12-2022

Power System Transients: Impacts of Non-Ideal Sensors on Measurement-Based Applications

Aaron Wilson
awilson@tennessee.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Aaron Wilson entitled "Power System Transients: Impacts of Non-Ideal Sensors on Measurement-Based Applications." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Yilu Liu, Major Professor

We have read this dissertation and recommend its acceptance:

Yilu Liu, Fangxing Li, Peter Fuhr, Yarom Polsky

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Power System Transients: Impacts of Non-Ideal Sensors on Measurement-Based Applications

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Aaron Jacob Wilson

December 2022

Dedication

There are many people in my life that that helped me cross the finish line. I want to dedicate this dissertation to:

My parents, who have supported me through all these years of college and graduate school,

My loving wife, who is a constant source of encouragement, as well as my mother- and father-in-law, who have been nothing short of wonderful.

My 95-years-young grandfather who is beyond ecstatic at the thought of a "doctor" in the family,

My extended family, for supporting me even though they ask "What are you studying again?".

Acknowledgments

This work would not have been possible without the wide net of support that has been cast by my colleagues, peers, friends, and supervisors. First and foremost, I would like to acknowledge my wonderful advisor, Dr. Yilu Liu, whose patience knows no bounds when putting up with my endless email questionnaires. I would also not have been given the opportunity to even pursue this Ph.D degree without her recommendations to the University of Tennessee and Oak Ridge National Laboratory.

I would also like to acknowledge Dr. Peter Fuhr, who saw enough in me to give me a chance to further my career with ORNL, for acting as a second advisor despite no requirements being placed on him to do so, and for being a friend to me through challenging times during this process.

Additionally, I would like to acknowledge my remaining committee members. Dr. Yarom Polsky of ORNL & UT Bredesen Center has been an inquisitive and helpful guide through this process, willing to have thoughtful conversations about interesting engineering-related topics. Dr. Fangxing Li, who always continues to impress, agreed to serve on my committee without a second thought, for which I am extremely grateful.

The last individual I wish to acknowledge is my former graduate school advisor, Dr. Donald Reising from the University of Tennessee at Chattanooga. I had no idea what career path to take until I was offered a path through graduate school at UTC. I don't know where I would have ended up without him.

Lastly, I would like to acknowledge all of my group members and colleagues at ORNL. Bill Monday, Jason Richards, Ali Riza Ekti, Elizabeth Piersall, David Pesin, Rick Moyers, Gary Hahn, Bruce Warmack, Jianming "Jamie Lian", Coreen Prysmont, Ryan Kerekes, Joseph Olatt, Chris Cooper, and Chris "Cozmo" Engebretsen have all been extremely supportive

throughout the entire process. Drs. Dan Lu and Hoang Tran of ORNL's Computing and Computational Sciences Directorate are responsible for the development of the "Prediction Intervals from Three Neural Networks" (PI3NN) algorithm used and referenced heavily in Chapter 5.

Abstract

The power system is comprised of thousands of lines, generation sources, transformers, and other equipment responsible for servicing millions of customers. Such a complex apparatus requires constant monitoring and protection schemes capable of keeping the system operational, reliable, and resilient. To achieve these goals, measurement is a critical role in the continued functionality of the power system. However, measurement devices are never completely reliable, and are susceptible to inherent irregularities; imparting potentially misleading distortions on measurements containing high-frequency components. This dissertation analyzes some of these effects, as well as the way they may impact certain applications in the grid that utilize these kinds of measurements. This dissertation first presents background on existing measurement technologies currently in use in the power grid, with extra emphasis placed on point-on-wave (PoW) sensors, those designed to capture oscillographic records of voltage and current signals.

Next, a waveform playback system, developed at Oak Ridge National Laboratorys Distributed Energy Communications & Control (DECC) laboratory was used for comparisons between various line-post-monitor PoW sensors when subjected to different high-frequency current disturbances. Each of the three sensors exhibited unique quirks in these spectral regions, both in terms of harmonic magnitude and phase angle. A goodness-of-fit metric for comparing an ideal reference sensor with the test sensors was adopted from the literature and showed the extremes to which two test sensors vastly under performed when compared to the third. The subsequent chapter analyzes these behaviors under a statistical lens, using kernel density estimation to fit probability density functions (PDFs) to error distributions at specific harmonic frequencies resulting from sensor frequency response distortions. The remaining two chapters of the dissertation are concerned with resultant effects on applications

that require high-frequency transient data. First, a detection algorithm is presented, and its performance when subjected to statistical errors inherent in these sensors is quantified. The dissertation culminates with a study on an artificial intelligence (AI)technique for estimating the location of capacitor switching transients, as well as learning prediction intervals that indicate the level of uncertainty present in the data caused by sensor frequency response irregularities.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The electric power grid is one of the most vast and complex systems in the world. The United States Energy Sector is one of sixteen Critical Infrastructure Sectors defined by Presidential Policy Directive 21. Every component is crucial to ensure that power is accessible to all customers at all times of the day. It is imperative that when equipment failures happen, due to either internal or external causes, contingencies and redundancies are in place to ensure the continued flow of power to critical customers, such as hospitals, schools, and many others.

This dissertation studies the effects of inaccurate measurement systems on applications used, or will be used, in the power system. It begins with a description of typical measurement technologies used in the power grid, and some of their inherent limitations. From there, two different crucial operations that involve the usage of high-frequency voltage and current signals are studied, as well as their susceptance to false or erroneous measurements.

1.1 Measurement of the Power System

Power is transported in the form of energized transmission and distribution lines connected to electrical loads. Typically, power is generated at a few tens of kilovolts (kV), before being stepped up through a power transformer to a suitable range for transmission level, which may be as high as 750 kV, [1]. The amount of power supplied to a customer base is dependent upon the amount of load connected to the system. The "load" in this case consists of the

end users and the number of items these users have connected to the grid. Heavy loads (lower impedances) will draw more current (and therefore power) than lighter loads (higher impedances).

There are, however, limitations to the amount of power a system can supply. Heavier loads can lead to issues such as voltage instability and collapse, which can lead to cascading failures across the larger grid as a whole; over-current, which may cause damage to equipment; and frequency stability issues, such as under- and over-frequency, caused by imbalance between generation and load. The need to continually monitor key points along the grid is increasingly becoming more and more necessary to ensure proper operation of all equipment, implement control functions, and provide situational awareness to utilities.

1.1.1 Common Power System Sensing Technologies

Instrument Transformers

One of the most common uses of power system measurement is protection systems. Specifically, relays are devices capable of sending signals to circuit breakers to open or close portions of a circuit when a problem arises. For example, if a short circuit occurs on one of the phases of a transmission line (such as a line-to-ground fault), the nearest relay behind the location of the fault will sense the excess current being drawn and will send a signal to its circuit breaker to open, thus isolating the circuit and ceasing the flow of current into the fault.

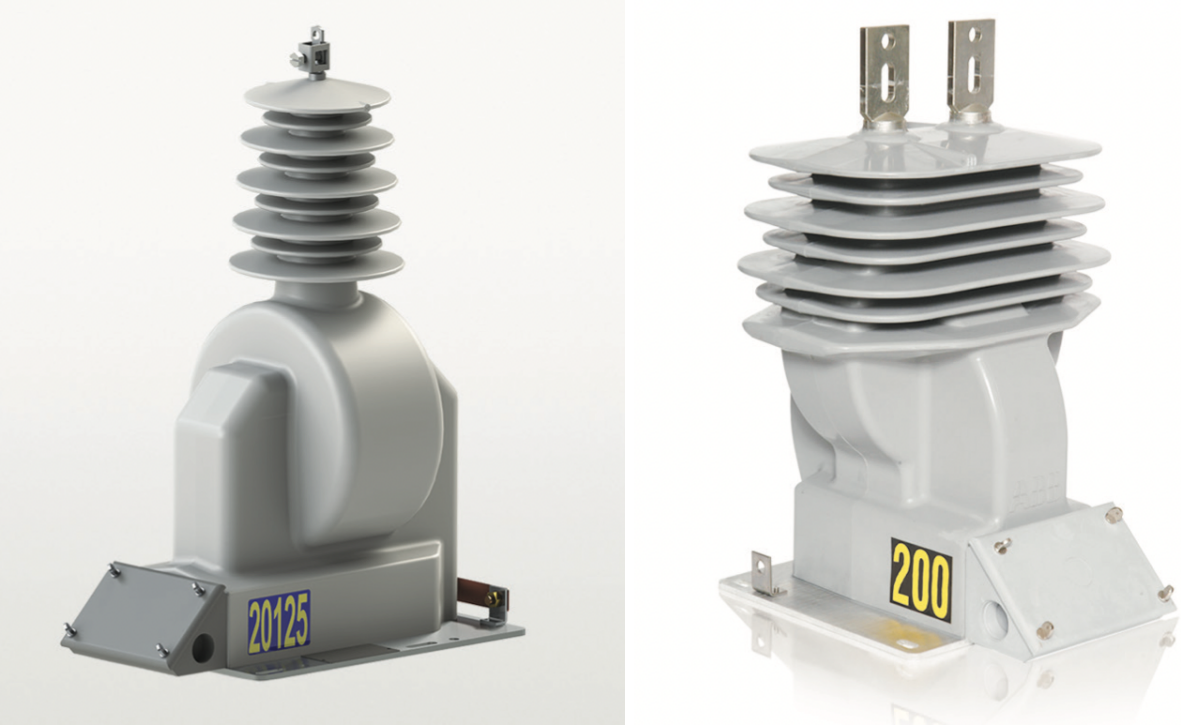
Relays are not interfaced directly with the transmission and distribution lines and substation bus-work. The voltages and currents used for measurement are "stepped down" through *instrument transformers*. Potential transformers (PTs) are shunt-connected devices designed to provide high-accuracy, low-voltage representations of the behavior of the grid's higher voltages. This isolates the metering and protection equipment from dangerous high-voltage levels of the grid while properly replicating the behavior of the grid voltage on the low-voltage secondary. Similarly, current transformers (CTs) are series-connected devices performing a similar function as PTs, the exception being that the current is stepped down instead of voltage.

Instrument transformers, in addition to being vital to protection systems, are also used for metering. Electric utilities bear the responsibility of ensuring that power distributed to customers is stable and healthy, meaning that the *power factor* never dips below prescribed limits. Metering circuitry is connected to the secondaries of instrument transformers placed throughout transmission and distribution networks, and provide data for utilities to produce accurate billing information for customers. Figure 1.1 shows some examples of common PT/CTs.

Phasor Measurement Units (PMUs)

Towards the end of the 20^{th} century, the integration of synchrophaser technology began to boom. Also known as phasor measurement units (PMUs), synchrophasers are used extensively for wide-area monitoring, control, and real-time situational awareness. Synchrophasers possess the unique ability to digitally sample the sinusoidal voltage and currents to produce accurate estimates of the instantaneous magnitudes and phase angles of the measured quantities. Some of the key advantages of PMUs include their ability to synchronize to an external timing signal such as GPS or IEEE 1588 Precision Timing Protocol (PTP), as well as their ability to report the instantaneous fundamental frequency of the waveforms of interest. The ability to capture frequency, as well as rate-of-change-of-frequency (ROCOF) is a critical element in the assessment of grid frequency stability.

One of the most common uses of synchrophasers are for wide-area situational awareness. One such example of this is the FNET/GridEye system. Operated and hosted at the University of Tennessee and Virginia Tech, FNET/GridEye employs many frequency-disturbance recorders (FDRs) across North America. These devices accurately measure the frequency at different distribution-level locations and reports their findings in real-time, and can be found at [2]. Figure 1.2 shows an example of the readings captured from these devices on February 6, 2022, at approximately 6:24 PM, UTC. The color bands along the top of the figure indicate how the various regions are drifting away from the fundamental frequency (60 Hz). It can be seen from the figure that at the indicated time, much of the west coast is experiencing between 0.07 and 0.09 Hz of over-frequency, whereas the east coast is experiencing between 0.05 and 0.07 Hz of under-frequency conditions.



(a) Potential transformer (PT) **(b)** Current transformer (CT)

Figure 1.1: Common instrument transformers

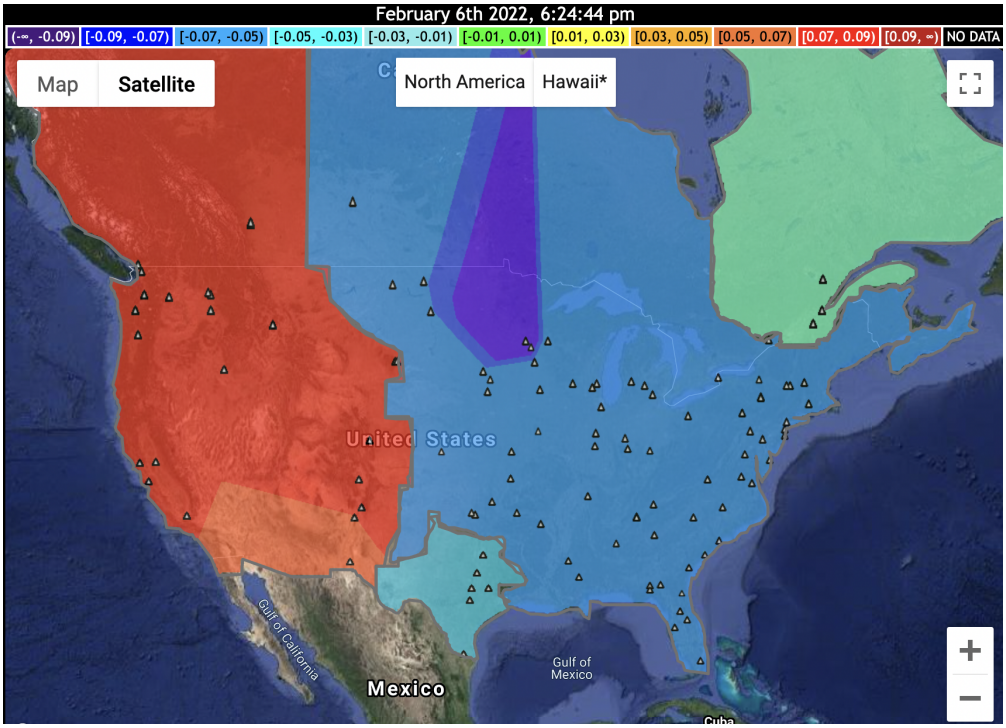


Figure 1.2: Live capture of FNET/GridEye system on February 6, 2022

Line-Post Monitors

Another emerging technology is the linepost sensor (LPs). These devices are typically connected directly to the line post or on the bus-bar at distribution-level voltages. These devices come in a wide variety of constructions and operating principles, and are typically capable of sampling point-on-wave voltages and currents well into the kHz ranges. These devices, while typically more both economically infeasible and bulky, provide measurement capabilities that are not found with synchrophasers and instrument transformers (specifically PTs). LPs are able to capture high voltages and currents at high sampling rates, without conversion to lower-resolution magnitudes and phases (as with synchophasers), and without transforming the voltages and currents down to lower levels (as with instrument transformers). Figure 1.3 shows some examples of commercial LPs.

1.2 Limitations of Conventional Measurements

All sensors are prone to producing distorted versions of the quantities that are being measured due to non-ideal characteristics of the sensors themselves. It is virtually impossible for a sensing mechanism to perfectly reproduce the measured phenomena. In the subsequent sub-sections, various sources of non-ideal signal behavior are discussed.

1.2.1 Instrument Transformer Accuracy

PTs and CTs are characterized by their *accuracy class*. Instrument transformer accuracy classes are defined by their *transformer correction factors* (TCFs). For PTs, [3]:

$$\text{TCF} = \text{RCF} + \frac{\gamma}{2600} \tag{1.1}$$

and for CTs:

$$\text{TCF} = \text{RCF} - \frac{\beta}{2600} \tag{1.2}$$

where $\text{RCF} = 1 - (\textit{Ratio Error}/100)$, *Ratio Error* is the degree to which the secondary voltage and/or current faithfully represents the primary voltage and/or current applied

through the transformer's turns ratio, and γ and β are the phase angles, in minutes, of the PT and CT, respectively.

Accuracy classes for metering are designated in such a way that the TCF of a transformer is within the specified burden (i.e. secondary-connected impedance) limits for metered (primary-connected) load between 0.6 and 1.0 power factor, lagging.

1.2.2 CT Saturation

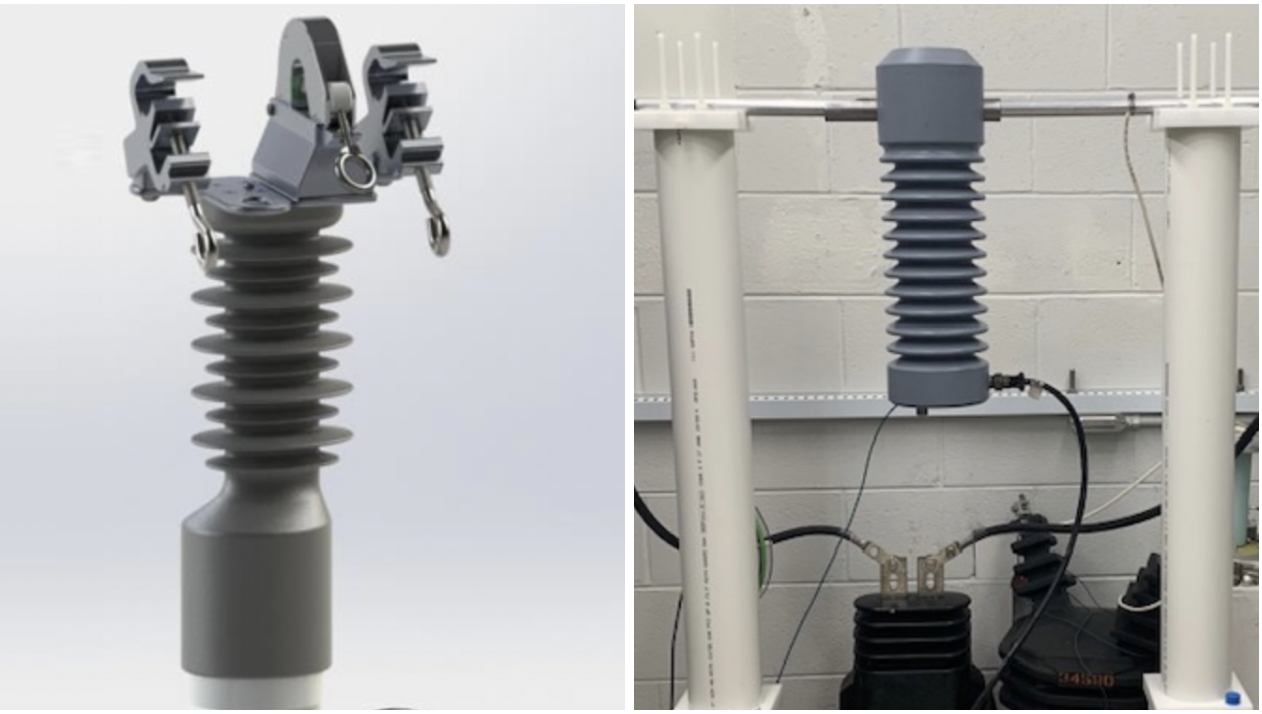
The simplest form of CTs involve wires wound tightly around an iron core. The core of a CT, when subjected to an alternating current in the primary winding, produces a magnetic field of intensity H, corresponding to an alternating magnetic flux ϕ in the core. If the secondary winding is connected to a burden, the alternating flux ϕ then induces an alternating voltage across the secondary winding, therefore producing a secondary alternating current, [4].

The induced magnetic field H causes the dipoles in the iron core, termed "flux density" (B), to align with the direction of the field, which is ultimately the cause of the alternating flux ϕ. The strength of H is directly proportional to the magnitude of the primary current. If the primary current's induced flux exceeds the number of dipoles available in the iron core, saturation will begin. Figure 1.4 shows example B-H curves; note the flattening that occurs on the extremities of the curve as H increases in either direction.

Saturation is one of the key drawbacks of CTs, and limit these devices' capability to measure higher currents.

1.2.3 Line-post Monitor Frequency Response

Frequency response characteristics are eventualities that arise with any electro-mechanical system. Higher frequency components, typically representative of transient behavior (Section 1.3), are most susceptible to distortions caused by sensor frequency response characteristics, as most sensors have a limited bandwidth, leading to attenuation of higher frequencies (Figure 1.5). Frequency response theory is presented in further detail in Chapter 2 Section 2.3.



(a) Eaton GridAdvisor Insight RG235 **(b)** Lindsey Current & Voltage Monitoring Insulator

Figure 1.3: Line post sensors

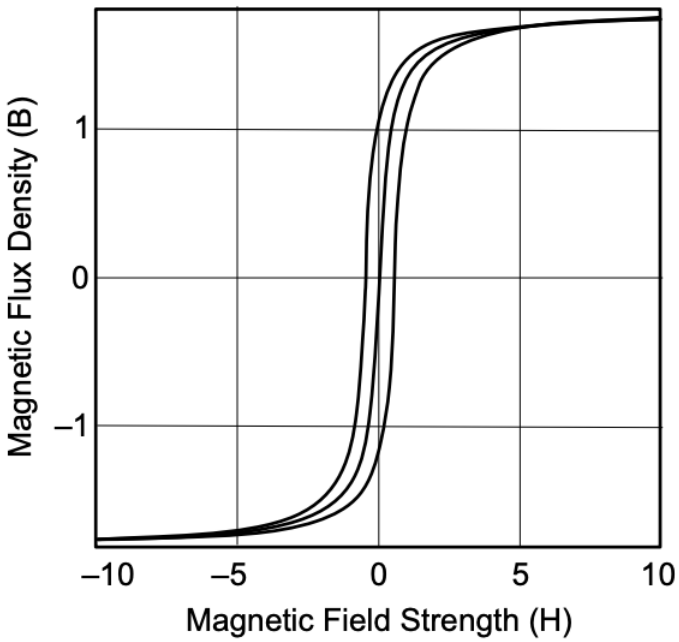


Figure 1.4: Example B-H curve, [4]

1.3 Power System Transients

The preceding sections of this chapter discuss common sensing and measurement techniques commonly used in power systems. Certain voltage and current phenomena are much more susceptible to inaccurate measurements and distortions than others. This section will introduce the concept of power system *transients*, discuss mathematical theory behind said transients, their sources, their effects on the grid, and the way that sensors will severely distort them.

AC voltages and currents in the power system are sinusoidal in nature, oscillating at 60 Hz in North America and 50 Hz in many other parts of the world. This frequency is often referred to as the *fundamental frequency*. Equipment is often rated for voltages and currents oscillating at the fundamental frequency. This means that the same equipment is not rated for higher-frequency operating points, and thus special care must be taken to limit the amount of high-frequency components present in the 60-Hz waveforms. One way to ensure this is to impose total harmonic distortion (THD) limits, as is the case with DC/AC inverters. Another means of analysis that typically must be performed are *transient* contingency studies. Transient over-voltages are a potentially dangerous phenomenon that can adversely effect the life of power system equipment, and designs that alleviate or mitigate the associated concerns should be observed.

1.3.1 Definition of "transient"

The general definition of the word "transient" per Merriam Webster is "passing especially quickly into and out of existence". There are two definitions of "transient" in the power systems context, per IEEE 1159-2009, [5]: *impulsive* transients and *oscillatory* transients. Impulsive transients are defined to be "a sudden non-power frequency change in the steady-state condition of voltage or current that is unidirectional in polarity (primarily either positive or negative)", whereas an oscillatory transient is "a sudden, non-power frequency change in the steady-state condition of voltage, current, or both, that includes both positive and negative polarity values."

The term "polarity" in the above definitions refers to the transient's number of zero-crossings. Impulsive transients represent surges in either the positive or negative direction with respect to the ordinate axis, and usually subside back to the original state before its next zero-crossing. The most common example of an impulsive transient is a lightning surge. Oscillatory transients, however, represent disturbances that contain one or more zero-crossings before subsiding, and are usually high-frequency in nature. Oscillatory transients are typically a result of sudden changes in the system state, such as switching operations, inverter-based resources (IBRs), and resonant responses.

