework for Power Quality Disturbance Events

Utility companies install and manage their own database software, relying on unpublished schemas, and have each developed proprietary–privately owned and controlled–software for accessing data [33] containing power quality disturbance events. Such proprietary storage limits the opportunities to share data and combine analyses of data from multiple sources, resulting in diminished ability to assess the robustness of large power grids. Recognizing such shortcomings, power companies have developed additional proprietary applications and adopted standards such as the IEC 61970 and IEC 61968, known as the Common Information Model (CIM). CIM defines class attributes and relationships in the Unified Modeling Language (UML). These semantic models are used to describe power system components, customer billing, and electricity markets and are expressed using Extensible Markup Language (XML) and Resource Description Frame (RDF) formats [34] for data exchange. Recent works [35, 36] propose using a CIM-oriented graph database framework to efficiently retrieve and store largely connected datasets in power systems.

The use of proprietary standards has limited the availability of PQ data to external researchers. Some data has been made publicly available through efforts such as the DOE/EPRI National Database Repository of Power System Events [37]. To further such initiatives and to promote the sharing of PQ data, power quality linked data (PQLD)–an open, documented database schema–is presented to catalog and describe PQ disturbance events. The schema is based

on open standards such as Linked Open Data (LOD) [38] and implemented using freeware software such as MongoDB and Python. While security-related arguments have been advanced in the past to promote proprietary software, counterarguments from the open source perspective have gained prominence in the past decade [39]. Swire presents an insightful cost-benefit analysis for assessing the impact of openness and security [40].

Chapter 2

Fault Detection and Location

2.1 Introduction

[1] Electrical faults refer to any abnormal current condition in power systems. These faults involve short-circuit conditions between phase conductor(s) and ground or between two or more conductors. For example, an untrimmed tree can touch uninsulated conductors. A conductor may come in contact with another conductor during stormy weather conditions. The most common fault type is single line-to-ground faults, covering 70% to 80% percent of all fault events, followed by line-to-line faults and three-phase faults [41].

Fault currents can cause damage to equipment and devices installed in a power system [41]. When a fault remains in the system for too long and the fault magnitude is high, the heat can damage equipment devices such as transformers, conductors, and capacitors. In addition, the system voltage level deviates from its nominal value during fault conditions because of the

[1]Parts of this chapter have been published in K. W. Min and S. Santoso, "DC offset removal algorithm for improving location estimates of momentary faults," *IEEE Trans. on Smart Grid,* vol. 9, no. 6, pp. 5503-5511, Nov. 2018 and K. W. Min, S. Santoso, and L. Biyikli, "Identifying fault clearing operations in distribution systems," in *Proc. IEEE Power Energy Soc. General Meeting*, July 2016, pp. 1-5. The author of this dissertation analyzed the data, developed the algorithms, and validated the analytical results in the papers.

fault current and the system equivalent impedance. The deviations can also cause malfunctions in voltage-sensitive loads such as motor drives. Typically, protection equipment devices are installed in the systems to detect and isolate the faults for these reasons. Most protective devices isolate faults based on inverse time-current characteristics, clearing the fault faster for higher-fault current. Fig. 2.1 illustrates a single line-to-ground fault with the magnitude of 530 amperes (A), which lasted for 12 cycles until a protective device cleared the fault.

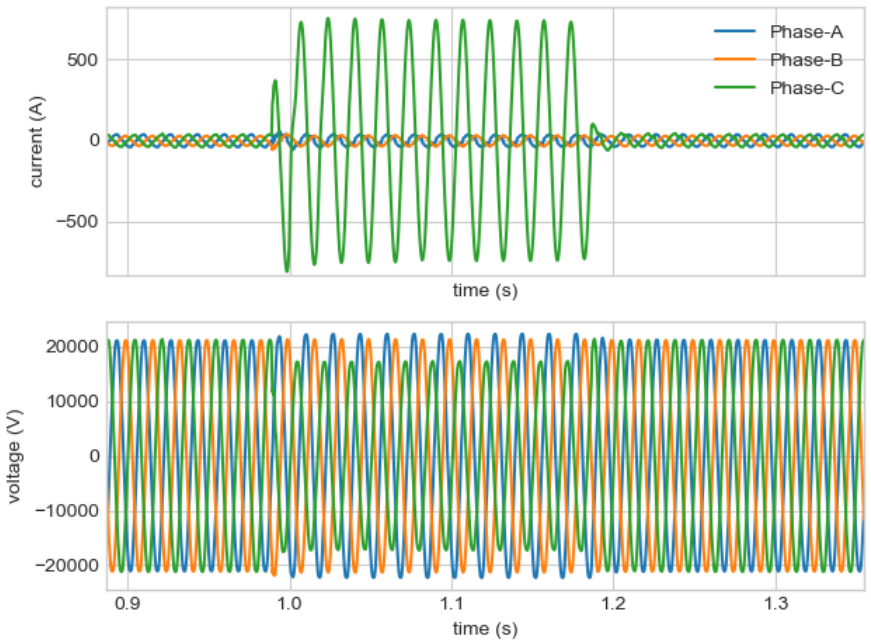


Figure 2.1: Fault current RMS (above) and instantaneous (below).

In this chapter, the data analytics algorithms for analyzing short-circuit faults are presented. These algorithms are implemented and automated to

detect and locate short-circuit conditions in power systems. The inputs of the fault analytics module are instantaneous phase A, B, and C voltage and current waveforms.

First, methods for estimating phasor inputs are described. The phasors are then used to calculate necessary features for the short-circuit analysis, such as RMS voltage and current variations, fault magnitude and duration, and the real and reactive power demands in the circuit.

Next, fault detection methods are described. The approach includes using pickup thresholds in RMS voltage and current and filtering out inrush currents, which could be mistakenly considered as fault events. In addition, the discrete wavelet transform is used to accurately detect the fault inception and the clearing times in the fault event recordings.

Finally, the fault-locating algorithm is proposed. The algorithm uses the RMS-wavelet method for fault detection and estimates voltage and current phasors using a nonlinear least squares algorithm. The application results show improved location estimates over the conventional methods, especially for momentary faults where the exponentially decaying DC offset has not fully decayed in the voltage and current measurements.

Short-circuit faults are closely related to protection device operations because the fault currents severely damage the system if not isolated as quickly as possible. The data analytics to identify and evaluate the operations of protection devices are described in Chapter 3.

2.2 Phasor Conversion

This section describes three methods that are used to estimate phasors from the instantaneous waveforms: the Fourier filter, the cosine filter, and the least squares algorithm. Each of these phasor computing methods is described in the following sections. Before analyzing the data files, the input voltage and current waveforms are resampled so the sampling frequencies are an integer multiple of 60 hertz (Hz). Typical target sampling frequencies are 16, 32, 64, 128, or 256 samples per cycle. For example, the raw datasets used in Chapter 3 have a sampling frequency of 500,000 samples per 30 seconds (sec), which corresponds to 277.78 samples per cycle. The datasets are resampled to 256 samples per cycle, or 15.36 kilohertz (kHz).

2.2.1 Fourier Filter

Fourier filters are finite impulse response filters whose coefficients are derived by sampling a cosine and a sine wave [42]. The filter is used to filter harmonics and DC offsets and calculate the phasor at the fundamental frequency. The filter output at time sample m is expressed as

$$i_{out,cosine}(m) = \frac{\sqrt{2}}{N} \sum_{k=1}^{N} i_{in}(m - N + k) \times \cos(\frac{2\pi(k-1)}{N}) \tag{2.1}$$

$$i_{out,sine}(m) = \frac{\sqrt{2}}{N} \sum_{k=1}^{N} i_{in}(m - N + k) \times \sin(\frac{2\pi(k-1)}{N}) \tag{2.2}$$

where N is the number of samples per cycle.

The fundamental phasor I_{cosine} can be calculated from the two samples, $i_{out,cosine}(m)$ and $i_{out,sine}(m)$ as in (2.3).

$$I_{fourier} = i_{out,cosine}(m) + ji_{out,sin}(m) \tag{2.3}$$

The phasor magnitude and the phasor angle are calculated as

$$|I_{fourier}| = \sqrt{i_{out,cosine}(m)^2 + i_{out,sine}(m)} \tag{2.4}$$

$$\angle I_{fourier} = \arctan \frac{i_{out,cosine}(m)}{i_{out,sine}(m)} \tag{2.5}$$

2.2.2 Cosine Filter

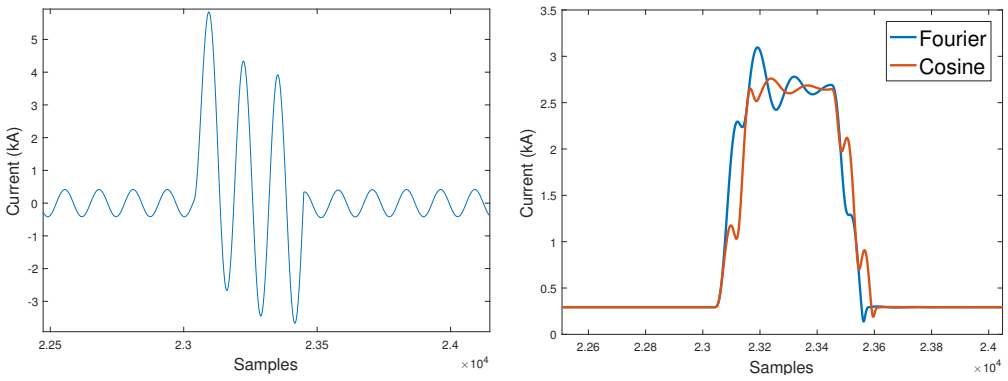
Cosine filters, popularly implemented in commercial relays, are finite impulse response filters whose coefficients are derived by sampling a cosine wave. Similarly to the Fourier filter, the cosine filter is used to filter harmonics and DC offsets and calculate the phasor at the fundamental frequency, but it often shows improved performance over the Fourier filter in phasor estimations [42]. The filter output at time sample m is expressed as

$$i_{out}(m) = \frac{\sqrt{2}}{N} \sum_{k=1}^{N} i_{in}(m - N + k) \times \cos(\frac{2\pi(k-1)}{N}) \tag{2.6}$$

where N is the number of samples per cycle.

The fundamental phasor I_{cosine} can be calculated from the two samples, $i_{out}(m)$ and $i_{out}(m - N/4)$ as in (2.7).

$$I_{cosine} = i_{out}(m) + ji_{out}(m - \frac{N}{4}) \tag{2.7}$$



(a) (b)

Figure 2.2: a) Input fault waveforms, b) RMS current calculated from Fourier and cosine filters

The phasor magnitude and the phasor angle are calculated as

$$|I_{cosine}| = \sqrt{i_{out}(m)^2 + i_{out}(m - \frac{N}{4})^2} \tag{2.8}$$

$$\angle I_{cosine} = \arctan \frac{i_{out}(m)}{i_{out}(m - \frac{N}{4})} \tag{2.9}$$

In contrast to the Fourier filter, both the real and the imaginary parts of the phasors are estimated by multiplying the input voltage and current measurements by the coefficients of a cosine function.

The Fourier and cosine filters assume the input signal is periodic. Therefore, these two filters cannot fully remove nonperiodic signals such as exponentially decaying DC offsets. For example, when we calculate the RMS current magnitudes to the instantaneous current samples in Fig. 2.2(a), the RMS outputs depicted in Fig. 2.2(b) fluctuate because the DC offset makes the input waveform aperiodic.

2.2.3 Least Squares Algorithm

Least squares algorithms fit multiple measurement points to a predefined function, which includes the sinusoidal function at the fundamental frequency. Other components such as an exponential decaying DC offset or the harmonics can be added to the predefined function. Any components that are not modeled in the predefined function are considered as noise. For example, during a short-circuit fault, the fault current consists of a sinusoidal at the fundamental frequency plus an exponential decaying DC offset. Therefore, defining fault current as $\beta_1 \cos(2\pi f t_i + \beta_2) + \beta_3 e^{-\beta_4 t_i}$, a least squares algorithm can be used to estimate the parameters (β_1, β_2, β_3, β_4) using the measurements y_i and the following quadratic loss function.

$$\begin{aligned} \underset{\beta_1,\beta_2,\beta_3,\beta_4}{\text{minimize}} \quad & \sum_i \left[y_i - \beta_1 \cos(2\pi f t_i + \beta_2) - \beta_3 e^{-\beta_4 t_i} \right]^2 \\ \text{subject to} \quad & \beta_4 > 0 \end{aligned} \tag{2.10}$$

The output parameters, β_1 and β_2, are then used in estimating the phasors at the fundamental frequency during the fault as (2.11), whereas parameters β_3 and β_4 represent the magnitude and the decaying constant of the DC offset.

$$I = \frac{\beta_1}{\sqrt{2}} \angle \beta_2 \tag{2.11}$$

Fig. 2.3 illustrates the fitted curve ($\beta_1 \cos(2\pi f t_i + \beta_2) + \beta_3 e^{-\beta_4 t_i}$), DC offset ($\beta_3 e^{-\beta_4 t_i}$), and symmetrical sinusoidal ($\beta_1 \cos(2\pi f t_i + \beta_2)$) using the fault waveforms extracted from Fig. 2.2(a). As opposed to the Fourier and the cosine filters, the least squares algorithm estimates a single current magnitude

(2.656kA) after removing the DC offset. The trust-region-reflective algorithm [43, 44] can be used to solve the optimization problem defined in (2.10).

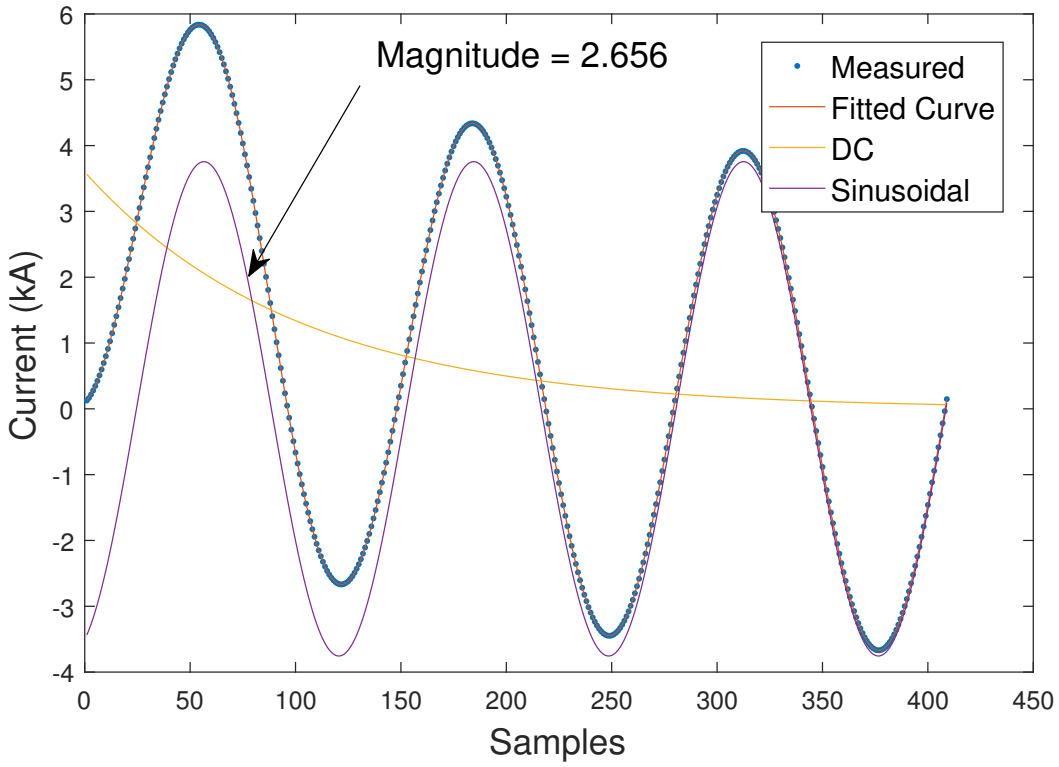


Figure 2.3: Fault current decomposed into exponentially decaying DC offset plus symmetrical sinusoidal.

2.3 Fault Detection, Type, and Relative Location

This section describes the methods for detecting fault events using thresholds in RMS voltage and current. The fault type (single line-to-ground, line-to-line, double line-to-ground, or three-phase fault) and the relative location (with respect to the monitoring location) are also estimated from this process. It is assumed the monitored circuit is radial.

First, voltage sag and interruption, or just voltage events, are detected using a threshold in RMS voltage. In IEEE Standard 1159-2009 [45], a voltage sag is defined as a drop in RMS voltage to between 0.1 and 0.9 per unit and a voltage interruption below 0.1 per unit. Therefore, 90% of the rated voltage is used as the voltage threshold to capture all voltage variation events as defined in [45]. Most of the voltage events are associated with faults, whether upstream or downstream of the monitoring location. Moreover, magnetizing or inrush currents from big loads such as motors and transformers may also cause the voltage level to decrease below 90%.

Similarly, a threshold in RMS current is used to detect faults occurring downstream from the monitoring location. This threshold should be higher than the maximum load current of the circuit. A current event is defined to have occurred when the RMS magnitude exceeds the current threshold value.

We can use the simple rules summarized in Tables 2.1 and 2.2 to categorize the type and relative location of the fault. The rules require both the current event and the voltage event to occur for a fault to be categorized as

a downstream fault. For an upstream fault, the current event should not be detected while the voltage event is detected. In addition, the zero-sequence current events are used to identify whether the fault involves the ground.

Table 2.1: Downstream fault categorization

Fault Type	Event Ia	Event Ib	Event Ic	Event Iz	Event Va	Event Vb	Event Vc
SLG A-ϕ	TRUE	–	–	TRUE	TRUE	FALSE	FALSE
SLG B-ϕ	–	TRUE	–	TRUE	FALSE	TRUE	FALSE
SLG C-ϕ	–	–	TRUE	TRUE	FALSE	FALSE	TRUE
LL AB-ϕ	TRUE	TRUE	–	FALSE	TRUE	TRUE	FALSE
LL BC-ϕ	–	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE
LL CA-ϕ	TRUE	–	TRUE	FALSE	TRUE	FALSE	TRUE
LLG AB-ϕ	TRUE	TRUE	–	TRUE	TRUE	TRUE	FALSE
LLG BC-ϕ	–	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
LLG CA-ϕ	TRUE	–	TRUE	TRUE	TRUE	FALSE	TRUE
LLL ABC-ϕ	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE
LLLG ABC-ϕ	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

Table 2.2: Upstream fault categorization

Fault Type	Event Ia	Event Ib	Event Ic	Event Iz	Event Va	Event Vb	Event Vc
SLG A-ϕ	FALSE	FALSE	FALSE	–	TRUE	FALSE	FALSE
SLG B-ϕ	FALSE	FALSE	FALSE	–	FALSE	TRUE	FALSE
SLG C-ϕ	FALSE	FALSE	FALSE	–	FALSE	FALSE	TRUE
LL(G) AB-ϕ	FALSE	FALSE	FALSE	–	TRUE	TRUE	FALSE
LL(G) BC-ϕ	FALSE	FALSE	FALSE	–	FALSE	TRUE	TRUE
LL(G) CA-ϕ	FALSE	FALSE	FALSE	–	TRUE	FALSE	TRUE
LLL(G) ABC-ϕ	FALSE	FALSE	FALSE	–	TRUE	TRUE	TRUE

Note that it is possible that a fault may occur without causing voltage events on any phases. This would not be categorized as a fault event according to the given rule. However, these types of faults are not of concern because they are not defined as disturbance events according to [45].

2.3.1 Inrush Current

In this section, two types of inrush currents–reenergizing inrush and voltage recovery inrush currents–are described. Inrush currents cause current and voltage events and can be misclassified as faults.

2.3.1.1 Reenergizing Inrush Current

When reclosers reenergize the circuit after a short interruption, magnetizing and inrush currents associated with the reenergizing of the transformers or big loads causes high current disturbances. These inrush currents gradually decrease as the load is successfully reenergized. An example waveform is shown in Fig. 2.4. The first disturbance is a fault, followed by the inrush event. The inrush can cause high current disturbances and exceed the current threshold. Therefore, the inrush event is separated from the fault event using the skewness of the current distributions.

Note that the high currents in inrush events gradually decrease as the load is successfully reenergized. On the other hand, the magnitude of the fault current is nearly consistent until it is cleared by the protective device. Therefore, the histograms of the two current disturbance events are compared to classify fault events from inrush events. The skewness of the inrush (usually a positive value) is greater than that in the fault event (usually a negative value). The implementation process is shown in algorithm 1, and the histograms are illustrated in Fig. 2.5 and Fig. 2.6.

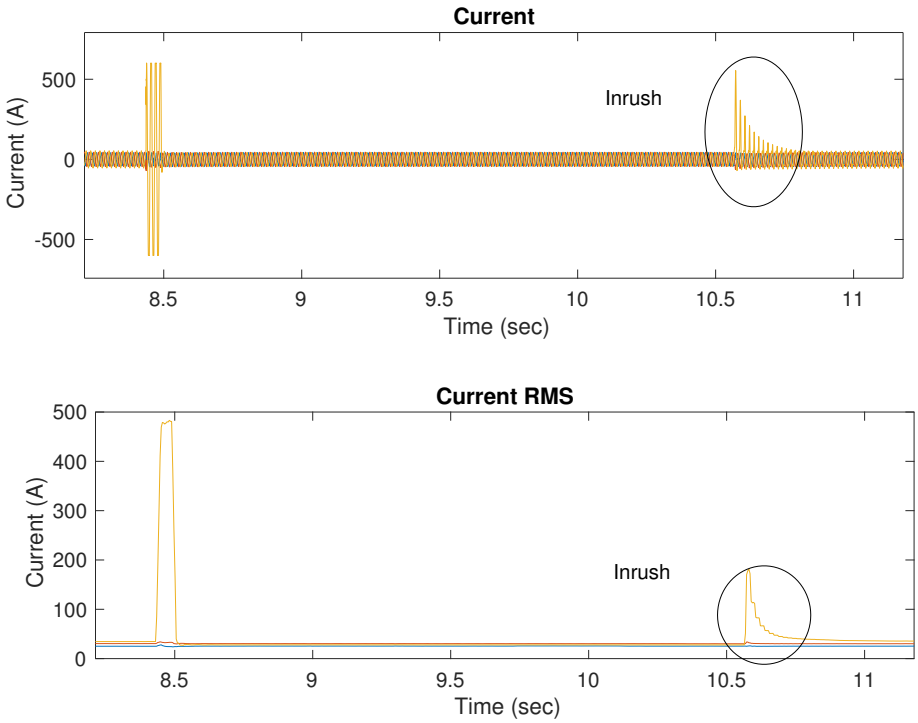


Figure 2.4: Fault event followed by successful load reenergizing.

Algorithm 1 Inrush Detection

1: **for** every detected current RMS event **do**
2: Normalize all values in the range of [0, 1]
3: Sample the portion in the range of $0.2 < i_{rms} < 1$
4: Calculate $\gamma_{event} = \frac{\frac{1}{n}\sum_{k=1}^{n}(i_k-\bar{i})^3}{(\sqrt{\frac{1}{n}\sum_{k=1}^{n}(i_k-\bar{i})^2})^3}$
5: **if** $\gamma_{event} < \theta$ **then**
6: Event is a fault
7: **else**
8: Event is an inrush
9: **end if**
10: **end for**

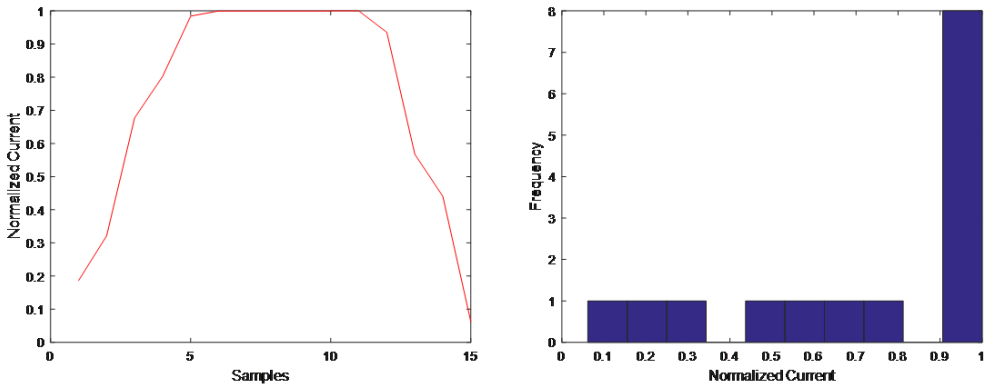


Figure 2.5: Fault current and its histogram.

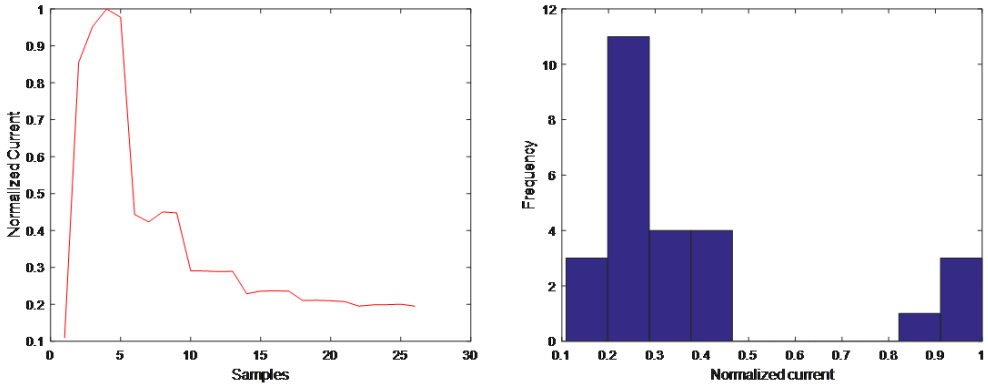


Figure 2.6: Inrush current and its histogram.

2.3.1.2 Voltage Recovery Inrush Current

Short-circuit faults cause voltage sags in the faulted phases. The voltage level returns to its nominal value when a protective device clears the fault. This recovery voltage causes transformer and motor inrush currents similar to those in the reenergizing scenario. The magnitude of the recovery inrush may exceed the threshold used to discriminate downstream and upstream faults. Fig. 2.7 shows a double line-to-ground fault on phase BC occurring upstream from the monitoring location. When the fault is cleared at time 25.09 sec, the magnitude of the recovery inrush current on phase B and phase C are 134.6 A and 211.7 A, respectively. The load currents on phase B and phase C before the fault are 41.31 A and 61.39 A, respectively.

Voltage recovery inrush current is detected using the same approach as for the reenergizing inrush current. The skewness parameter is used to identify the event.

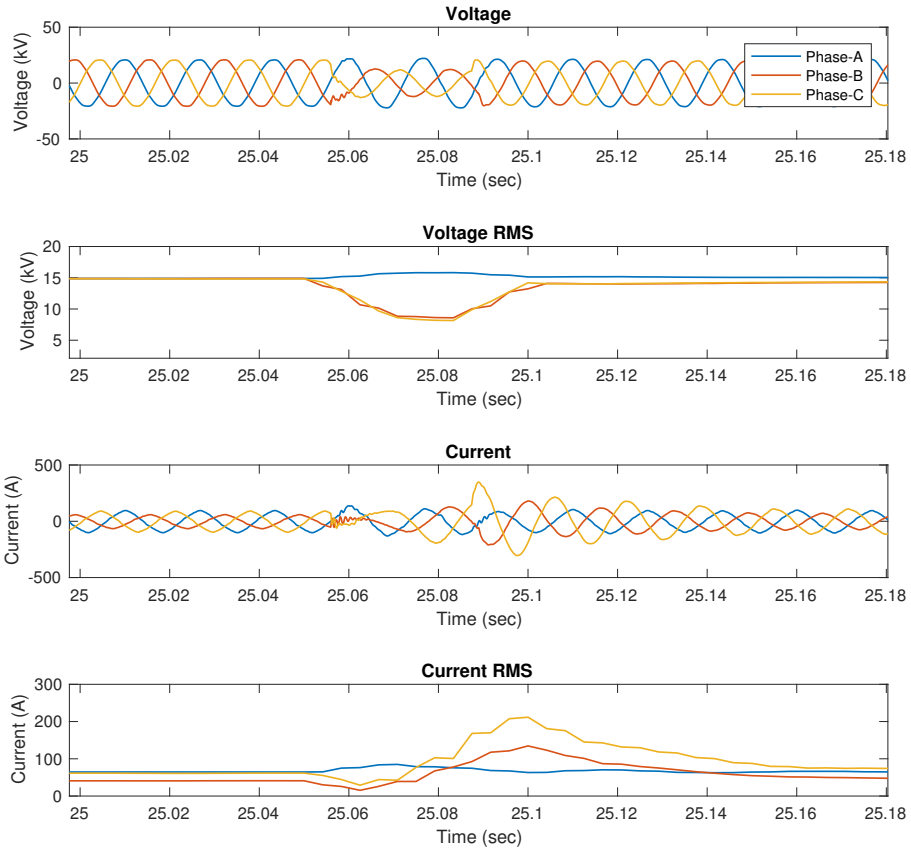


Figure 2.7: Voltage recovery inrush current exceeding three times the load current.

2.3.2 Fault Detection

In this section, two methods used to detect the fault inception time and the fault clearing times–RMS-based and RMS-wavelet methods–are presented.

2.3.2.1 RMS-Based Method

The simplest way to calculate fault inception and the clearing time is to have a pickup threshold and compare it with the RMS current magnitudes. The first and last samples exceeding the RMS current threshold are assumed to be the fault inception time and fault clearing time, respectively. These two indices are denoted as n_{ps} and n_{pe} throughout the chapter. Although this approach is affected by the window size of RMS calculation, the method can estimate the fault inception time and clearing time in any unexpected scenarios such as in the presence of current transformer (CT) saturation or evolving faults. Therefore, this method is preferred when the reliability of the estimation is an important issue. Fig. 2.8 shows the procedure.

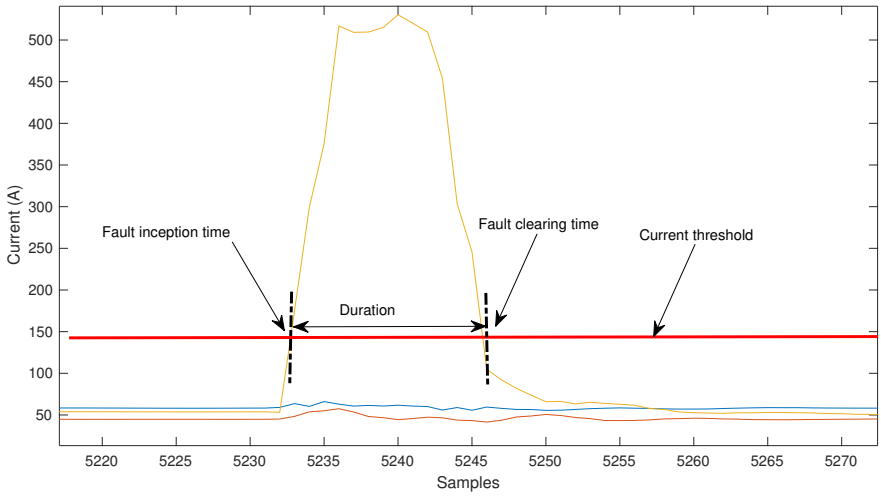


Figure 2.8: RMS-based fault duration calculation.

2.3.2.2 RMS-Wavelet Method

RMS-based fault detection is mostly successful in detecting the faults. However, the method requires a window for converting the time-domain waveforms into the RMS domain. In order to estimate the exact fault inception and clearing times, the offsets caused by the window size must be corrected. Therefore, the RMS-wavelet method is proposed for detecting fault events. Note that the discrete wavelet transform used alone can detect other power quality events as well as fault events. The discrete wavelet transform is used only to correct the time offsets of the fault events detected from the RMS pickup value. The time index where the wavelet coefficients have higher values among the neighbors is chosen to be the corrected location of the fault inception time or clearing time. Discrete wavelet transform can be implemented using high-pass and loss-pass filters as in (2.12)-(2.13), where g and h denote low-pass and high-pass filters, respectively. The outputs of the high-pass filter are referred to as detail coefficients, and the outputs of the low-pass filter are referred to as approximation coefficients.

$$y_{high}[n] = \sum_{k} x(k)g(2n-k) \tag{2.12}$$

$$y_{low}[n] = \sum_{k} x(k)h(2n-k) \tag{2.13}$$

Equations (2.12) and (2.13) show a single-level decomposition of signal x. This decomposition can be cascaded to get higher-frequency resolution. For power quality disturbance analysis, first-level detail coefficients are usually

sufficient for localizing the events [46]. Selection of the mother wavelet is an important task when wavelet transform is applied for power quality event detection. It has been presented in [46, 47] that Daubechies wavelets are a good candidate for power quality applications. The Daubechies wavelets with filter length of four, or db2, are used for better localization of fault inception time and clearing time. Using the wavelet coefficients, the exact fault inception (n_{ws}) and clearing times (n_{we}) can be estimated by using (2.14) with n_{ps} and n_{pe} as the input parameter n_p, respectively.

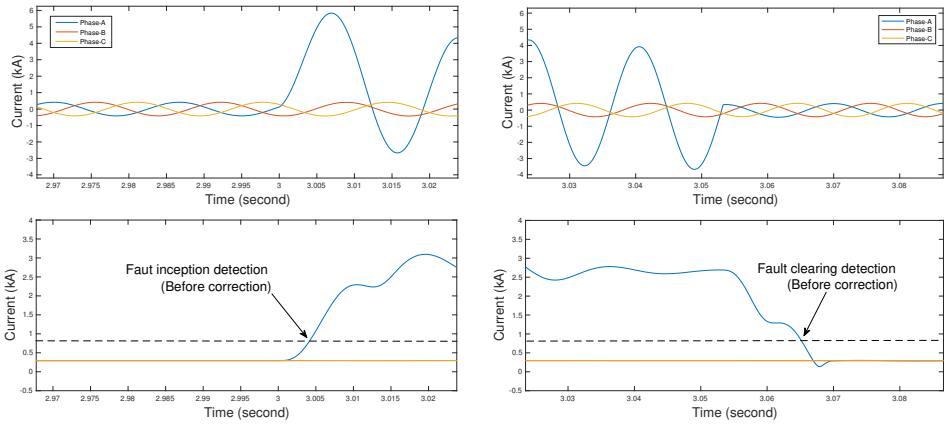
$$n_w = n_p - n_0 + 2\Delta n_w \tag{2.14}$$

$$\Delta n_w = \arg\max_n |\phi(n)|^2 \tag{2.15}$$

$$\phi(n) = \sum_k i(k)' g(2n-k) \tag{2.16}$$

$$i(n)' = \begin{cases} i(n+n_p-n_0+1) & \text{if } 0 \le n \le n_0 - 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

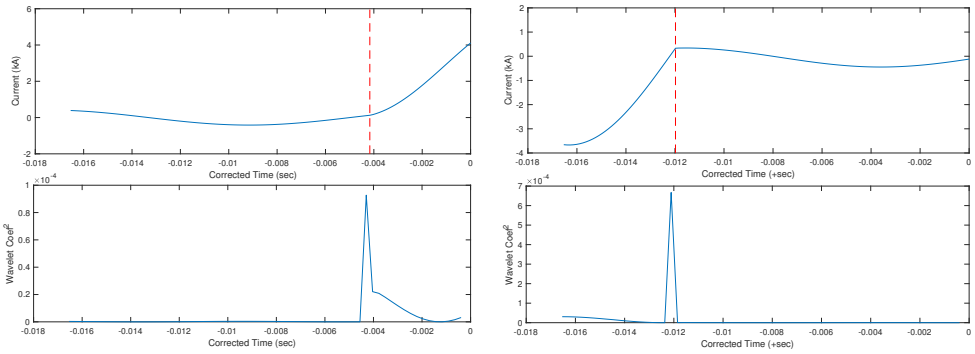
In Fig. 2.9(a), the first time index that exceeds the RMS pickup value is at 3.0042 sec. The wavelet transform is performed at current samples extracted from t = 3.0042 - t_0 (t_0 is chosen to be the window size used in the RMS computation) sec to t = 3.0042 sec. In Fig. 2.10(a), the wavelet coefficient has the highest value when the corrected time is -4.2 ms. The fault inception time ($t_{f,start}$) is adjusted to 3.0000 sec. A similar approach is performed when the current drops below the RMS pickup value to correct the fault clearing time ($t_{f,end}$). From Fig. 2.9(b) and Fig. 2.10(b), the fault clearing time ($t_{f,end}$) is estimated to be at 3.0533 sec after applying the wavelet transform. The fault



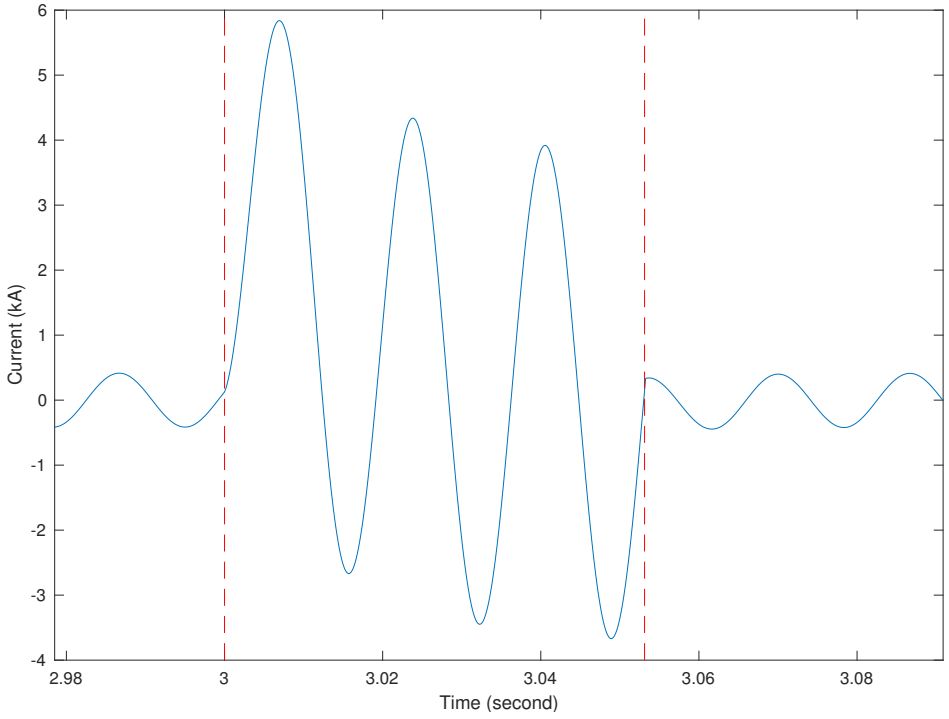
(a) (b)

Figure 2.9: Fault current detection in RMS domain: (a) first index above threshold, instantaneous (top) and RMS (bottom); (b) last index above threshold, instantaneous (top) and RMS (bottom).

inception time ($t_{f,start}$) and the clear time ($t_{f,end}$) are then passed to the curve-fitting module. Fig. 2.10(c) shows the detected fault waveform. Table 2.3 lists the first and the last time indices exceeding the RMS pickup value when RMS was calculated using the Fourier filter and the cosine filter, and the corrected times using the wavelet transform.



(a) (b)



(c)

Figure 2.10: Corrected fault inception and clearing time using wavelet transform: (a) fault inception time; (b) fault clearing time; (c) detected fault waveform.

Table 2.3: Corrected fault inception time and clearing time

	Fourier	Cosine	Corrected time
Fault inception time	3.0042	3.0046	3.0000
Fault clearing time	3.0652	3.0694	3.0533

2.4 Fault Location

Impedance-based fault-locating methods require the voltage and the current phasors at the fundamental frequency as the input. A common approach to estimating the fault locations is to select a cycle during a fault and use the current and voltage phasors at the chosen time sample as the inputs to the fault-locating algorithms. This section proposes the use of the RMS-wavelet fault detector and the least squares phasor estimating algorithm described in Section 2.2.3 and Section 2.3.2.2 to determine the location to the fault. This approach results in improved location estimates for momentary faults where the data are limited because of the short duration and the phasor estimation is complicated by the exponentially decaying DC offsets. The proposed method uses variable window size in calculating phasors and estimates a single fault location that is more accurate than the multiple locations estimated by the Fourier and cosine filters. The method is validated using simulated and actual field data.

2.4.1 Problem Description

In the event of a momentary fault in a transmission or distribution feeder, impedance-based fault location algorithms are widely used by utilities and in microprocessor-based relays to estimate the distance to the fault [48]. These phasor-based algorithms require the input of fundamental frequency voltage and current measurements captured during a fault to estimate the apparent impedance from the monitoring location to the faulted point. Given the line impedance in ohms per unit distance, the corresponding distance to the fault can be easily obtained. For example, consider the Takagi method popularly implemented in commercial relays [5]. The distance to the fault, d, is estimated as

$$d = \frac{\text{imag}(V \times I_{sup}^{*})}{\text{imag}(Z_{line,1} \times I_s \times I_{sup}^{*})} \tag{2.18}$$

where: $I_{sup} = I - I_{preflt}; I_s = I + (\frac{Z_{line,0}}{Z_{line,1}} - 1)I_0;$
the asterisk * denotes a complex conjugation operator.

From (2.18), the voltage and current phasors during a fault play an important role in accurately estimating the distance to the fault. Phasor computation, however, is complicated by the presence of an exponentially decaying dc offset, which makes the fault current asymmetrical in the first few cycles, as shown in Fig. 2.11. To filter out the DC offset and preserve

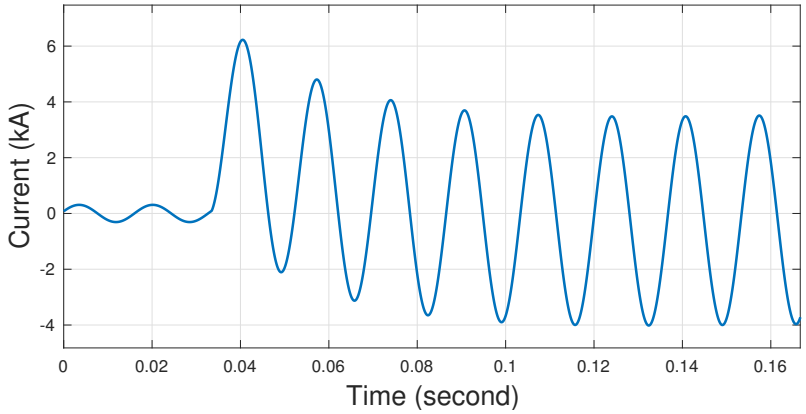


Figure 2.11: Asymmetrical fault current.

the accuracy of the fault location algorithms, Fourier and cosine filters are commonly used by phasor estimation algorithms in relays [10]. The Fourier filter requires one cycle of waveform data to extract the magnitude and phase angle of fundamental frequency voltage and current. The cosine filter needs one cycle and an additional quarter cycle to calculate the fundamental frequency magnitude and the phase angle. The phasor outputs, I, of the Fourier and cosine filters at time index m are determined using (2.3) and (2.7).

Fig. 2.12 shows the fault current magnitude of the waveform shown in Fig. 2.11 obtained by applying the Fourier and cosine filters. It can be observed that after the DC offset has fully decayed, the fault current is 2.65 kA. However, when DC offset is present, the worst-case fault current estimates using the Fourier and cosine filters are 3.02 kA and 2.71 kA, respectively. From this perspective, the cosine filter does a better job of eliminating the DC offset.

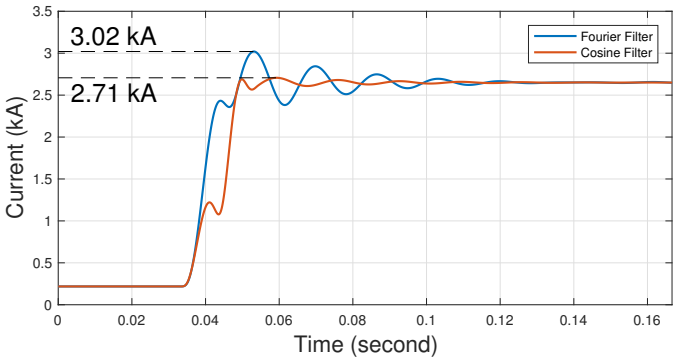


Figure 2.12: Fault current magnitude estimated by Fourier filter and cosine filter.

In summary, a decaying DC offset affects the accuracy of the voltage and current phasors computed using the Fourier and cosine filters, particularly in the case of momentary faults. This in turn degrades the accuracy of impedance-based fault location algorithms. Therefore, the problem addressed in this section can be stated as follows: given the three-phase voltage and current waveforms recorded by a relay at the monitoring location, eliminate the DC offset and use only the symmetrical AC fault voltage and current to determine the location of momentary faults.

2.4.2 Exponential Decaying DC Offset in Fault Waveforms

When a fault occurs in a transmission or distribution feeder, the fault current consists of an AC symmetrical component and a DC offset that decays exponentially with a time constant. Consider a simple distribution feeder shown in Fig. 2.13. For a three-phase fault, the current during fault is as follows [41]:

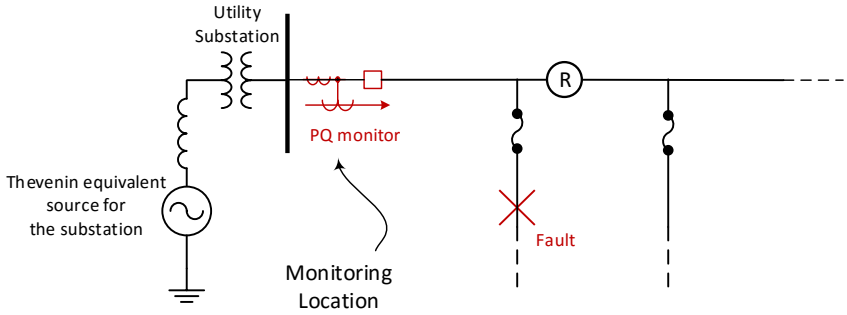


Figure 2.13: Simple distribution feeder.

$$i(t) = Ae^{-\frac{wt}{X/R}} + \sum_n B_n \sin(nwt + \theta_n) \tag{2.19}$$

$$= i_{dc}(t) + i_{ac}(t) \tag{2.20}$$

where, $i_{dc} = Ae^{-wt/(X/R)}$, $i_{ac} = \sum_n B_n \sin(nwt + \theta_n)$

The i_{ac} in (2.20) represents the symmetrical or steady-state AC fault current, while i_{dc} represents the exponentially decaying DC offset. Since current in an inductor cannot change instantaneously, DC offset (i_{dc}) appears when the value of i_{ac} at fault incidence is different from the value of load current before the fault. Therefore, the initial magnitude of the DC offset depends on the fault incidence angle. Then the DC offset decays with a time constant dependent on the X/R ratio of the system at the faulted point. The higher the X/R ratio, the longer it takes for the offset to decay.

The proposed approach to eliminating the DC offset in fault location estimates is as follows. First, the exact fault inception time and fault clearing time are estimated using the RMS-wavelet method presented in Section 2.3.2.2. Then, the faulted section of the waveforms is fitted to an exponential decaying

function plus a sinusoidal using the nonlinear least squares algorithm defined in Section 2.2.3. Finally, the phasors are estimated and applied in fault-locating algorithms.

2.4.3 Application of DC Offset Removal Algorithm Using Simulated Fault Data

This section validates the efficacy of the proposed method using simulated data acquired from a time-domain simulation in [49]. A simple two-bus transmission system was modeled at the nominal voltage rating of 69 kV, as shown in Fig. 2.14. Single line-to-ground faults are simulated 10 miles away from Terminal G, where the measurements are recorded. Three cases are considered in this section: bolted fault (case 1), fault with resistance (case 2), and fault with duration less than one cycle (case 3). The modeling parameters are summarized in Table 2.4. Since the Fourier and cosine filters are most commonly used in commercial relays, the location estimates of the proposed method are compared with the results of these two filters.

For cases 1 and 2, the fault duration used in the simulation is 10 cycles. Note that in the actual application of the proposed method, it is likely that only a few cycles of fault waveforms are available. Longer fault duration was used in the two cases to show the results both when DC offset is present in the waveforms and after it dies out. Throughout the demonstration, the Takagi method [5] is used in fault location evaluation.

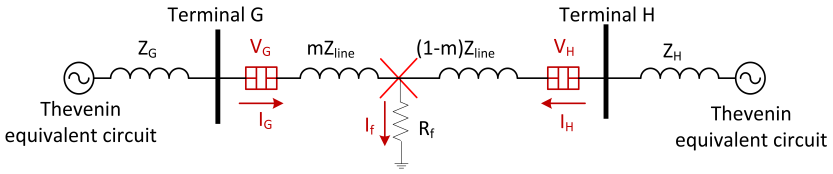


Figure 2.14: Simple two-bus transmission system.