1.3.2 Sources of transients in the power system

Switching Operations

Many behaviors in the power system involve switching actions, which are characterized by sudden changes in the steady-state operation of the system. Common examples of switching operations include circuit breakers opening to interrupt high amounts of fault current, capacitor banks energizing and de-energizing to adjust the amount of reactive power supplied to the system, and load shifting from one circuit to another. Because the power system is comprised of many different resistive (R), inductive (L), and capacitive (C) elements, switching operations are often modeled in terms of RLC circuit combinations.

Resistive-inductive (RL) circuit models are often used to model power system components, including transformer windings, rotating machines, and simple transmission lines. A simplified RL circuit is given in Figure 1.8. To examine the behavior of the current passing through the resistor R and inductor L, Kirchhoff's voltage law (KVL) is applied to the loop to obtain:

$$E_{max} \sin(\omega t + \phi) = Ri + L\frac{di}{dt} \tag{1.3}$$

The solution to this differential equation is the sum of both a general and particular solution. Setting the left-hand side of (1.3) equal to zero and solving for $i(t)$, the general solution becomes

$$i_g(t) = C_1 e^{(R/L)t} \tag{1.4}$$

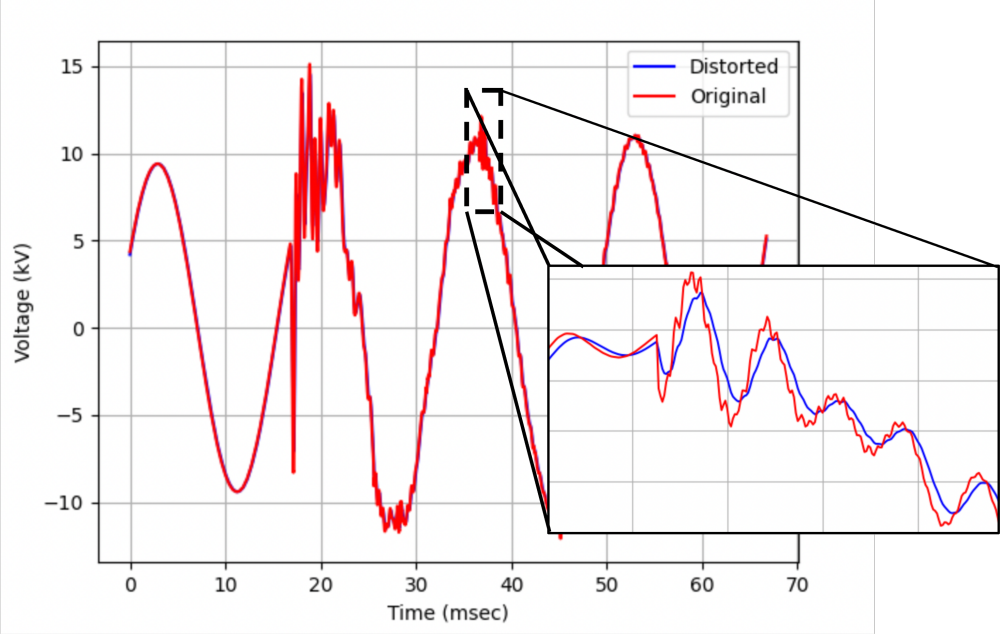


Figure 1.5: Example of capacitor switching transients distorted by sensor with low-pass characteristics. The red curve shows a simulated switching transient voltage signal, and the blue curve represents the same signal as produced by a sensor with known frequency response.

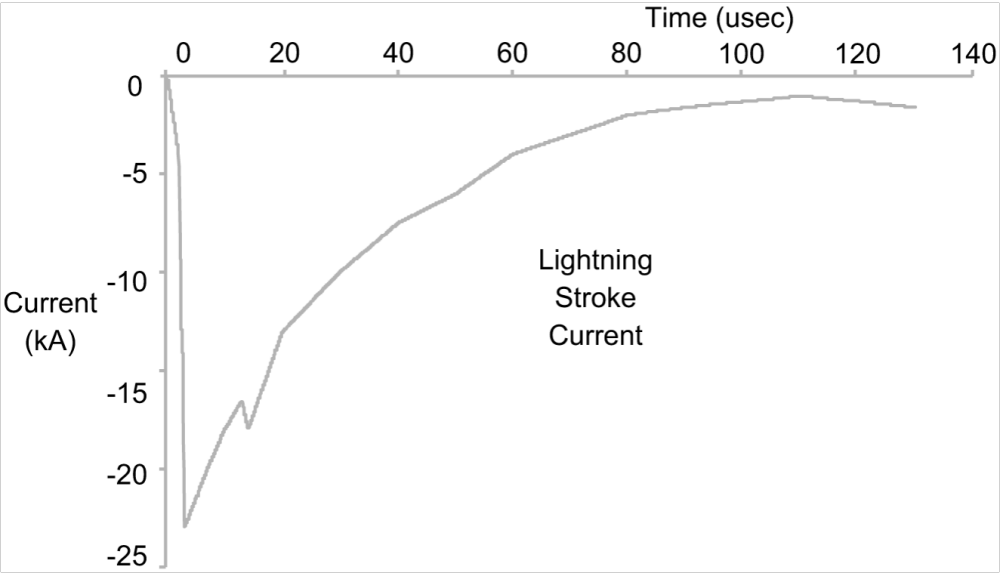


Figure 1.6: Example impulsive transient caused by a lightning strike

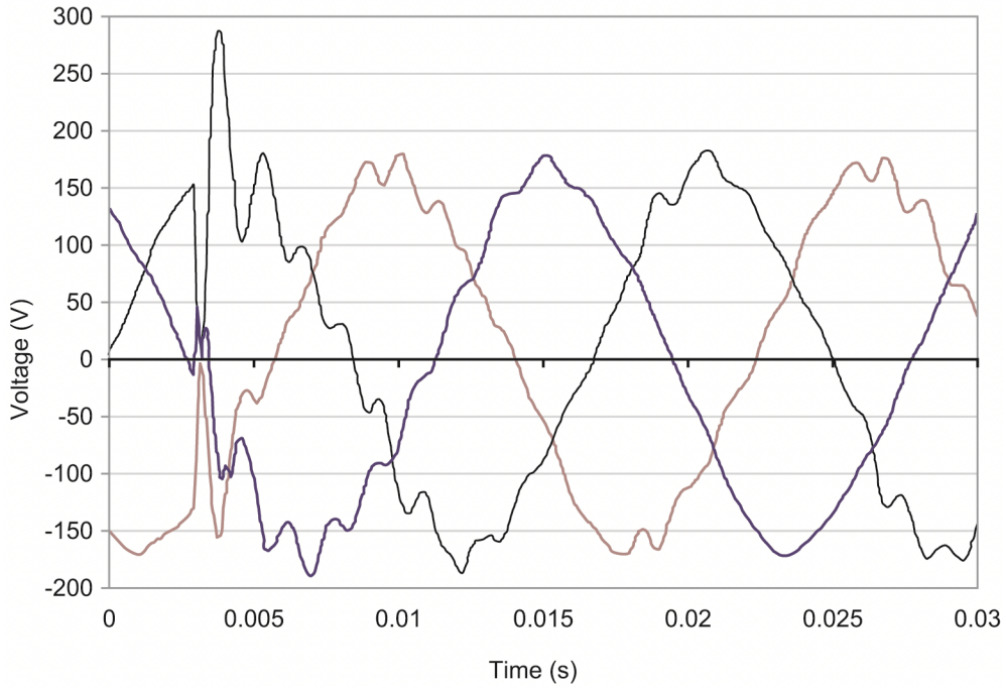


Figure 1.7: Example oscillatory transient caused by capacitor-bank energization

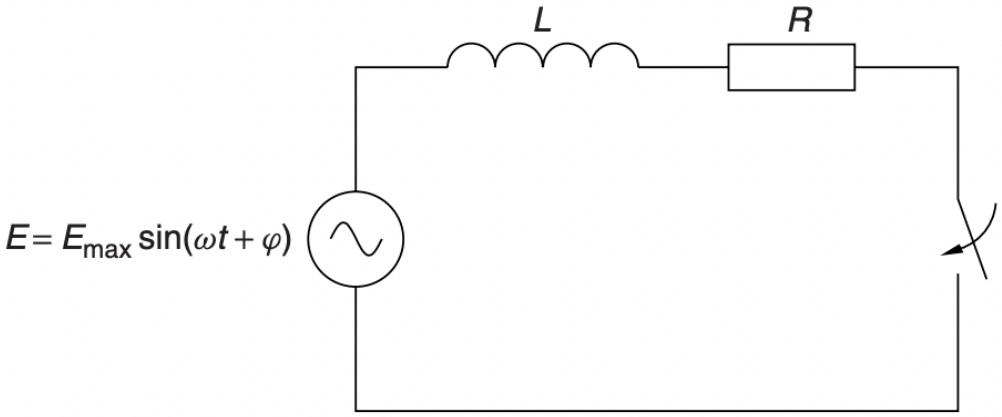


Figure 1.8: Simplified RL Circuit [6]

The constant C_1 may be obtained by setting (1.4) equal to zero and letting $t = 0$. This is because at $t = 0$, the magnetic flux in the inductor is equal to zero, and when the switch closes at $t = 0$, it is still zero because the current through an inductor may not change instantaneously.

The particular solution $i_p(t)$ may be found as in [6] to be:

$$i_p(t) = \frac{E_{max}}{\sqrt{R^2 + \omega^2 L^2}} \sin\left[\omega t + \phi - \tan^{-1}\left(\frac{\omega L}{R}\right)\right] \tag{1.5}$$

Combining (1.4) and (1.5) yields

$$\begin{aligned} i(t) = i_g(t) + i_p(t) = \\ e^{(-R/L)t}\left\{\frac{-E_{max}}{\sqrt{R^2 + \omega^2 L^2}} \sin\left[\phi - \tan^{-1}\left(\frac{\omega L}{R}\right)\right]\right\} + \\ \frac{E_{max}}{\sqrt{R^2 + \omega^2 L^2}} \sin\left[\omega t + \phi - \tan^{-1}\left(\frac{\omega L}{R}\right)\right] \end{aligned} \tag{1.6}$$

The first term in (1.6) represents a decaying DC component with initial phase $\theta = \phi - \tan^{-1}\left(\frac{\omega L}{R}\right)$. The second term represents the oscillating portion of the current. If $\theta = 0$, or an integer multiple of π, the DC component is zero and the current is oscillating in its steady-state. If the switch closes at a time that causes $\theta = \pm\, 90°$, the total transient current will be at a maximum.

An LC circuit, which models a capacitor switching scenario, is shown in Figure 1.9. The current through the capacitor may once again be solved for by applying KVL:

$$E = L\frac{di}{dt} + \frac{1}{C}\int i dt \tag{1.7}$$

Assuming the voltage across the capacitor at time $t = 0$, the current is then found to be (as described in [6]):

$$i(t) = E\sqrt{\frac{C}{L}} \sin(\omega_0 t) \tag{1.8}$$

where $\omega_0 = \sqrt{LC}$ is the natural frequency of the circuit.

The *surge impedance* of the circuit is $Z_0 = \sqrt{L/C}$ and determines the peak of the transient current. The transient voltage across the capacitor may be shown to be:

$$V_c(t) = E - [E - V_c(0)] \cos(\omega_0 t)) \tag{1.9}$$

When $V_c(0) = 0$, the transient voltage oscillates between $[-2E, \ +2E]$, and for $V_c(0) = -E$, $V_c(t)$ oscillates between $[-3E, \ +3E]$, indicating an overvoltage of *three times* the nominal value, E.

A series RLC circuit is shown in Figure 1.10. KVL to obtain the current through this circuit yields:

$$E_{max} \sin(\omega t + \phi) = L\frac{di}{dt} + \frac{1}{C}\int i dt + Ri \tag{1.10}$$

Solving for $i(t)$, as in [6], yields:

$$i(t) = \left(C_1(t)e^{\lambda_1 t} + C_2(t)e^{\lambda_2 t}\right) + \frac{E_{max}}{\sqrt{R^2 + \left(\frac{1}{\omega C} - \omega L\right)}} \sin\left[\omega t + \phi + \tan^{-1}\left(\frac{\frac{1}{\omega C} - \omega L}{R}\right)\right] \tag{1.11}$$

When $(R/2L)^2 > 1/LC$, the transient current is considered *overdamped*. When $(R/2L)^2 = 1/LC$, then the transient is considered *critically damped*. Finally, when $(R/2L)^2 < 1/LC$, the transient current is *oscillatory*. Illustrations of these three cases are shown in Figure 1.11.

Capacitor Bank Energization and De-energization

Reactive loads used in the power system are primarily inductive, meaning that the current drawn by the load lags the supplied voltage. This causes the amount of real power supplied by the system to decrease while the amount of reactive power increases. As reactive power is not "usable" by loads, the power factor, which is analogous to the "efficiency" of the power system supplying the load, decreases. The most common way of improving the power factor

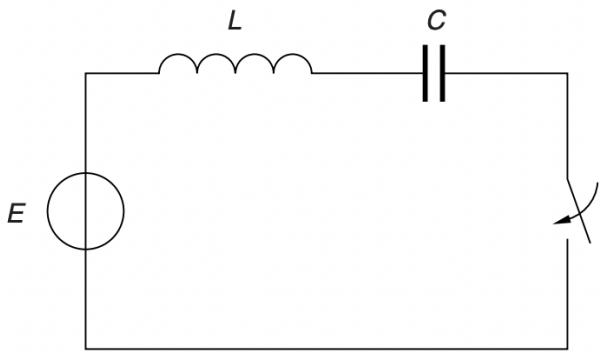


Figure 1.9: Simplified LC Circuit, [6]

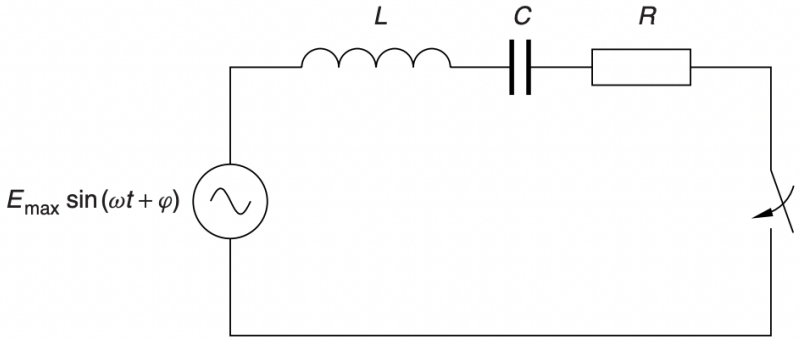


Figure 1.10: Simplified RLC Circuit

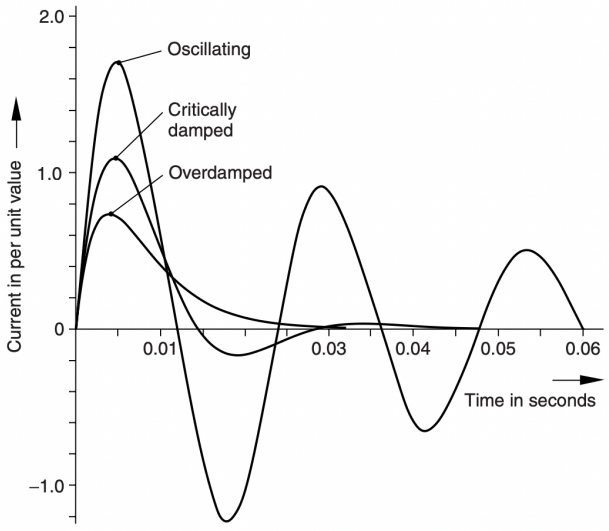


Figure 1.11: Illustration of Damping Cases for RLC Circuit, [6]

of a system is to supply capacitive reactive power through the use of switched capacitor banks.

When a capacitor bank is in its discharged state, it is at ground potential (zero volts). Upon initiating a switching operation of the capacitor bank, at which point contact with the line occurs, the line voltage briefly falls close to zero volts due to the capacitor's resistance to instantaneous changes in voltage. Because of this sudden voltage drop, the current drawn through the system will see the still-charging capacitor bank as a short-circuit, causing a current inrush. This inrush current can reach up to the tens of thousands of amperes range, and may oscillate at a frequency up to the tens of kilohertz. This inrush current will in turn cause a transient overvoltage, which may reach as high as 2 per unit, [5]. The frequency of oscillation depends on the impedances of the lines, loads, and capacitor banks. Figure 1.7 shows an example of an over-voltage with oscillatory ringing as a result of a capacitor bank being switched into operation.

Traveling Wave Phenomena

When transmission lines are sufficiently long, they are typically modeled using distributed parameter elements, meaning that electrical phenomena (voltage and current) are mathematically modeled in terms of both space and time, rather than the typical time-dependent variables seen in traditional circuit analysis. Assuming the transmission line may be broken into individual "snippets" of length Δx, each "snippet" is modeled as in Figure 1.12, where R and L are the series resistance and inductance per unit length, and G and C are the shunt conductance and capacitance per unit length, respectively. It can be shown, [7], that the voltage and current may be represented by the sum of incident (v_I) and reflected v_R waves:

$$v(x,t) = v_I e^{-\gamma x} + v_R e^{\gamma x} \tag{1.12}$$

$$i(x,t) = \frac{v_I}{Z_0} e^{-\gamma x} + \frac{v_R}{Z_0} e^{\gamma x}, \tag{1.13}$$

where $\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \alpha + j\beta$ is the propagation constant and Z_0 is the characteristic impedance of the line. When a fault occurs on a transmission line, high-frequency transients are transmitted in both directions away from the fault. When this

"incident" wave reaches a boundary (either physical - as with an open circuit or material change, or electrical - as with a change in impedance) a portion of the wave is transmitted through the boundary while the rest of the wave is reflected back. The reflection ρ and transmission τ coefficients are given as

$$\rho = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{1.14}$$

$$\tau = \frac{2Z_1}{Z_1 + Z_0} \tag{1.15}$$

where Z_1 is the impedance of the new material or portion of the circuit.

1.3.3 Effects of transients on the power system

Transient Over-voltage (TOV)

Although transients occur in the microseconds-to-milliseconds time scale, the potential for severe overvoltages and overcurrents still poses risks for equipment and personnel. *LC* circuits, as in Figure 1.9, are typically used to model high-voltage circuit breakers. Transient over-voltage (TOV) studies are crucial in the design of such circuit breakers. These breaker designs need to be able to both withstand maximal TOV surges (1.9) and include protection mechanisms for reducing the severity of the TOV.

Transient Recovery Voltage (TRV)

Transient recovery voltage (TRV) is the voltage that appears across the terminals of a circuit breaker during the interruption of current. When current is interrupted in a circuit breaker, specifically in sulfur hexafluoride (SF_6) breakers, an arc is generated. Over time, typically on the order of microseconds, the conductivity of the arc decreases as the instantaneous current reaches zero. The TRV is a result of a balance in energy contained in the arc immediately after extinction of the current. TRV may also be thought of as the voltage induced across circuit breaker terminals in an attempt to keep the load and source voltages constant after the extinction of the current through the breaker.

The frequency of oscillation of TRV surges is:

$$f_{\text{TRV}} = \frac{1}{2\pi\sqrt{LC}} \tag{1.16}$$

where L is the short-circuit inductance of the supply-side circuit and C is the supply capacitance.

If TRV is not properly considered when manufacturing and installing circuit breakers, scenarios involving TRV can exceed the rated *gap voltage*, causing re-ignitions or re-strikes.

1.4 Dissertation Outline

This chapter presented conventional measurement capabilities currently used in the power grid. These measurement devices (e.g. sensors), all have inherent imperfections in their ability to completely replicate observed electrical phenomena. This is especially apparent when these sensors are subjected to high-frequency disturbances in the form of transients.

The rest of this dissertation is structured in two parts: Chapters 2 and 3 analyze some of the inherent effects on high-frequency transients due to imperfect sensors, whereas Chapters 4 and 5 examine just a few of the ways that these imperfections manifest in actual applications.

Chapter 2 provides basic background on the theory of linear systems, which will be the prevailing assumption used in all subsequent experiments; namely that sensors used to measure point-on-wave (PoW) quantities may be faithfully approximated using a linear time-invariant (LTI) model. From there, various commercial-grade PoW sensors are used to evaluate the production of various current disturbances containing high-frequency components. In Chapter 3, the results in Chapter 2 are expanded upon and analyzed statistically, using a different set of disturbance types. Kernel density estimation (KDE) is used to approximate the probability density functions (PDFs) of the magnitude and phase errors due to these sensor irregularities.

Chapter 4 begins to look at the effects of these distortions on applications for the grid that rely on accurate measurements. This chapter in particular examines the effect of sensor distortions on the accurate *detection* of high-frequency components of transient

signals in the form of the *energy detector*. This dissertation culminates in Chapter 5, in which a deep learning technique is used to identify the location of switching transient signals. More importantly, an uncertainty quantification (UQ) component is introduced that is able to reflect confidence (or lack thereof) of the model's predictions when subjected to sensors distortions of the kind introduced in Chapter 2. Finally, Chapter 6 concludes the document and discuss future work ideas and opportunities leveraging the work presented in this dissertation.

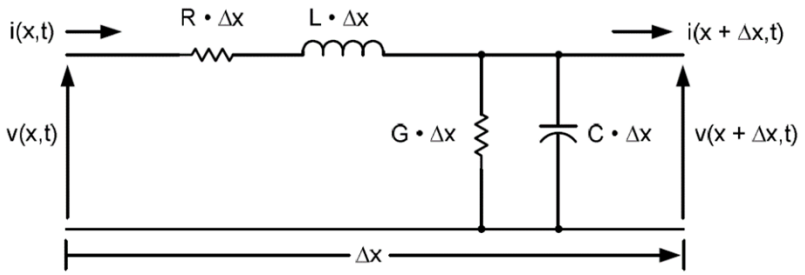


Figure 1.12: Distributed Parameter Model of Transmission Line Segment

Chapter 2

Non-Ideal Sensors' Effects on Power System Transients

This chapter presents a comparison of various common electrical disturbance events; namely arcing, capacitor switching, and three-phase inrush, and the ways in which three different sensing technologies are able to capture the finer details of the waveforms (i.e. high-frequency content). The experiment utilized a data acquisition-and-playback feature that uses existing data files and plays them through the various sensors via a current amplifier and arbitrary function generator. It is important to note that the results of this experiment are not meant to be treated as a recommendation for one sensing technology over another. This experiment was conducted using Oak Ridge National Laboratory's (ORNL) Distributed Energy Communications & Controls (DECC) Laboratory and three commercially-available medium-voltage (MV) line-post sensors.

This chapter first provides a literature survey on measurement problems in the power system, followed by a brief refresher on the theory of linear systems. This then feeds into a section on frequency response analysis. The proceeding section describes the mathematical foundations for comparison between signals captured by both a reference sensor and the equipment under test (EUT). Following this, a description of the playback-sensing system developed at ORNL is presented, and the final section presents the results of the study along with associated analysis and discussion.

2.1 Background and Literature Review

As the power grid expands both in terms of customers and technology, so do the number of complexities associated with protection and metering applications. Increasing numbers of inverter-based resources (IBRs) penetrating the grid lead to a variety of issues that need to be addressed, such as bi-directional power flows and increased harmonics. Thus collection and accurate interpretation of voltage and current harmonic content is rapidly becoming of paramount importance. For this reason, robust sensing technology is needed to ensure a high level of confidence in captured waveforms.