Table 2.4: Modeling parameters

Modeling Parameter	Value
$Z_{G,1}$	$0.2616 + j\,3.7409\ \Omega$
$Z_{G,0}$	$0.7848 + j\,11.2226\ \Omega$
$Z_{H,1}$	$0.6512 + j\,3.6930\ \Omega$
$Z_{H,0}$	$1.9535 + j\,11.0791\ \Omega$
$Z_{line,1}$	$0.1588 + j0.5185\ \Omega$/mile
$Z_{line,0}$	$0.5260 + j1.5075\ \Omega$/mile
Sampling Frequency	128 samples per cycle
Line Length	30 miles
Fault Location (m)	10 miles (from Terminal G)
Fault Duration (t_f)	10 cycles / 0.8 cycles
Fault Resistance (R_f)	$0\ \Omega / 5\ \Omega$

2.4.3.1 Case 1: Bolted Fault ($R_f = 0$)

The fault current and voltage waveforms used in the demonstration are shown in Fig. 2.15. The phasors are calculated using the Fourier, cosine, and proposed methods. The calculated phasors were given as inputs to the Takagi method to estimate the fault location. The results are shown in Fig. 2.16 and Table 2.5. The percentage error is defined as

$$\% \text{ Error} = \frac{\text{Estimated Fault Location} - \text{Actual Location}}{\text{Actual Fault Location}} \times 100 \qquad (2.21)$$

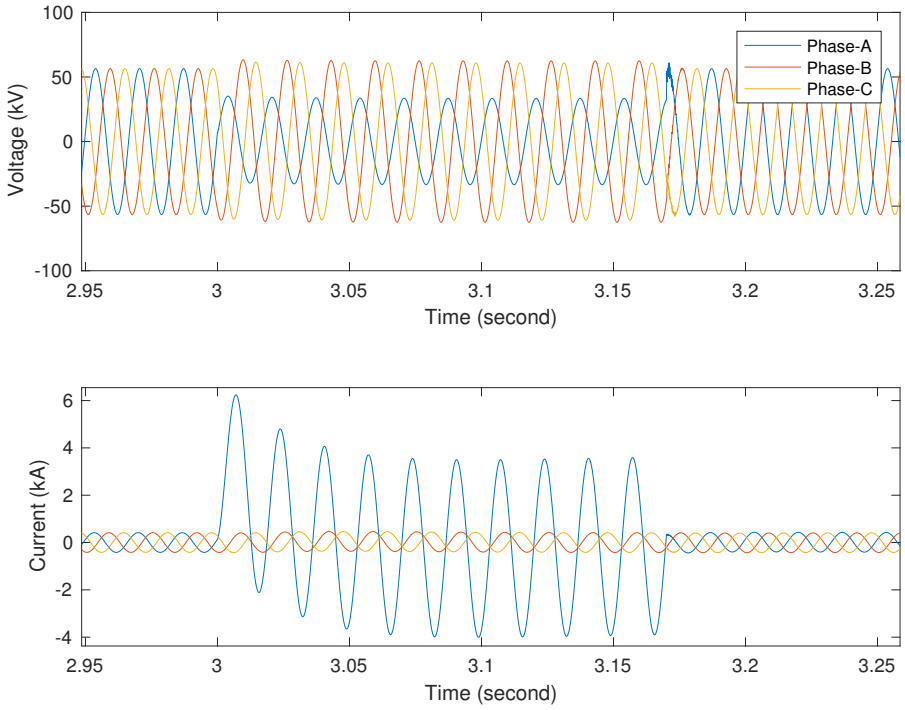


Figure 2.15: Simulated fault waveforms for test case $R_f = 0$.

Table 2.5: Location estimates case $R_f = 0$

Actual Location	10 miles		
Estimated Location	Fourier	Cosine	Proposed
Min. Est. (% Error)	8.7265 (-12.74)	9.7786 (-2.21)	9.9947 (-0.05)
Max. Est. (% Error)	11.1387 (11.39)	10.1649 (1.65)	

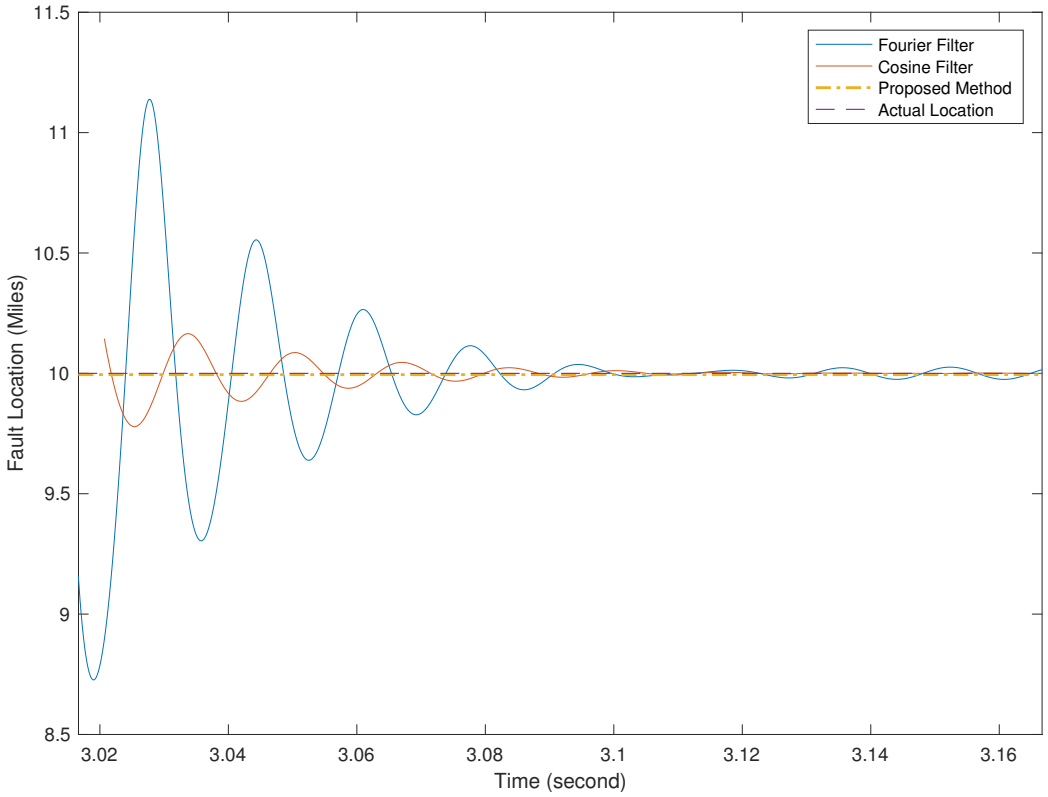


Figure 2.16: Fault location estimates comparison for test case $R_f = 0$.

A sliding window approach was used for the fault locations estimated by the Fourier and cosine filters. Since they use one cycle and one and a quarter cycles of data points, respectively, these two filters have a wide range of location estimates, which rely on the location of the sliding window. On the other hand, the proposed method adjusts its window size to the entire fault waveforms' duration, and therefore it yields a single-value estimate of the fault location.

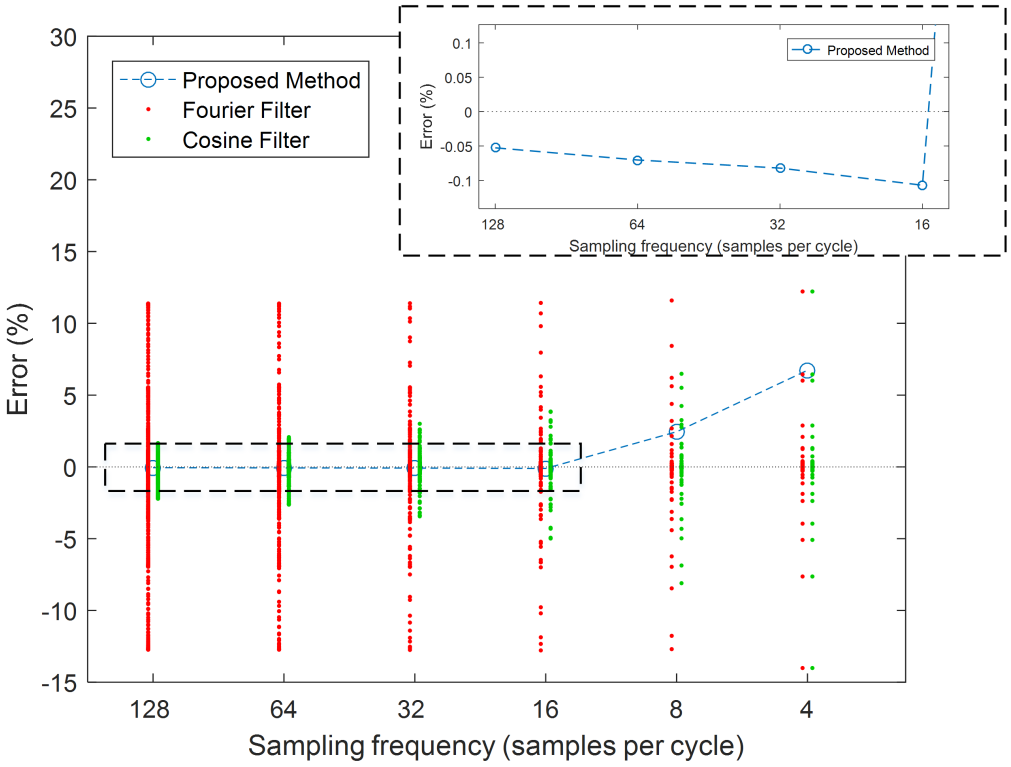


Figure 2.17: Error estimates and sampling frequency.

As seen in Fig. 2.16, the location estimates derived from the Fourier and cosine filters have oscillations and are centered at the actual fault location, 10 miles. In the worst case, the Fourier filter overestimated and underestimated the fault location by 11.39% and −12.74%, respectively. The location estimates of the cosine filter oscillate in a range resulting in maximum and minimum errors of 1.65% and −2.21%. On the other hand, the fault location estimated by the proposed method was 9.9947 miles, and the error is only −0.05%.

It should be noted that when there is no DC offset in the fault waveforms, the Fourier and cosine filters produce location estimates identical to the proposed method. This can be seen from the latter portion of the illustration (Time $>$ 3.1 sec) in Fig. 2.16, where the fault locations estimated by all three methods are identical.

Fig. 2.17 depicts the location estimation results according to the sampling frequency of the measurement device. A higher sampling frequency enables more samples to be involved in phasor estimations, and therefore the phasors and the estimated fault locations are more accurate. In this evaluation, the error estimate of the proposed method is 5% for the sampling rate of 4 samples per cycle. The error reduces to -0.11% when the sampling rate increases to 16 samples per cycle, then marginally improves as the sampling frequency is further increased. The error estimate is -0.05% when the sampling rate is 128 samples per cycle.

2.4.3.2 Case 2: Fault with Resistance ($R_f = 5\Omega$)

The location results for case 2 are shown in Fig. 2.18 and Table 2.6. Note that the chances of the DC offset being present in the waveforms are less for higher-resistance faults. The increase in fault resistance reduces the X/R ratio during the fault, and the DC offset dampens more quickly than in the bolted fault case. In Fig. 2.18, the Fourier and cosine filters produce location estimates identical to the proposed method at Time $>$ 3.05 sec, whereas for the bolted fault case in Fig. 2.16, this occurs at Time $>$ 3.1 sec.

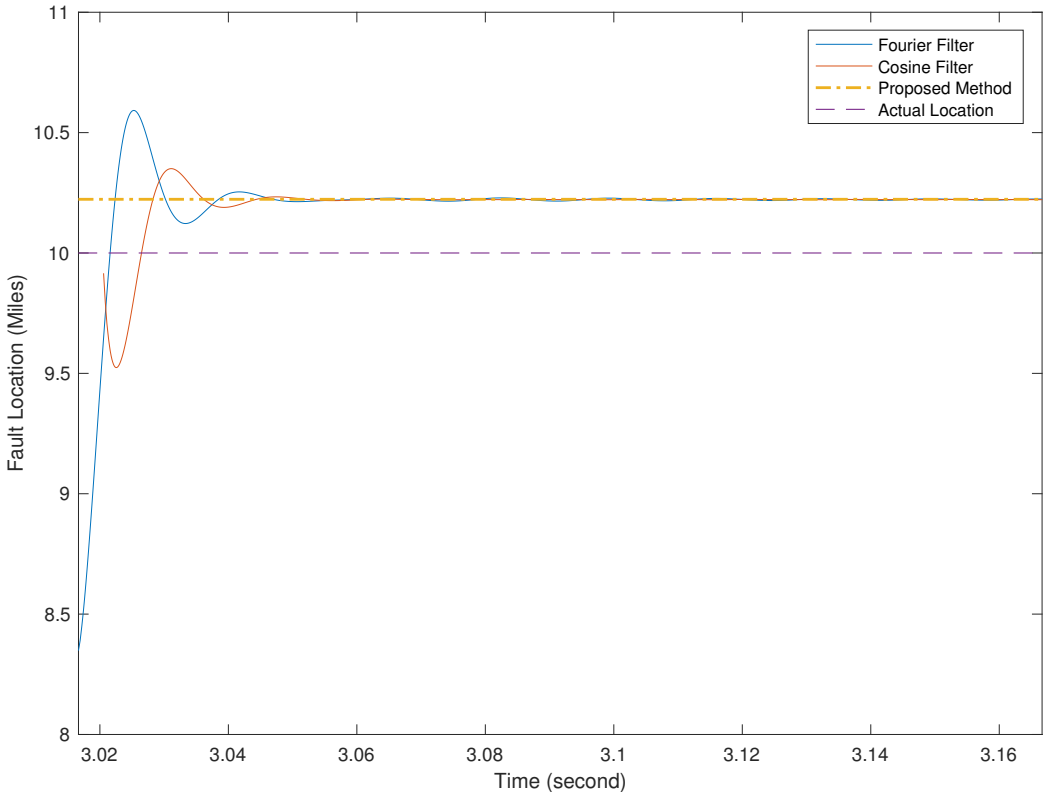


Figure 2.18: Fault location estimates comparison for test case $R_f = 5$.

Similarly to the case where $R_f = 0\Omega$, the Fourier and cosine filters produce a wide range of location estimates. As shown in Table 2.6, the error ranges of the Fourier and cosine filters are from -16.72% to 5.92% and from -4.76% to 3.50%, respectively. The proposed method, on the other hand, successfully removes the DC offset and produces an accurate estimation result, with the error being only 2.23%.

It should be noted that the errors originate from two sources in this simulation environment: one from the inaccurate phasor calculation due to

Table 2.6: Location estimates case $R_f = 5$

Actual Location	10 miles		
Estimated Location	Fourier	Cosine	Proposed
Min. Est. (% Error)	8.3283 (-16.72)	9.5239 (-4.76)	10.2230 (2.23)
Max. Est. (% Error)	10.5922 (5.92)	10.3504 (3.50)	

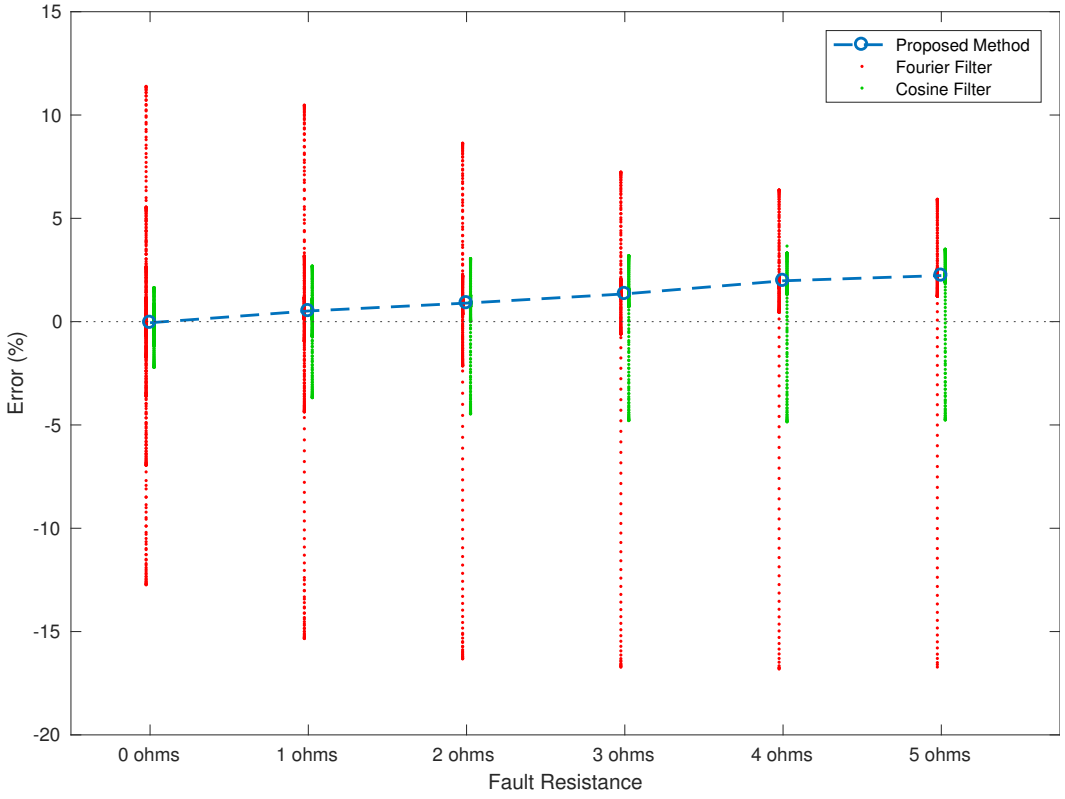


Figure 2.19: Error estimates and fault resistance.

the presence of the DC offset, and the other from the assumptions made in the fault-locating algorithm, the Takagi method. The Takagi method assumes a homogeneous system, where the local and the remote source impedances have the same impedance angle as the power line [50]. This assumption is violated, as the simulated system is nonhomogeneous. Fig. 2.19 shows that the increase in fault resistance increases the errors caused by the system nonhomogeneity.

2.4.3.3 Case 3: Fault with Duration Less Than One Cycle

In this scenario, a bolted fault with a short duration of less than one cycle is simulated. The voltage and the current waveforms are shown in Fig. 2.20. Recall that the Fourier and cosine filters require one cycle and one and a quarter cycles of data points, respectively, during the fault to calculate the fundamental phasors. Since the fault duration used in this case is less than one cycle, the requirements are not fulfilled. Therefore, for this case only, both the data points during the fault and after the fault clearance are used in calculating phasors using the Fourier and cosine filters, as shown in Fig. 2.20.

The location results for the Fourier, cosine, and proposed methods are shown in Table 2.7. The estimates of both the Fourier and cosine filters are severely affected by the short fault duration. The cosine filter, which requires one and a quarter cycles of data points, provides the least accurate estimate, with 60.58% error. On the other hand, the proposed method adjusts its window size and uses only the data points during the fault to estimate the phasors, giving a very accurate fault location result with only -1.67% error.

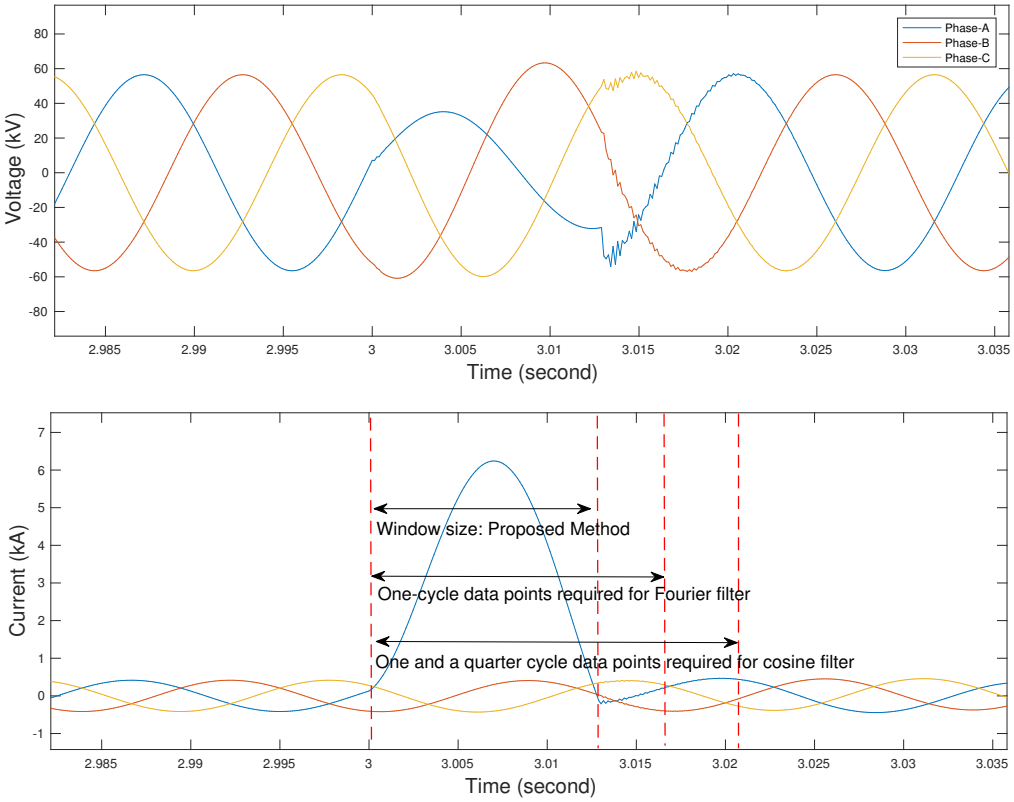


Figure 2.20: Simulated fault waveforms for short-duration fault case.

Table 2.7: Location estimates case short-duration faults $R_f = 0$

Actual Location	10 miles		
	Fourier	Cosine	Proposed
Estimated Location	11.9027	16.0583	9.8331
(% Error)	(19.03)	(60.58)	(-1.67)

2.4.4 Application of DC Offset Removal Algorithm Using Field Data

In this section, actual field data collected from transmission and distribution systems are used in the demonstration. The datasets include single

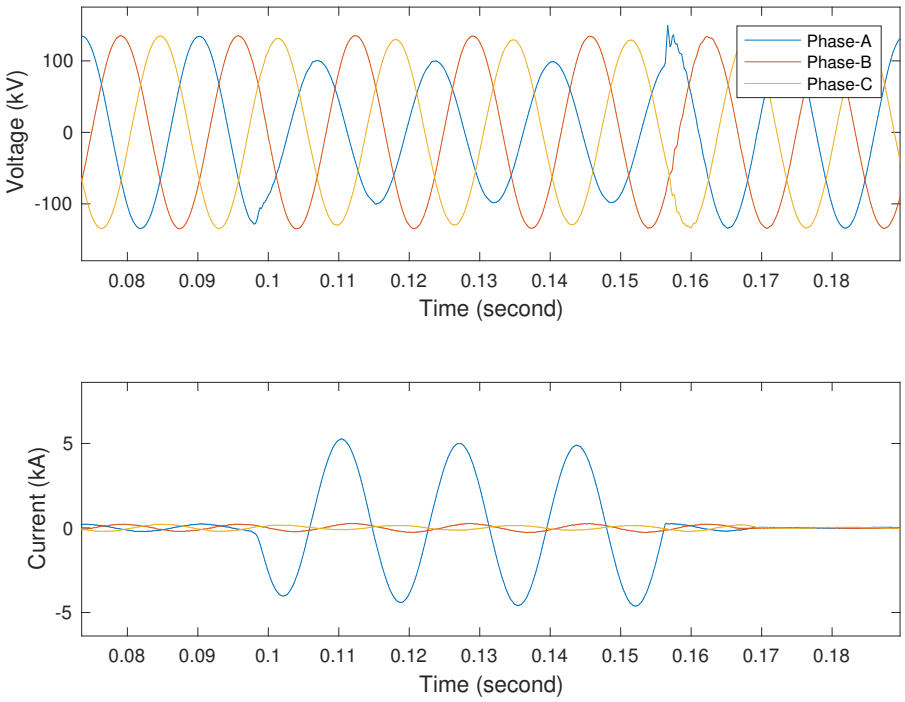


Figure 2.21: Utility test case 1: recorded waveforms.

line-to-ground faults, which were cleared by the protective device within a few cycles. The faults are cleared before the DC offsets completely die out, which is a good condition for evaluating the proposed method.