As presented in Chapter 1, all electrical sensing equipment possesses some degree of distortive characteristics, due to phenomena such as resonance, temperature, and saturation, among others. Typically, sensors used for protection and metering (such as instrument transformers) are designed to be extremely accurate for steady-state sinusoidal conditions at the fundamental frequency (50 or 60 Hz). For example, standards such as IEC 61869-2, [8], and IEC 61869-3, [9], for inductive current transformers (CTs) and potential transformers (PTs), provide accuracy limits at the fundamental frequency. These sensor designs present problems for applications monitoring harmonic content, limiting the ability to make accurate decisions based upon the retrieved frequency information. IEC 60044-8 [10], describes accuracy limits for instrument transformers in terms of *ratio* and *phase* errors. While these metrics are typically used for characterization of the effectiveness of the individual transformer effects, the ratio error may be modified into a simple percent error calculation.

Existing literature characterizes sensors in accordance with the ratio and phase errors discussed above, [11, 12, 13, 14, 15]. Crotti et al. in [11] use a voltage-divider setup along with a National Instruments (NI) PCI Extension for Instrumentation (PXI) data acquisition system to characterize voltage and current transformers as well as their effects on phasor measurement unit (PMU) measurements. In [13], techniques for metrological characterization of CTs and VTs is performed using generated signals at the fundamental frequency plus individual harmonic components.

In [14], Cataliotti et al. perform characterization of clamp-on CT's under non-sinusoidal situations. A compensation method for non-sinusoidal condition-based distortions for

CT's and VT's is presented in [12], by using current and voltage harmonics obtained in sinusoidal conditions to "correct" values obtained under non-sinusoidal conditions. In [15], a compensation method based on the "Best Linear Approximation" theory is used to correct for distortions of harmonic signals under non-sinusoidal conditions.

2.2 The Theory of Linear Systems

As discussed in Chapter 1, all sensing mechanisms will possess some degree of non-ideal distortive characteristics. Such characteristics often manifest in the form of *frequency* distortions. Frequency distortions are a direct consequence of a sensor or other measurement device's frequency response, which characterizes the gain, attenuation, and phase shift of each frequency component induced by the measurement device in question. In this chapter, common current waveform signatures are examined under different sensor responses and compared in terms of the effects on their harmonic components.

Many real-world systems may be modeled as a *linear system*. In brief, a linear system is one that satisfies both the additive and scaling properties. Mathematically, given two inputs $x_1(t)$ and $x_2(t)$, a system $T[\cdot]$ is linear if it produces outputs $y_1(t)$ and $y_2(t)$ in the following manner, [16]:

$$y_1(t) + y_2(t) = T\left[ax_1(t) + bx_2(t)\right] = aT\left[x_1(t)\right] + bT\left[x_2(t)\right] \tag{2.1}$$

for arbitrary constants a and b at all times t.

It is often convenient to model systems as both linear and *time-invariant*. The time-invariance property dictates that an input at any time $t = n - k$ will produce system response $h(n-k)$. In other words, the system's response function $h(t)$ does not change over time.

LTI systems give rise to the concept of *convolution*, namely that a system output is a sum of the input $x(t)$ multiplied against shifted iterations of the system response $h(t)$.

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \tag{2.2}$$

The convolution integral is a foundational operation in signal processing and provides a useful tool for system analysis. When dealing with discrete-time systems, this integral is represented as a sum:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \tag{2.3}$$

It describes the filtering operation of finite-impulse-response (FIR) filters. Systems represented as FIR filters possess the property of *linear phase*, meaning that all frequency components of a given input signal $x(t)$ are shifted by the same amount.

Sensor systems acting on sinusoidal signals, however, may not possess this linear phase property. Thus they need to be modeled as an infinite impulse response (IIR) filter:

$$y(n) = \sum_{k=0}^{N} b_k x(n-k) - \sum_{l=1}^{M} a_l y(n-l) \tag{2.4}$$

IIR filters are recursive in nature, and thus systems with response functions $h(t)$ are never bounded in t as FIR systems are.

2.3 Frequency Response

System responses $h(t)$ are more usefully studied in their frequency domain representations, $H(\omega)$, where $H(\omega)$ is simply the Fourier transform of $h(t)$. This representation is often referred to as the system's *frequency response*, and is complex-valued. The frequency response function of a system contains both magnitude and phase-shift information as functions of frequency ω, calculated respectively as:

$$|H(\omega)| = \sqrt{\Re(H(\omega))^2 + \Im(H(\omega))^2} \tag{2.5}$$

$$\angle H(\omega) = \arctan\left[\frac{\Im(H(\omega))}{\Re(H(\omega))}\right] \tag{2.6}$$

where $\Re(x)$ and $\Im(x)$ are the real and imaginary components of x. The magnitude (sometimes referred to as the "gain") $|H(\omega)|$, signifies how much the system $H(\cdot)$ amplifies

or attenuates input signals at frequency ω. Likewise, the phase $\angle H(\omega)$ contains information on how the system shifts the initial phase of each frequency component of the input signal.

Commercial Sensor Frequency Response Curves

Figure 2.1 shows an example of non-ideal sensor frequency response vs one that is closer to the ideal. The near-ideal magnitude curve shows a near-unity gain across the entire range of frequencies studied. However, the non-ideal curve shows a sharp resonance at roughly the 2 kHz mark, followed by a sharp downward trend. Frequencies passing through the non-ideal sensor in the immediate band around 2.5 kHz will be amplified by almost 50 times. Similarly, all frequencies below 2.5 kHz experience a phase shift between 140 and 190 degrees.

2.4 Mathematical Formulation for Comparisons

It is necessary to employ metrics that accurately capture the degree of similarities and differences between signals. For that reason, three techniques are discussed next: percent error, phase difference, and goodness-of-fit, [17].

Under steady-state conditions, a distorted current in the power system can be represented by a periodic function consisting of a 60-Hz fundamental wave written I_1 amps and N harmonics, written I_h amps, $h = 2, \ldots N$, with the form:

$$i(t) = \sqrt{2}\left[I_1 \sin(\omega t + \phi_1) + \sum_{h=2}^{N} I_h \sin(h\omega t + \phi_h)\right] \tag{2.7}$$

where $\omega = 2\pi \times 60$ rad/s, and ϕ_h is the phase angle of harmonic h. A sensor measuring the same signal will produce:

$$\hat{i}(t) = \sqrt{2}\left[\hat{I}_1 \sin(\omega t + \hat{\phi}_1) + \sum_{h=2}^{M} \hat{I}_h \sin(h\omega t + \hat{\phi}_h)\right] \tag{2.8}$$

where ($\hat{}$) indicates estimated parameters, and $M < N$ due to sensor frequency response limitations (e.g. bandwidth and sampling rate). However, this model does not hold for non-steady-state situations.

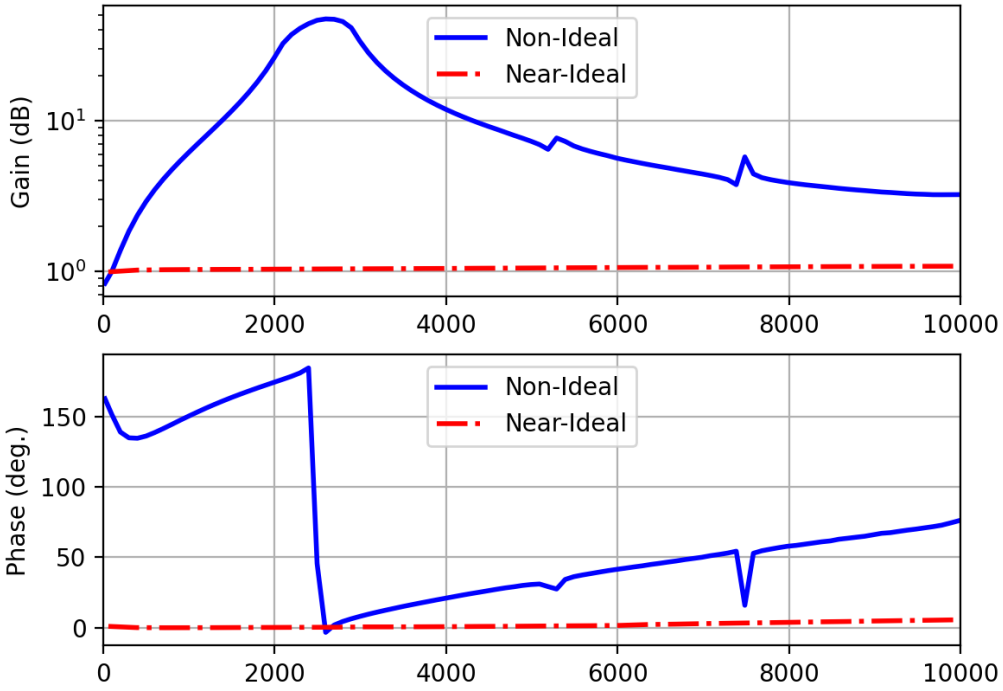


Figure 2.1: Commercial-grade near-ideal (red) and non-ideal (blue) sensor frequency responses

2.4.1 Percent Error

When characterizing sensor performance against the various harmonics, it is beneficial to use the standard percent error formula:

$$e_{ih} = \left| \frac{\hat{I}_h - I_h}{I_h} \right| \times 100\% \tag{2.9}$$

The expression in (3.5) describes the percent difference between input current harmonic amplitude I_h and measured current harmonic amplitude $\hat{I}_h$. The percent error formula is a common means of performance evaluation, and encompasses modified versions for gauging performance of instrument transformers, as described by the *ratio error* in [10].

2.4.2 Phase Difference

Along with (3.5), the relative phase displacement between actual and measured harmonic currents $\Delta\phi_{ih}$ may be calculated from [10] as:

$$\Delta\phi_{ih} = \phi_h - \hat{\phi}_h \tag{2.10}$$

Harmonic amplitudes I_h and $\hat{I}_h$ and phase angles ϕ_h and $\hat{\phi}_h$ are obtained using the Fast Fourier transform (FFT).

2.4.3 Goodness-of-Fit (GoF)

Riepnieks and Kirkham in [17] present a metric for evaluating the accuracy of PMUs based on a comparison of the measurement model (the mathematical definition of the quantities being measured) with the observed signal. We have adapted their "goodness-of-fit" (GoF) metric by comparing a time series of measurements recorded from the sensor with that from a reference or ground-truth sensor:

$$GoF = 20 \log_{10} \frac{A}{\sqrt{\frac{1}{N-m} \sum_{k=1}^{N} (i_k - \hat{i}_k)^2}}, \tag{2.11}$$

where A is the amplitude of the signal model, N is the number of points in the time-domain record, m is the number of parameters being measured (in this case, $m = 1$), i_k is the k^{th} sample of the reference time-domain current, and $\hat{i}_k$ is the k^{th} sample of the sensor time-domain current. The GoF is expressed in decibels (dB) in order to ensure the dynamic range is compressed, and the root-mean-square error is in the denominator of (3.12), yielding greater numbers for closer fits.

2.5 Waveform Playback System

At ORNL's DECC lab, a signal playback system has been developed. This system employs an National Instruments (NI) PXIe 5423 arbitrary waveform generator (AWG) that can produce voltages emulating digital signals stored in comma-separated-value (csv) text files.

This AWG produces analog voltages in the ± 2.4 V peak-to-peak range. These produced voltages are then converted to current and amplified through an AE 7228 power amplifier, modified to provide accurate frequency response within ±1% between 60 Hz and just under 5 kHz. The current is then stepped up using a KOR-11 15kV 400:5 T200CT whose frequency response was measured to be flat at least up to 10kHz. On the high side of the step-up transformer, a reference sensor is used to measure the actual current being fed into the equipment under test (EUT, i.e. the sensors being evaluated). This is because a phase delay of 20 μs is induced between the EUT signals and WG output signals due to the intermediate equipment.

The Ultrastab 866 Precision Current Transducer (current ratio of 1500:1) was chosen to serve as the reference sensor due to its extremely flat frequency response up to 100 kHz, and is connected connected to a 10-ohm burden resistor to measure currents of several hundred amps with accuracy better than 0.1 percent. The current signals also pass through the EUT and the measurements are fed back into the data acquisition system (DAQ) for side-by-side comparison with the reference sensor's readings. See Figure 2.2 for a diagram of the playback-sensing system.

The event recordings were sampled at a rate of 256 samples per cycle, or 15,360 Hz. To remain faithful to the original recordings, the DAQ system samples the received reference

and EUT signals at the same rate., and synchronizes them. Three sensors were used while conducting the experiments, denoted hereafter as S_1, S_2, and S_3. Each sensor works on a different operating principle, and thus have different eccentricities.

2.5.1 Event Descriptions

In total there are 5 events used for study. Each of the 5 events consists of three current phases (A, B, & C) with the exception of one, which only includes phases B & C. Phase A from this particular event did not present any deformities found in the other two phases, so it was discarded.

Fig. 4.3 shows current-vs.-time plots of the events under study. These events consist of arcing (Fig. 2.3a), capacitor switching (Figs. 2.3b-2.3c), and three-phase inrush (Figs. 2.3d-2.3e). All of the noticeable high-frequency contributions appear at the 2-second mark. In Fig. 2.3a, phase C (orange, thick line) shows a noticeable "blip" at the inception of the arc, whereas phase B (blue, thin line) shows an increase in amplitude for both the fundamental and harmonic components present in the peaks and valleys of the waveform. The capacitor switching events show clear high-frequency content in all three phases, the most extreme of which happen in the "spikes" of phases B (orange, middle-thickness line) and C (green, thick line). Similarly, the inrush events show high-frequency activity on all three phases.

2.6 Results and Discussion

Representative plots of sensor performance superimposed upon one another are shown in Fig. 2.4. In Fig. 2.4a, current phase C for the arcing event is shown. It can be seen that sensor S_1 (solid thin line, blue) possesses high-frequency components all along the waveform that are not seen in the reference (dashed, red) waveform. Sensor S_2 appears upon inspection to follow the reference waveform the closest, but it is not without noisy additions as well. Sensor S_3 has trouble responding to instantaneous events and appears to filter out higher harmonics. Figs 2.4b-2.4c show example sensor responses to capacitor switching and inrush events, respectively, which contain higher concentrations of harmonic content. S_1 appears

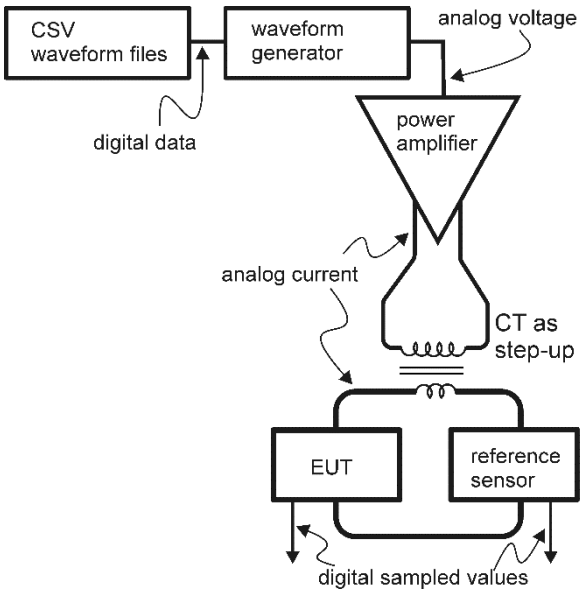
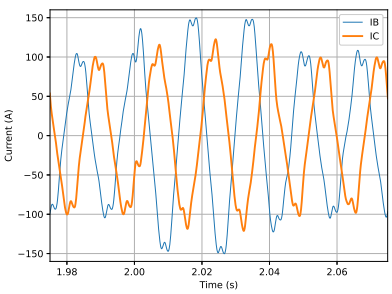
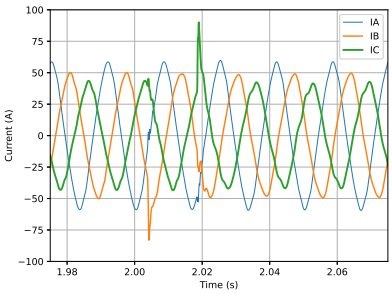


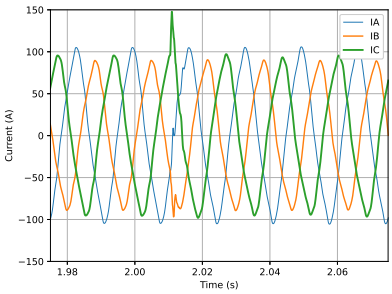
Figure 2.2: Waveform playback and DECC sensor system



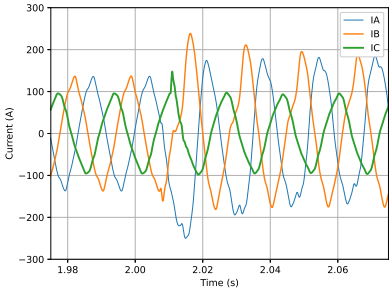
(a) Arcing



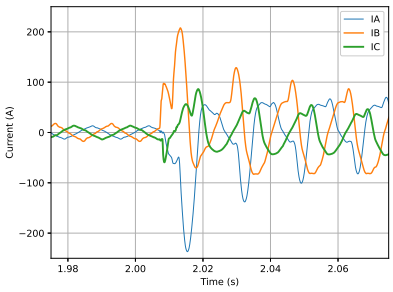
(b) Capacitor Switching 1



(c) Capacitor Switching 2



(d) Inrush 1



(e) Inrush 2

Figure 2.3: Events used for study

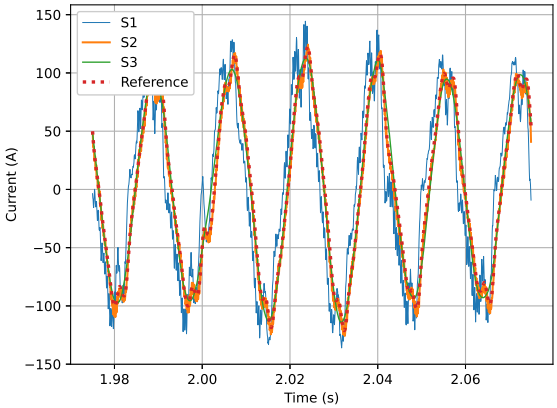
to be extremely sensitive to sudden high-frequency changes in the wave shape. This is due to a very strong resonance in the kHz range.

Tables 2.1-2.3 show amplitude error and phase difference, as computed using (3.5) and (2.10), for each of the events shown in Fig. 2.4. A six-cycle window (0.1 seconds) is used in the FFT calculation, with FFT size equal to 4096 to ensure a small-enough frequency resolution for extracting harmonic amplitudes and phases. Table 2.4 shows the fundamental amplitude errors and phase differences for these events. Additionally, Table 2.4 shows the *mean harmonic errors* and *phase differences* for each event across the 2^{nd} through 35^{th} harmonics.

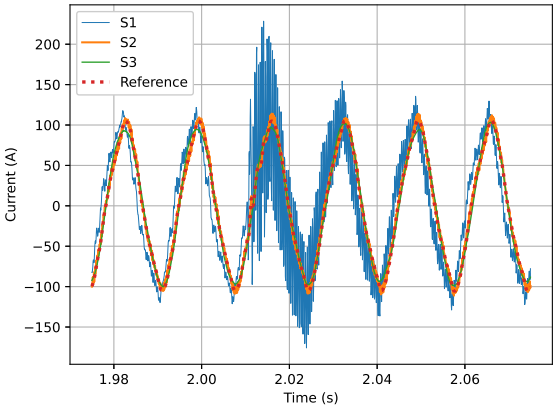
It is clear from the tables that S_1 presents a significant phase drift across all frequencies and events. S_1 appears to actually *lead* the reference signal during steady-state conditions, for reasons still unknown to the authors. Sensor S_3 performs best under slow-changing (i.e. steady-state) conditions, and does not respond well to sudden changes in amplitude or frequency. S_2 appears to perform the best across the board for all events.

Using the GoF metric defined in (3.12), each sensor's performance may be characterized as in Fig 2.5. The y-axis lists each of the events under study, the x-axis is the six-cycle (0.1 second) time window used, and each block represents a single cycle's worth of GoF computed using (3.12). It is clear from these figures that S_1 performs poorly for the tested events.

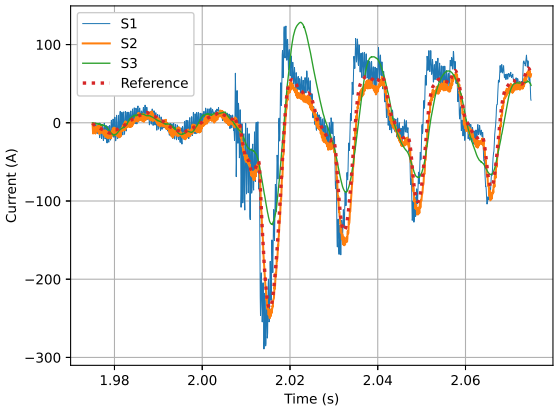
S_3 appears to work well for capacitor switching events, but is lackluster for capturing the inrush and arcing events studied. S_2 works best for capturing arcing, inrush event IN1, and the capacitor switching events, but fails to faithfully reproduce the signals contained in IN2. This is likely due to the fact that IN2 has much lower load current levels (15-20 amps as opposed to 100+ amps in the other events), and thus the noise present in the signal is amplified due to the sensor's dynamic range (< 30 kA).



(a) Arcing phase C



(b) Capacitor switching 2 phase A



(c) Inrush 2 phase A

Figure 2.4: Representative sensor responses

Table 2.1: Arcing phase C (AR-IC) harmonic amplitude/phase errors

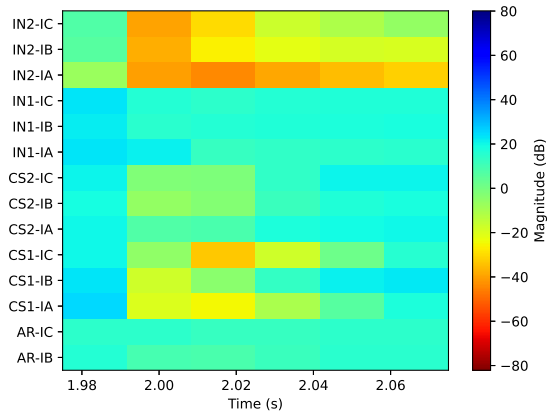
Parameter	**Harmonic No.**						
	1	3	5	7	11	13	19
S_1 (%)	2.3	30.3	108.4	178.6	70.3	246.3	287.5
S_2 (%)	0.1	2.4	0.1	0.1	0.8	3.1	18.6
S_3 (%)	0.8	42.3	83.4	85.9	40.5	47.2	42.5
S_1 (°)	-22.8	-30.6	-59.2	-64.9	-119	35.7	-111
S_2 (°)	-0.3	-0.5	0.1	0.4	-6.8	-0.1	12.3
S_3 (°)	-1.3	-75.9	130.8	-149.2	16.8	-37.1	2.3

Table 2.2: Capacitor switching phase A (CS2-IA) harmonic amplitude/phase errors

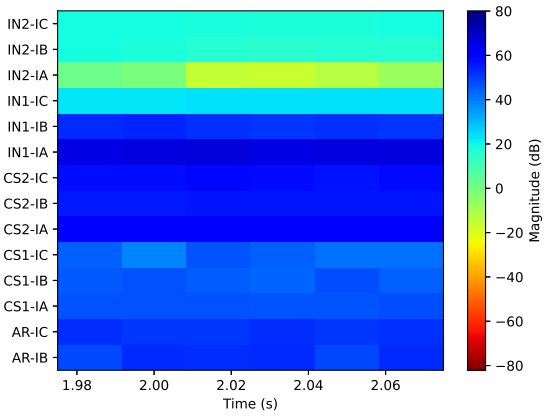
Parameter	**Harmonic No.**						
	1	3	5	7	9	11	13
S_1 (%)	2.7	19.7	42.5	78.7	9.5	230.5	135.2
S_2 (%)	0.4	2.1	1.3	1.4	5.8	3.3	3.4
S_3 (%)	0.3	47.1	31.6	21.4	32.5	60.2	16.1
S_1 (°)	-22.8	-36.3	94.6	91	-68.3	-26.6	87.2
S_2 (°)	-0.2	-1	-0.1	0.8	0.3	-1.4	-3.1
S_3 (°)	0.2	18.1	59.3	56.9	54.4	-1.2	48.3

Table 2.3: Inrush 2 Phase A (IN2-IA) harmonic amplitude/phase errors

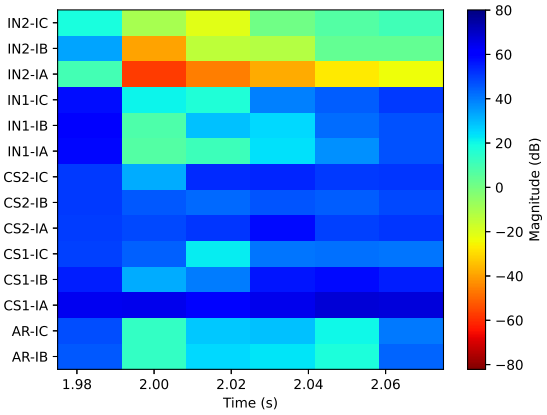
Parameter	**Harmonic No.**						
	1	3	5	7	9	11	13
S_1 (%)	3.8	47.7	100.1	122.6	59.7	132.4	158.1
S_2 (%)	1.8	0.1	5.9	4.6	11.9	5.8	3.8
S_3 (%)	0.6	71.7	75.8	70.9	50.9	64.8	66.4
S_1 (°)	-21.7	-44.3	-48.5	-77.7	-78.3	56.3	-53
S_2 (°)	2.5	1.8	2.6	0.2	2.5	-1.1	108.3
S_3 (°)	1.1	23.1	-1	-119.6	-70.7	-146.8	36.45



(a) S_1 GoF heatmap



(b) S_2 GoF heatmap



(c) S_3 GoF heatmap

Figure 2.5: GoF for each sensor around the event of interest

Table 2.4: Fundamental and average harmonic amplitude errors and phase differences for events not shown in Fig. 2.4. Values in the two right-most columns are averages over all non-fundamental frequencies.

Event	**Fund. (%)** S_1 S_2 S_3	**Fund. (°)** S_1 S_2 S_3	**Harm. (%)** S_1 S_2 S_3	**Harm. (°)** S_1 S_2 S_3
AR-IB	2.7 0.3 1.6	-22.9 -0.4 -1.4	66.8 5.6 11.1	-17.2 3.9 3.3
CS1-IA	3.1 0.4 1.3	-22.9 -0.6 0.3	142.5 8.3 8.1	-26.2 2.2 -0.2
CS1-IB	2.2 0.1 1.4	-23 -0.2 -1.6	524.7 14.5 46.3	-52.1 -3.3 -3.8
CS1-IC	3 1.2 0.8	-23.3 0.8 -1	538.3 14.9 23.4	-51.7 -4.4 21.9
CS2-IB	2.9 0.1 2.4	-23 -0.3 -1.7	631 30.4 37.4	-34.9 0.5 6.7
CS2-IC	3.1 0 0	-22.8 -0.3 -2.9	197.6 7,4 18.2	-15 -0.5 5.1
IN1-IA	3 1.3 2.5	-22.5 -0.3 -1.1	211.7 7.2 68.9	-41.8 -1.1 -18.9
IN1-IB	2.6 0 1.2	-22.8 -0.3 -1.7	72 6 24.3	0.9 0.5 6.1
IN1-IC	3.4 0 0.3	-22.7 -0.3 -2.3	46 23.1 16.9	-24.8 -0.3 0.5
IN2-IB	3.3 0.1 3.9	-23.1 -0.2 -3.1	607.2 15.4 31.5	-58.4 -9.3 18.2
IN2-IC	5.1 0.1 0.5	-22.5 -0.1 -3.3	390.7 11.8 35.5	-48 -4.7 32.8

Chapter 3

Statistical Profiling of Non-Ideal Sensor Responses to Transient Events

Chapter 2 presented the groundwork for this dissertation by examining frequency response characteristics of medium-voltage sensors. In this chapter, these characteristics, and their effects on high-frequency transient disturbances, are studied in a more in-depth manner. Probability density functions (PDFs) of harmonic magnitude and phase error statistics are estimated using kernel density estimation (KDE).

The advancement of power system observability has seen a dramatic shift in recent decades through the deployment of synchrophasors, or phasor measurement units (PMUs). PMUs provide the advantage of directly measuring voltage magnitude, phase angle, frequency, and rate-of-change-of-frequency (ROCOF); quantities that would otherwise be treated as latent variables through other measurement systems, requiring other real-time-capable signal processing solutions to extract. In spite of these recent developments, PMUs are typically only capable of measuring fundamental quantities; that is, information contained in the frequencies beyond 50/60 Hz are lost.

3.1 Existing Work

Much work has been done on the analysis of errors resulting from PMU measurement chains. In [18], the Gaussian assumption for PMU error is re-assessed, and corrected using Gaussian

Mixture Modeling (GMM), a technique for describing distributions comprised of more than one Gaussian "mode". The flaw in the Gaussian assumption is further examined in [19]. The authors in [20] explore various ways in which PMU error affect different applications, such as power system disturbance location, oscillation detection, island detection, and dynamic line rating. Analysis of PMU error for state estimation is also a burgeoning field of study [21, 22, 23, 24].

There have been works published on the estimation of harmonic phasors in the form of $\mathbf{A}_h = A_h e^{j\phi_h}$, in which harmonic h amplitude A_h, phase ϕ_h, frequency f_h, and rate-of-change-of-frequency (ROCOF) $ROCOF_h$ are the variables of interest. In [25] and [26], harmonic phasors are modeled as complex exponential functions and solved for via a least-squares approach applied to sampled frequency-domain models of the harmonic phasors. However, as the sampling rate increases, the necessary computational burden increases dramatically, to the point of being unable to perform the required matrix inversions on an 8-core machine.

Other techniques on extracting harmonic information from signals have been proposed as well. The authors in [27] employ a series of frequency-modulated finite impulse response (FIR) filters to estimate instantaneous harmonic parameters. A variation on the estimation of signal parameters using rotational invariance technique (ESPRIT) was proposed in [28], and in [29] in which the exact model order is estimated from the data rather than having to be configured and tuned by hand. Literature on *accurate* harmonic measurement in general is extremely sparse.

This study analyzes high-frequency *transient* electrical current waveforms captured by equipment dubbed hereafter as "Point-on-Wave" (PoW) sensors. At present, most studies concerning non-power frequency content of power system phenomena are reduced to total harmonic distortion (THD) and power quality index computations [30]. Harmonics are detrimental to the power system, and measures are typically taken to eliminate or reduce their effects rather than study them. However, complete removal of harmonics in the power system is almost impossible, and potential actionable information may be lost.

PoW sensors capture oscillographic representations of the measured phenomena, typically sending the resultant analog measurements to a device capable of digitally sampling at high rates. Two different commercially-available PoW sensors are compared with one another,

as well as against a "reference" sensor, representing the ideal. Statistical analysis of the harmonic amplitude and phase error for each sensor over a variety of transient current waveforms is performed, including the estimation of non-parametric probability density functions (PDFs) at each chosen harmonic. Note that equipment manufacturers will be kept anonymous in this document to avoid any perceived endorsement of one technology over another.

The motivation behind studying individual harmonic error probability distributions is a simple one. In parameter or state estimation applications, measurements are typically modeled mathematically in the form:

$$Y_{meas}(t) = Y_{true}(t) + err(t), \tag{3.1}$$

where *err* is an *error* term indicating random deviations between Y_{meas} and Y_{true}. This term is usually considered a lumped parameter, including contributions from systematic and random errors. In the power system scope, Y is usually an electrical parameter of interest, such as voltage magnitude, phase angle, or frequency. The work presented in this study examines error of electrical current harmonics as a result of high-frequency disturbances and characterizes the behavior of harmonic amplitude and phase error over a variety of dominant harmonics frequencies.

Individual harmonic error distribution analysis has not been studied in-depth. We hope this study will continue the advancement of the body of knowledge to be used for improving power system observability and situational awareness.

The contributions of this chapter may be summarized as follows:

- Harmonic analysis of three distinct power system current disturbance types, two of which are products of real-world scenarios: a current inrush event and a microgrid switching "online" to the grid (i.e. a switch closing). The third disturbance is simulated based on an electromagnetic transients program (EMTP) model of a wind farm experiencing a short-circuit three-phase fault at its terminals, as measured at said terminals. This event was included in this study to highlight extremely high harmonics well into the 10's of kHz by modifying the simulation step to a very small value.

- Direct comparison of two commercial-grade distribution PoW sensors against a near-idealized reference, in terms of amplitude percent error, amplitude residual error, and phase difference.
- Analysis of the Gaussian (i.e. normal) distribution assumption of harmonic errors using the Anderson-Darling test.
- Use of a non-parametric probability density function estimation technique, known as *kernel density estimation* (KDE), to learn generalized distributions of harmonic amplitude and phase errors. A mean-squared-error goodness-of-fit (GoF) metric is used to quantify the results of this estimation process.

3.2 Harmonic Extraction

To capture non-fundamental frequency information, it is necessary to capture the frequency-domain information of the received measurement waveform(s). This is typically accomplished through the computation of the Discrete Fourier Transform (DFT) via the Fast Fourier Transform (FFT) algorithm. For a measured discrete-time signal $x[n]$, its DFT representation $X(k)$ is computed as:

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{\frac{-j2\pi kn}{N}} \tag{3.2}$$

where the frequency at bin k may be computed as $f_k = k\frac{F_s}{N}$, F_s is the sampling frequency, N denotes the length of the FFT vector, and $j = \sqrt{-1}$. Because the frequency vector $\hat{\mathbf{f}}$ is discretized, estimation of the fundamental frequency is dependent on finding the nearest frequency bin $\hat{k}$, to the query frequency, in this case, 60 Hz:

$$\hat{k} = \underset{k}{\operatorname{argmin}} \left[\left(\hat{\mathbf{f}} - 60 \right)^2 \right] \tag{3.3}$$

$$f_{\text{fun}} = \hat{\mathbf{f}}[\hat{k}], \tag{3.4}$$

where $\hat{k}$ denotes the estimated FFT bin closest to 60 Hz, and f_{fun} is the estimated fundamental frequency obtained from the FFT. This value may be confirmed by finding the maximum value in $|X(k)|$, because it is reasonable to expect the fundamental frequency of power signals will be the dominant feature in the DFT magnitude spectrum. Harmonics of this fundamental frequency can then be *estimated* by taking multiples of f_{fun}. To extract the corresponding harmonic frequency FFT bins, simply replace the query frequency (60 Hz in (3.3)) with $h \times f_{fun}$, where $h = 2, \ldots, H$ is the harmonic order, and H denotes the maximum number of harmonics in the signal.

3.3 Statistical Analysis of Harmonics

Often error analysis is performed under the assumption of a Gaussian distribution - that is, error is typically assumed to take the form of additive white Gaussian noise (AWGN). In this section, error metrics for harmonic magnitudes and phases are presented. Two metrics for harmonic amplitudes are first discussed: percent error and residual error. Phase error is computed using a simple difference.

Given a sensor under test (SUT), and an ideal reference sensor measuring the same quantity side-by-side, the percent error at a specific harmonic amplitude h may be quantified as, [31]:

$$err^h_{I,\%} = \frac{\left|I^h_{sut} - I^h_{ref}\right|}{I^h_{ref}}. \tag{3.5}$$

This quantity may be multiplied by 100 if it is desired to be expressed in percent. Otherwise, $0 \leq \ err^h_I \ \leq 1$ is unitless, and provides a relative measure of the deviation of a measured harmonic amplitude I^h_{sut} from the reference harmonic amplitude I^h_{ref}. The residual harmonic amplitude error may be calculated using:

$$err^h_{I,res} = I^h_{ref} - I^h_{sut}. \tag{3.6}$$

Similarly, the phase may be compared with a simple difference:

$$\Delta\phi^h_I = \phi^h_{I,ref} - \phi^j_{I,sut} \tag{3.7}$$

where $\phi^h_{I,ref}$ and $\phi^j_{I,sut}$ represent reference and measured phase angles at harmonic order h, respectively.

It is useful to determine the distributions of $err^h_{I,\%}$, $err^h_{I,res}$, and $\Delta\phi^h_I$. Knowing these distributions of may allow measurement devices to make corrections if errors are suspected. These distributions may also give insight as to if a particular sensor, measuring a particular harmonic h, possesses systemic error (loosely equivalent to a *bias*), or if the error seems purely random. The natural assumption is to assign Gaussian distribution to error quantities, but as will be shown later in this work, that is not necessarily the case at each harmonic!

3.3.1 The Anderson-Darling Test

A common problem in statistical inference is determining a distribution, or family of distributions, that a given sample has come from. It is often not sufficient to simply visualize a histogram of data, and more rigorous methods are required to fully *quantify* the "goodness-of-fit" of a distribution family to a given sample.

In [32], T. W. Anderson and D. A. Darling proposed a test statistic used for accomplishing this. Given an ordered sample $x_1 \leq x_2 \leq \cdots \leq x_n$ with cumulative distribution function $F(x)$, compute

$$W_n^2 = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\,G(x_j), \tag{3.8}$$

for

$$G(x_j) = \left[\log\left(F(x_j) + \log\left(1 - F(x_{n-j+1})\right)\right)\right]. \tag{3.9}$$

Stephens in [33] notated the statistic W_n^2 for various distributions at various significance levels. For a test against a normal distribution with unknown parameters at a significance level of $p = 0.05$, the "threshold" is 0.787. This means that if computation of W_n^2 yields a number greater than this threshold, the test will *reject* the hypothesis that the sample x came from a normal distribution. MATLAB provides a simple function, `adtest()`, that accepts a series of numbers as an input, and yields a 0 if the input sample likely came from a normal distribution at a significance level of $p = 0.05$ or less and a 1 otherwise.

3.3.2 Kernel Density Estimation (KDE)

Often a random sample does not appear to come from a known family of distributions. In this case, nonparametric techniques are usually applied to estimate the distribution function $f(x)$. One of the more common approaches to this problem is that of *Kernel Density Estimation* (KDE). A density function f describing the distribution of a random variable X may be approximated as $\hat{f}_h(x)$ using the kernel density estimator:

$$\hat{f}_h(x) = \frac{1}{hn} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{3.10}$$

for some kernel function K and bandwidth, or *smoothing* parameter h. In many applications the standard normal kernel is assumed:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp \frac{-(x^2)}{2} \tag{3.11}$$

KDE essentially overlays the kernel function K over the data histogram, computes the kernel function on the values of x_i within the kernel, shifts the kernel function, and sums the result (the summation in (3.10)), yielding a continuous function approximating the true density $f(x)$. The bandwidth parameter h controls the width of the kernel function. Ideally h would be as small as possible, however too small an h will result in overfitting. Similarly, too large an h will result in a curve that is *too* smooth. An example is shown in Figure 3.1.

Goodness-of-Fit using Root Mean Squared Error

A simple yet effective metric for gauging the goodness-of-fit (GoF) of a probability density estimate $\hat{f}$ is a simple mean-squared error calculation between the empirical cumulative distribution function (ECDF) calculated from the data, $\hat{F}(x)$, and the estimated CDF $F(x)$, computed given the estimated density function $\hat{f}(x)$:

$$GoF = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(F(x_i) - \hat{F}(x_i)\right)^2} \tag{3.12}$$

Ideally, the GoF for an estimated distribution function $\hat{F}(x)$ will be as close to zero as possible, indicating little deviation between $F(x)$ and $\hat{F}(x)$. Note that the GoF metric here differs from that presented in Chapter 2.

3.4 Experimental Setup

Using the experimental setup from Chapter 2, each event was "played back" through the sensor suite 100 times, yielding 100 comparisons between a sensor's produced signal and the reference sensor signal. In this chapter, only sensors S_2 and S_3 from Chapter 2 are employed, as S_1 showed extreme deviations from typical operation. These two sensors are hereafter denoted as S_1 and S_2, respectively.

3.4.1 Event Descriptions

Three event types were studied: a current inrush event (denoted as E_1), a microgrid close-in event E_2, and an event depicting a fault on the terminals of a wind farm connected to a distribution system E_3. Events E_1 and E_2 are from real world data, whereas the wind-fault event was simulated in PSCAD, [34]. Events E_1 and E_2 were sampled at a different rate (20 kHz and 30.72 kHz, respectively) due to the nature of their originating measurement sources, and the sampling rate for the simulated event E_3 was chosen to be 200 kHz to obtain as many harmonics from the wind fault waveform as possible, as well as ensuring that the Nyquist frequency (i.e 100 kHz) matched the frequency response limit of the reference sensor.

Figure 3.2a-3.2c depict single-phase current waveforms produced by S_{ref}, S_1, S_2. It can be seen that in most cases, S_2 produces a significant amount of noise. It can also be seen by examination alone that, while S_1 appears to follow S_{ref} more closely at lower frequencies (Figure 3.2a), it begins to deviate more at higher frequencies, (Figures 3.2b-3.2c).

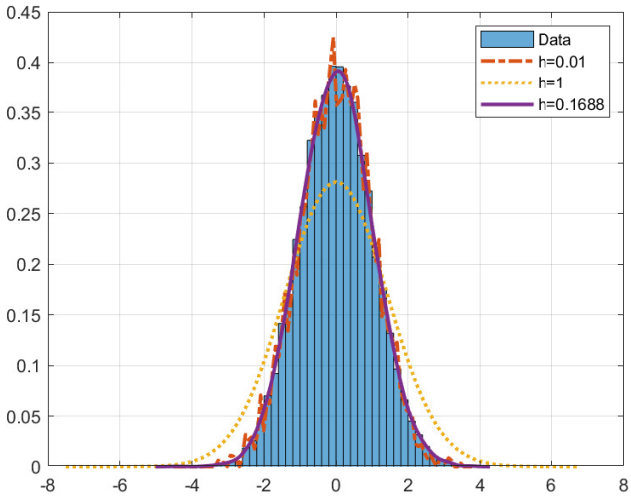
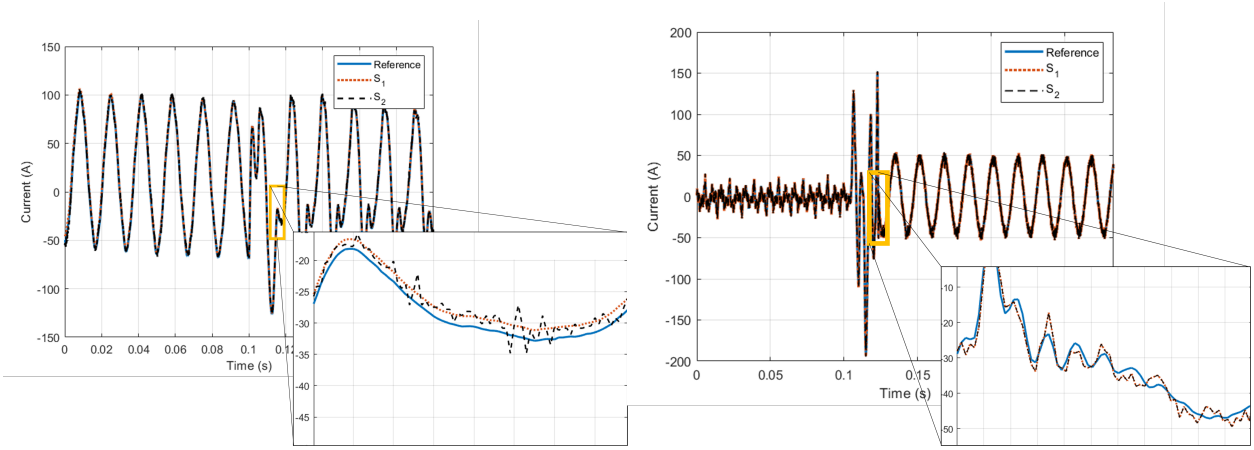
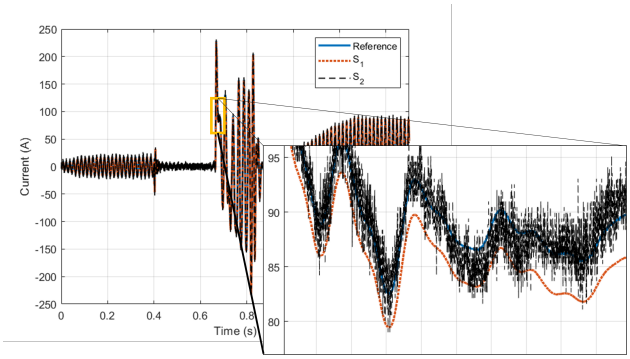


Figure 3.1: Example of KDE on data drawn from the standard normal distribution. The blocks indicate the histogram of the raw data. The solid line indicates the best "fit". The dash-dotted line shows a case in which the value of h is too small. The lighter dotted line shows a case of h being too large.



(a) Inrush Current Waveform

(b) Microgrid Close-in Current Waveform



(c) Wind Fault Current Waveform

Figure 3.2: Event Waveforms

3.4.2 Selecting Harmonics

For a given event waveform, it is not likely that every frequency component between 0 and the Nyquist frequency $F_s/2$ will be present in the examined signal. For this reason, harmonics were hand-picked from the prominent "peaks" in the waveform frequency spectra. Table 3.1 lists the chosen harmonics. Note that the fundamental frequency (60 Hz) is excluded due to the large existing body of knowledge and design characteristics included to ensure peak performance at this frequency.

3.5 Results

As mentioned in the previous section, 100 trials for each event play-through were conducted using both sensors independently. After taking the FFT of each signal's trial, the selected harmonic amplitudes and phases were extracted.

3.5.1 Anderson-Darling: Testing for Normality

At this point, the distribution of each harmonic was tested using the Anderson-Darling test, as described in Section 3.3. The x_j's used in (3.8) represent each of the three computed error metrics. Tables 3.2-3.4 show the Anderson-Darling test results for both sensors. Clearly for frequencies below the 50^{th} harmonic (approximately), as produced by S_1, exhibit normally-distributed behavior (indicated with a **1** in the tables) with few exceptions (harmonics 7 for both residual amplitude and percent error, and 61 for just percent error). However, it should be noted that the AD test does not definitively *prove* that a distribution follows "normal" behavior; it computes the probability (the p-value) that the assumption of a normal distribution is true. In other words, if $p < 0.05$, there is enough evidence to reject the hypothesis that the given distribution is normal, implying that the result indicating allowable rejection of the null hypothesis is *statistically significant* with $\alpha = 100\% \times (1 - p) = 95\%$ confidence.

Both amplitude metrics computed from S_2 samples show non-normal behavior at harmonics 37 and 49. The phase as captured by S_2 has an interesting mix of normal and

Table 3.1: Selected harmonics for each event of interest

Inrush	**Microgrid Close-in**	**Wind Fault**
2	3	7
3	5	14
4	7	149
5	9	195
6	12	232
7	13	270
11	17	389
13	18	427
17	19	464
23	20	501
27	21	600
31	23	659
37	25	697
49	27	734
61	29	854
	31	892
	33	1086
	35	1124
	37	1161
	39	1281
	41	1318
	43	1360
	47	1393
	49	1550
	51	1588
		1625

non-normal distributions, the most notable standout being the 7^{th} harmonic exhibiting non-normal qualities. However, the pattern shifts towards non-normal as the frequencies increase, as with the wind fault case. Example non-normal distribution plots will be presented in Section 3.5.2.

3.5.2 Distribution Fitting via KDE

As described in Section 3.3, KDE was used to estimate a continuous distribution from the data samples obtained at each harmonic, for all three error metrics: percent amplitude error, residual amplitude error, and phase error. Figures 3.3a-3.3b depict examples of predicted normal and non-normal distributions, respectively. To the naked eye, Figure 3.3a does not appear to be normally distributed, however there was not enough sufficient evidence in the data to reject this hypothesis when performing the AD test on this particular harmonic. Figure 3.3b shows a case of a harmonic amplitude's distributions failing the AD test, and the skewness of the distribution clearly reflects this. Also included in this figure are the RMSE values for the estimated distributions. For more examples of estimated distributions exhibiting non-normal behavior, see the harmonic error distributions for the wind fault case in Appendix A.