2.4.4.1 Utility Test Case 1

Test case 1 considers a single line-to-ground fault that occurred on the phase A line of a 161 kV transmission system. The transmission line is 21.15 miles long, and the positive-sequence and zero-sequence line impedances are

$z1 = 0.1504 + j0.7883$ Ω/mile and $z0 = 0.7192 + j2.4798$ Ω/mile, respectively. The load current was 148 A before the fault. When a fault occurs, the fault current magnitude of 3.4 kA caused the breaker at the substation to trip after 3.5 cycles. The voltage and current waveforms are recorded by the digital fault recorder (DFR) at the substation and are shown in Fig. 2.21. The sampling frequency of the DFR is 100 samples per cycle. The actual fault location was known to be 14.90 miles away from the substation.

Fig. 2.22 shows the illustrative results of the least squares algorithm. In Fig. 2.22(a), the RMS-wavelet method was applied to the fault current waveform to obtain the exact fault inception time and clearing time. Then these time indices were used to extract fault waveforms in phase A current, phase A voltage, and the zero-sequence (neutral) current waveforms. These waveforms were then decomposed into the DC offset and the sinusoidal waveform shown in Fig. 2.22(b)-2.22(d).

Fig. 2.23 and Table 2.8 show the fault location estimation results of the Fourier, cosine, and proposed methods. As described in Section 2.4.3, the Fourier and cosine filters output a wide range of location estimates compared with the proposed method. The maximum estimate errors of these two filters were 5.27% and 2.69%, respectively. On the other hand, the fault location estimated using the proposed method was 14.8428 miles, with -0.38% error.

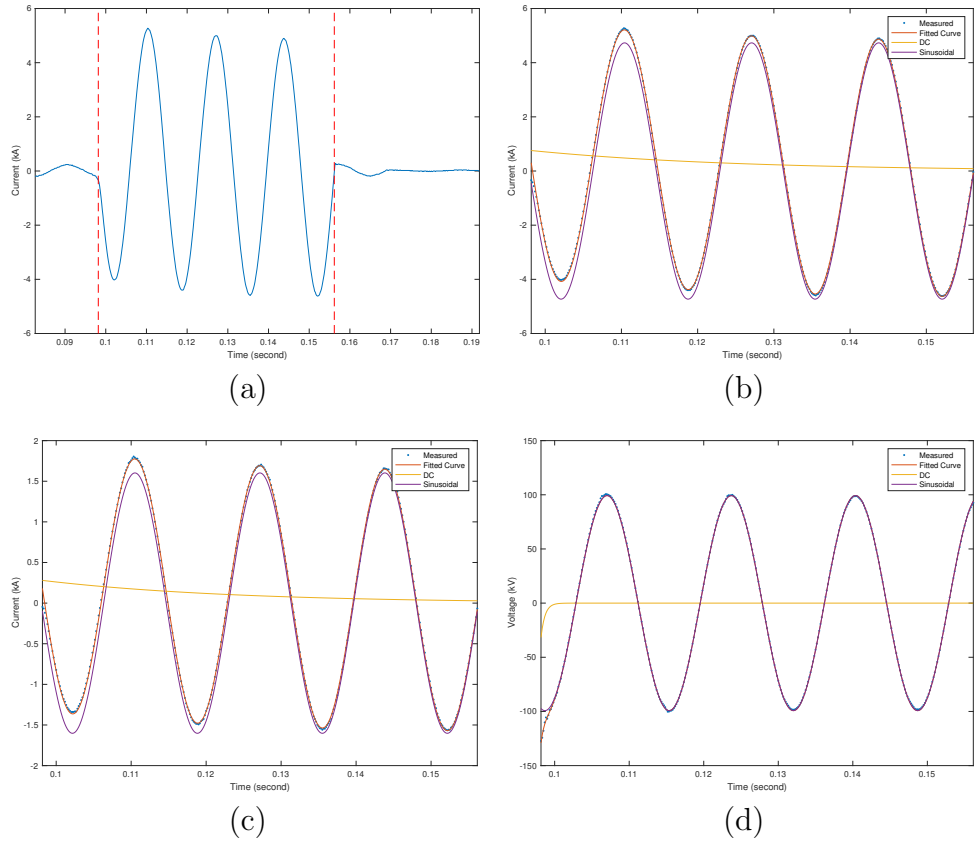
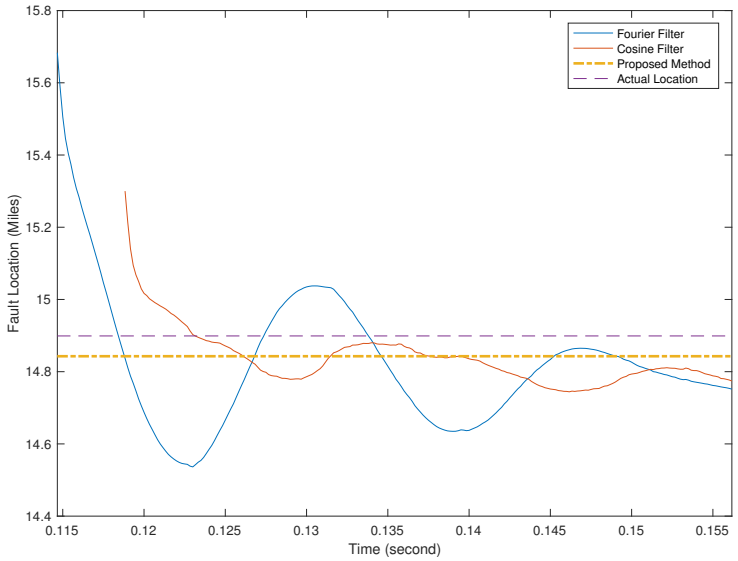


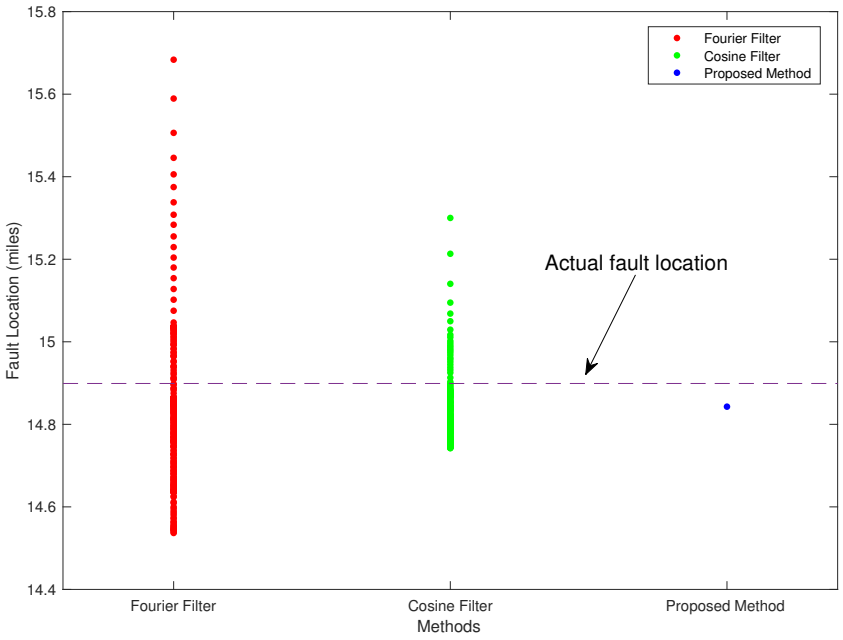
Figure 2.22: Utility test case 1: DC offset removal: (a) fault detection; (b) curve-fitting phase A current; (c) curve-fitting zero sequence current; (d) curve-fitting phase A voltage.

Table 2.8: Location estimates utility test case 1.

Actual Location	14.90 miles		
Estimated Location	Fourier	Cosine	Proposed
Min. Est. (% Error)	14.5364 (-2.43)	14.7443 (-1.04)	14.8428 (-0.38)
Max. Est. (% Error)	15.6835 (5.27)	15.3001 (2.69)	



(a)



(b)

Figure 2.23: Utility test case 1: fault location estimates: (a) estimates over time (sample); (b) ranges of estimates. Unlike estimates from Fourier and cosine filters, the proposed method produces a steady single-value estimate closest to the actual fault location.

2.4.4.2 Utility Test Case 2

Test case 2 considers a single line-to-ground fault that occurred on the phase B line of a 25 kV distribution system. The positive sequence and zero sequence line impedances are $z1 = 0.1308 + j0.5546\ \Omega$/mile and $z0 = 0.4029 + j1.8619\ \Omega$/mile, respectively. The load current was 36 A before the fault. The fault current magnitude was 2.3 kA and lasted for 2.5 cycles before the protection device cleared the fault. The waveforms recorded by a digital relay at the substation are illustrated in Fig. 2.24. The sampling frequency of the relay is 32 samples per cycle. The actual fault location was 2.67 miles from the substation.

The fault location results of the Fourier, cosine, and proposed methods after repeating the same procedure and eliminating the DC offset are shown in Table 2.9. In this scenario, the proposed method successfully removes the DC offset but does not significantly improve the location estimates because the Takagi method underestimates the fault location. The cause of underestimation may come from the tapped loads along the feeder, nonhomogeneous line impedance, or other factor. The largest underestimation errors for the Fourier and cosine filters are -9.90% and -7.34%, respectively, and the proposed method underestimates the fault location by -5.56%. The estimation results are depicted in Fig. 2.25.

The output of the impedance-based fault-locating algorithms is the distance between the fault and the monitoring location where the voltage and currents are recorded. Line crews can be sent to known fault locations to clear

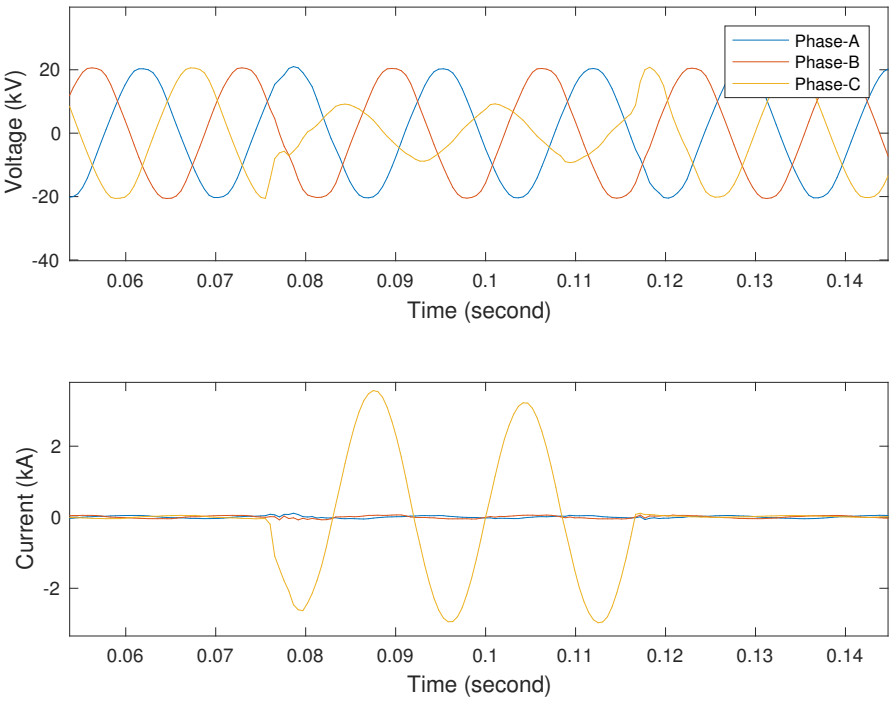
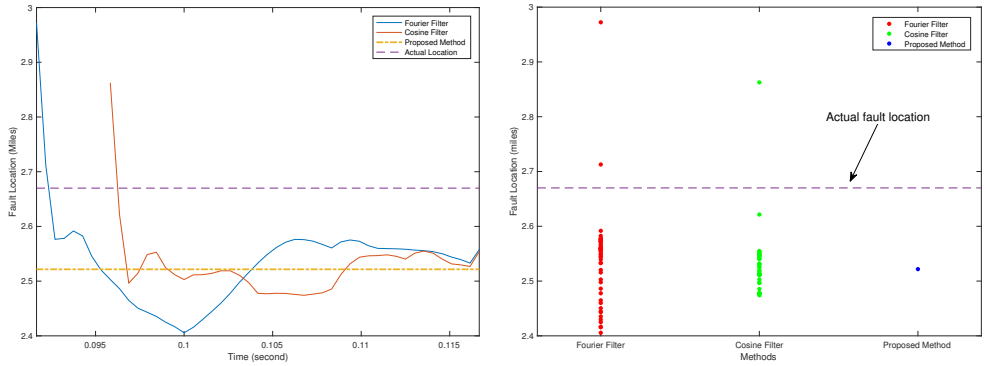


Figure 2.24: Utility test case 2: recorded waveforms.

the root causes of the faults and prevent their recurrence. In a distribution system, more than one possible fault location may exist because of lateral branches connected to the primary feeder. In such scenarios, customer outage calls or fault indicators mounted on the lines can be used to narrow down the possible fault locations[51, 52, 53, 54].



(a) (b)

Figure 2.25: Utility test case 2: fault location estimates: (a) estimates over time (sample); (b) ranges of estimates. The proposed method produces an accurate single-value estimate.

Table 2.9: Location estimates utility test case 2.

Actual Location	2.67 miles		
Estimated Location	Fourier	Cosine	Proposed
Min. Est. (% Error)	2.4056 (-9.90)	2.4741 (-7.34)	2.52 (-5.56)
Max. Est. (% Error)	2.9722 (11.32)	2.8625 (7.21)	

2.5 Summary

This chapter has presented data analytics algorithms for detecting and locating faults in the monitored circuit. The proposed method consists of exact detection of the fault inception and clearing times using an RMS-wavelet method and phasor estimations using the nonlinear least squares method. Results indicate the success of the proposed approach in eliminating the DC

offset, and the fault location estimates were observed to be closer to the actual location of the fault, especially for low-resistance faults. The proposed approach is evaluated using a time-domain simulation implemented in [49] and actual fault events.

The error estimates of the proposed method are less than 5.6% in all considered cases, whereas the errors when using conventional Fourier and cosine filters can be as high as 16.7% and 7.3%, respectively. In addition, the method can be applied to short-duration (less than one cycle) fault events where the Fourier and cosine filters cannot be applied, because these filters require a minimum of one cycle or one and a quarter cycles of data points, respectively, to calculate the fundamental phasors.

Chapter 3

Identification and Evaluation of Overcurrent Protection Devices

3.1 Introduction

[1] Overcurrent protection devices are installed in distribution systems to isolate the faulted section from the electric power network. When a fault occurs in a radial distribution system, protective devices such as a recloser or a fuse isolate the fault to prevent electric devices such as transformers, conductors, and capacitors from overheating and to quickly restore power [55, 41]. The speed at which the devices operate must be quick to reduce the duration of power quality problems such as voltage sags and interruptions. The protective devices are coordinated with other devices located downstream and upstream in the system to minimize the number of customers affected by fault conditions and improve system reliability. Normally, the protective devices are coordinated such that the device closest to the fault operates faster than the devices located farther upstream, with the exception of the fuse-saving scheme, where the upstream recloser operates faster than the fuse [41].

[1]Parts of this chapter have been published in K. W. Min, S. Santoso, and L. Biyikli, "Identifying fault clearing operations in distribution systems," in *Proc. IEEE Power Energy Soc. General Meeting*, July 2016, pp. 1-5. The author of this dissertation analyzed the data, developed the algorithms, and validated the analytical results in the paper.

This chapter describes the data analytics algorithms for evaluating overcurrent protection devices in distribution systems. First, rule-based screening module is presented. This tool is used to identify the type of protective device clearing the short-circuit fault. Necessary features that characterize fault-clearing devices are calculated from the input voltage and current waveforms captured at intelligent electronic devices (IED) such as digital relays, digital fault recorders, and power quality monitors. The rule-based expert system then uses these features to determine whether the fault was cleared by recloser or fuse or if it was self-cleared. In the second part of this chapter, a methodology is proposed to estimate the empirical inverse time-current characteristics (TCC) of overcurrent relay/reclosers installed in distribution systems. The datasets recording -clearing operations of overcurrent relay/reclosers are used for this purpose. The empirical TCC curves can be used to evaluate opening intervals of breakers and identify the device (TCC curve) clearing the fault.

3.2 Protection Devices and Coordination in Distribution Systems

The most common types of protective devices in distribution systems are fuses and reclosers. A fuse is a device that is relatively inexpensive and maintenance-free compared with other devices, such as reclosers. Fault is isolated when a fuse link (typically tin, silver, or copper) melts from the heat caused by the fault current. The melting time is determined by the magnitude of the fault current and the TCC of the device. Typically, the TCC curve is

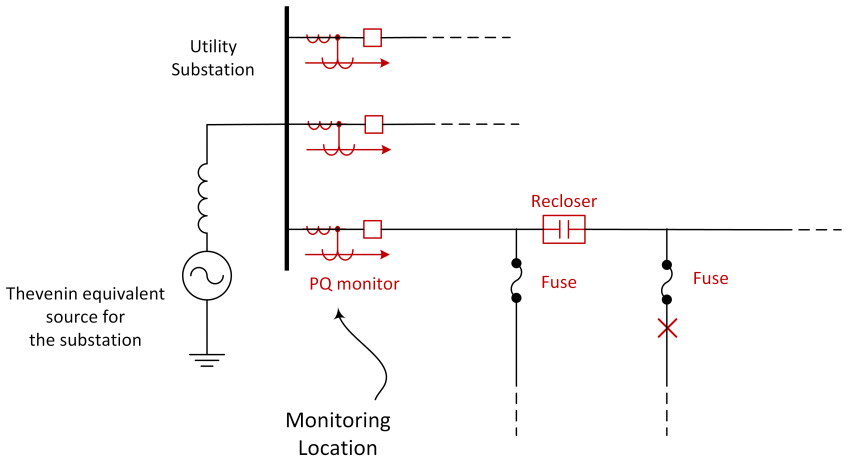


Figure 3.1: Recloser and fuse installed in the distribution system.

inversely proportional to the magnitude of the fault current. The interruption time is shorter for high fault current and longer for low fault current.

A recloser is a special type of circuit breaker used to clear momentary faults in a distribution system. The device repetitively opens and recloses a predetermined number of times (three or four) at fault conditions, until the fault current extinguishes. If the fault is not cleared until the last operation, the recloser locks out. More than 75% of the faults in distribution systems are temporary. Therefore, in most cases, a recloser can successfully clear the fault and reenergize the circuit on its first operation. Reclosers can operate on fast or delayed TCC curves. Typical sequences are one fast and three delayed or two fast and two delayed operations. If the recloser fails to clear the fault in its fast operations, the delayed curve allows other protective devices in the system to clear the fault. Reclosers also operate on TCC curves that are coordinated with other protection devices located upstream and downstream.

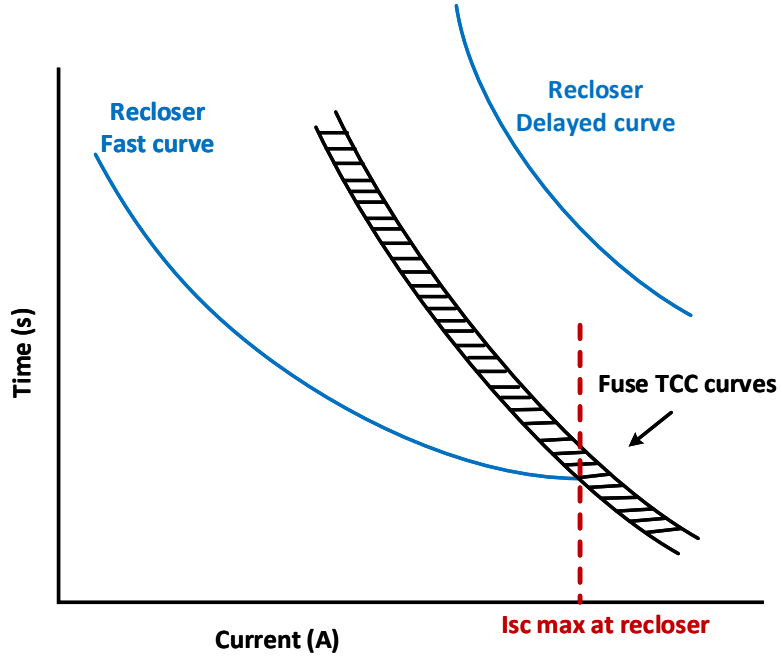


Figure 3.2: Recloser-fuse coordination in a fuse-saving scheme.

Protection devices in distribution systems are coordinated to minimize the service outage and number of affected customers from the fault. In general, protection devices are coordinated so that the equipment closest to the fault operates to isolate the fault from the system [41]. A fuse-saving scheme is one in which the recloser and the fuse located downstream from this recloser are coordinated so that the recloser clears the fault before the fuse melts. Fig. 3.2 shows the TCC curves of the reclosers and fuse in a fuse-saving scheme. The recloser and fuse are coordinated so that both the minimum melting time and total clearing time of the fuse lies in between the fast and slow curves of the recloser [56]. This coordination ensures that the recloser clears the fault for momentary faults, and the fuse clears the fault for permanent faults.

3.3 Identifying Fault-Clearing Operations in Distribution Systems

In this section, the algorithm for identifying fault-clearing operations is presented. The inputs of the algorithm are three-phase voltage and current waveforms recording recloser and fuse operations. The output of the algorithm is the classification result: whether the fault was cleared by a recloser or a fuse. Recloser and fuse operations can be characterized using features such as load demand, inrush current, and fault duration. These features are calculated for each recorded fault event and are used to identify which type of device has cleared the fault.

3.3.1 Real Power Load Demand Difference

The difference in real power load demand is an important characteristic that is used in determining the type of device clearing the fault. Assume a fault was detected from t_{start} to t_{end}. This analysis defines real power load demand calculated before t_{start} as $P_{prefault}$ and after t_{end} as $P_{postfault}$. The real power load demand calculated in phase C in Fig. 2.4 is shown in Fig. 3.3. Changes in the real power load demand before and after the fault event and the inrush event are illustrated.

3.3.2 Inrush Current

As described in Section 2.3.1.1, inrush event occurs when a recloser reenergizes the circuit after an interruption in the system. An example wave-

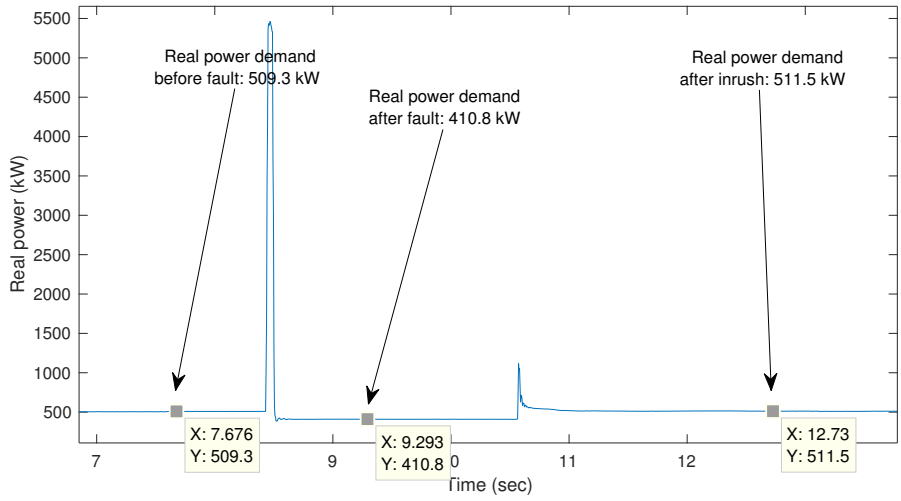


Figure 3.3: Real power demand in fault events.

form was shown in Fig. 2.4. The first disturbance is a fault. After the fault is cleared, high inrush current can be found. If this type of disturbance is recorded, it can be used as a strong indicator of the recloser reenergizing the circuit.

3.3.3 Fault Duration

Fault duration is calculated using the fault inception time and the clearing time, described in Section 2.3.2. Fault duration can be computed as the time difference between the fault clearing time and the inception time as show in (3.1).

$$t_f = t_{f,end} - t_{f,start} \tag{3.1}$$

The durations calculated by RMS and RMS-wavelet method are denoted by $t_{f,rms}$ and $t_{f,wavelet}$, respectively.

3.3.4 Case Study

Four fault events recorded in a distribution system are used to demonstrate the process to identify the fault-clearing device. Each event represents self-clearing faults, recloser-cleared faults (successful reenergizing and lockout), and fuse-cleared faults. Each dataset used consists of 30-sec voltage and current measurements with fault event(s) recorded. The sampling frequency is 500,000 samples per 30-sec, which corresponds to approximately 277.78 samples per cycle. The analysis is made after resampling the data to 256 samples per cycle. No other information such as line impedance, line length, TCC curves, or location of installed protective devices is used in this demonstration. Only the major rules are explained. However, the concept is similar for all the other cases.