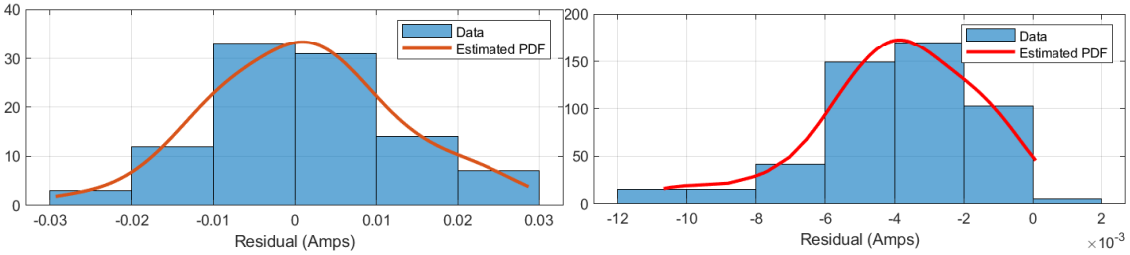
3.5.3 Goodness-of-Fit

Each of the computed distributions were then tested against the empirical data by using (3.12). The GoF results for both sensors' harmonic distributions over all three events may be seen in Tables 3.5-3.7. It can be seen that for both sensors, the RMSE tends to lie around the $0.02 - 0.03$ mark, meaning that, on average, the probability of the harmonic amplitude or phase error X being less than or equal to some value x differs by $0.02 - 0.03$ between the empirical data CDF $F(x)$ and the estimated CDF, $\hat{F}(x)$.

Bold items in Tables 3.5-3.7 indicate higher RMSEs between the two sensors' estimated PDFs for a given metric. For example, in Table 3.5, KDE seemed to perform worse on phase error for S_2. However, in Table 3.6, S_2's estimated PDFs for phase error outperformed those of S_1. This goes to show that different sensors will yield different error distributions

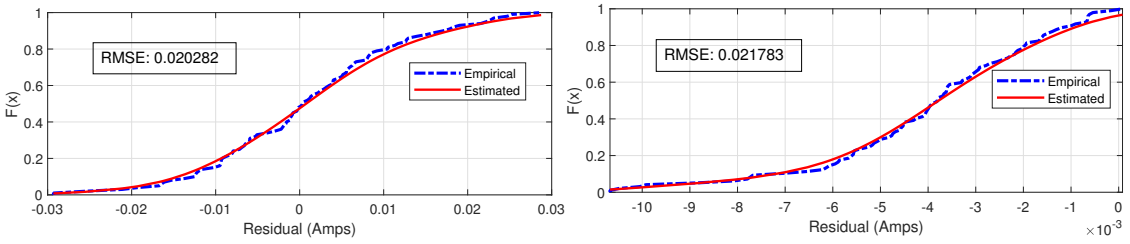
Table 3.2: AD results for inrush event distribution as seen by S_1 and S_2

Harmonic #	Amplitude (%)		Amplitude (res.)		Phase	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	**1**	0	**1**	0	0	**1**
11	0	0	0	0	0	0
13	0	0	0	0	0	0
17	0	0	0	0	0	0
23	0	0	0	0	0	**1**
27	0	0	0	0	0	**1**
31	0	0	0	0	0	**1**
37	0	**1**	0	**1**	0	**1**
49	0	**1**	0	**1**	0	**1**
61	**1**	0	0	0	0	**1**



(a) Example Normal Fit: Histogram and Estimated PDF

(b) Example Non-Normal Fit: Histogram and Estimated PDF



(c) Example Normal Fit: Empirical and Estimated CDFs

(d) Example Non-Normal Fit: Empirical and Estimated CDFs

Figure 3.3: Normally-distributed (a), (c) estimated residual amplitude error for harmonic # 7 as seen by S_2 for the wind fault transient event, and non-normally-distributed (b), (d) example harmonic # 1625.

Table 3.3: AD results for microgrid close-in event distribution as seen by S_1 and S_2

Harmonic #	**Amplitude (%)**		**Amplitude (res.)**		**Phase**	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
3	0	0	0	0	0	0
5	0	0	0	0	0	0
7	0	0	0	0	0	0
9	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
23	0	0	0	0	0	0
25	0	0	0	0	0	0
27	0	0	0	0	0	0
29	0	0	0	0	0	0
31	0	0	0	0	0	0
33	**1**	0	**1**	0	0	0
35	0	0	0	0	0	**1**
37	0	0	0	0	0	0
39	0	**1**	0	**1**	0	**1**
41	0	0	0	0	0	0
43	0	0	0	0	0	**1**
47	0	0	0	0	0	0
49	0	**1**	0	**1**	0	0
51	0	0	0	0	0	0

Table 3.4: AD results for wind-fault event as seen by S_1 and S_2

Harmonic #	**Amplitude (%)**		**Amplitude (res.)**		**Phase**	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
7	0	0	0	0	**1**	0
14	0	0	0	0	0	0
149	**1**	**1**	0	**1**	0	**1**
195	**1**	**1**	0	0	0	**1**
232	**1**	**1**	0	0	0	**1**
270	**1**	**1**	**1**	**1**	**1**	0
389	**1**	**1**	0	**1**	**1**	0
427	**1**	**1**	**1**	**1**	**1**	0
464	**1**	**1**	0	**1**	0	**1**
501	**1**	**1**	0	**1**	**1**	**1**
600	0	**1**	0	**1**	**1**	0
659	**1**	**1**	**1**	**1**	**1**	0
697	**1**	**1**	0	0	**1**	**1**
734	**1**	**1**	0	0	0	**1**
854	**1**	**1**	0	**1**	**1**	**1**
892	0	**1**	0	0	**1**	0
1086	**1**	**1**	**1**	0	**1**	0
1124	**1**	**1**	**1**	0	**1**	0
1161	**1**	**1**	**1**	0	**1**	**1**
1281	**1**	**1**	0	**1**	**1**	**1**
1318	0	**1**	**1**	**1**	**1**	0
1360	**1**	**1**	**1**	0	**1**	1
1393	**1**	**1**	**1**	0	**1**	**1**
1550	**1**	**1**	0	0	**1**	0
1588	**1**	**1**	**1**	**1**	**1**	0
1625	**1**	**1**	**1**	**1**	**1**	0

over different harmonics, and that there is no "one-size-fits-all" solution to learning error characteristics.

3.6 Conclusions

A fully situationally-aware power system is a goal that, while seemingly impossible to achieve, is something worth pursuing. In this chapter, high-frequency transient power system current disturbances and their distorted representations are analyzed through both statistical and probabilistic lenses over a wide variety of harmonic frequencies. Harmonic amplitude error, quantified in terms of percent and residual, largely showed characteristics of normally-distributed behavior per the Anderson-Darling test for normality in the lower (i.e. less than the 50^{th} harmonic) frequencies.

As the harmonic frequency moves beyond this level, the error distributions tend to drift away from normal behavior, as seen in the wind-fault event results. RMSE was used as an indicator for goodness-of-fit between estimated distribution functions and empirical data, showing the validity of the presented approach. The results presented in this chapter go against the assumption that measurement error may be treated as a normally-distributed quantity.

Table 3.5: RMSE for Inrush Harmonic Distributions

Inrush						
Harmonic	**%**		**res.**		**Phase**	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
2	0.020	**0.022**	0.019	**0.022**	**0.022**	0.020
3	0.019	**0.023**	0.019	**0.023**	**0.024**	0.023
4	0.017	**0.019**	0.017	**0.019**	0.019	**0.022**
5	0.017	**0.023**	0.017	**0.023**	0.023	**0.023**
6	**0.021**	0.020	**0.022**	0.020	0.018	**0.021**
7	**0.030**	0.027	**0.030**	0.027	0.020	**0.028**
11	**0.023**	0.023	**0.023**	0.022	0.021	**0.022**
13	0.024	**0.026**	0.024	**0.026**	0.019	**0.020**
17	0.020	**0.020**	0.020	**0.021**	0.019	**0.019**
23	**0.026**	0.021	**0.026**	0.022	0.022	**0.037**
27	0.021	**0.023**	0.022	**0.023**	**0.026**	0.022
31	**0.023**	0.021	**0.023**	0.022	0.022	**0.062**
37	0.022	**0.027**	0.020	**0.026**	0.022	**0.035**
49	0.020	**0.026**	0.019	**0.023**	0.022	**0.026**
61	**0.031**	0.022	**0.028**	0.022	0.020	**0.039**

Table 3.6: RMSE for Microgrid Close-in Harmonic Distributions

Microgrid Close-in						
Harmonic	**%**		**res.**		**Phase**	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
3	**0.024**	0.021	**0.024**	0.021	**0.024**	0.021
5	**0.019**	0.019	0.019	**0.020**	**0.022**	0.021
7	0.020	**0.022**	0.020	**0.023**	**0.025**	0.022
9	0.019	**0.021**	0.019	**0.021**	**0.020**	0.019
12	0.021	**0.025**	0.021	**0.026**	0.017	**0.020**
13	**0.020**	0.019	**0.020**	0.019	**0.024**	0.021
17	0.019	**0.020**	**0.020**	0.019	0.020	**0.022**
18	0.021	**0.025**	0.021	**0.024**	**0.022**	0.021
19	0.020	**0.024**	0.020	**0.023**	**0.021**	0.020
20	0.019	**0.022**	0.019	**0.020**	0.028	**0.031**
21	0.022	**0.023**	0.022	**0.023**	0.019	**0.023**
23	0.018	**0.023**	0.018	**0.022**	**0.023**	0.020
25	0.021	**0.026**	0.021	**0.026**	0.020	**0.021**
27	**0.019**	0.019	**0.019**	0.019	0.018	**0.025**
29	**0.024**	0.021	**0.025**	0.019	**0.021**	0.020
31	0.018	**0.019**	**0.018**	**0.019**	**0.025**	0.023
33	**0.022**	0.020	**0.023**	0.020	**0.023**	0.018
35	**0.025**	0.019	**0.025**	0.019	0.022	**0.023**
37	**0.020**	0.019	**0.020**	0.020	**0.021**	0.020
39	0.023	**0.024**	0.023	**0.024**	0.017	**0.031**
41	0.022	**0.029**	0.022	**0.029**	0.020	**0.022**
43	**0.021**	0.021	**0.021**	0.021	**0.025**	0.025
47	0.021	**0.022**	0.021	**0.022**	**0.020**	0.019
49	**0.022**	0.022	0.022	**0.022**	**0.022**	0.022
51	0.022	**0.027**	0.022	**0.027**	0.019	**0.021**

Table 3.7: RMSE for Wind Fault Harmonic Distributions

Wind Fault						
Harmonic	**%**		**res.**		**Phase**	
	S1	**S2**	**S1**	**S2**	**S1**	**S2**
7	**0.023**	0.020	**0.023**	0.020	**0.028**	0.024
14	0.020	**0.020**	0.020	**0.020**	**0.024**	0.022
149	0.027	**0.028**	0.026	**0.026**	0.024	**0.024**
195	0.025	**0.034**	0.021	**0.022**	**0.025**	0.021
232	**0.028**	0.022	0.020	**0.024**	0.021	**0.029**
270	0.023	**0.029**	0.021	**0.023**	0.023	**0.024**
389	0.024	**0.028**	**0.023**	0.022	**0.026**	0.023
427	0.023	**0.030**	**0.026**	0.022	**0.025**	0.022
464	0.026	**0.026**	0.020	**0.025**	0.019	**0.028**
501	**0.027**	0.020	0.024	**0.024**	0.023	**0.025**
600	0.021	**0.029**	0.024	**0.025**	**0.031**	0.022
659	**0.029**	0.025	**0.024**	0.019	**0.031**	0.023
697	**0.023**	0.021	0.020	**0.025**	**0.023**	0.021
734	**0.026**	0.022	0.018	**0.022**	0.025	**0.029**
854	0.027	**0.030**	**0.027**	0.025	**0.035**	0.022
892	0.021	**0.028**	0.022	**0.026**	**0.029**	0.022
1086	**0.026**	0.025	**0.023**	0.019	**0.039**	0.024
1124	**0.028**	0.027	**0.027**	0.021	**0.039**	0.026
1161	**0.021**	0.020	**0.022**	0.020	**0.035**	0.028
1281	**0.026**	0.023	0.020	**0.022**	**0.043**	0.025
1318	0.025	**0.026**	0.021	**0.022**	**0.025**	0.021
1360	**0.031**	0.030	**0.033**	0.020	**0.036**	0.030
1393	**0.030**	0.025	0.024	**0.026**	**0.040**	0.030
1550	0.020	**0.027**	**0.025**	0.019	**0.025**	0.019
1588	0.025	**0.031**	**0.022**	0.021	**0.058**	0.023
1625	**0.030**	0.025	**0.028**	0.022	**0.038**	0.024

Chapter 4

Detection of Corrupted Signals Using the "Energy Detector"

The previous two chapters examined the way common off-the-shelf sensors are susceptible to high-frequency distortions in their produced output waveforms. In this chapter, the dissertation shifts its focus towards the examination of how these sensor irregularities effect applications in the power grid, specifically with respect to the *detection* of high-frequency events.

4.1 The "Detection Problem"

Modern protection and monitoring systems utilized in the power industry rely on accurate detection of disturbances that may be indicative of undesirable behavior in the system. In traditional relaying mechanisms, configurable current thresholds and sequence component transforms provide simple means of detecting faults in order to send a trip signal to the associated circuit breaker(s) to protect the system from drawing dangerous levels of fault current.

Traditional faults are relatively simple to detect using existing hardware and software systems. However, the power grid is quickly transitioning away from the familiar unidirectional and radial system into a decentralized network of independently-operated generating sources (e.g. microgrids) working in tandem with the existing infrastructure.

The push for using renewable energy sources to supplement generation capacity poses the challenge of operators not presently being equipped to recognize bidirectional power flows and the subsequent voltage and current phenomena that result.

There are a wide variety of detection algorithms available in the literature. Many of these algorithms, however, assume a certain (often very high) level of available computational capability. For example, many works make use of the Discrete Wavelet Transform (DWT) for detection of abrupt signal changes, an algorithm notorious for its computational complexity due to a series of convolution operations at every iteration, [35, 36, 37, 38, 39, 40].

Time-frequency (TF) methods are popular in the literature, including the above-mentioned DWT method. Another popular detection technique employs the Short-Time Fourier Transform (STFT) and its variations, [41, 42, 43, 44, 45]. The STFT computes the Fourier Transform over a specified window of the signal of interest. This process is repeated until the entire signal has been analyzed, yielding a matrix of complex coefficients that may be used to extract the frequency information of the signal as it varies with time. One fundamental limitation of the STFT is the time-frequency uncertainty principle which states $\Delta t \Delta \omega \geq 2\pi$, [46]. This relation reveals that for an increase in time resolution Δt, the frequency resolution $\Delta \omega$ will decrease, and vice versa, as time and frequency are fundamentally inversely-proportional quantities, therefore introducing uncertainty in analysis using results of the STFT and its variants.

Other detection techniques require certain assumptions about the system under observation. In [47], a technique involving successive differences of points along the waveform cycle is presented. This method assumes a fixed fundamental frequency, however it is widely known that the power frequency in the U.S. can vary at 60 ± 0.1 Hz, [2]. Over time, the detection algorithm will be subject to frequency drift which will lead to inaccurate results.

This chapter presents a novel yet simple parameter-agnostic power system event detection algorithm to address the deficiencies defined above, as well as its performance against transient signals that have been corrupted by the sensor irregularities discussed in previous chapters. The *energy detector*, based on the theory developed in [48] and proposed initially in [49], provides a simple yet computationally-efficient means of power system event detection that is free of the dependence on signal parameter assumption.

Contributions of this chapter include:

- Presentation of the *energy detector* for high-frequency power system event detection
- Evaluation of the energy detector's performance when subjected to multiple types of the same event when subjected to sensor frequency response errors.
- Comparison against other common detection algorithms, in most cases showing superior performance in its capability to detect the precise start time.

4.2 The Energy Detector

This section presents the *Energy Detector (ED)*, a parameter-agnostic algorithm capable of quickly detecting the beginnings and endings of transient signals.

4.2.1 Signal Model

A voltage or current wave may typically be represented as a simple sine or cosine expression with constant amplitude A, frequency f, and initial phase θ:

$$s(t) = A\cos\left(2\pi f t + \theta\right) \tag{4.1}$$

It is assumed that none of these parameters change during the observation period (i.e. measurement window).

Often these signals are captured by some sort of data acquisition (DAQ) system, the first operation of which is that of analog-to-digital conversion (ADC). This process involves digital sampling the incoming signal $s(t)$ at a finite rate f_s. The resultant digital signal at sample index k may be approximated by

$$s[k] = A\cos\left(2\pi f\frac{k}{f_s} + \theta\right) \tag{4.2}$$

4.2.2 Algorithm Derivation

A captured signal $r[k]$, through use of a sensor/DAQ system contains an unknown noiseless signal $s[k]$ (i.e. voltage or current in this case), additive noise $n[k]$, system linear time-invariant (LTI) impulse response approximation $h[k]$, and contaminated disturbance $f[k]$, and can be expressed by the following equation [48]:

$$\begin{aligned} r[k] &= s[k] * h[k] + n[k] + f[k] \\ &= s[k] * h[k] + n[k] + q[k]p_{T_q}[k - N_p] \end{aligned} \tag{4.3}$$

where "$*$" denotes the linear convolution operator, $q[k] \neq 0$ denotes the random process (e.g. transient) of interest, p_{T_q} is the windowing function that captures the duration of the transient (short–term) portion of the waveform, and N_p is the delay which determines the temporal location of transient. The transient in this context ($q[k]p_{Tq}[k - N_p]$) is treated as a random process superimposed on the steady-state (i.e. 60 Hz) waveforms. This problem may be boiled down to a simple hypothesis test where $H_0 : \ p_{Tq}[k] = 0$ and $H_1 : \ p_{Tq}[k] \neq 0$.

The decision statistic for the energy detector output $\psi[k]$, can be expressed by the equation:

$$\psi[k] = \sum_{i=0}^{N-1} |r[k - i]|^2 \geq \gamma, \tag{4.4}$$

where N is the number of samples taken per "bin", or windowing iteration. The conventional energy detector calculates the associated energy of the received signal over a given duration N and compares the resulting energy value(s) with a selected threshold.

For this work, the appropriate threshold value γ is determined empirically. After computing $\psi[k]$, the energy values may be sorted from lowest to highest, as in Fig. 4.1. The point after the "elbow" at which the curve begins to flatten is chosen as γ. Threshold selection may also be computed analytically. For derivation of this technique, please see Appendix B.

4.3 Experimental Setup

The experimental setup from chapters 2 and 3 was once again employed. Each event described in the next subsection was "played back" through the EUT using the system depicted in Fig. 2.2 100 times. This was done to simulate the energy detector's capability of detecting the same types of events over time - as sensors will have inherent variability that cannot always be accounted for.

4.3.1 Event Descriptions

Three events were used in this study, depicted in Fig. 4.3. Events depicted in Figs. 4.3a-4.3b are real-world recordings, and Fig. 4.3c shows a simulated event. The capacitor switching event of Fig. 4.3a shows a sharp high-frequency current spike at the point of inception and was sampled at 20 kHz. The event shown in Fig. 4.3b is from a small residential-scale microgrid housed at Oak Ridge National Laboratory switching online to the main grid, where the current was measured at the inverter and sampled at 30.72 kHz (512 samples/cycle). The last event, Fig. 4.3c shows a simulated fault on the terminals of a wind farm taken from the PSCAD example given in [34], sampled at 200 kHz.

4.3.2 Sensor Descriptions

Two sensors were independently used to serve as the EUT. The Lindsey 9670 35 kV class line post monitor (Figure 4.2) was first used, possessing $\leq 1\%$ accuracy in the current gain up to 6 kHz as well as induced phase error of less than 10°. The second sensor, a G&W CVS-36-O 36 kV class, possesses a more erratic magnitude response, reaching a maximum of 5% magnitude error and 10∘ phase error up to 6 kHz. The G&W sensor is rated for up to 30 kA (as opposed to approximately 1 kA for the Lindsey sensor) and thus has a much higher noise floor in the range of currents being used in this study.

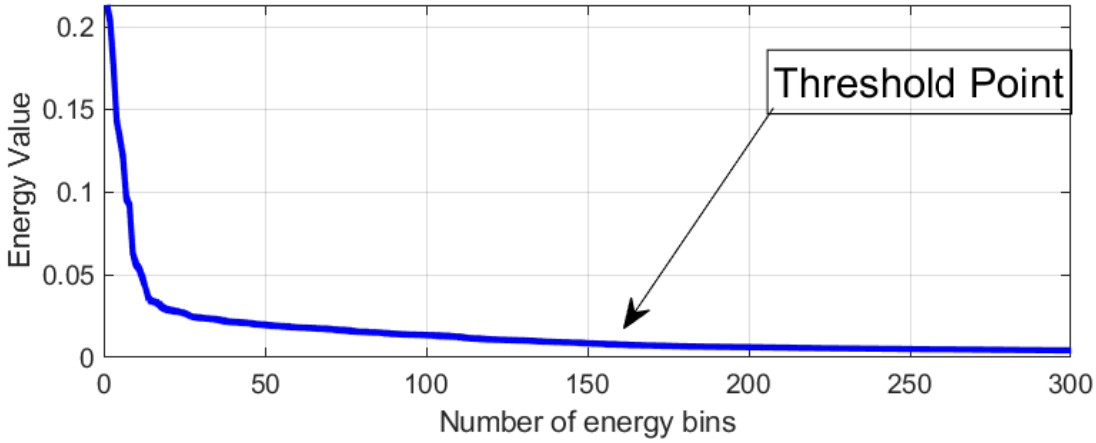
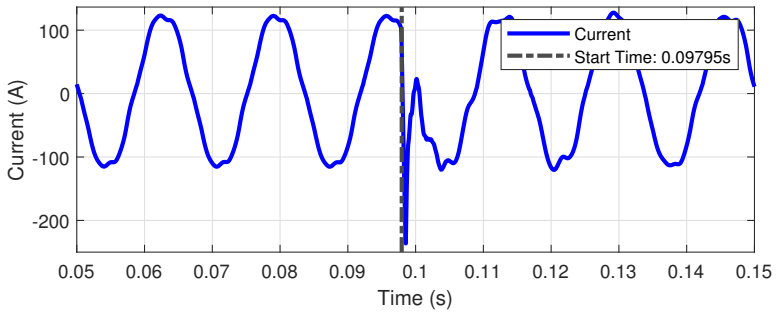
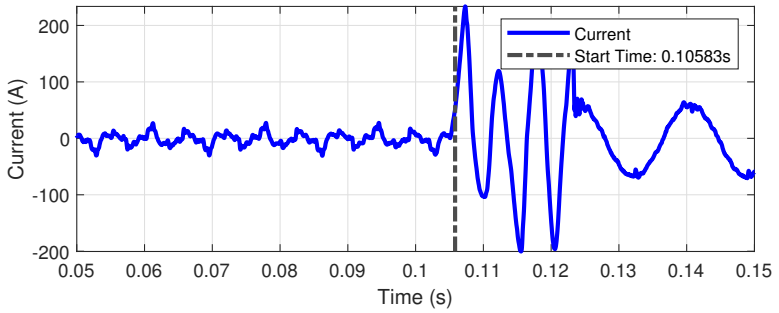


Figure 4.1: Elbow point

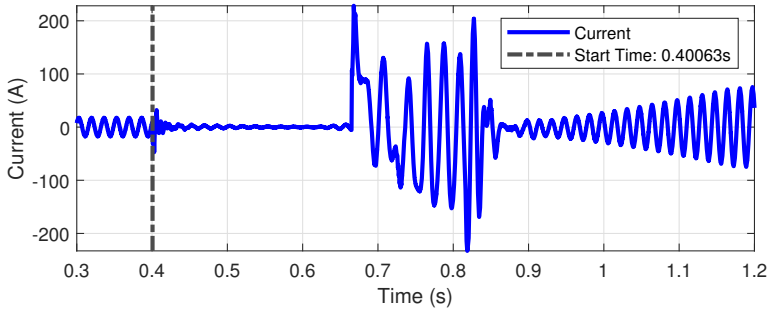
Figure 4.2: Laboratory setup with Lindsey sensor connected to driving CT



(a) Switching event current waveform - sampled at 20 kHz.



(b) Microgrid close-in current waveform - sampled at 30.72 kHz.



(c) Wind fault current waveform - sampled at 200 kHz.

Figure 4.3: Events under study

4.4 Results

Each event was recorded by the DAQ system 100 times after being passed through each sensor individually. This was done to create a statistical "profile" of the detection start times. Fig. 4.4 displays histograms of these detection times.

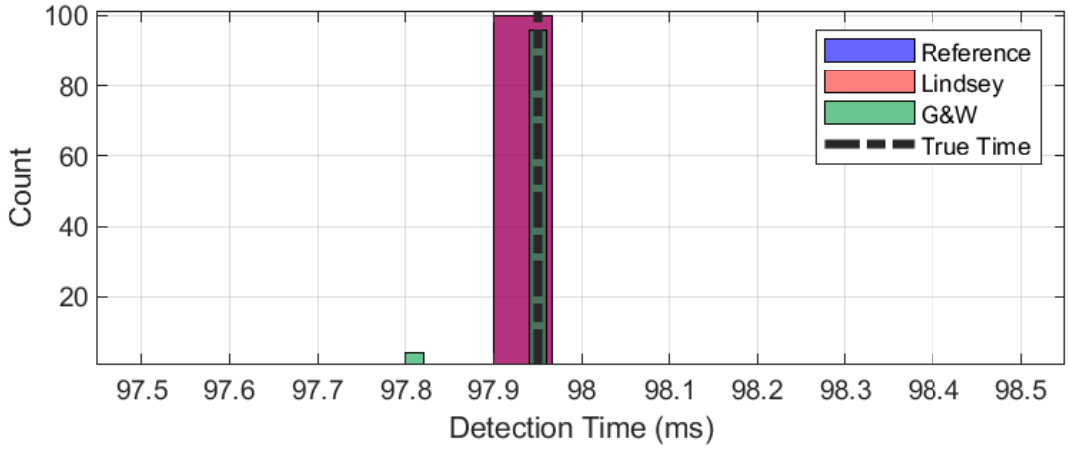
For the switching event depicted in Fig. 4.3a, the true start time is at 97.95 ms. As shown in Fig. 4.4a, both the reference and Lindsey sensors' waveforms are accurately detected by the ED at the precise start of the event over all 100 trials. Waveforms produced by the G&W sensor cause early detections four times.

The second event under study, the microgrid operation depicted in Fig. 4.4b, shows much more dispersion amongst its detection times. The reference and G&W sensors interestingly both exhibit more false positives than the Lindsey sensor. This is likely due to the amount of noise present in the first 0.1 seconds of the signal. There is very little fundamental-frequency current present in this time interval, which makes consistent energy estimation difficult.

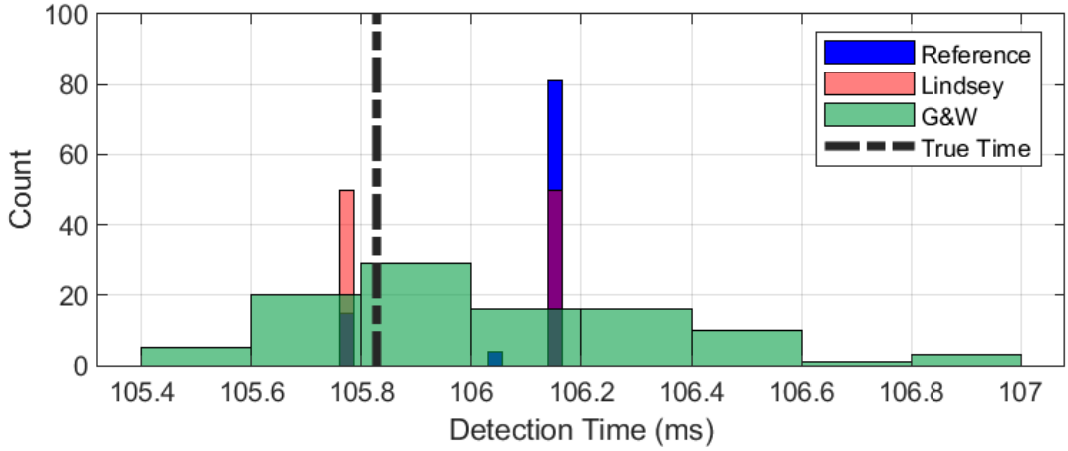
The last event shown in Fig. 4.3, resulted in a very wide dispersion of detection times for the G&W sensor (Fig. 4.4c), whereas the Lindsey sensor was able to produce waveforms easily detectable at the start time correctly for each trial. The reference sensor, on the other hand, shows some minor distribution amongst its detection times within a 2 ms window (zoomed-in portion of Fig. 4.4c). The G&W sensor, as described in Section 4.3.2, has a much higher allowable current range, yielding a higher noise floor capable of corrupting the signals produced by this device.

4.4.1 Comparison with other detection algorithms

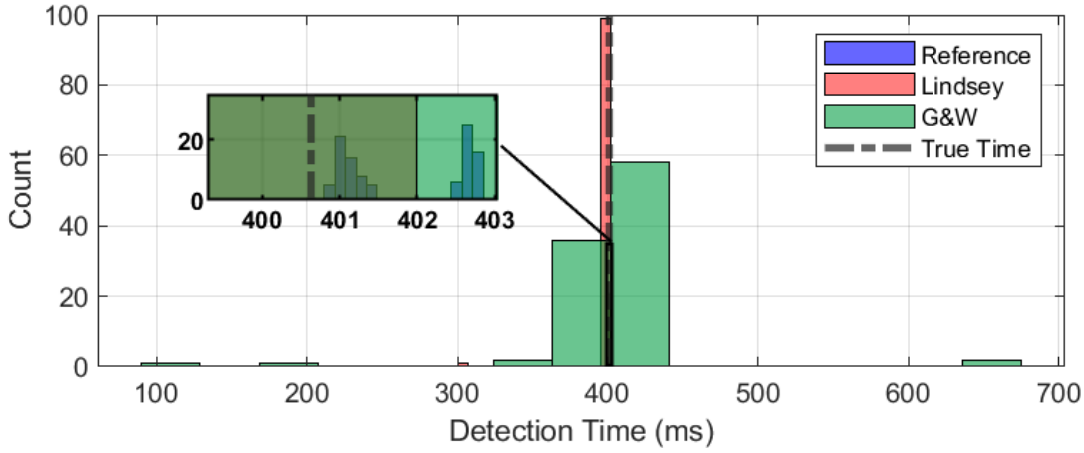
Here,the performance of the energy detector against two other detection algorithms proposed in the literature is examined. The two detection algorithms used for comparison are the discrete wavelet transform (DWT), as discussed in Section 4.1, and the ultra-fast transient (UFT) detector, [47]. The algorithms were compared using two metrics: mean percent error between the detected start time and the actual start time (black lines in Fig. 4.3) across all 100 trials, and average deviation between the detected start time and actual start time in terms of samples. It should be noted that as the sampling rate increases, the number



(a) Distributions of detected switching event start times for all three sensors.



(b) Distributions of detected microgrid close-in event start times for all three sensors.



(c) Distributions of detected wind fault event start times for all three sensors.

Figure 4.4: Detection Time Distributions

of samples to cover the same amount of time will increase. For example, 50 microseconds sampled at 20 kHz equates to a single sample, however it takes 10 samples to represent the same duration of time when sampled at 200 kHz.

For the switching and close-in events (Tables 4.1 and 4.2 respectively), the ED detector vastly out-performed both the UFT and DWT, both in terms of percent error and average sample deviation. This is likely due to the sensitivity to noise that the UFT and DWT algorithms possess. Given the switching event in particular, the mean percent error for the ED is shown to be essentially zero, for both the reference and Lindsey sensor signals. The G&W sensor percent error in this case was still very good at -0.01%. Note that both performance metrics can be positive or negative, indicating a late detection (negative for sample deviation, positive for percent error) or early (vice versa). In the case of the microgrid close-in event, the ED percent error never exceeds $\pm 2\%$. The maximum average sample deviation in this case was yielded from the reference sensor readings at 49.34 samples, or 1.6 ms (sampled at 30,720 Hz).

The wind fault event shows a case in which the ED under-performed with respect to the UFT. This illustrates a pitfall of the ED, namely the need to select a threshold γ (see (4.4)). This could potentially be overcome by adjusting N in (4.4) to be dependent upon the sampling rate. Because the sampling rate has increased drastically (200 kHz in this case), the number of energy bins calculated has grown considerably, making the choice of threshold challenging. It should still be noted however that the percent error never exceeds 1%. The maximum average sample deviation was -245.03 samples, corresponding to an average *late* detection by 1.2 ms at 200 kHz, which is still less than a tenth of a cycle.

4.5 Conclusions

This study presented the application of the *energy detector* for high-frequency power system event detection for three typical current disturbance types: capacitor switching, microgrid switching, and the reaction of a wind farm to a short-circuit fault at its terminals. It should be noted that this technique is waveform agnostic, meaning that it is not exclusive to current. It was shown that the detector exhibits strong performance in detecting these

Table 4.1: Switching Event Detection Comparisons

Switching			
	Ref.	**Lindsey**	**G&W**
ED Mean % Err.	**-1.417E-14**	**-1.417E-14**	**-0.01**
UFT Mean % Err.	0.1	0.1	0.1
DWT Mean % Err.	-0.641	-0.714	-0.714
ED Mean Sample Diff.	**0**	**0**	**0.12**
UFT Mean Sample Diff.	-2.05	-2.04	-2.04
DWT Mean Sample Diff.	12.56	14	14

Table 4.2: Microgrid Event Close-in Detection Comparisons

Microgrid Close-In			
	Ref.	**Lindsey**	**G&W**
ED Mean % Err.	**-1.52**	**0.12**	**0.19**
UFT Mean % Err.	-6.43	-6.41	-6.41
DWT Mean % Err.	16.082	16.03	16.026
ED Mean Sample Diff.	**49.34**	**-4**	**-6.34**
UFT Mean Sample Diff.	208.98	208.5	208.5
DWT Mean Sample Diff.	-522.8	-521	-521

Table 4.3: Wind Fault Event Detection Comparisons

Wind Fault			
	Ref.	**Lindsey**	**G&W**
ED Mean % Err.	0.31	-0.18	0.07
UFT Mean % Err.	**-2.37E-04**	**-1.25E-03**	**-1.25E-03**
DWT Mean % Err.	0.692	0.688	-0.56
ED Mean Sample Diff.	-245.03	140.65	-52.88
UFT Mean Sample Diff.	**0.19**	**1**	**1**
DWT Mean Sample Diff.	-554.32	-551	449

events through distortions caused by various commercial-grade sensors. One sensor in particular, the G&W sensor, possesses a higher noise floor, therefore negatively influencing the detector's performance. It was also shown that the ED mostly out-performs other common detection algorithms in the identification of the precise starting point of the events under study. It should be noted that start times for each of the 100 trials were not picked out individually. The start time references (Fig. 4.3) were taken from the original files, prior to being introduced to the sensors. It is possible that the sensors distort the actual start times for each of the 100 trials, therefore potentially distorting the error.

Chapter 5

Uncertainty Quantification for Switching Transient Signal Location

This chapter presents a machine learning application for capacitor switching transient location, as well as the effects seen on its predictions by sensor distortions. It is widely known that machine learning algorithms are extremely sensitive to bias and variance issues arising from differences in training data samples and testing data samples. In this chapter, a scenario is recreated in which a model is trained offline using simulated data from a distribution grid, and then validated using "actual" data, collected from the "field". This "actual" data is emulated by applying a digital filter transfer function representative of one of the sensors seen in the preceding chapters. It can be easily seen that results are negatively effected in a big way if these data irregularities are not considered when initially training the model.

5.1 The Transient Location Problem

The power system is one of the most complex systems in the world. The responsibility of those that maintain and operate these systems is to ensure that uninterruptible, stable power flow is provided to their customers. Evolving customer demands, along with increasing penetration of renewables and distributed energy resources (DERs), require that the power system be able to efficiently and *dynamically* control and monitor its state at all locations

and times. These problems present the need for capabilities that enhance situational awareness (SA), in many cases involving retrofitting new technology into existing systems and equipment.

Total power system SA is not an easy task. It involves the installation of measurement devices, communication systems, and processing equipment, as well as proper understanding of the produced data. Power system electromagnetic behavior (i.e. voltage and current) can be either easily-detectable (as in short-circuit current faults), or very subtle (as in high-impedance arcing or short-duration transients). As such, being able to differentiate between these subtle and unsubtle disturbances is extremely important.

The most famous example of failed SA technology is the 2003 Northeast Blackout. In August of 2003, much of the northeastern part of the United States, along with parts of Ontario and Quebec, Canada, were subjected to a blackout, affecting an estimated 50 million people. Due to a software error in a utility's alarm system, operators were not notified about a 345 kV transmission line that was tripped due to contact with foliage. Due to the lack of a proper alarm, the load was not properly re-distributed amongst the other lines and was handed off to the 138 kV lines, which were not equipped to handle all of the extra burden, causing a series of cascading failures, [50].

Similarly, in the Texas Interconnection, a failed surge protector on a combustion turbine connected to a step-up transformer led to a fault during start-up for testing in May of 2021. The associated circuit breaker tripped within three cycles, clearing the fault, and restoring the system back to normal operation after an estimated loss of 192 MW. In addition to these losses, a number of solar PV and wind plants saw reductions in active power production caused by the fault; however, it was determined that this loss of renewable generation was *not* caused by the fault itself. In the North American Electric Reliability Council's (NERC) September 2021 report on this disturbance (dubbed the "Odessa" disturbance), it was found that this simultaneous loss of renewable generation during the fault time was caused by "inverter-level or feeder-level tripping or control system behavior within the resources", [51]. The term "Smart Grid" originated from the idea of a fully autonomous power system. Alongside traditional equipment such as transformers, lines, generators, protection equipment, and more are power electronics, distributed generation, control

devices, and communications infrastructure. These "extra layers" provide much of the grid's current ability to "self-govern". However, the increased penetration of DERs and other power electronics-based equipment means special attention must be paid to their potential vulnerabilities, such as susceptibility to *transient over-voltage* (TOV).

TOV is a phenomenon in which short-duration voltage surges occur as a result of various disturbances, such as lightning strikes and switching behavior. If equipment is not rated to handle repeated exposure to these TOVs, the risk of damage to said equipment is significantly increased, therefore leading to a potential loss of crucial communication and control. As such, the need for adequate SA to handle decision-making processes based on these TOV signals is needed.

One of the more comment pieces of equipment in any power system is the capacitor. Capacitors are installed at various points in transmission and distribution grids for a variety of reasons, including: voltage support, power factor correction, and reduction of harmful harmonics. As such, a common source of TOV is *capacitor switching transients*.

Utility engineers will often employ control strategies at the capacitor bank location to ensure the switching operation at each phase occurs at the voltage zero-crossing, therefore nearly eliminating the amount of current inrush through the capacitor. Additionally, filters may be placed near the capacitor bank to mitigate the production of harmonics. Therefore, in many cases, the presence of TOV signatures in voltage waveforms may be indicative of equipment malfunction, insufficient design and planning, or other issues that need addressing.

For the reasons mentioned above, this type of TOV signal is chosen for study in this paper. The ability for monitoring equipment to not only identify that a TOV signal has occurred, but also *where* it occurred provides valuable insight into the state of the associated equipment. Engineers would then be able to pin-point exactly which capacitor bank(s) is (are) contributing to the disturbance, and therefore know how to implement or adjust control mechanisms to mitigate these issues. An example oscillatory capacitor switching TOV signal is shown in Figure 1.5.

There have been a number of studies done on the location of such TOV signals. In [52], signal energy is extracted to determine the direction of power quality (PQ) disturbances relative to the measurement location. Similarly, the cross time-frequency distribution with

Wigner kernel is used in [53] to determine the location of a switched capacitor transient relative to the measurement location.

The studies conducted in [54, 55, 56] use the slope and polarity of the detected TOV itself to determine which feeder of a distribution circuit originated the transient. A probabilistic neural network is used to classify features obtained from the S-transform of capacitor switching transients in [57]. The authors in [58] and [59] use disturbance indices to determine the location of capacitor switching transients. This index varies according to when there is a significant-enough change in energy typically characteristic of TOV signals.

5.2 Uncertainty quantification for TOV Signal Location using Machine Learning

This work focuses on using convolutional neural network (CNN) models to classify, and therefore locate, time-frequency representations of capacitor switching voltage transient signals. CNNs are powerful in feature extraction and have achieved state-of-the-art performance on a wide variety of classification tasks in smart grid domains. However, CNNs are data-driven models and their prediction can be affected by inherent noise in the data and uncertainty. Additionally, the data-driven CNN models usually produce reasonable predictions when the data in the unseen testing sets have similar features to those in the training set and can suffer from large extrapolation errors when the test data differs from the training data, leading to overconfident predictions. Overconfident incorrect predictions can be harmful, resulting in wrong decision-making and catastrophic failure in the smart grid system. Therefore, it is crucial to identify whether the CNN predictions are reliable and if the trained model is suitable for the unseen test data.

Uncertainty quantification (UQ) can help address the challenges of assessing the trustworthiness of ML model predictions and model reliability when applied to changing conditions and corrupted data. For training data, a well-calibrated UQ method can produce an uncertainty bound that precisely encloses a specified portion of the data consistent with the desired confidence level to consider the data uncertainty and assess the model prediction's

trustworthiness. For unseen test data, where the predicted values are not labelled, a high-quality UQ method can produce a larger prediction uncertainty as this data has a greater difference from the training set, indicating that the ML model is outside of the training support and its predictions should be treated with extra caution to avoid overconfidence.

However, UQ for ML model predictions is challenging and the development of a high-quality UQ method, which produces precise in-distribution uncertainty and identifies out-of-distribution (OOD) samples, is even more challenging. Generally speaking, there are two types of UQ-for-ML methods: prediction interval (PI) approaches, which quantify uncertainty using intervals, and non-PI approaches, which quantify uncertainty using a distribution. The non-PI approaches can be further divided into Bayesian and non-Bayesian methods. Bayesian methods place priors on NN weights and then infer predictive posterior distribution from the weight distribution [60, 61]. Its' results are sensitive to the choice of prior distributions, and it has been criticized for slow training, overconfident predictions, and being impractical for large-scale ML applications. Non-Bayesian methods include evidential learning [62] that places priors directly over the likelihood function and some ensemble methods that do not use priors, such as deep ensembles, Monte Carlo dropout, and anchored ensembling [63, 64, 65]. Recently, some methods used deterministic deep learning for uncertainty estimation with some special NN architecture designs, such as the spectral-normalized neural Gaussian process. These non-Bayesian methods usually involve a Gaussian assumption which might not be satisfied in power system applications where data noise and irregularities can manifest in the higher harmonics as skewed and non-Gaussian. They could also suffer from an overestimation of the uncertainty in training data caused by the symmetric uncertainty bound from the Gaussian assumption and result in an underestimation of the uncertainty in extrapolation [66].

The PI methods provide a lower and upper bound for a prediction such that the target falls between the bounds with a certain percentage. PIs directly communicate uncertainty, providing more understandable information for decision-making. Additionally, they usually do not involve distributional assumptions, making them particularly useful for scientific ML. The most common techniques to construct PIs are the delta method (also known as analytical method) [67, 68], methods that directly predict the variance (maximum likelihood method

and ensemble method) [69, 70] and quantile regression method [71, 72]. Most recent PI methods are developed on the high-quality principle—a PI should be as narrow as possible, whilst capturing a specified portion of data. Khosravi et al. [73] developed the Lower Upper Bound Estimation method, incorporating the high-quality principle directly into the NN loss function for the first time. Inspired by [73], the QD approach in [74] defined a loss function that can both generate a high-quality PI and optimize the loss using stochastic gradient descent. Built on QD, the PIVEN method in [75] adds an extra term in the loss to enable the calculation of point estimates and the PI method in [76] further integrates a penalty function to the loss to improve the training stability of QD.

Recently developed PI methods [74, 75, 76] tend to design sophisticated loss functions to obtain a well-calibrated PI. Although these works have achieved promising results, their performance is sensitive to unusual hyperparameters introduced into their customized loss functions. Since these hyperparameters are not commonly used, we have very little knowledge and experience about how to properly choose them. In practice, these hyperparameters usually need fine tuning [76] to achieve the desired performance, which makes these methods less practical and less robust when deployed. Additionally, existing PI methods usually lack a sufficient OOD identification capability, resulting in unreasonably narrow PIs for OOD samples.

In this chapter, a new PI method is employed and integrated it with a CNN model for determining the source location of capacitor switching TOV signals with a UQ component. The method is called PI3NN, which calculates prediction intervals based on three independent neural networks [77]. The first NN calculates the mean prediction, and the following two NNs produce the upper and lower bounds of the interval. After the three NNs' training, and given a certain confidence level, PI3NN uses a root-finding algorithm to precisely determine the uncertainty bound that covers the desired portion of the data consistent with the confidence level. Additionally, PI3NN proposes a simple but effective initialization scheme for OOD identification. PI3NN is computationally efficient given the training of three networks; and for a different given confidence level, it just needs to perform the root finding step to calculate the shifting coefficients to precisely determine the corresponding interval. Additionally, PI3NN uses the standard mean squared loss and

does not introduce extra hyperparameters, which enables robust prediction performance and mitigates tedious parameter turning. Furthermore, PI3NN has an OOD identification capability which can produce a wider uncertainty for the predictions outside of the training data. Last but not least, PI3NN is generalizable to various network structures and applicable to different data with no distributional assumptions, which makes it suitable for a wide range of ML-based power system applications.

In previous works [77], PI3NN has been integrated into fully-connected, multilayer perceptron (MLP) networks and its superior performance against several baselines using a range of diverse datasets has been demonstrated. In this effort, this newly-developed method is integrated with CNNs for capacitor switching TOV source location. The CNN model used in this study has a substantially different architecture from the MLP networks used in past works. In this implementation, the convolutional and fully-connected layers of the CNN are first separated as two distinct sets of networks. For the first convolutional network, the the feature information is extracted from its outputs and used as the inputs for the fully-connected network. Then, the PI3NN technique is performed on this fully-connected network, treated as an MLP problem. This design improves training reliability, reduces the computational costs, and most importantly, reduces the requirement of large training data.

In this study, two-dimensional time-frequency representations of TOV signals are collected from post-processed sensor data. CNN networks first use convolutional layers to extract the time-frequency features and then use fully-connected dense layers to map these features to class labels for classification. The sensors used in this study are commercial-grade, and as such typically produce highly-accurate signal representations at the fundamental frequency, but are susceptible to corrupting voltage and current signals with high frequency components (e.g. Figure 1.5). Additionally, these sensors could suffer from component degradation which can further compound these frequency-response irregularities. This suggests that the TOV signals contain noise and as such the distribution of the data may change with time. This data, along with OOD uncertainties, greatly affect CNN model predictions and their influence should be considered and evaluated when deploying CNNs for classification. In this work, the PI3NN technique is used for CNN model prediction

uncertainty quantification. The important UQ information not only enables trustworthy predictions, but also allows utility engineers to know how ML model prediction accuracy may degrade and allow stakeholders to abstain from decisions due to low confidence.