3.3.4.1 Case 1: Self-Clearing Fault Identification

On July 11, 2015, a fault event was captured in the power quality monitor. The waveforms are shown in Fig. 3.4. The duration of the event, calculated from the wavelet transform and RMS waveform, is 1.04 cycles and 1.72 cycles, respectively. The short duration indicates that the event was a possible self-clearing fault. The next step is to observe whether the real power demand decreased after the fault event. The results tabulated in Table 3.1

show that the real power load demand did not decrease after the fault event. In fact, it increased by about 16 kW (approximately 4.45%). It can be concluded with high confidence that the event was a self-clearing fault. The rules used are that the fault durations should be very short (less than 2 cycles) and the real power load loss is negligible. Note that self-clearing faults are defined as temporary faults that self-extinguish before any protective device operates.

Table 3.1: Case 1: feature analysis

Feature	**Value**
$P_{prefault}$	313.31 kW
$P_{postfault}$	327.37 kW
$t_{f,wavelet}$	1.04
$t_{f,rms}$	1.7188
Device	Self-clear

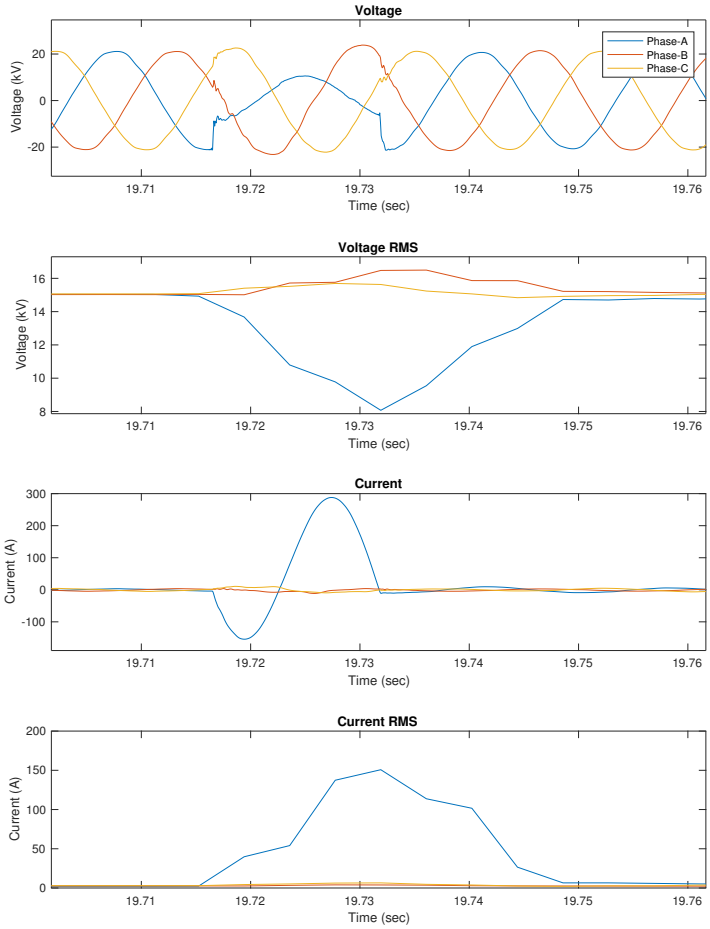


Figure 3.4: Self-clearing fault.

3.3.4.2 Case 2: Recloser Operation Identification (Successful Reenergizing)

On November 19, 2014, three consecutive fault events were recorded in the monitor. The features are analyzed as listed in Table 3.2. For identifying

Table 3.2: Case 2: feature analysis

Feature	**Event 1**	**Event 2**	**Event 3**
$P_{prefault}$	921. 73 kW	936.53 kW	721.00 kW
$P_{postfault}$	727.32 kW	720.25 kW	720.31 kW
$t_{f,rms}$	2.6055	15.4844	15.7031
Device	Recloser	Recloser	Recloser

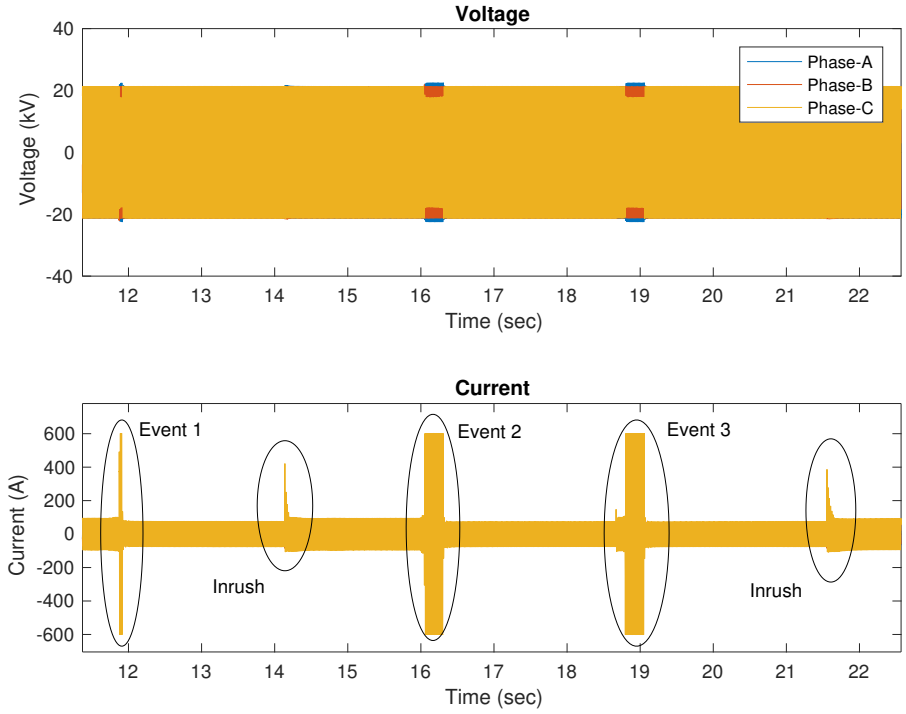


Figure 3.5: Recloser operation.

recloser operation, the detection of inrush events can be used. Two inrush events are found in Fig. 3.5, after fault events 1 and 3. Therefore, events 1, 2, and 3 are all identified as recloser-cleared faults.

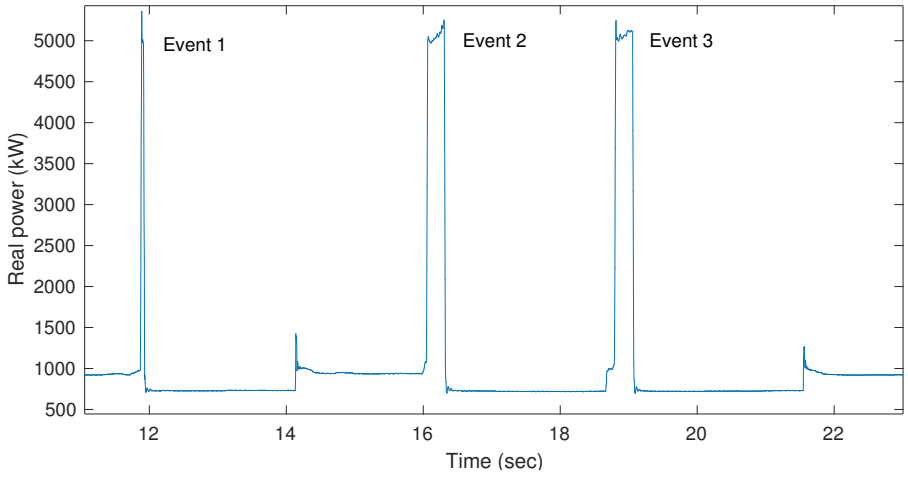


Figure 3.6: Recloser operation - real power series.

The changes in real power load demand can be used to analyze recloser operation in more detail. In events 1 and 2, the load demand is reduced from 921 kW to 727 kW and from 936 kW to 720 kW, respectively. This corresponds to the amount of load disconnected from the system due to the recloser operation. For event 3, the real power change is nearly zero. This characterizes the recloser attempting to reenergize the circuit but failing to clear the fault.

It should be noted that the inrush is indicative of the recloser successfully reenergizing the circuit. Inrush waveforms are not recorded if the recloser fails to reenergize and locks out.

3.3.4.3 Case 3: Recloser Operation Identification (Lockout)

If the fault is permanent, the recloser shifts to its delayed curve after one or two fast operations. This allows the fuse to clear the fault. However, if the fault is not cleared until the last operation, the recloser locks out. The following rules can be used to detect such a recloser lockout. First, the recloser can only lock out after the last recorded fault event. Second, the inrush event should not be recorded after the fault event. Third, the real power load demand should not decrease or increase after the last fault event. Finally, the duration of the last fault event should be longer than that of the first fault event.

Consider the fourth fault event recorded on August 5, 2015, as shown in Figs. 3.7 and 3.8. The duration of the fault current was 8.75 cycles, which is longer than the duration of the first recorded fault event, 2.75 cycles. Therefore, the recloser is assumed to have been in delayed operation. The real power load loss after the fourth fault event is calculated as -2.21%, which is assumed negligible. The power was not reenergized after the last recorded event (no inrush recorded). It can be concluded that the recloser locked out after the last fault event.

Note that the fault current magnitude of the fourth current is greater than that of the first three events. Therefore, the fourth fault event was of shorter duration than the second and third events because of the inverse time-current characteristics of the recloser.

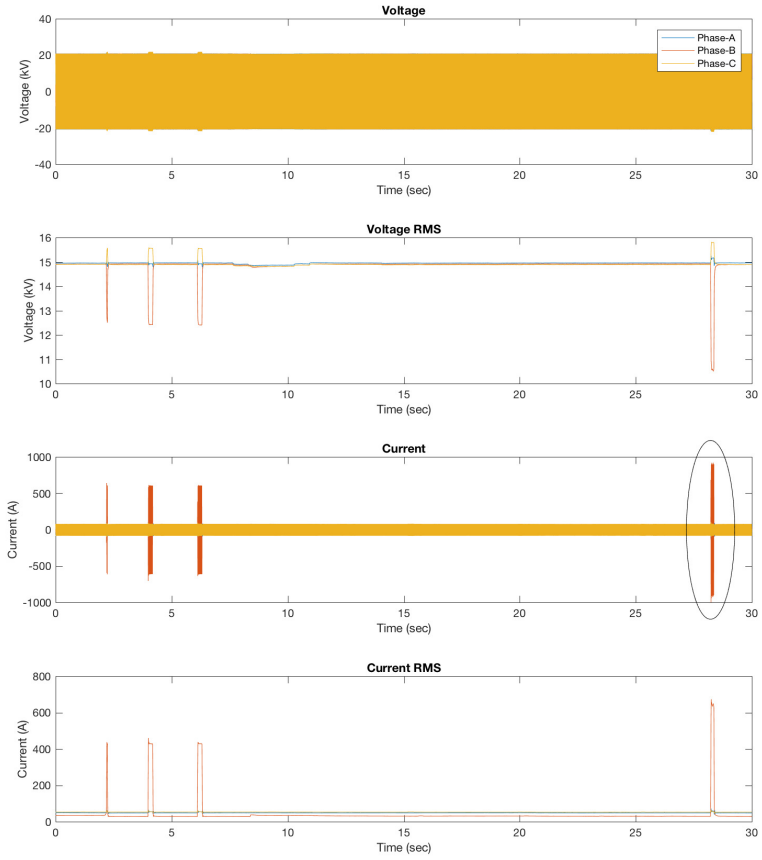


Figure 3.7: Fault event on August 5, 2015. Voltage and current waveforms.

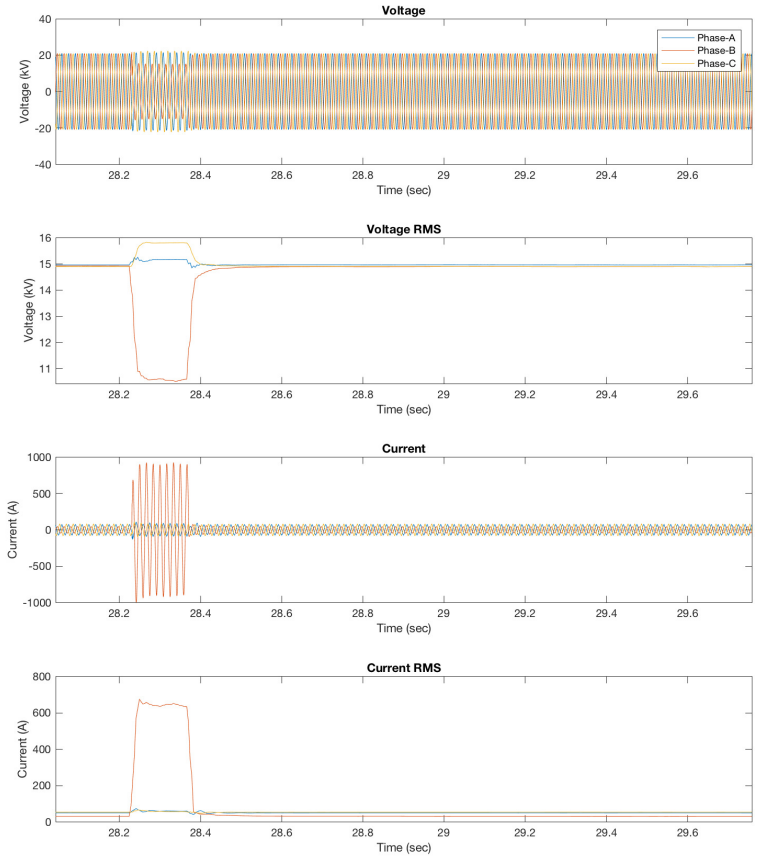


Figure 3.8: Fault event on August 5, 2015. The last recorded event.

Table 3.3: Case 3: feature analysis

Feature	**Event 1**	**Event 2**	**Event 3**	**Event 4**
$P_{prefault}$	504.68 kW	427.43 kW	427.72 kW	451.90 kW
$P_{postfault}$	427.65kW	427.68 kW	433.26 kW	441.91 kW
$t_{f,rms}$	2.75	12.5	12.25	8.75
I_{rms}	437.63 A	459.98 A	437.84 A	763.68 A
Device	Recloser	Recloser	Recloser	Recloser

3.3.4.4 Case 4: Fuse-Saving Scheme Identification

On February 16, 2015, two fault events occurred. The features are calculated and shown in Table 3.4. Although inrush is not recorded in this dataset, two consecutive fault events having durations of 2.39 and 9.93 cycles indicate that the recloser cleared the first fault event in the fast curve and shifted to its delayed operation. The real power load loss was 247.1 kW after the recloser cleared the first fault event.

The device clearing the second event is identified as a fuse. The first rule used is to compare the fault duration of the first and second events and ensure that the recloser shifted to its delayed operation. Then, if there was an increase in the load demand immediately after the fault current clearance, the device is assumed to be a fuse. In event 2, the real power demand increased to 1251.kW from 1089.8 kW immediately after the fault event. The increase (161.3 kW) can be assumed as the amount of loads reenergized, located upstream from the melted fuse but downstream from the reclosed recloser. Note that the load level is still lower than the amount before event 1. This amount of load

Table 3.4: Case 4: feature analysis

Feature	**Event 1**	**Event 2**
$P_{prefault}$	1336.9 kW	1089.8 kW
$P_{postfault}$	1089.5 kW	1251.1 kW
$t_{f,rms}$	2.39	9.93
Device	Recloser	Fuse

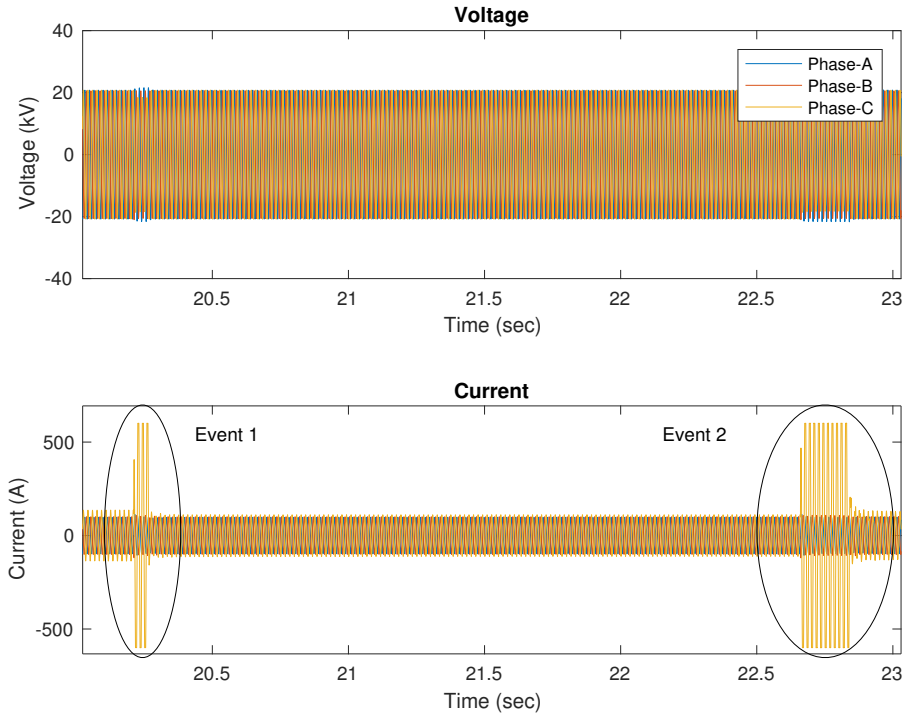


Figure 3.9: Fuse-saving scheme.

(85.8 kW) corresponds to the load located downstream from the operated fuse, which was disconnected from the system.

3.4 Empirical Estimation of Inverse Time-Current Characteristics in Distribution Systems

This section describes a methodology for formulating and estimating the empirical inverse time-current characteristics (TCC) of overcurrent relays/reclosers installed in utility distribution circuits. The algorithm makes use of three-phase voltage and current measurements for the estimation. First, the algorithm estimates the magnitude of the current flowing through the device and the times of fault-clearing operations. This makes the algorithm less sensitive to load currents and the location of the power quality monitor. Then a nonlinear least squares algorithm is formulated to estimate the TCC curve parameters of the reclosers clearing the fault. The estimated TCC parameters are used to construct an empirical TCC curve that can be used in various applications. Two potential applications are presented in this section. They include evaluating breaker opening intervals to monitor misoperations of the breakers. In addition, the empirical TCC curve is used to identify the type of fault-clearing device. The efficacy of the proposed algorithm is validated using simulated data, event reports generated from a digital relay test bench, and field events collected from a 24.9 kV distribution circuit.

IEEE Standard C37.112-2018 [29] defines an analytical equation of the inverse-time characteristics (3.2) to characterize the relay operating characteristics.

$$t(I) = \frac{A}{M^p - 1} + B \tag{3.2}$$

where M refers to multiples of the pickup current, and parameters A, p, B

determine the curve shapes. For example, constants $A = 28.2, p = 2.0$, and $B = 0.1217$ are used to define a TCC curve with extremely inverse curve characteristics. These constants are used in microprocessor-based protective relays to emulate the characteristics of electromechanical relays and to maintain coordination with conventional devices.

The trip signal is set at $t = T$ when the integral of $1/t(I)$ reaches value one, as shown in (3.3). This provides coordination between protection devices for any varying fault current magnitudes.

$$\int_0^T \frac{1}{t(I)} dt = 1 \tag{3.3}$$

3.4.1 Algorithm Detail

In this section, a nonlinear least squares algorithm is formulated to estimate the unknown parameters A, p, B and the inverse TCC of reclosers. The inputs are three-phase voltage and current measurements taken from the substation. Additionally, data preprocessing techniques are presented that transform the input voltage and current to the inputs required by the least squares formulation. Data preprocessing involves estimating the phasor current flowing through the recloser and extracting the fault currents associated with fault-clearing operations. It is assumed that the voltage and current measurements are taken from the substation.

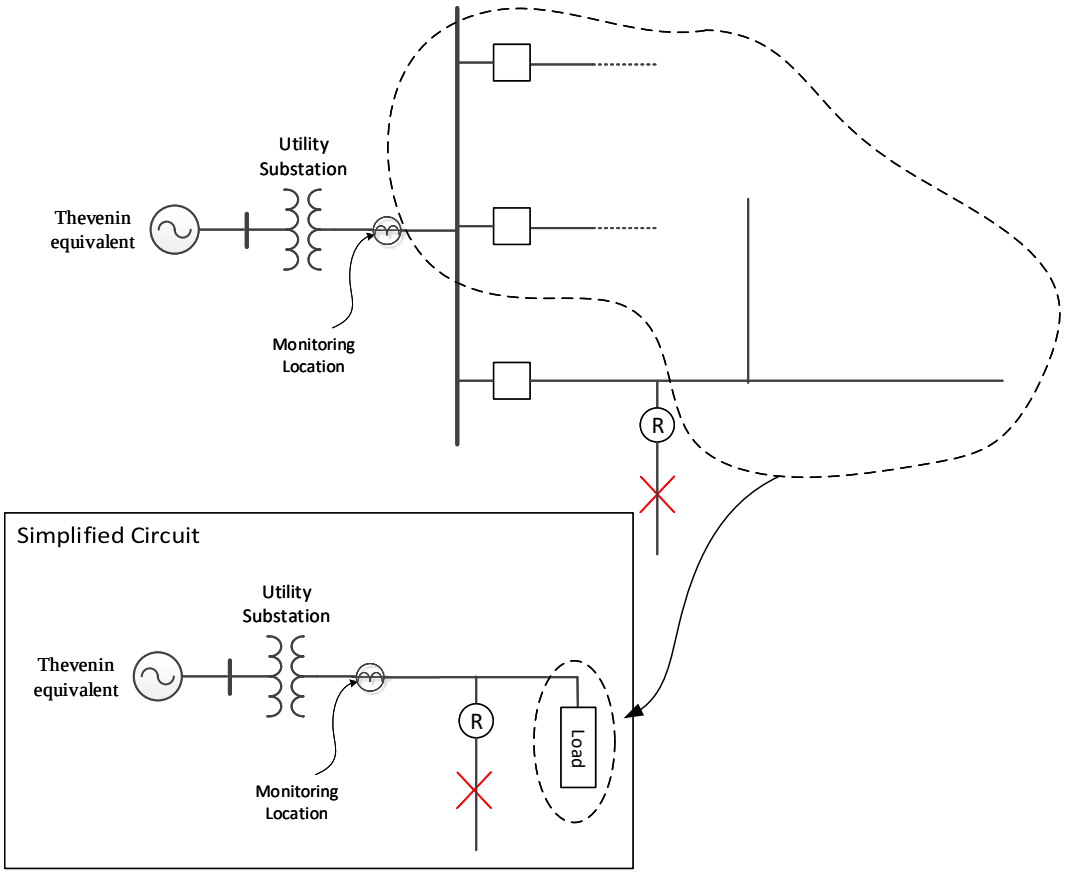


Figure 3.10: Simplified distribution circuit.