The main contributions of this chapter are as follows.

- Integration of a novel UQ method developed by the authors in [77] into CNN models to facilitate SA in power distribution systems for accurate, credible predictions, thereby avoiding catastrophic failures caused by overconfident predictions.
- Design of a CNN model identifying up to 96% correct transient switching locations, spanning five utility distribution feeders.
- Reasonable quantification of the CNN model's prediction uncertainty by both indicating when the model results can be trusted and if the system suffers from degradation.

The rest of this chapter is structured as follows. Section 5.3 introduces the UQ method used for ML-based robust TOV signal location. In section 5.4 the simulation environment, signal pre-processing stages, CNN model design, and integration of the UQ method into the CNN model are all discussed. Results are presented in section 5.5, and section 5.6 concludes the chapter.

5.3 PI3NN: Prediction Intervals from Three Neural Networks

PI3NN uses a combination of the three separately trained NNs to learn the mean prediction (i.e. point estimation), and the lower and upper bounds of the PI. PI3NN not only has the nice properties that state-of-the-art PI methods have—such as requiring no distributional assumption and producing tight PI bounds—but also embraces some exclusive advantages. For instance, it introduces no extra hyperparameters, enabling robust training; and it is also able to capture domain shift and reasonably quantify uncertainty on OOD samples.

PI3NN is generalizable to a variety of network structures. In this work, this newly-developed method is integrated with CNNs for capacitor switching TOV source location. In

the following, the general procedure of PI3NN for a fully-connected MLP network is described and then its capability of OOD identification discussed. In Section 5.4.4, the integration of PI3NN is introduced into CNN for robust and credible classification.

Procedure of PI3NN

The key idea of the PI3NN method is to construct the PI by training three neural networks separately using the standard mean-squared error (MSE) loss and using root-finding methods to define the upper bound $U(\boldsymbol{x})$ and lower bound $L(\boldsymbol{x})$. Let the mean, upper-, and lower-bound neural networks be denoted by $f_{\boldsymbol{\omega}}(\boldsymbol{x})$, $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$, respectively. Here $\boldsymbol{x}$ represents an input to the neural network, and the functions $f_{\boldsymbol{\omega}}(\boldsymbol{x})$, $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ represent the functions, approximated by each neural network, that act on $\boldsymbol{x}$ to produce the mean and upper and lower bounds. The PI3NN method constructs the PI in the following three steps.

Step 1: Train $f_{\boldsymbol{\omega}}(\boldsymbol{x})$ to approximate the mean of $f(\boldsymbol{x})$, $\mathbb{E}[f]$. This completely follows the standard NN-based regression process using the MSE loss. The trained $f_{\boldsymbol{\omega}}(\boldsymbol{x})$ serves two purposes. The first is to provide a baseline to generate data for training $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ in **Step 2**; the second is to provide a point estimation of $\mathbb{E}[f]$. In this step, the well-established regularization techniques are used (e.g. the conventional L_1 and L_2 penalties) to avoid over-fitting.

Step 2: Train $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ to learn the uncertainty profile. The trained $f_{\boldsymbol{\omega}}(\boldsymbol{x})$ is used as a baseline to generate two separate data sets, denoted by $\mathcal{D}_{\text{upper}}$ and $\mathcal{D}_{\text{lower}}$, respectively. Mathematically,

$$\begin{aligned} \mathcal{D}_{\text{upper}} &= \big\{(\boldsymbol{x}_i, y_i - f_{\boldsymbol{\omega}}(\boldsymbol{x}_i)) \;\big|\; y_i \geq f_{\boldsymbol{\omega}}(\boldsymbol{x}_i), i = 1, \ldots, N\big\}, \\ \mathcal{D}_{\text{lower}} &= \big\{(\boldsymbol{x}_i, f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) - y_i) \;\big|\; y_i < f_{\boldsymbol{\omega}}(\boldsymbol{x}_i), i = 1, \ldots, N\big\}, \end{aligned} \tag{5.1}$$

where $\mathcal{D}_{\text{upper}}$ and $\mathcal{D}_{\text{lower}}$ include data points above and below $f_{\boldsymbol{\omega}}(\boldsymbol{x})$, respectively. The number of data points in $\mathcal{D}_{\text{upper}}$ and $\mathcal{D}_{\text{lower}}$ should be comparable when the MSE loss for training $f_{\boldsymbol{\omega}}(\boldsymbol{x})$ achieves a sufficiently small value. Next, $\mathcal{D}_{\text{upper}}$ is used to train $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, and $\mathcal{D}_{\text{lower}}$ is used to train $l_{\boldsymbol{\xi}}(\boldsymbol{x})$. To ensure that the outputs of $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ are positive, the operation

$\sqrt{(\cdot)^2}$ is added to the output layer of both networks. The two NNs are trained *separately* using the standard MSE loss, i.e.,

$$\begin{aligned}\boldsymbol{\theta} &= \text{argmin}_{\boldsymbol{\theta}} \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_{\text{upper}}} (y_i - f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) - u_{\boldsymbol{\theta}}(\boldsymbol{x}_i))^2, \\ \boldsymbol{\xi} &= \text{argmin}_{\boldsymbol{\xi}} \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_{\text{lower}}} (f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) - y_i - l_{\boldsymbol{\xi}}(\boldsymbol{x}_i))^2.\end{aligned} \tag{5.2}$$

Unlike the sophisticated losses in [63, 62, 76], the three NNs of PI3NN are trained using the standard MSE without introducing unusual hyperparameters, which promises more robust training.

Step 3: construct the PI via root-finding methods. Note that $u_{\boldsymbol{\theta}}(\boldsymbol{x})$, $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ do not directly represent the upper and lower bounds of the PI. Instead, they only approximate the difference between the data and $f_{\boldsymbol{\omega}}$ described by the datasets $\mathcal{D}_{\text{upper}}$ and $\mathcal{D}_{\text{lower}}$. In this work, the upper and lower bounds of the PI are defined as:

$$\begin{aligned}U(\boldsymbol{x}) &= f_{\boldsymbol{\omega}}(\boldsymbol{x}) + \alpha u_{\boldsymbol{\theta}}(\boldsymbol{x}), \\ L(\boldsymbol{x}) &= f_{\boldsymbol{\omega}}(\boldsymbol{x}) - \beta l_{\boldsymbol{\xi}}(\boldsymbol{x}),\end{aligned} \tag{5.3}$$

where α and β are two unknown scalars. For a given quantile $\gamma \in [0, 1]$, the bisection method is used to determine the value of α and β by finding the roots of the following equations:

$$\begin{aligned}Q_{\text{upper}}(\alpha) &= 0, \\ Q_{\text{lower}}(\beta) &= 0\end{aligned} \tag{5.4}$$

where

$$\begin{aligned}Q_{\text{upper}}(\alpha) &= \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_{\text{upper}}} \mathbf{1}_{y_i > U(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) - \frac{N(1-\gamma)}{2} \\ Q_{\text{lower}}(\beta) &= \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_{\text{lower}}} \mathbf{1}_{y_i < L(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) - \frac{N(1-\gamma)}{2},\end{aligned} \tag{5.5}$$

where N is the number of samples in $\mathcal{D}_{\text{train}}$ and $\mathbf{1}(\cdot)$ is the indicator function, defined by

$$\mathbf{1}_{y_i>U(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) = \begin{cases} 1, & \text{if } y_i > U(\boldsymbol{x}_i), \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\mathbf{1}_{y_i<L(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) = \begin{cases} 1, & \text{if } y_i < L(\boldsymbol{x}_i), \\ 0, & \text{otherwise,} \end{cases}$$

which count how many training samples are outside the interval $[L(\boldsymbol{x}), U(\boldsymbol{x})]$.

When the root-finding problems in (5.5) are exactly solved (i.e., $Q_{\text{upper}}(\alpha) = Q_{\text{lower}}(\beta) = 0$), the number of training samples falling in $[L(\boldsymbol{x}), U(\boldsymbol{x})] = [f_{\boldsymbol{\omega}} - \beta l_{\boldsymbol{\xi}}, f_{\boldsymbol{\omega}} + \alpha u_{\boldsymbol{\theta}}]$ will be exactly $N\gamma$. In this way, the prediction interval method can produce an accurate uncertainty bound that precisely encloses a specified portion of data with a narrow interval width. Moreover, this prediction interval calculation *does not impose any distributional assumptions to enable a general application.*

Identifying out-of-distribution (OOD) samples

When using the trained model $f_{\boldsymbol{\omega}}(\boldsymbol{x})$ to make predictions for $\boldsymbol{x} \notin \mathcal{D}_{\text{train}}$, it is required that the UQ method can accurately identify the OOD samples and reasonably quantify their uncertainty, i.e., for $\boldsymbol{x} \notin \mathcal{D}_{\text{train}}$, the PI's width increases with the distance between $\boldsymbol{x}$ and $\mathcal{D}_{\text{train}}$. PI3NN achieves OOD identification by properly initializing the output layer biases of $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$. Specifically, the following operations are added into the **Step 2** before training $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$.

- Define $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$ as fully-connected ReLU networks, and initialize their weights and biases using the default option.
- Compute the mean outputs $\mu_{\text{upper}} = \sum_{i=1}^{N} u_{\boldsymbol{\theta}}(\boldsymbol{x}_i)/N$ and $\mu_{\text{lower}} = \sum_{i=1}^{N} l_{\boldsymbol{\xi}}(\boldsymbol{x}_i)/N$ using the training set $\mathcal{D}_{\text{train}}$ and initial weights and biases.
- *Modify the biases of the output layers of $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$ to $c\,\mu_{\text{upper}}$ and $c\,\mu_{\text{lower}}$, where c is a relatively big number (e.g., $c = 15$ in this study).*

- Follow the rest of **Step 2** to train $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\theta}}$ using the MSE loss.

Through the above initialization strategy, the outputs of networks $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ will be larger for the OOD samples than the in-distribution data. Then after calculating the positive values of α and β in **Step 3**, it will correspondingly produce the larger uncertainty bounds $[L(\boldsymbol{x}), U(\boldsymbol{x})]$ for the OOD samples to indicate that their predictions are of low confidence. We use the PI width to measure the size of the uncertainty bound, which is defined as,

$$\text{PI width} = U(\boldsymbol{x}) - L(\boldsymbol{x}) = \alpha u_{\boldsymbol{\theta}}(\boldsymbol{x}) + \beta l_{\boldsymbol{\xi}}(\boldsymbol{x}). \tag{5.6}$$

The key ingredient in this OOD identification strategy is the modification of the biases of the network output layer. It is known that a MLP dense network is formulated as a piece-wise linear function. The weights and biases of hidden layers define how the input space is partitioned into a set of linear regions; the weights of the output layer determine how those linear regions are combined; and the biases of the output layer act as a shifting parameter. These network weights and biases are usually initialized with some standard distributions, e.g., uniform $\mathcal{U}[0, 1]$ or Gaussian $\mathcal{N}[0, 1]$, as default options. Setting the output layer biases to $c\mu_{\text{upper}}$ and $c\mu_{\text{lower}}$ with a large value of c will significantly lift up the initial outputs of $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$. During the training, the loss in (5.2) will encourage the decrease of $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ only for in-distribution data (i.e., $\boldsymbol{x}_i \in \mathcal{D}_{\text{train}}$), not for OOD samples. Therefore, after training, $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ will be larger in the OOD region than in the in-distribution region (see Figure 1 in [77] for an illustration). Correspondingly, the PI width of the OOD samples will be larger compared to that of the training data, based on which we identify the data/domain shift. Note that the exact value of c does not matter much, as long as it is a large positive value, e.g., $c = 15$ in this study. For training data, PI3NN will produce prediction intervals precisely enclosing $\gamma \times 100\%$ portion of data for a given confidence level $\gamma \in [0, 1]$ no matter how large the c value is, although a larger c in the network initialization may take a slightly longer training time for convergence.

For the unseen testing data, if they are in-distribution with similar input features as the training set, PI3NN will produce uncertainty bounds with a similar width as the training data despite the large c value. If the test data are OOD (i.e. outside of the training set

support), PI3NN will produce a larger PI width than that of the training data. The larger the c value is, the wider the PI width. Then, by comparing the PI widths of the test data with those of the training data, we diagnose whether the unseen test data are in-distribution or OOD to quantify the trustworthiness of the ML model predictions. For OOD samples, it is not expected to accurately predict them, due to data-driven ML model deficiency, but more importantly it is to identify them to avoid overconfident predictions and provide a guidance for data collection to improve the predictability.

5.4 Experimental Setup

5.4.1 Distribution grid model

A series of four 12.47 kV distribution feeders, modeled after actual feeders belonging to a partner utility company, were constructed in PSCAD transient simulation software. An anonymized version of the grid topology is shown in Figure 5.1. Note that the topology contains hundreds of electrical nodes, but the specifics of this information has been excluded from the paper. Four parallel-connected, wye-grounded capacitor banks were placed at each of the ten chosen nodes, encircled in the figure. The capacitance values were chosen such that applying all four banks at once improves the power factor at that node to as close to 1.0 as possible. The values of each capacitor bank are given in Table 5.1.

A base-case power flow simulation, without any shunt capacitors switched on, was used to obtain the nominal real (P) and reactive (Q) power flows at each bus. The capacitance values were then computed from the following well-known relationship governing reactive power supplied by a capacitor:

$$C = \frac{Q}{2\pi f V^2} \tag{5.7}$$

where C is the capacitance, V is the RMS voltage, and f is the system frequency (60 Hz).

In addition to the capacitor banks, measurement points (i.e. "sensors") were placed at each node to collect voltage measurements during the simulation, which are then exported to files for off-line use. Each sensor captures voltage waveforms over a few different operating variables:

- Load profiles for each feeder over a 24-hour period were obtained from the utility. Simulations were carried out at the minimum, maximum, and mean load values applied equally across each load tap of the associated feeder.

- At each bus, the transient "switch-on" period was varied at 20 points along one cycle of the voltage waveform, yielding best-case, worst-case, and in-between scenarios for each bus.

- Simulations were performed on each of the preceding two operating conditions for one, two, three, and four capacitor banks switched on simultaneously.

Three load profiles, 20 switch-on times, and four capacitor bank combinations result in 240 unique operating conditions for each bus. For each operating condition, simulated one bus at a time, the sensors at *each node* record a waveform for the duration of the simulation. This then leads to a total of 2,400 sets of three-phase voltages for each bus's disturbance contributions. Multiplying this number by 10, to account for *all* buses' operating scenarios, yields to a data set of size $24,000 \times N \times 3$, where N is the length of the captured disturbances, and the third dimension is to account for all three voltage phases.

5.4.2 Data Structure and Preprocessing

The simulations were run with a time step of $\Delta t = 20$ μs, or 50 kHz. To reduce the amount of data needed to train the neural network, the three phases were combined using the α-modal voltage obtained from the $\alpha\beta\gamma$ transformation for three-phase signals, [78]:

$$V_\alpha = \frac{2}{3} V_A - \frac{1}{3} \left(V_B - V_C \right), \tag{5.8}$$

where V_A, V_B, and V_C represent the phase A, B, and C voltages, respectively. This signal transformation runs the slight risk of information loss due to destructive interference, however it is believed that this loss is minimal and doesn't result in a reduction in performance of the prediction model.

After collecting all $24,000$ samples and performing the $\alpha\beta\gamma$ transformation, the data was curated to remove samples that did not possess enough identifying harmonic content. This

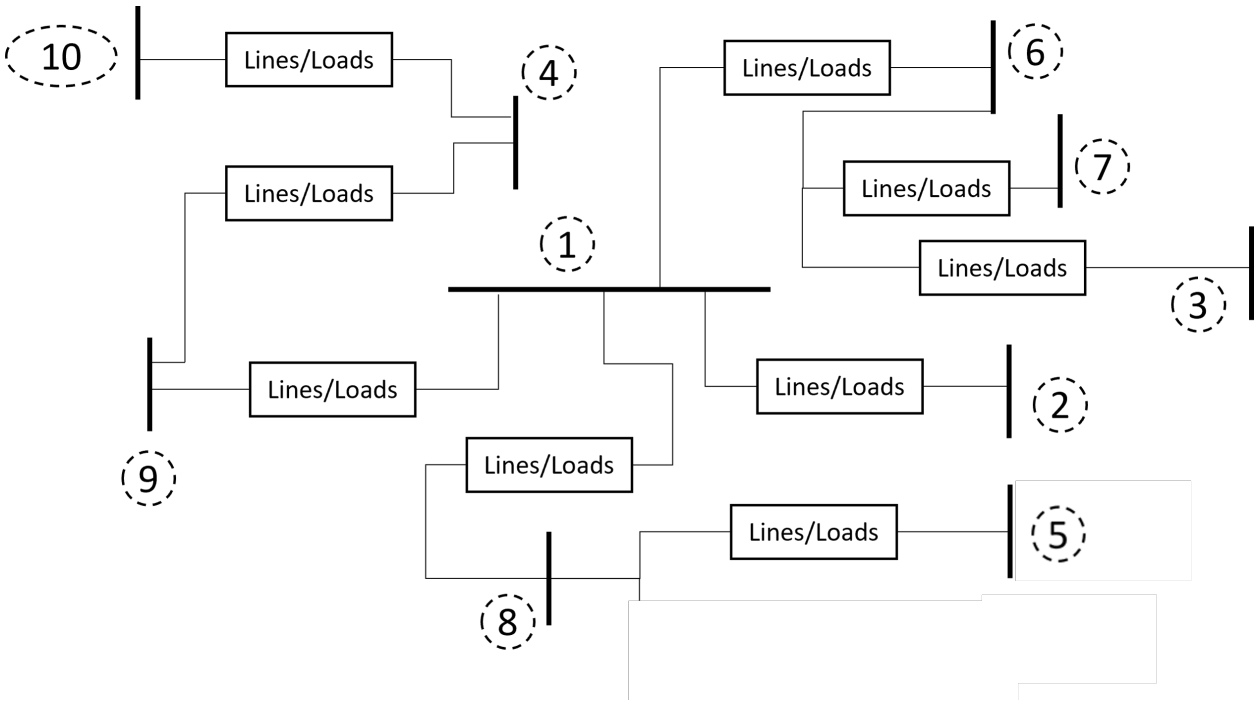
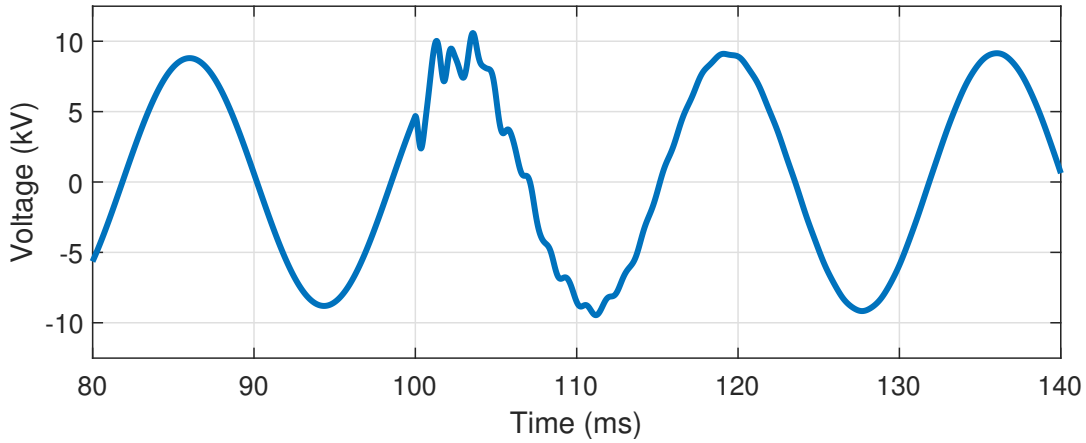


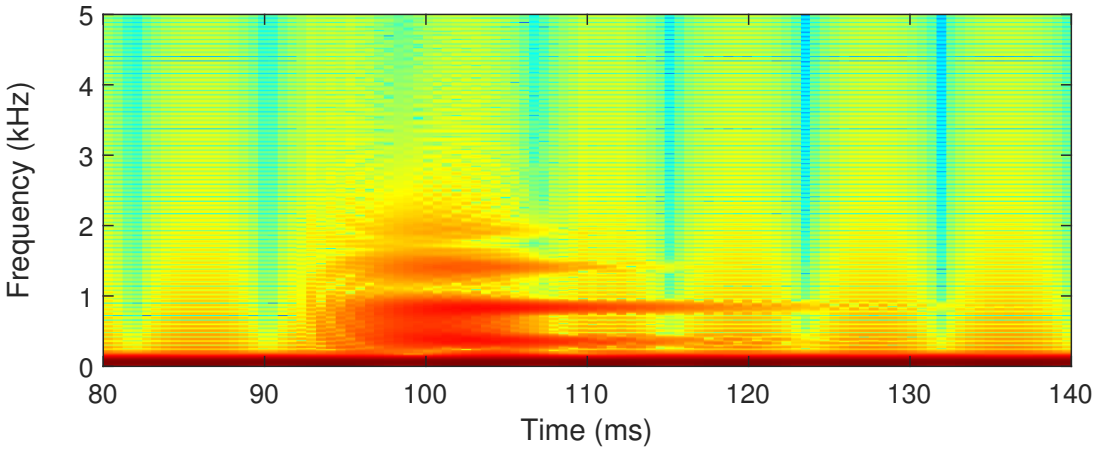
Figure 5.1: Anonymized Distribution Grid

Table 5.1: Derived capacitance values

Bus	$\boldsymbol{Q}$ needed (MVAr)	$\boldsymbol{C}$ (per capacitor, $\boldsymbol{\mu}$F)
1	14.27	122
2	1	8.5
3	2.159	17
4	0.3	2.55
5	0.3	2.55
6	3	25.5
7	0.03	0.256
8	3	25.5
9	7.4	63.1
10	0.03	0.256



(a) Example V_α



(b) Spectrogram transformation of V_α

Figure 5.2: Example V_α and its spectrogram representation

was accomplished by first applying min-max normalization to each sample after computing its gain in decibels:

$$\begin{aligned} \boldsymbol{X}_{dB}^{i} &= 20 \log_{10} \left| X^{i} \right| \\ \boldsymbol{X}_{n}^{i} &= \frac{\boldsymbol{X}_{dB}^{i} - \min \boldsymbol{X}_{dB}^{i}}{\max \boldsymbol{X}_{dB}^{i} - \min \boldsymbol{X}_{dB}^{i}} \end{aligned} \tag{5.9}$$

where X^i denotes the i^{th} α-transformed voltage signal. If $\boldsymbol{X}_n^i$ does not possess any frequency content above the 7^{th} harmonic (420 Hz) exceeding 35% of the fundamental frequency's amplitude, that signal is discarded from the signal set. The threshold of 35% was determined empirically after examining different thresholds and their associated waveform plots. This is due high-frequency transients being filtered out before they reach remote sensors. For example, per Figure 5.1, if a transient with high-frequency components originated at bus 5, it may simply appear as a normal sine wave at bus 10. These signals were pruned from the signal set to remove as much ambiguity as possible. In total, only $1,346$ samples were discarded, or 5.6%. The remaining data set consisted of $22,654$ samples.

After transforming the voltage signals into their α-modal components and curating the signal set, the signals are then transformed into *spectrograms*. Spectrograms are two-dimensional representations of one-dimensional signals, capturing both time and frequency information simultaneously. The short-time Fourier transform (STFT) with a Gaussian window was used to compute the signal spectrograms. The resulting coefficients are complex-valued due to the FFT computation. Thus, these coefficients are then squared to obtain real values and accentuate larger values, while also attenuating smaller ones. Figure 5.2b shows an example spectrogram plot of the TOV signal shown in Figure 5.2a. The blue regions indicate regions where the Gaussian window used in the spectrogram "zeroes out" the extremeties of the signal at each shift of the window (approximately 9-10 ms in the figure). Note how the energy in frequencies up to 3 kHz increases in the spectrogram at the same time that transient behavior is observed in the time-domain signal (roughly 100 ms in Figure 5.2a).