3.4.1.1 Recloser Current Estimation

First, the cosine filter [42] is used to filter harmonics and DC offsets and calculate the phasor at the fundamental frequency. The filter output at time sample m is expressed as (2.7). Using the cosine filter, three-phase voltage and current phasors at the substation are calculated. The next step is to estimate the line current flowing through the recloser. This step is necessary because the measurements are taken from the substation, and the local current

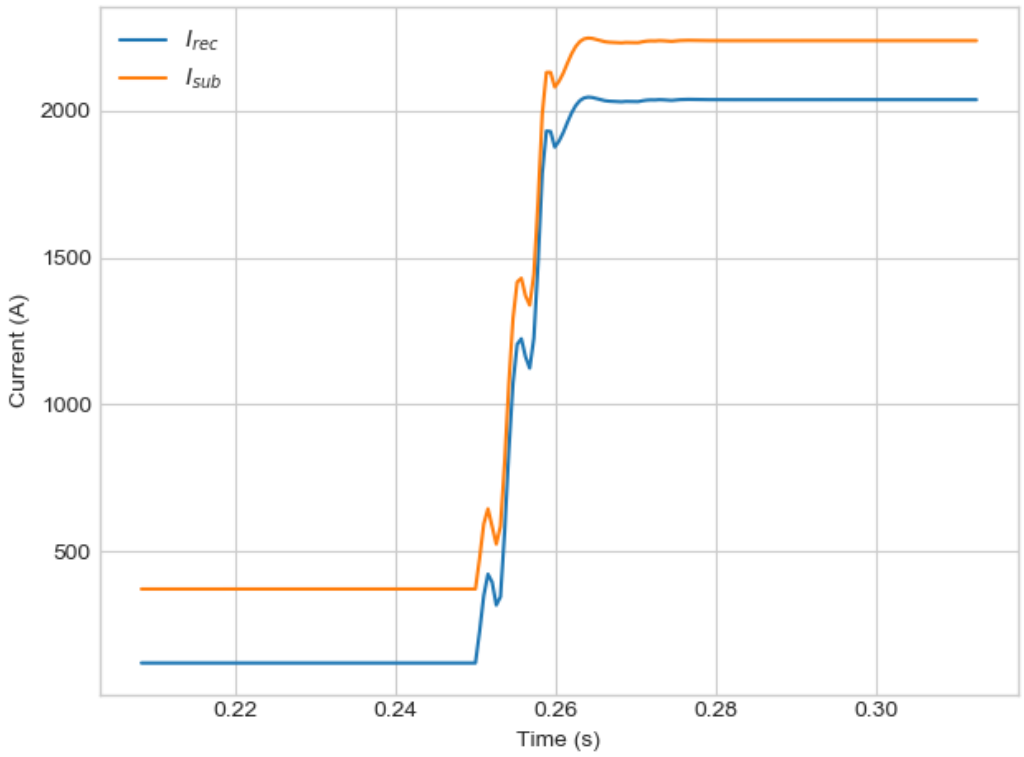


Figure 3.11: Current seen from the substation and the recloser during fault.

measurements actuating the reclosers are not directly measured. Consider a single line-to-ground fault occurring downstream from a recloser, as illustrated in Fig. 3.10. The circuit in Fig. 3.10 is modeled in PSCAD, and the fault currents measured from the substation and the recloser are shown in Fig. 3.11. Note that the current flowing through the recloser is less than the fault current measured from the substation because of load currents.

To estimate the load currents, we assume that the loads upstream from the recloser are lumped as a single impedance load at the end of the feeder,

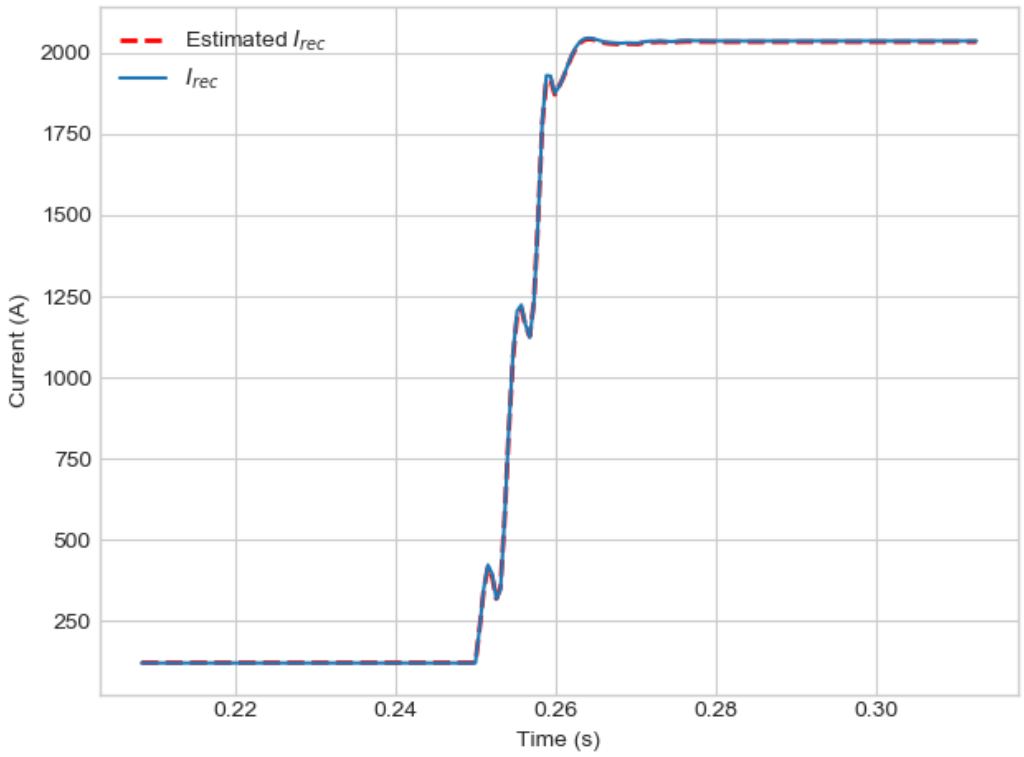


Figure 3.12: Recloser current estimated from the substation.

as shown in Fig. 3.10. This assumption is valid because the line impedance is generally much smaller than the load impedance [57].

Then zero, positive, and negative sequence currents flowing through the device are estimated, respectively using (3.4)-(3.6). The zero-sequence current measured from the substation is assumed to be identical to the zero-sequence current flowing through the device. The positive and negative sequence currents flowing through the recloser are estimated by subtracting the positive and

negative sequence currents measured from the substation by the load current.

$$I_{r0} = I_0 \tag{3.4}$$

$$I_{r1} = I_1 - \frac{V_1}{Z_{load}} \tag{3.5}$$

$$I_{r2} = I_2 - \frac{V_2}{Z_{load}} \tag{3.6}$$

where I_0, I_1, I_2 denotes the zero-sequence, positive-sequence, and negative-sequence currents at the substation and I_{r0}, I_{r1}, I_{r2} through the recloser.

Z_{load} can be estimated using the post-fault voltage and current measurements as shown in (3.7).

$$Z_{load} = \frac{V_{post,1}}{I_{post,1}} \tag{3.7}$$

I_{rA}, I_{rB}, I_{rC} can be calculated by transforming the symmetrical components, I_{r0}, I_{r1}, I_{r2}, to phase components using (3.8), where the operator $a = 1\angle 120°$.

$$\begin{bmatrix} I_{rA} \\ I_{rB} \\ I_{rC} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{bmatrix} \begin{bmatrix} I_{r0} \\ I_{r1} \\ I_{r2} \end{bmatrix} \tag{3.8}$$

Fig. 3.12 shows the estimated recloser current using substation measurements as the inputs. The figure illustrates that the estimated recloser current matches the actual current flowing through the recloser.

3.4.1.2 Recloser Operating Time Estimation

The fault detection methods described in Sections 2.3.2.1 and 2.3.2.2 are used to estimate recloser operating times. It is assumed that the recloser has first sensed the fault current and sent the trip signal at the index n_{ps} (the first time index exceeding the RMS current pickup) and the index n_{we} (the last time index exceeding the RMS pickup value, corrected by the discrete wavelet transform), respectively.

3.4.1.3 Recloser Curve Estimation Using Nonlinear Least Squares Algorithm

Using N number of fault events captured from a distribution system, a nonlinear least squares algorithm can be formulated to estimate the TCC parameters, as in (3.9). The objective function is formulated on the basis of (3.3), which is the standard equation in IEEE Std C37.112-2018 [29].

$$\arg\min_{A,p,B} \sum_{i} [f(\beta, X_i) - y]^2 \tag{3.9}$$

where,

$$\beta = \begin{bmatrix} A & p & B \end{bmatrix}^T \tag{3.10}$$

$$X = \begin{bmatrix} I_{f,1}(n) & I_{f,2}(n) & ... & I_{f,N}(n) \end{bmatrix}^T \tag{3.11}$$

$$y = \begin{bmatrix} 1 & 1 & ... & 1 \end{bmatrix}^T \tag{3.12}$$

$$f(\beta, X_i) = \sum_{j=1}^{N} \left[\frac{1}{t_{inverse}(X_{i,j})} \right] \times \Delta \tag{3.13}$$

$$t_{inverse} = \frac{A}{X^p - 1} + B \tag{3.14}$$

$$\Delta = \text{RMS sampling interval} \tag{3.15}$$

This formulation returns the parameters (3.10), which can be used to construct the empirical TCC curve. Recall from (3.3), the integration of the inverse TCC curve should yield one. This characteristic is defined in (3.13-3.15), where the input (3.11) is integrated to the value of ones (3.12). Note that in (3.11), the ith row vector of X is extracted from a single fault event, using the estimated quantities $I_{r,i}$, $n_{ps,i}$, and $n_{we,i}$ defined in Sections 3.4.1.1 and 3.4.1.2.

$$I_{f,i}(n) = \begin{cases} I_{r,i}(n + n_{ps,i} - 1) & 1 \leq n \leq n_{we,i} - n_{ps,i} + 1 \\ 0 & n_{we,i} - n_{ps,i} + 1 < n \leq M \end{cases} \tag{3.16}$$

where M corresponds to the maximum value of $n_{we,i} - n_{ps,i} + 1$ among all fault events.

Output parameters A, p, B are then used to construct the empirical recloser operating TCC curve. In this demonstration, (3.9) is solved using the Levenberg-Marquardt algorithm implemented in Python [58].

Reclosers in distribution systems can operate on either the phase current or the zero-sequence current. To estimate the TCC curve parameters for ground current, I_{r0} in (3.4) can be used as I_r.

3.4.2 Distribution System Applications

In this section, field events recording fault-clearing operations are used to estimate the TCC curve parameters of the reclosers installed in the system. Fault measurements are captured by a power quality monitor located at the substation of a 24.9 kV distribution system. Two types of reclosers are installed in the system: hydraulically controlled reclosers (Type I), operating in fast (Curve I-F) or delayed (Curve I-D) operations, and an electronically controlled recloser (Type II, Curve II), which is located upstream from a Type I recloser. Hydraulic reclosers operate in single-phase tripping mode–that is, a recloser on each phase senses the fault current and trips individually–and Type II reclosers operate in three phases. To demonstrate the proposed method, we have collected 44 fault events from this system; 29 faults are cleared using Curve I-F, 11 faults are cleared using Curve I-D, and 4 faults are cleared using Curve II.

Using the algorithm proposed in Section 3.4.1, three recloser TCC curves (Curve I-F, I-D, and Curve II) and the parameters (A, p, B) of each TCC curve are derived from the fault events. Table 3.5 shows the estimated parameters (A, p, B) of Curve I-F, Curve I-D, and Curve II clearing the fault events, which are estimated from (3.9). Fig. 3.13 shows the empirical recloser curve constructed using the estimated parameters. The scatter plot for each of the fault events is also depicted in this figure.

Note that the scatter plot was generated assuming constant fault current magnitude to illustrate the algorithm in simple and intuitive manners.

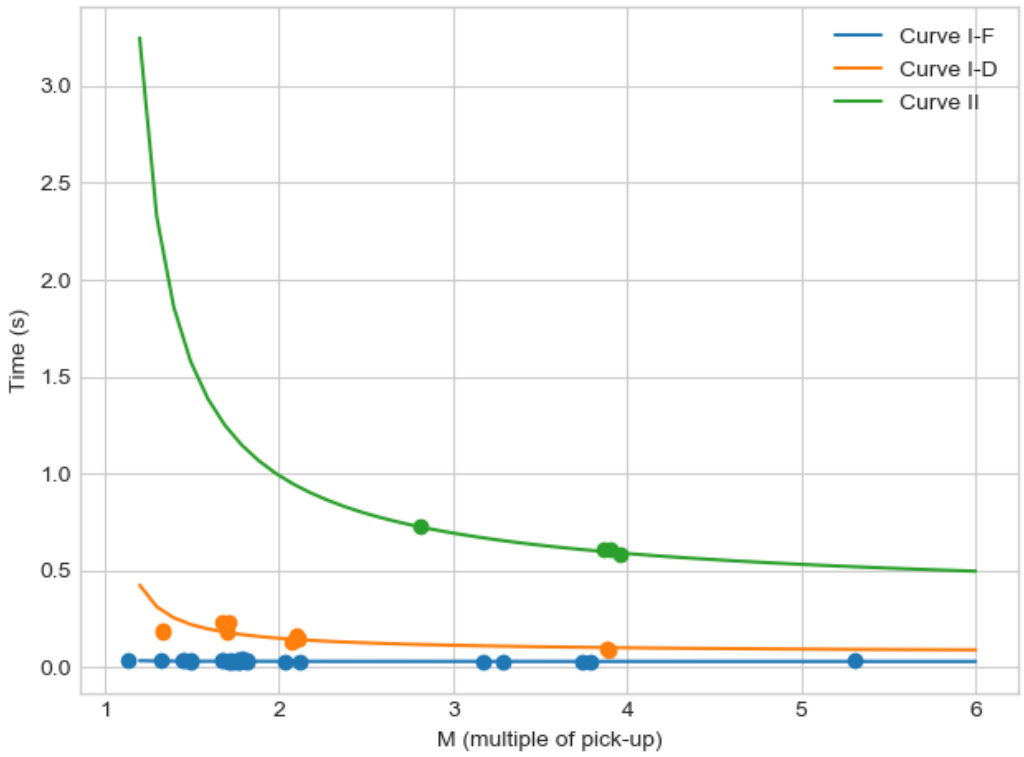


Figure 3.13: Empirical estimation of inverse time-current characteristics of Curve I-F, Curve I-D, and Curve II.

The proposed method does not make such an assumption, so the algorithm is applicable for varying fault current as well.

It should also be noted that having more fault data results in more reliable modeling of recloser curves. In this demonstration, only four fault events are collected for Curve II operation. It is possible that the modeled curve is overfitted because of this limited number of the datasets. Nonetheless, the results are shown along in this section for reference.

Table 3.5: Estimated TCC parameters using the proposed method.

	A	p	B
Curve I-F	-1.9665e-06	-2.7737e-03	2.7914e-02
Curve I-D	7.9850e-05	1.17356e-03	4.9850e-02
Curve II	0.0101	0.01803	0.1874

3.4.3 Evaluation of Recloser Operation

This section briefly demonstrates how the empirical recloser curve models are used to evaluate recloser operations in distribution systems.

3.4.3.1 Comparison of Empirical Recloser Curve with Manufacturer Specifications

For the Type I recloser used in this demonstration, the manufacturer provides the maximum operating time for the fast operation and the average operating time (±10%) for the delayed operation. Fig. 3.14 shows the empirical recloser Curve I-F constructed using the proposed algorithm and the operating time specified by the manufacturer. The empirical Curve I-F is below the manufacturer's maximum operation time, and therefore, the recloser is operating as expected. Similarly, Fig. 3.15 shows the comparison between the empirical Curve I-D and the delayed curve specified by the manufacturer. The empirical delayed curve also matches the expected operating time of the recloser with a small mismatch. Therefore, it can be concluded that the recloser is behaving as expected in the delayed operation.

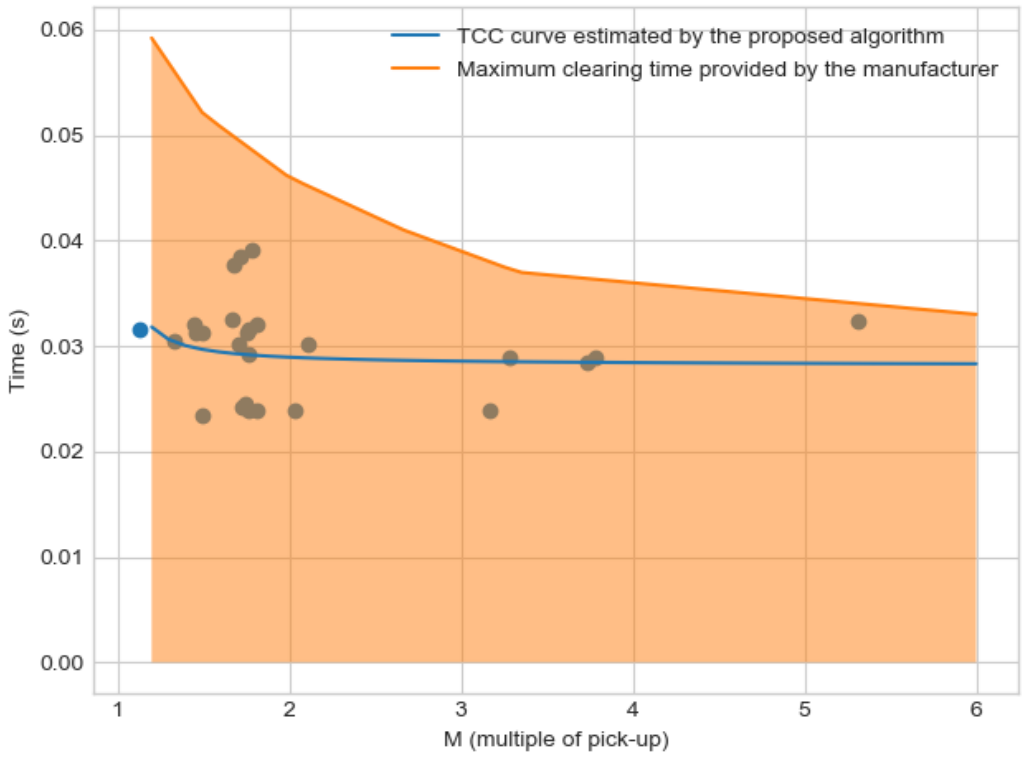


Figure 3.14: Evaluation of recloser fast curve using empirical Curve I-F.

3.4.3.2 Evaluating Individual Recloser Operation

The ith residual in the objective function (3.9) is used to evaluate the individual operation of the reclosers. Simple rules can be used on the residuals to evaluate recloser operations.

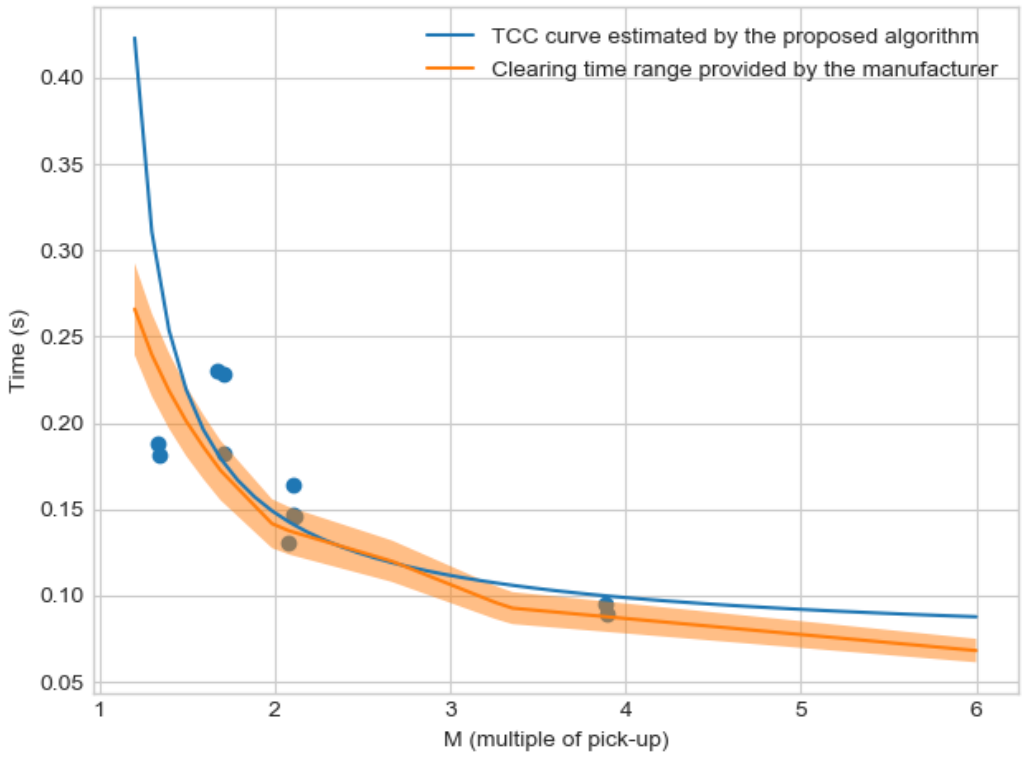


Figure 3.15: Evaluation of recloser slow curve using empirical Curve II-D.

$$\text{if } r_i = f(\beta, X_i) - 1 < 0,$$

recloser operated faster than normal

$$\text{if } r_i = f(\beta, X_i) - 1 > 0,$$

recloser operated slower than normal

where $\beta = [A\ p\ B]^T$ are the estimated parameters of the recloser curve and function f is defined in (3.13). Normal operation is defined as the recloser operating using the timings in the empirical recloser curve.

If the value of r_i is positive, the recloser is operating slower than normal. Similarly, if the value of r_i is negative, the recloser is operating faster than normal. In Fig. 3.16, box plots of the residuals for Curve I-F, Curve I-D, and Curve II are shown. A single dot in a box plot represents the ith residual, r_i, of a recloser operation. The bottom and top of the box represent the first and third quartiles, respectively, and the ends of the whiskers represent the $\text{Q1} - 1.5\text{IQR}$ and $\text{Q3} + 1.5\text{IQR}$, where Q1, Q3, and IQR are the first and third quartiles and the interquartile range Q3 - Q1, respectively. These ranges can be used to provide further details on delay times of the recloser operations, such as that the data points located farther from the median correspond to the recloser operating slower or faster. For , in Fig. 3.16 Curve I-F, six fault events are found between $\text{Q3} + 1.5\text{IQR}$ and Q3. These faults are slower to clear than the other fault events.

Outliers are defined as data points that lie above $\text{Q3} + 1.5\text{IQR}$ or below $\text{Q1} - 1.5\text{IQR}$. These data points lie outside the whiskers in the box plot. Although there were no outliers in Fig. 3.16, outliers are considered to be fast or slow operations that are very unlikely to happen and could be used to detect recloser misoperations.

3.4.4 Identifying Fault-Clearing Devices

This section uses the one vs. all (OvA) method, a popular strategy in multiclass classification problems, to identify the fault-clearing recloser curves in distribution systems using the empirical recloser curves. A single binary

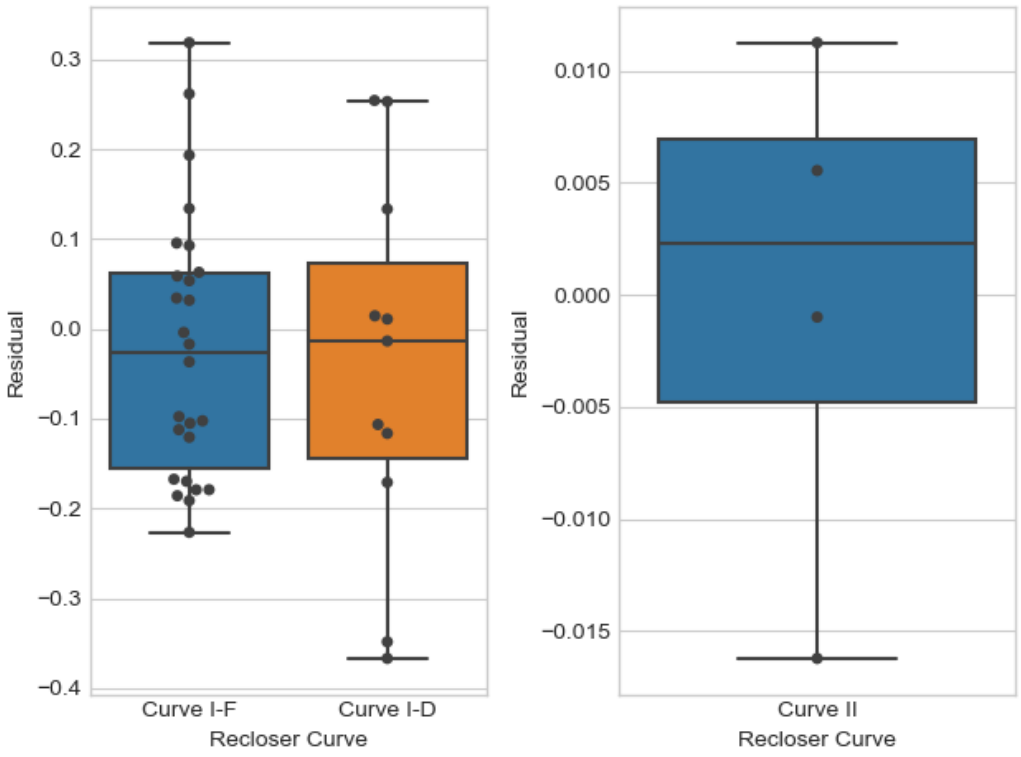


Figure 3.16: Residual box plots of Curve I-F, Curve I-D, and Curve II.

classifier determines whether a fault is cleared from a specific recloser curve. A binary classification rule, made using the empirical recloser curve model, is as follows:

if $r_i = f(\beta, X_i) - 1 <$ Q3 and

$r_i = f(\beta, X_i) - 1 >$ Q1,

fault is cleared from the recloser curve (high)

else if $r_i = f(\beta, X_i) - 1 >$ Q1-1.5IQR or

$r_i = f(\beta, X_i) - 1 <$ Q3+1.5IQR,

fault is cleared from the recloser curve (moderate)

else if $r_i = f(\beta, X_i) - 1 <$ Q1-1.5IQR or

$r_i = f(\beta, X_i) - 1 >$ Q3+1.5IQR,

fault is not cleared from the recloser curve

3.4.4.1 Validation using Field Data

Fig. 3.17 shows a single line-to-ground fault on phase C that occurred in the 24.9 kV distribution system. This fault was cleared by the delayed operation of the recloser (Curve I-D). Now, assume that we do not know which recloser has cleared this fault but have the empirical inverse time-current characteristics of the three recloser curves within the system. The classification rule defined in the previous section can be used to identify which recloser has cleared this fault. First, calculate the values of r_i using the three empirical recloser curve parameters (A, p, B of Curve I-F, Curve I-D, and Curve II). The three values are 5.23, 0.01, -0.85, respectively. Fig. 3.18 shows these residual

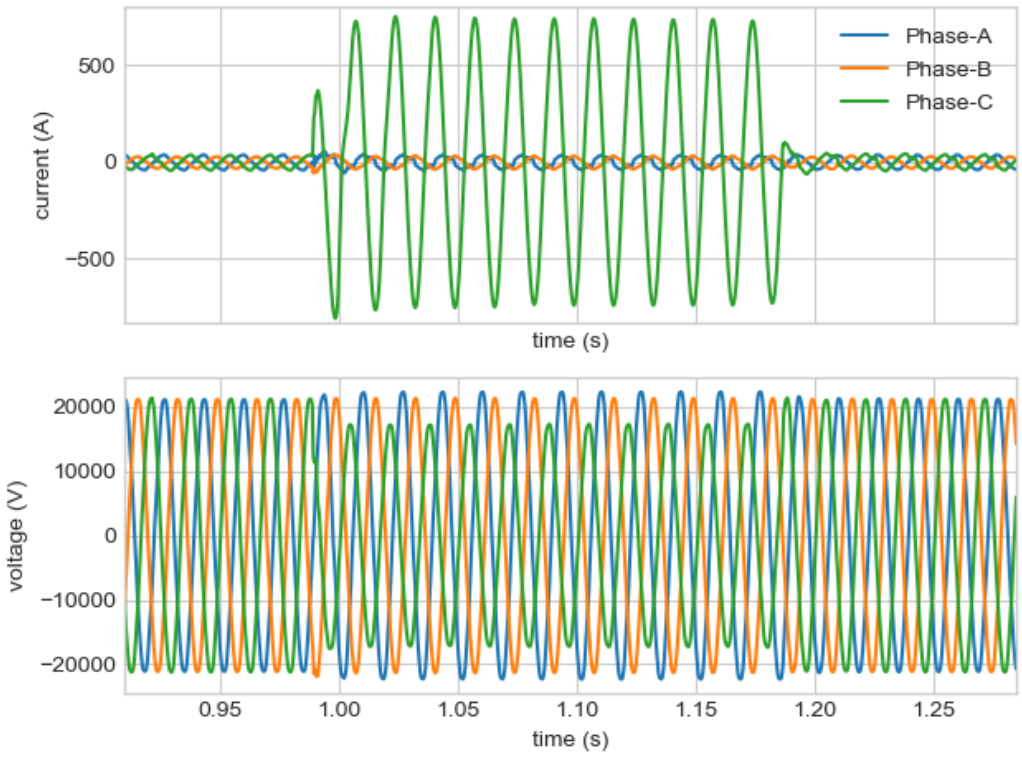


Figure 3.17: Single line-to-ground fault on phase C. Voltage and current waveforms recorded from the substation.

values on top of the three box plots of the recloser curves. Note that the first residual value, 5.23, is greater than Q3 + 1.5IQR of the box plot of Curve I-F. From the classification rule, this fault event is not cleared by Curve I-F.

On the other hand, the second residual value, 0.01, is within the whiskers of the box plot of Curve I-D. It can be concluded that this fault has been cleared by recloser curve I-D with a high confidence level. Similarly, the third residual value, -0.85, is below the Q1 - 1.5IQR of the box plot of Curve II, and therefore this fault event is not cleared by Curve II.

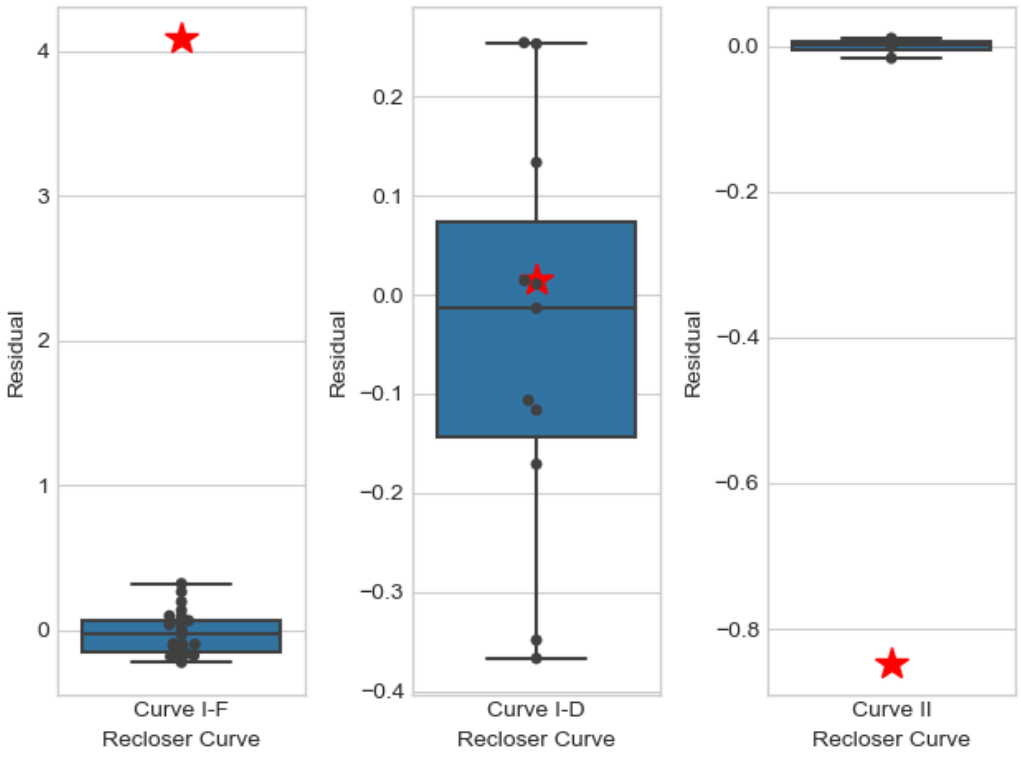


Figure 3.18: Residual box plot classification results of the fault data that occurred in the distribution system.

3.4.4.2 Validation Using Digital Relay Event Report

This section validates the identification algorithm using a test bench designed with a signal generator and digital relay. This test bench allows the simulation of various fault scenarios at any fault magnitudes without physically short-circuiting the lines and also provides the labels to validate the identification results. To emulate multiple fault scenarios, the output channels of the signal generator are connected in serial with the digital relay. The time-overcurrent elements in the digital relay are programmed to time the

fault clearing using the empirical curves Curve I-F and Curve I-D. After simulating multiple fault scenarios, the event reports stored in the digital relay are downloaded and used as the input to the classification algorithm. The event reports include fault measurements (voltage and current) and the status of the time-overcurrent element, which can be used for this validation. The faults currents are simulated at different fault magnitudes (1.5 to 10 times the current pickup). Evolving faults are also used in this study.

Fig. 3.19 shows the box plots of the three recloser curves in the system (Curve I-F, Curve I-D, and Curve II). On top of these three box plots, the residuals of the recloser operations emulated from the test bench are depicted. Note that the blue stars represent the emulated recloser operations of Curve I-F, and the red stars represent the operations of Curve I-D. According to the classification rule, it is shown that the blue stars (emulated recloser operations of Curve I-F) are within the box plot whiskers of Curve I-F, and the red stars (emulated recloser operations of Curve I-D) are within the box plot whiskers of Curve I-D. This result validates that the classification rule can be used to identify the recloser curve clearing the fault.

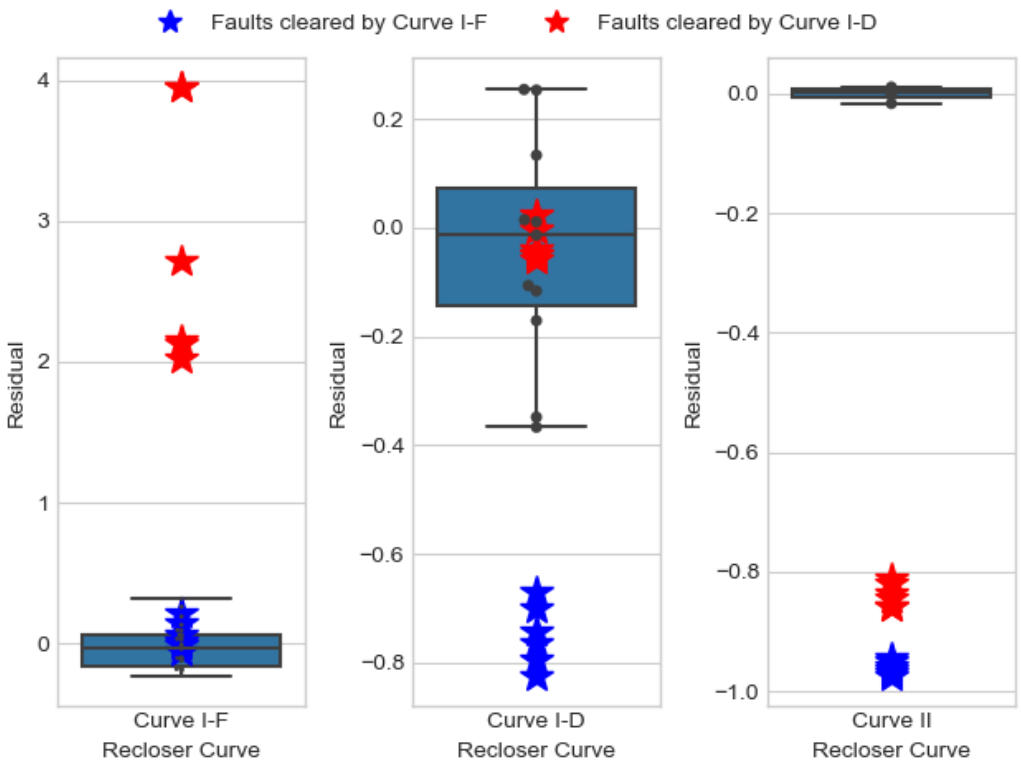


Figure 3.19: Residual box plot classification results using digital relay data implementing Curve I-F and Curve I-D.

For further validation, the fault datasets recording the fault-clearing operations of Curve I-D are modified to emulate a new recloser in the feeder. The empirical curve of the recloser, Test Curve, is estimated from the modified fault datasets. Fig. 3.20 shows the two estimated empirical curves: Curve I-D and Test Curve.

The event reports of Curve I-D generated from the test bench system are used to evaluate whether Curve I-D operations can be identified from the Test Curve operations. From Fig. 3.21, the residuals calculated from the Test

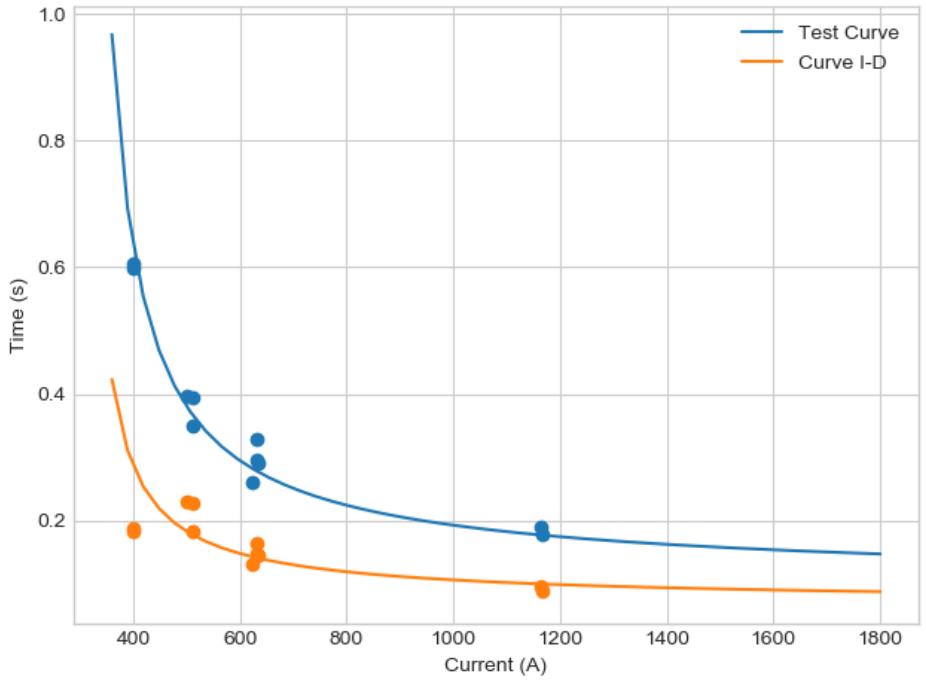


Figure 3.20: Empirical estimation of inverse time-current characteristics of Curve I-D and Test Curve.

Curve model are much smaller than the Q1-1.5IQR of the Test Curve box plot. On the other hand, the residuals calculated from the Curve I-D model are within the first and third quartiles of the Curve I-D box plot. Therefore, input fault events are successfully classified as Curve I-D operations.

It should be noted that the final classification result may be more than one recloser curve or none. These scenarios could occur if the time intervals between the TCC curves are very small or a recloser misoperates. In this case, the fault-clearing curve can be identified using the smallest r_i, or making a decision is rejected.

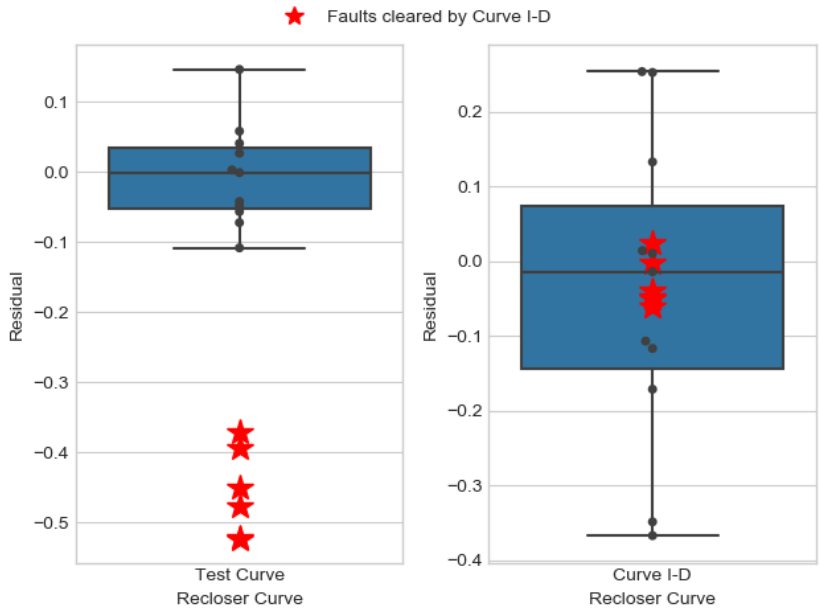


Figure 3.21: Identifying Curve I-D operations from Test Curve operations.

3.5 Summary

This chapter has presented algorithms for evaluating and identifying overcurrent protection devices. First, the rule-based algorithm was developed to classify whether a fault was cleared by a recloser or a fuse. Necessary features associated with the operations of the recloser and fuse were explained and implemented. Next, a methodology for estimating empirical TCC curves was described. The empirical curve was used to further narrow down the specific device that clear the fault. The empirical curve can help narrow down the fault locations, evaluate protection coordination, and detect possible misoperation of reclosers.

Chapter 4

An Extensible, Open Framework for Power Quality Disturbance Events

4.1 Introduction

[1] Although power quality (PQ) disturbance events such as RMS variations and transients occur in transmission and distribution systems, these datasets are mostly managed through proprietary solutions in different data formats. This chapter presents a simple, yet effective, metadata database schema to manage voluminous PQ disturbance events in power systems. A database constructed using the schema can be used to store the metadata providing descriptive and analytical analysis of the disturbance events. The proposed schema defines five classes to store the metadata associated with PQ disturbance events: event, time-series, description, software analysis, and IEEE disturbance classification. The classes that form the schema are presented, then are demonstrated using actual disturbance events captured from a distribution system.

[1] Parts of this chapter have been published in, K. W. Min, A. F. Bastos, S. Santoso, and U. Karadkar, "An extensible, open framework for power quality disturbance events," in *Proc. IEEE/PES Transmission and Distribution Conf. and Expo. (T&D)*, Apr. 2018, pp. 1-9. The author of this dissertation designed the framework and analyzed the data in the paper.

The remainder of this chapter is organized as follows. Section 4.2 presents an overview and the approach adopted for developing our schema. The classes defined in the schema and their properties are described in Section 4.3. Finally, Section 4.4 illustrates sample power quality disturbances modeled using this schema.

4.2 Overview and Approach

The PQLD schema is designed to foster the sharing of data regarding power disturbance events described in the IEEE 1159-2009 standard [45]. To facilitate broad adoption, the schema lowers the barrier to entry by minimizing the mandatory field–and hence, the amount of data–that must be provided by data sources (such as power companies and existing publicly available data sources). The schema takes an incremental approach to data publishing–sources may share minimal data initially, and both sources and third parties may enhance published data by posting their analyses. The schema includes elements to describe signal-based event characteristics as well as human-generated ground truth labels in order to maximize its utility in developing and validating PQ software analytics tools. Most conventional information or database models in power systems have adopted entity-relationship (ER) models [34, 35, 36, 59], which are best suited for structured data–records that possess identical properties. However, characteristics of disturbance events vary depending on their root causes. For example, a short-circuit fault causes RMS variations of voltage and currents, while harmonic resonance may occur

after a capacitor is switched on. Thus, ER-based power quality schemas either are non-normalized or have significant empty fields, as the data fields are sparsely populated. Document-based schemas provide a much better model for these semistructured events. Therefore, the open schema used here adopts a document model, allowing for flexibility of description in event characteristics.

4.3 Power Quality Linked Data Schema

Fig. 4.1 illustrates the overview of the PQLD schema classes. The schema consists of five classes: event, time-series, description, software analysis, and IEEE disturbance classification, which together describe the characteristics of an event. Each class stores relevant metadata using meaningful properties, which record the details related to an event. All recorded events are stored using event objects. Thus, the event class serves as the base class, and objects of other classes are used as necessary.

Tables 4.1 through 4.5 present the most significant properties of each class. The complete schema is available through the project wiki at `https://wikis.utexas.edu/display/pqmetadata/`. To succinctly convey key aspects of a much larger schema, the tables include representative columns that best convey the characteristics for each class. Different tables thus include different columns; however, the master schema includes all these columns and more to adequately document the necessary and sufficient conditions for each property of a power event. The columns in the tables are as follows: "Label" includes human-readable descriptions of the event properties. The "Property"

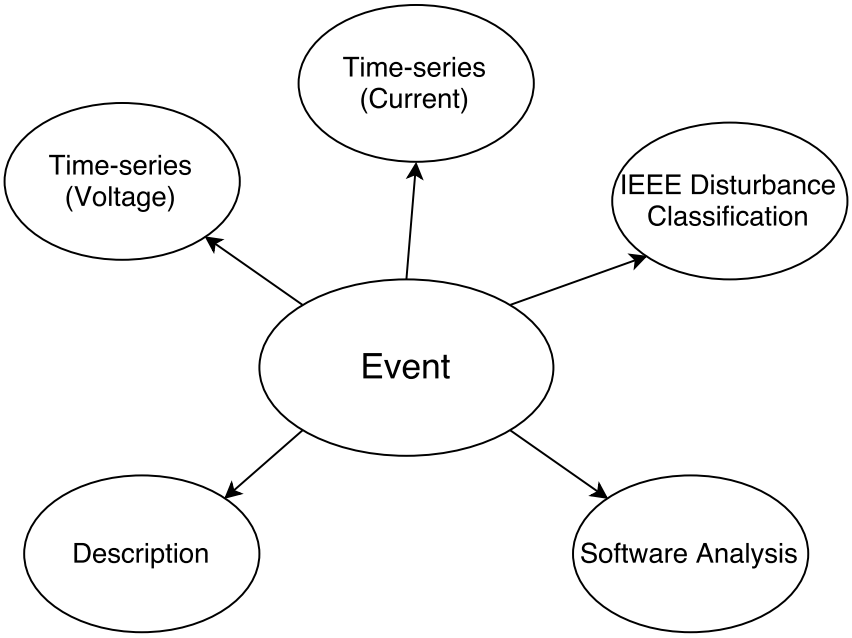


Figure 4.1: Overview of the presented schema structure.

column provides the computational counterparts to the human-readable labels. These are the property descriptors that appear in the database, and values are associated with these property titles. "Source" lists who is responsible for recording each property. "Vocab Schema" articulates the controlled vocabularies of permissible values. "Obligation" indicates the cardinality of the properties–whether a property may or must appear for an event to be considered valid, and how many instances of the property are permissible.

Controlled vocabularies list acceptable values for categorical data, thus preventing the possibility of different events describing semantically similar values differently, as well as guarding against the inclusion of invalid values. The subsections below describe the key characteristics of each table.

4.3.1 Event

This core class describes essential characteristics of an event, such as the location of the power quality monitor, the time the event file was created, rated voltage and system frequency, and the root cause. In addition, voltage, current, description, analysis, and IEEE classification are aggregated into the event class. Note that a data provider such as a power company needs to provide only those details that cannot be obtained post facto. Thus, the data provider is not required to provide the IEEE classification, which can be calculated by an analyst later.

4.3.2 Time-Series

Time-series data such as instantaneous voltage and current measurements are recorded using this class. As current and voltage data share the same characteristics, the "headers" property helps identify which of these attributes the data belongs to. This class is reusable and, in fact, is used for recording both the current and voltage properties in Table 4.1.

For example, an array [Va, Vb, Vc] indicates that the time series stores line voltage data, in the phase order indicated. While the measured values must originate from the data provider, an analyzing organization may further populate properties such as the count of data values in the series as well as the sampling rate at a later time. Unlike the labels, which use phrases, property names are encoded in lower camel case.

Table 4.1: Class Event

Label	Source	Vocab Schema	Obligation
Location	Data Provider	Substation, feeder, service entrance, ...	0-1
Nominal Voltage	Data Provider or Analyzer		0-1
System Frequency	Data Provider or Analyzer	50, 60	0-1
Event Time	Data Provider or Analyzer		0-1
Root Cause	Data Provider or Analyzer	Short-circuit fault, Cap. switching, ...	0-1
Voltage	Data Provider		1
Current	Data Provider		1
Description	Data Provider		0-1
Analysis	Data Provider or Analyzer		0-1
IEEE Classification	Data Provider or Analyzer		0-n

4.3.3 Description

This class records metadata that are associated with the root cause of an event. Properties such as weather, device clearing the fault, internal agent, and external agent are included in this class. Internal agents refer to electrical components within the system that caused the event or have failed as a result of the event. External agents refer to factors outside the electrical system that caused the event. For example, an animal such as a squirrel may climb into a

Table 4.2: Class Time-series

Label	Source	Property	Obligation
Header	Data Provider or Analyzer	.headers	1
Time-series	Data Provider	.timeSeries	1
Count of data samples	Data Provider or Analyzer	.count	1
Sampling rate	Data Provider or Analyzer	.samplingRate	1

Table 4.3: Class Description

Label	Source	Vocab Schema	Obligation
Weather	Data Provider	Clear day, wind, rain, snow, ...	0-n
Isolation equipment	Data Provider	Breaker, recloser, fuse, ...	0-1
Internal agent	Data Provider	Transformer, insulation, lightning arrester, ...	0-1
External agent	Data Provider	Animal, tree, lightning, ...	0-1

pole transformer, causing a short-circuit fault. In this case, the squirrel and the transformers are the internal and external agents.

Table 4.4: Class Software Analysis

Label	Source	Vocab Schema	Obligation
Minimum RMS voltage	Data Provider or Analyzer		0-1
Maximum RMS voltage	Data Provider or Analyzer		0-1
Maximum RMS current	Data Provider or Analyzer		0-1
Peak instantaneous current	Data Provider or Analyzer		0-1
Peak instantaneous voltage	Data Provider or Analyzer		0-1
Real power variation	Data Provider or Analyzer		0-1
Reactive power variation	Data Provider or Analyzer		0-1
Resonant frequency	Data Provider or Analyzer		0-1
Power frequency variation	Data Provider or Analyzer		0-1

Table 4.5: Class IEEE Classification

Label	Source	Vocab Schema	Obligation
Category label	Data Provider or Analyzer	Table II in [45]	1
Category ID	Data Provider or Analyzer	Table II in [45]	1

4.3.4 Software Analysis

This class describes the quantitative analytics of the PQ disturbance event. It is possible to complete all the properties in this class during the anal-

ysis phase and record only abnormal values as desired by the analysts. When analytics of the disturbance events are not available from the data provider, third-party analytics modules [7, 12, 60] can be used to provide the analysis.

4.3.5 IEEE Classification

The events are categorized per IEEE Standard 1159-2009 [45] in IEEE classification class. For example, if a short-circuit fault has caused a voltage sag to 0.8 per unit for the duration of two cycles, this event is categorized as an instantaneous voltage sag as defined by the standard. While the obligation for the classification properties is 1, the obligation for the IEEE classification property in the event class is 0-n, indicating that an event may not yet be classified or may have several IEEE classifications. However, each category instance must have both an ID and a label.

4.4 Sample Distribution Events

This section describes encoding of disturbance events in the PQLD schema. The root causes for these events are a short-circuit fault and capacitor switching. Following the obligation field in the schema specification, note that the document-based data representation allows metadata entries to simply omit property names when the corresponding values are not available or the properties are not applicable. For example, the minimum RMS voltage is not provided in the case of capacitor switching because it does not provide any useful information.

Table 4.6: Demonstration results: short-circuit fault

Class	Label	Example (Fault)
Event	Location	Substation
	Nominal Voltage	25 kV
	Event Time	3/09/2010 23:05:30 PM
	Root Cause	Short-circuit fault
	System Frequency	60 Hz
Voltage	Header	['Va', 'Vb', 'Vc']
	Time-series	[Va(n), Vb(n), Vc(n)]
	Count of data samples	512
	Sampling rate	32
Current	Header	['Ia', 'Ib', 'Ic']
	Time-series	[Ia(n), Ib(n), Ic(n)]
	Count of data samples	512
	Sampling rate	32
Description	Isolation equipment	Recloser
	External agent	Lightning
Analysis	Minimum RMS voltage	6.05 kV
	Maximum RMS voltage	14.78 kV
	Maximum RMS current	2384 A
	Peak instantaneous current	3152 A
IEEE Classification	Category label	Sag, instantaneous
	Category ID	2.1.1.

4.4.1 Short-Circuit Fault

Fig. 4.2 shows a single line-to-ground fault that occurred on phase C caused by lightning in a 25 kV distribution system. A short-circuit fault

results in a high increase in the current on the faulted phase(s). This increase in current and the Thevenin equivalent source impedance at the monitoring location induced voltage sag or swell on the faulted and healthy phases. In this event, the current magnitude on phase C increased to 2.38 kA and lasted for approximately 9.7 cycles until a recloser isolated the fault. During the fault, the voltage magnitude in the faulted phase dropped below 0.9 per unit, and thus the event is categorized as an instantaneous sag according to IEEE Standard 1159-2009. The associated metadata values of the short-circuit fault are stored following the presented schema, as shown in Table 4.6.

4.4.2 Capacitor Switching

Fig. 4.3 shows the oscillatory transient due to the energizing of a 600 kvar three-phase capacitor bank, located downstream from the PQ monitor. Phase C waveforms are most affected by the oscillatory transient; the transient voltage and current reached a maximum of 22.64 kV and 317.3 A, respectively. This transient event lasted approximately 0.15 cycle, with a resonant frequency of 540 Hz. The reactive power flow decrease at the substation was 187.04, 189.04, and 189.32 kvar in phases A, B, and C, respectively. The energizing of this bank increased the RMS voltage by 0.2834%, decreased the RMS current by 5.2960%, and moved the power factor from 0.9395 to 0.9994. This event is classified as a low frequency oscillatory transient according to IEEE Standard 1159-2009. The associated metadata values of the capacitor switching are stored following the presented schema, as shown in Table 4.7.

Table 4.7: Demonstration results: capacitor switching

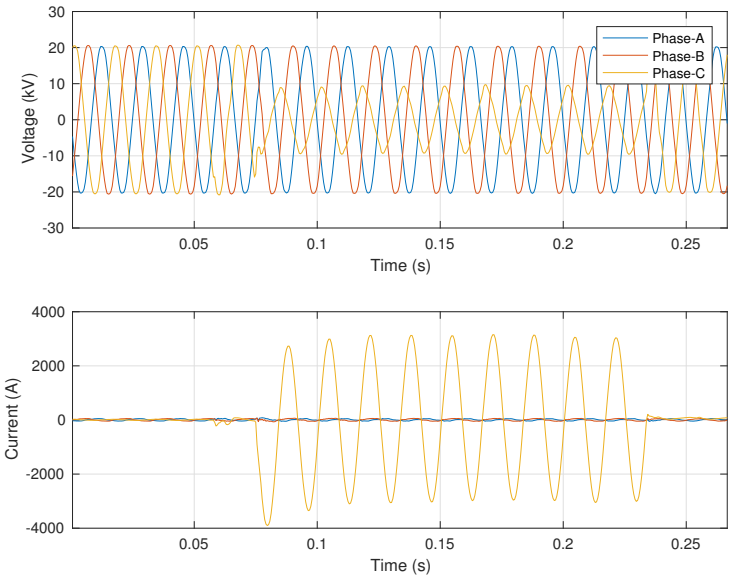
Class	Label	Example (Cap. Switching)
Event	Location	Substation
	Rated Voltage	25 kV
	Event Time	9/29/2015 3:59:00 PM
	Root Cause	Capacitor Switching
	System Frequency	60 Hz
Voltage	Header	[‘Va’, ‘Vb’, ‘Vc’]
	Time-series	[Va(n), Vb(n), Vc(n)]
	Count of data samples	250,000
	Sampling rate	128
Current	Header	[‘Ia’, ‘Ib’, ‘Ic’]
	Time-series	[Ia(n), Ib(n), Ic(n)]
	Count of data samples	250,000
	Sampling rate	128
Analysis	Peal instantaneous voltage	22.64 kV
	Peak instantaneous current	317.3 A
	Reactive power variation	[187.04, 189.04, 189.32] kvar
	Resonant frequency	540 Hz
IEEE Classification	Category label	Low-frequency, oscillatory
	Category ID	1.2.1.

4.5 Summary

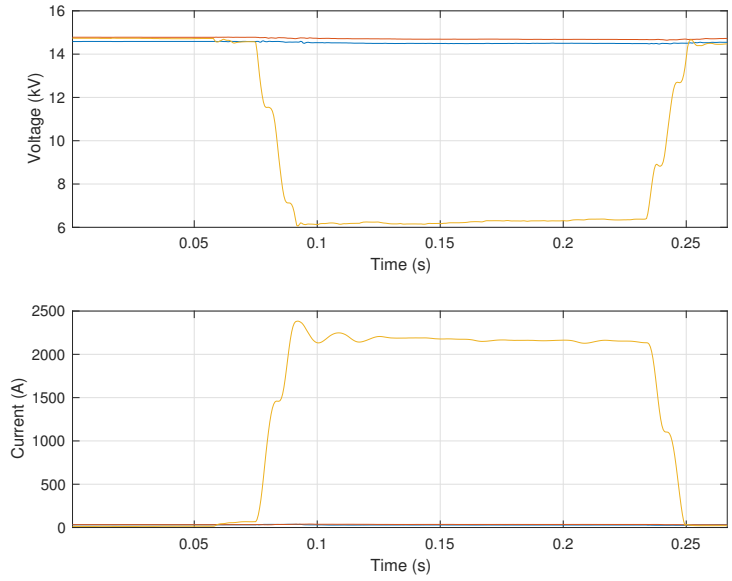
The PQLD schema is designed to promote low-overhead, incremental description and sharing of PQ disturbance events. The schema is standards-

based, balances human readability and computation support, and is openly available. The document-oriented schema supports the conditional expression of properties that are most relevant for each recorded event and avoids the representation of unnecessary properties.

The use of this schema in observed events has been demonstrated by populating properties in the presented classes. The schema supports the expression of instantaneous measurements, descriptive analysis, and quantitative analysis, as well as the real-world causes for power quality disturbances. The sample set of described events continues to broaden, and the linked data model supports the unique identification of published events.

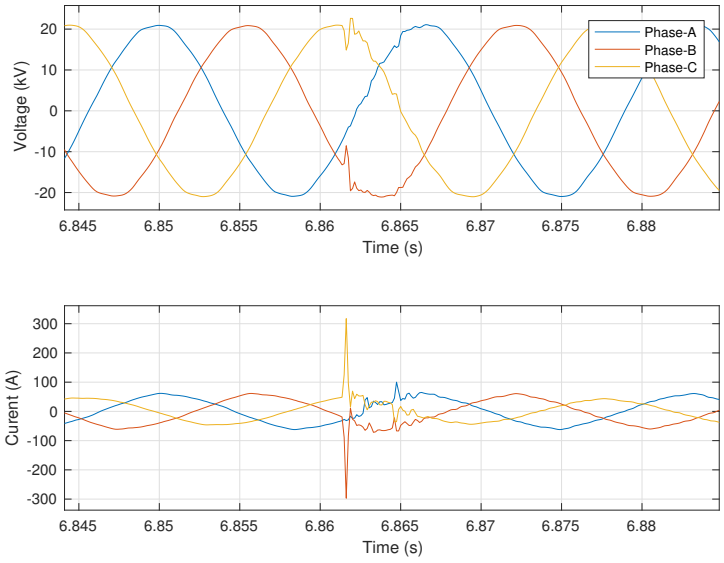


(a)

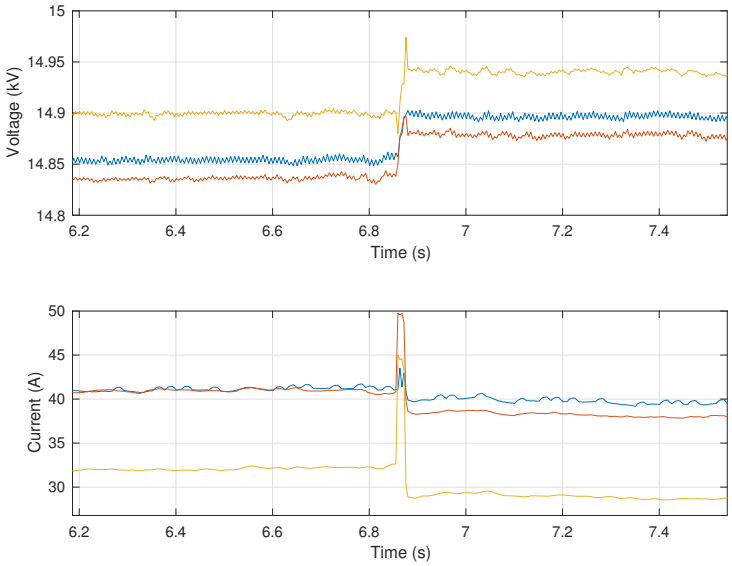


(b)

Figure 4.2: Demonstration fault event: (a) instantaneous and (b) RMS waveforms.



(a)



(b)

Figure 4.3: Demonstration capacitor-switching event. (a) instantaneous and (b) RMS waveforms.

Chapter 5

Conclusion

This dissertation aims to develop the data analytics modules to analyze short-circuit faults and overcurrent protection devices clearing the faults. Chapter 2 presents the data analytics for short-circuit faults. The analysis includes detection and categorization of fault events and estimating the fault location. The fault-locating method described in this chapter is used to remove the effects of DC offset, which can cause location estimate errors when used with conventional phasor-estimating algorithms such as the Fourier and cosine filters. Chapter 3 describes the data analytics for evaluating and identifying overcurrent protection devices. The chapter presents an algorithm that identifies the type of device clearing the fault: whether a fault is cleared by a recloser or a fuse. The status of the recloser, whether the recloser has successfully cleared the fault or locked out, can also be estimated through this process. In addition, a TCC curve estimation algorithm is proposed to evaluate and identify fault-clearing devices. An empirical TCC curve is estimated using multiple fault events collected from a distribution circuit. The empirical TCC curve can be used to evaluate the timings of the recloser operations with respect to the manufacturer specifications and to narrow down which specific recloser curve timing has been used to clear the fault. Finally, Chapter 4

presents the PQLD schema for power quality disturbance events. The schema facilitates storing and sharing of power quality measurements data, with descriptive and analytical analysis, and the root causes of the disturbance events.

Bibliography

[1] M. H. J. Bollen, I. Y. H. Gu, S. Santoso, M. F. Mcgranaghan, P. A. Crossley, M. V. Ribeiro, and P. F. Ribeiro, "Bridging the gap between signal and power," *IEEE Signal Process. Mag.*, vol. 26, no. 4, pp. 12–31, July 2009.

[2] K. Hur and S. Santoso, "Estimation of system damping parameters using analytic wavelet transforms," *IEEE Trans. Power Del.*, vol. 24, no. 3, pp. 1302–1309, July 2009.

[3] K. Hur, S. Santoso, and I. Y. H. Gu, "On the empirical estimation of utility distribution damping parameters using power quality waveform data," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 175–175, Jan. 2007. [Online]. Available: https://doi.org/10.1155/2007/95328

[4] M. M. Saha, J. J. Izykowski, and E. Rosolowski, *Fault location on power networks*. Springer Science & Business Media, 2009.

[5] T. Takagi, Y. Yamakoshi, M. Yamaura, R. Kondow, and T. Matsushima, "Development of a new type fault locator using the one-terminal voltage and current data," *IEEE Trans. Power App. Syst.*, vol. PAS-101, no. 8, pp. 2892–2898, Aug. 1982.

[6] L. Eriksson, M. M. Saha, and G. D. Rockefeller, "An accurate fault locator with compensation for apparent reactance in the fault resistance resulting from remore-end infeed," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 2, pp. 423–436, Feb. 1985.

[7] K. W. Min and S. Santoso, "DC offset removal algorithm for improving location estimates of momentary faults," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5503–5511, Nov. 2018.

[8] S. Das, S. N. Ananthan, and S. Santoso, "Estimating zero-sequence line impedance and fault resistance using relay data," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1637–1645, Mar. 2019.

[9] B. Xia, Y. Wang, E. Vazquez, W. Xu, D. Wong, and M. Tong, "Estimation of fault resistance using fault record data," *IEEE Trans. Power Del.*, vol. 30, no. 1, pp. 153–160, Feb. 2015.

[10] M. S. Sachdev *et al.*, "Understanding microprocessor-based technology applied to relaying," in *Tech. Rep. WG I-01*, 2009.

[11] S. Santoso and T. Short, "Identification of fuse and recloser operations in a radial distribution system," *IEEE Trans. Power Del.*, vol. 22, no. 4, pp. 2370–2377, Oct. 2007.

[12] K. W. Min, S. Santoso, and L. Biyikli, "Identifying fault clearing operations in distribution systems," in *Proc. IEEE Power Energy Soc. General Meeting*, July 2016, pp. 1–5.

[13] K. W. Min, A. F. Bastos, S. Santoso, and U. Karadkar, "An extensible, open framework for power quality disturbance events," in *Proc. IEEE/PES Transmission and Distribution Conf. and Expo. (T&D)*, Apr. 2018, pp. 1–9.

[14] G. Benmouyal, "Removal of DC-offset in current waveforms using digital mimic filtering," *IEEE Trans. Power Del.*, vol. 10, no. 2, pp. 621–630, Apr. 1995.

[15] J. Gu and S. Yu, "Removal of DC offset in current and voltage signals using a novel Fourier filter algorithm," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 73–79, Jan. 2000.

[16] S. Kang, D. Lee, S. Nam, P. Crossley, and Y. Kang, "Fourier transform-based modified phasor estimation method immune to the effect of the DC offsets," *IEEE Trans. Power Del.*, vol. 24, no. 3, pp. 1104–1111, July 2009.

[17] Y. Guo, M. Kezunovic, and D. Chen, "Simplified algorithms for removal of the effect of exponentially decaying DC-offset on the Fourier algorithm," *IEEE Trans. Power Del.*, vol. 18, no. 3, pp. 711–717, July 2003.

[18] M. R. D. Zadeh and Z. Zhang, "A new DFT-based current phasor estimation for numerical protective relaying," *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2172–2179, Oct. 2013.

[19] Y. Cho, C. Lee, G. Jang, and H. Lee, "An innovative decaying DC component estimation algorithm for digital relaying," *IEEE Trans. Power Del.*, vol. 24, no. 1, pp. 73–78, Jan. 2009.

[20] M. Sachdev and M. Nagpal, "A recursive least error squares algorithm for power system relaying and measurement applications," *IEEE Trans. Power Del.*, vol. 6, no. 3, pp. 1008–1015, July 1991.

[21] P. Banerjee and S. C. Srivastava, "An effective dynamic current phasor estimator for synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 625–637, Mar. 2015.

[22] S. Das and T. Sidhu, "A simple synchrophasor estimation algorithm considering IEEE standard C37.118.1-2011 and protection requirements," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 10, pp. 2704–2715, Oct. 2013.

[23] S. Santoso, J. Lamoree, W. Grady, E. Powers, and S. Bhatt, "A scalable PQ event identification system," *IEEE Trans. Power Del.*, vol. 15, no. 2, pp. 738–743, Apr. 2000.

[24] S. Santoso, E. Powers, W. Grady, and A. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier. i. theoretical foundation," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 222–228, Jan. 2000.

[25] Z.-L. Gaing, "Wavelet-based neural network for power disturbance recognition and classification," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp.

1560–1568, Oct. 2004.

[26] P. Janik and T. Lobos, "Automated classification of power-quality disturbances using SVM and RBF networks," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1663–1669, July 2006.

[27] W.-M. Lin, C.-H. Wu, C.-H. Lin, and F.-S. Cheng, "Detection and classification of multiple power-quality disturbances with wavelet multiclass SVM," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2575–2582, Oct. 2008.

[28] F. Fahrisi, D. O. Anggriawan, I. Sudiharto, Suryono, and A. Tjahjono, "Overcurrent relay unconventional curve modeling in the power systems application using adaptive neuro fuzzy inference system," in *Proc. Int. Electron. Symp. on Eng. Technol. and Appl. (IES-ETA)*, Sept. 2017, pp. 192–197.

[29] "IEEE standard for inverse-time characteristics equations for overcurrent relays," *IEEE Std C37.112-2018 (Revision of IEEE Std C37.112-1996)*, pp. 1–25, Feb. 2019.

[30] A. A. Razi-Kazemi, M. Vakilian, K. Niayesh, and M. Lehtonen, "Circuit-breaker automated failure tracking based on coil current signature," *IEEE Trans. on Power Del.*, vol. 29, no. 1, pp. 283–290, Feb. 2014.

[31] S. M. Strachan, S. D. J. McArthur, B. Stephen, J. R. McDonald, and A. Campbell, "Providing decision support for the condition-based main-

tenance of circuit breakers through data mining of trip coil current signatures," *IEEE Trans. Power Del.*, vol. 22, no. 1, pp. 178–186, Jan. 2007.

[32] S. Das and S. Santoso, "Utilizing relay event reports to identify settings error and avoid relay misoperations," in *Proc. IEEE/PES Transmission and Distribution Conf. and Expo. (T&D)*, May 2016, pp. 1–5.

[33] A. W. McMorran, G. W. Ault, I. M. Elders, C. E. T. Foote, G. M. Burt, and J. R. McDonald, "Translating CIM XML power system data to a proprietary format for system simulation," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 229–235, Feb. 2004.

[34] A. W. McMorran, "An introduction to IEC 61970-301 & 61968-11: The common information model," *University of Strathclyde*, vol. 93, p. 124, 2007.

[35] G. Ravikumar, S. A. Khaparde, and Y. Pradeep, "CIM oriented database for topology processing and integration of power system applications," in *Proc. IEEE Power Energy Soc. General Meeting*, July 2013, pp. 1–5.

[36] G. Ravikumar and S. A. Khaparde, "A common information model oriented graph database framework for power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2560–2569, July 2017.

[37] "Electric power research institute, DOE/EPRI national database repository of power system events," accessed August 2017 at http://pqmon.

epri.com/disturbance_library/index.html.

[38] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data-the story so far," *Int. Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.

[39] J.-H. Hoepman and B. Jacobs, "Increased security through open source," *Communications of the ACM*, vol. 50, no. 1, pp. 79–83, 2007.

[40] P. P. Swire, "A model for when disclosure helps security: What is different about computer and network security," *J. on Telecomm. & High Tech. L.*, vol. 3, p. 163, 2004.

[41] T. Short, *Electric Power Distribution Handbook*. Abingdon: CRC Press, 2003. [Online]. Available: http://cds.cern.ch/record/994949

[42] E. O. Schweitzer and D. Hou, "Filtering for protective relays," in *Proc. IEEE WESCANEX*. IEEE, 1993, pp. 15–23.

[43] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, vol. 6, no. 2, pp. 418–445, 1996.

[44] T. F. Coleman and Y. Zhang, "Optimization toolbox: User's guide (R2016b)," *MathWorks: Natick, MA, USA*, 2016.

[45] "IEEE recommended practice for monitoring electric power quality," *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, pp. c1–81, June 2009.

[46] S. Santoso, E. Powers, W. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Del.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.

[47] N. S. D. Brito, B. A. Souza, and F. A. C. Pires, "Daubechies wavelets in quality of electrical power," in *Proc. 8th Int. Conf. Harmon. Qual. Power*, vol. 1, Oct. 1998, pp. 511–515.

[48] K. Zimmerman and D. Costello, "Impedance-based fault location experience," in *Proc. 58th Annu. Conf. Protective Relay Engineers*, Apr. 2005, pp. 211–226.

[49] C. Muller and R. Jayasinghe, "PSCAD user's guide," *Manitoba HVDC Research Centre Inc*, 2010.

[50] S. Das, S. Santoso, A. Gaikwad, and M. Patel, "Impedance-based fault location in transmission networks: theory and application," *IEEE Access*, vol. 2, pp. 537–557, 2014.

[51] J. H. Teng, W. H. Huang, and S. W. Luan, "Automatic and fast faulted line-section location method for distribution systems based on fault indicators," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1653–1662, July 2014.

[52] Y. Jiang, C. C. Liu, M. Diedesch, E. Lee, and A. K. Srivastava, "Outage management of distribution systems incorporating information from

smart meters," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4144–4154, Sept. 2016.

[53] M. Farajollahi, M. Fotuhi-Firuzabad, and A. Safdarian, "Deployment of fault indicator in distribution networks: A MIP-based approach," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 2259–2267, May 2018.

[54] I. Dzafic, R. A. Jabr, S. Henselmeyer, and T. Donlagic, "Fault location in distribution networks through graph marking," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1345–1353, Mar. 2018.

[55] R. C. Dugan, M. F. McGranaghan, and H. W. Beaty, *Electrical power systems quality*, 1996.

[56] H. V. Padullaparti, P. Chirapongsananurak, M. E. Hernandez, and S. Santoso, "Analytical approach to estimate feeder accommodation limits based on protection criteria," *IEEE Access*, vol. 4, pp. 4066–4081, 2016.

[57] D. Novosel, D. Hart, Y. Hu, and J. Myllymaki, "System for locating faults and estimating fault resistance in distribution networks with tapped loads," US Patent number 5,839,093, Nov. 17 1998.

[58] E. Jones, T. Oliphant, P. Peterson *et al.*, "SciPy: Open source scientific tools for Python," 2001–, [Online; accessed Mar. 19 2018]. [Online]. Available: http://www.scipy.org/

[59] W. Dabbs, D. Sabin, T. Grebe, and H. Mehta, "PQView-a power quality data management and analysis system," *IEEE Computer Applications in Power*, 1994.

[60] A. F. Bastos and S. Santoso, "Identifying switched capacitor relative locations and energizing operations," in *Proc. IEEE Power Energy Soc. General Meeting*, July 2016, pp. 1–5.