The final pre-processing stage involves duplicating the data set twice, each with varying levels of additive white Gaussian noise (AWGN) superimposed on the data. Each duplicated set was incorporated with noise to ensure signal-to-noise (SNR) ratios of 35 and 20 dB.

5.4.3 Classification model architecture

A CNN was chosen for this work due to CNNs' strong capability for feature extraction. The CNN architecture consists of three convolutional layers, each with a 2×2 max pooling layer, four fully-connected layers, an output layer with 10 units (representing the chosen "bus"), and dropout layers placed before the second-to-last fully-connected layer and immediately preceding the output layer, with dropout probabilities of 0.25 and 0.3, respectively. Every layer uses the rectified linear unit (ReLU) activation function. The loss function used during training was the Kullback-Leibler (KL) Divergence.

Filter sizes of 5 were chosen to ease some of the computational burden on the spectrogram images, and due to the assumption that features on the spectrograms will likely be spread over larger regions, thus reducing the need for smaller filter sizes to capture minute details. The network structure (number of layers, sizes, etc.) was chosen based on trial-and-error, in which the authors studied a number of different architectures. Dropout layers were incorporated to ensure the network was not over-fitting.

Training was performed on an NVIDIA DGX A100 system running Linux Ubuntu SMP with 8 GPU cores. The Adam optimizer with learning rate set to 0.001 was used. During training, the spectrograms are min-max normalized, to ensure all values lie in [0, 1]. A batch size of 256 was chosen during training, and training lasted for 1000 epochs. A 75/25% train/test split was used.

5.4.4 Setup for PI3NN

The CNN model can be viewed as a combination of two components: the feature extraction portion (ending at the flattening layer) and the classification portion (starting from the first fully connected layer). The convolution and pooling layers perform feature extraction from the data, and the fully connected layers then act as a classifier on top of these features. To develop a PI3NN framework for this classification problem, a transfer learning approach is adopted, which re-uses the CNN model for the feature extraction part and replaces the fully-connected classification part with PI3NN to provide the prediction with quantified uncertainty. Thus, the inputs of the PI3NN are feature vectors of size $23,040$, obtained

by funneling the images through the feature extraction portion of the pre-trained CNN model. The outputs are 10-D vector created by one-hot-encoding of original labels. It is worth noting that because the PI3NN framework does not require retraining of the feature extraction, which carries most of the cost for classification problems, it could be trained to calculate the PI and provide the uncertainty estimation for a given confidence level quickly, within minutes for the test case in this paper. Once trained, the PI3NN model can be deployed to produce predictions on new input queries in real-time.

Three neural networks are trained in this framework: $f_{\boldsymbol{\omega}}$ to predict the output vectors, and $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$ to approximate upper and lower variance of those predictions. These networks use the same architecture as the classification portion of the CNN models; i.e., five fully-connected layers with the ReLU activation function. Moreover, the biases of the output layers of $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$ are initialized to $c\,\mu_{\text{upper}}$ and $c\,\mu_{\text{lower}}$, where $c = 15$ to encourage OOD identification. The MSE loss function was used as described in Section 5.3. The loss function was applied directly to the predictions of the classification models, without any correcting/calibrating step. However, in case of imbalanced datasets, calibrated outputs with, e.g., Platt scaling [**?** **?**] or Isotonic Regression [**?**], can be beneficial to our framework. The Adam optimizer with learning rate 0.001 is used to train $f_{\boldsymbol{\omega}}$, allowing for a maximum of $8,000$ training epochs. For $u_{\boldsymbol{\theta}}$ and $l_{\boldsymbol{\xi}}$, stochastic gradient descent with learning rate 0.01 was used, allowing for a maximum of $4,000$ training iterations. The neural network architectures for the base CNN and the PI3NN framework built on pre-trained CNN model are shown in Figure 5.3.

5.4.5 Sensor Frequency Response Distortions

To capture the effects of frequency distortions caused by common medium-voltage sensors, the amplitude and phase responses were obtained in a laboratory from one of said sensors. This particular sensor has a -3 dB bandwidth of roughly 2600 Hz and a maximum phase distortion of 30° at 10 kHz.

The curated dataset of three-phase voltage waveforms, described in the preceding sections, is passed through an infinite impulse response (IIR) digital representation of this

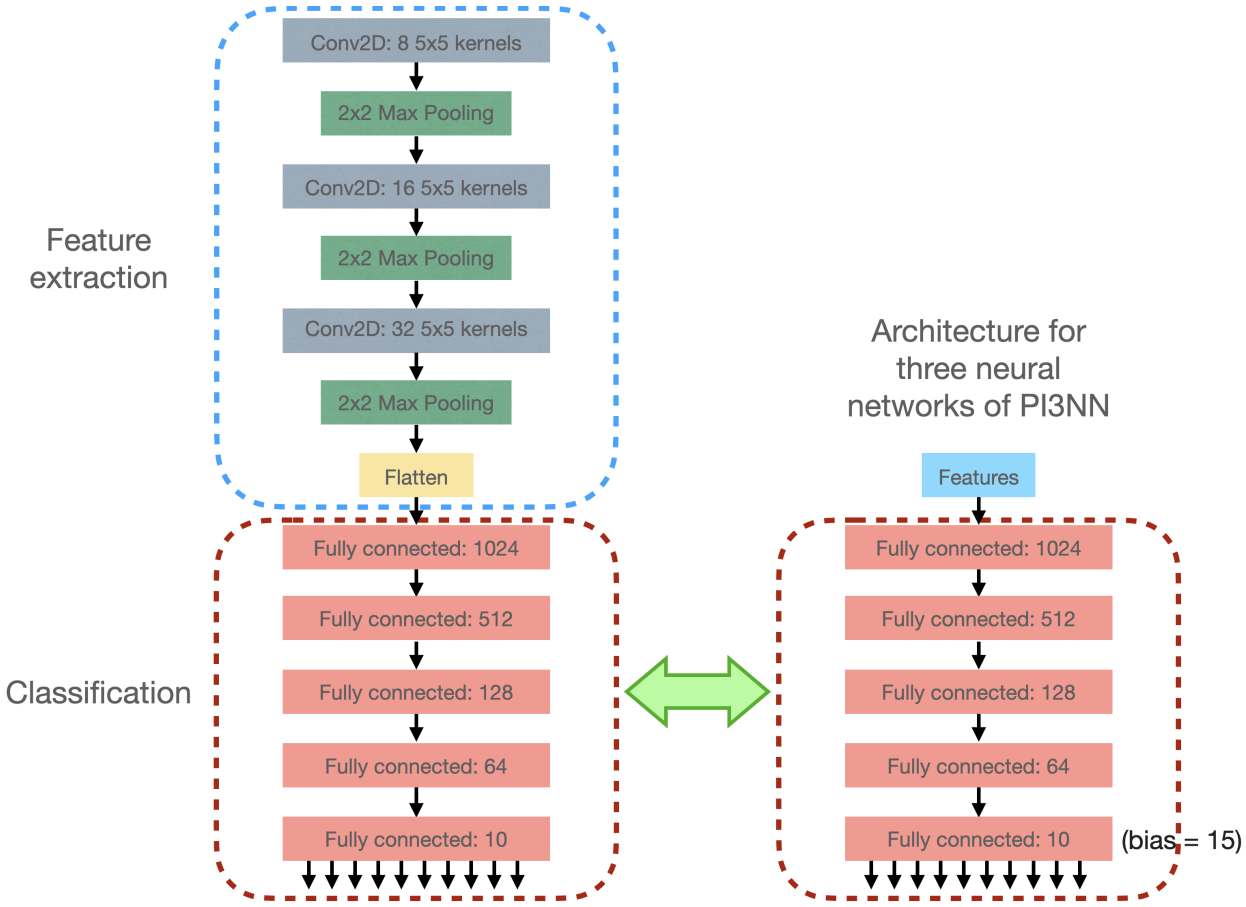


Figure 5.3: (Left) The neural network architectures for the base CNN. (Right) PI3NN framework is built on the pre-trained CNN model and accepts feature vectors as the inputs (right).

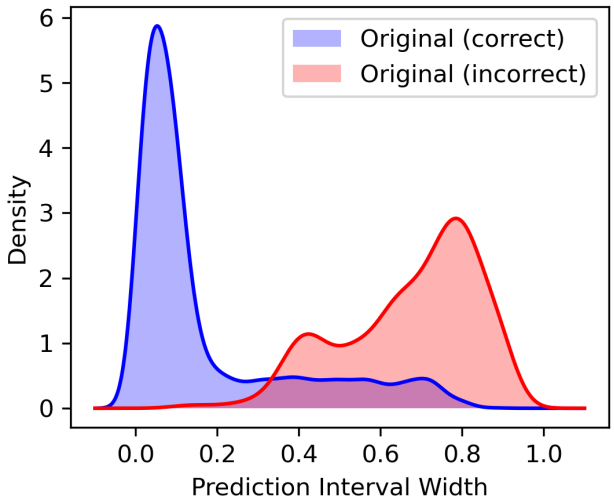
particular sensor's frequency response. It should be noted that the signals themselves are filtered, not the spectrogram images.

The CNN model is first trained and tested using the original dataset, and the subsequent set of filtered data represents the OOD data to be validated against the PI3NN UQ technique. It was anticipated that data that has been corrupted (i.e. filtered) will tend to force the model to make predictions that are *less confident* than those corresponding to data it has already seen. Therefore, the PI width should be larger, on average, than that of an original data sample.

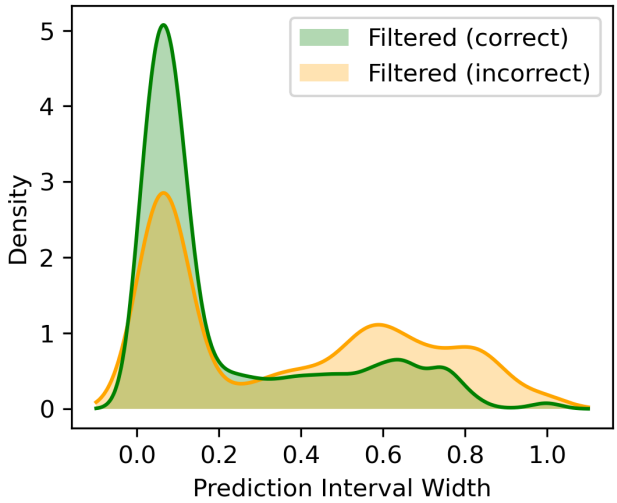
5.5 Results and Analysis

The initial CNN model, prior to the application of UQ via PI3NN, achieved roughly 96% accuracy for both the training and testing sets. When applied to the set of filtered data, this model achieved 70% accuracy, indicating that the distortion introduced to the data in the filtering process hinders the prediction capability of the CNN model. We train the PI3NN framework on the whole set of original data and tested it on the corrupted (i.e. "filtered") data. The method has 96.81% and 69.84% accuracy on the original and corrupted data, respectively. This agrees with the base CNN classifier. The accuracy-per-class of PI3NN obtained upon completion of the training stage of the mean output network, $f_{\boldsymbol{w}}(\boldsymbol{x})$, is shown in Table 5.2 (left), where the results on the training and testing set are presented in the middle and right columns of the table, respectively. Also shown are the results in another multi-class metric, the F1-score, in Table 5.2 (right).

It can be seen that buses 7 and 10 have subpar accuracies of 90.46% and 73.52% each (and similarly, relatively low F1-scores at 82.82% and 80.45%), indicating that these buses' transient voltage signals are not easily recognized by the model during training. This is likely due to the low magnitude of oscillations induced by the capacitors installed at those buses (see Table 5.1). Additionally, many of these data samples have been excluded from the data set following the curation stage described in Section 5.4.2 due to their limited frequency content.



(a) Original (non-filtered) signal data PDFs



(b) Filtered signal data PDFs

Figure 5.4: Probability density functions (PDFs) of the PI width for the correct and incorrect predictions, collected over all buses.

Following the training of $f_{\boldsymbol{w}}(\boldsymbol{x})$, networks $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $l_{\boldsymbol{\xi}}(\boldsymbol{x})$ are trained to estimate the uncertainty of the PI3NN mean output. The uncertainty of the prediction is characterized by the PI width, (5.6). Smaller PI widths indicates higher confidence in a predicted sample's class, whereas larger widths indicates higher uncertainty. Here, a PI width with quantile $\gamma = 0.95$ is considered and the *distribution* of PI widths over each bus (class) is investigated. Note that this analysis does not significantly change if γ takes other moderately large values ($\gamma > 0.7$).

The capability of PI3NN is next evaluated in estimating the certainty of output predictions. Figure 5.4 shows the probability density functions (PDF) of the PI width for the correctly labeled and incorrectly labeled signals collected over all buses. It can be seen that the PDFs shift to the right for incorrectly labeled signals, indicating that the PI widths are larger and that the model is less certain about these signals. For the original signal set, the PDFs are highly separable: most of correct predictions have a PI width less than 0.2 and all of the incorrect predictions have a PI width larger than 0.2. The difference is less prominent for the filtered signal set. Here, the support of both PDFs are overlapped. Both have peak at 0.1, however the PDF for incorrect predictions (orange) also has a second peak at 0.6. It can be observed that the PI widths of correct predictions are distributed analogously to the original case. However, the model seems more certain about its incorrect predictions. This is likely because a considerable portion of samples from buses 1, 3, 6, 8, and 9 are now incorrectly labeled due to the filtering process, and the model assigns them new labels with a similar level of certainty. These findings suggest that in the feature space, data belonging to these problematic buses may not be well-separated. In other words, such as a small distortion (from sensor-based filtering) can move them completely into the decision boundary of other buses.

Figure 5.5 shows the distribution of prediction interval widths for each class, for both the original and filtered signal sets. Buses 1, 3, 6, 8, and 9 are the closest, in terms of electrical distance, to the "center" of the grid, with bus 1 named as the source bus (Figure 5.1). As transient signals are captured at every bus for *each switching instance*, transients that occur at bus 5 are near-unrecognizable by the time they reach, for instance, bus 10. Those closest

to the electrical center of the grid have less distance to travel, and thus the filtering effects of the grid itself are lessened.

For this reason, the aforementioned buses have smaller variance around the mean of their prediction interval distributions. Similarly, their corresponding filtered signal sets show little shift in mean along with only a slightly larger variance. This indicates that the model, when switching signals from these buses are captured by the non-ideal sensor used in this study, are more likely to produce confident predictions.

Coinciding with Figure 5.5, Table 5.3 shows the per-class mean and standard deviation PI width for both original and filtered signal sets. It can clearly be seen that these first- and second-order PI statistics (in bold) for the filtered data sets indicate shifted, wider distributions. This naturally leads to the conclusion that signals corrupted by sensor irregularities lead to less confident predictions by the CNN model. For each bus, the mean of the distribution has been shifted, indicating an average higher uncertainty of the filtered data set. Further confirming this are the standard deviations of each bus; the distribution of the filtered signal set are wider than those of the corresponding original signal set.

5.6 Conclusions

To ensure continuous and reliable situational awareness of the power system, it is important for responsible parties to take appropriate action and put the necessary technologies in place. Traditional ML-based methods for monitoring and autonomous decision-making typically only provide point estimates for their predictions, and the developed models' confidence levels are almost entirely ignored. This paper presents a methodology for not only predicting the source location of capacitor transients, but for also providing prediction intervals, thereby gauging the overall confidence (or lack of) in the model's predictions. The value of the technique was verified by exposing the trained model to transients that had been corrupted by a non-ideal sensor's frequency response function, taken from an actual commercial-grade medium voltage distribution system sensor. It is shown that the PI3NN technique allows for users to quantify the amount of uncertainty, or even risk, associated with a given prediction,

Table 5.2: Accuracy (left) and F1-score (right) by bus

Bus	Original	Filtered
1	100%	93.37%
2	97.88%	78.71%
3	99.96%	63.49%
4	94.09%	59.96%
5	97.67%	58.60%
6	100%	66.71%
7	90.46%	59.98%
8	100%	56.47%
9	100%	99.41%
10	73.52%	41.73%

Bus	Original	Filtered
1	100%	95.52%
2	98.89%	65.32%
3	99.98%	66.46%
4	95.32%	59.71%
5	95.32%	63.31%
6	100%	61.71%
7	82.82%	53.05%
8	100%	69.99%
9	100%	97.69%
10	80.45%	45.57%

Table 5.3: Width of PI

Bus	Original Mean	Filtered Mean	Original Std.	Filtered Std.	Original Skewness	Filtered Skewness
1	0.045	**0.067**	0.036	**0.071**	1.6	**4.473**
2	0.217	**0.298**	0.215	**0.260**	**1.688**	1.031
3	0.093	**0.130**	0.068	**0.111**	1.476	**2.801**
4	0.198	**0.275**	0.203	**0.247**	**1.427**	0.817
5	0.221	**0.308**	0.228	**0.278**	**1.501**	0.857
6	0.065	**0.095**	0.051	**0.091**	1.525	**3.561**
7	0.257	**0.326**	0.189	**0.214**	**1.007**	0.66
8	0.071	**0.101**	0.055	**0.096**	1.556	**3.218**
9	0.071	**0.102**	0.056	**0.096**	1.535	**3.192**
10	0.277	**0.354**	0.211	**0.238**	**1.063**	0.662

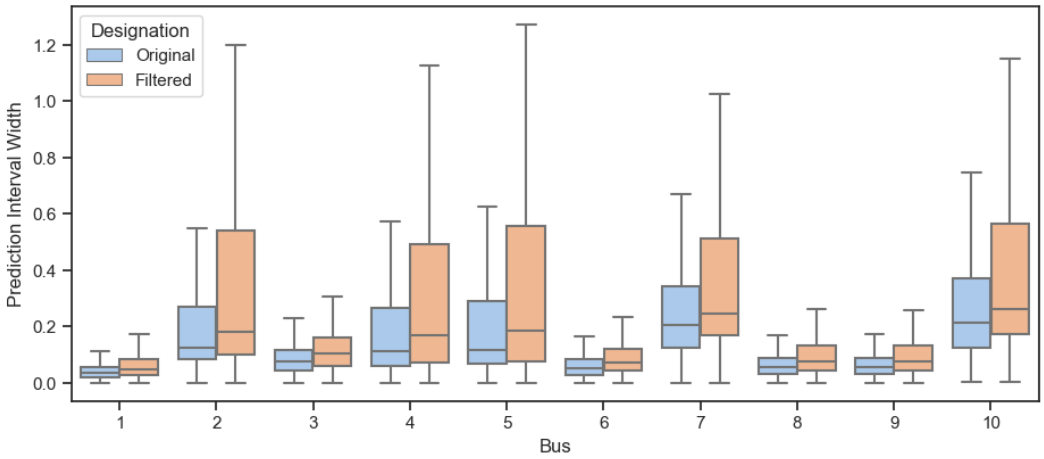


Figure 5.5: PI width distributions

contributing to the overall awareness of the power system's operating state at any given time.

These results can be further reinforced by examining multiple sensors with varying frequency response characteristics to quantify the degree of uncertainty introduced into the model's predictions and their correlations with the associated frequency response functions.

Chapter 6

Conclusions and Future Work

Applications that require the use of high-frequency waveform phenomena produced by power system sensors traditionally do not consider the effects of sensor distortions on said waveforms. This dissertation shows some of the effects on various advanced power system applications, including detection and classification problems. It was first shown that frequency response characteristics of line-post sensors have significant impacts on the high-frequency components of transient events (Chapter 2). It was shown that using percent error, phase difference, and goodness-of-fit metrics that various sensors possessing these frequency response irregularities have trouble reproducing certain events of interest (Tables 2.1-2.3).

These findings are further expanded upon in Chapter 3, in which statistical evaluation of the performance of two of these sensors is presented. For various current transient disturbances, including a capacitor switching event, a microgrid close-in, and a simulated fault on the terminals of a wind farm, the harmonic error distribution tended to drift away from a normal distribution, per the Anderson Darling Test for Normality. This contrasts with the common assumptions of the way error of measured quantities is distributed. The plots of percent error distributions, as estimated using KDE (see plots in Appendix A) show significant deviations from the normal distribution in the higher harmonics.

Chapters 2 and 3 illustrate some of the effects on measurement of high-frequency signals using commercial off-the-shelf line-post monitor sensors. The remainder of this dissertation examined the way these irregularities effect applications that require accurate representation of high-frequency transients.

Detection is the fundamental operation that must be successfully completed before any other application can proceed. Inaccurate pin-pointing of the start and end of transient signals can lead to inaccurate event diagnoses, improper parameter computations (i.e. fundamental and harmonic active and reactive power quantities), and false protection equipment statuses. For this reason, a detection algorithm was developed in this dissertation to address some of the concerns brought to light in Chapters 2 and 3.

In Chapter 4, an algorithm used for automated detection of these types of signals is presented in the form of the *energy detector*, requiring no assumptions of the underlying signal model. It was shown that this algorithm exhibits good performance for detection of high-frequency transients when subjected to signals that have been distorted by these types of sensors. Additionally, a comparison was performed against two other detection algorithms found in the literature: the "ultra-fast transient" (UFT) detector, and a detector based on the maximal-overlap discrete wavelet transform (MODWT). After a signal is successfully detected, it may be used for other purposes properly.

The final chapter of this dissertation examines the way sensor frequency response distortions affect applications that use these detected signals. In particular, a machine learning-based application is studied, with the goal of quantifying the uncertainty of predicted results and how that uncertainty shifts when subjected to distorted signals. In Chapter 5, an uncertainty quantification (UQ) technique for deep learning is presented in the context of the location of capacitor switching transients. These are a very common form of transient disturbance, as described in Chapter 1.3.2.

A convolutional neural network (CNN) is used as the "feature extractor" on time-frequency representations of capacitor switching voltage event waveforms. From there, the PI3NN technique is used for determining prediction intervals on the estimated location of a given signal. This essentially means that, once the CNN predicts the location that the signal originated from, associated uncertainty bounds are placed on that result, analogous to confidence intervals. It was shown that the performance of this technique when subjected to "clean" (i.e. ideal) signals vs. those after being subjected to typical sensor distortions results in much larger prediction interval widths, indicating an increase in uncertainty associated with predicted signal locations.

6.1 Future Work

The contents of this dissertation were primarily concerned with the study of frequency response-induced errors in high-frequency transient signals. There are of course other sources of error, such as those resulting from external temperature variations and life cycle degradation of sensors and equipment, to name just a few. It will be necessary to characterize *dynamic* and *shifting* errors on these types of signals for real-world deployment of advanced AI applications. Some other potential applications outside of the scope of this dissertation that should be considered include:

Fault location via traveling waves:

An emerging technique for the location of faults in transmission and distribution systems involves detecting very high-frequency traveling waves at extremely precise moments in time. With multiple measurement apparatuses deployed throughout an electrical network capable of capturing the various reflected and incident waves resulting from a fault, the location is able to be deduced from the relative detection times of these waves. Filtering characteristics described in this dissertation will of course have a significant impact on the collection of these high-frequency waves, and methods need to be developed to account for these distortions.

Measurement device placement:

Eliminating sources of error such as those presented in this dissertation may not always be feasible. Additionally, it is also well-known that high-frequency signals will filter out more quickly than lower-frequency ones over large electrical distances due to the nature of the capacitive and inductive properties of the lines and connected loads. It may be necessary to identify locations that these network-induced distortions are minimized, and use these locations to place PoW sensors.

Synthetic Data Generation

One of the potential outcomes of the results presented in Chapter 3 is advanced synthetic data generation. The derived non-parametric distributions may be used to augment synthetic

data such that they represent data corrupted by these kinds of sensors. This could be accomplished by developing a generative model.

This process could be further aided by taking into account uncertainties not necessarily arising from the sensors themselves, but also from electrical properties used in the data generating processes (e.g. EMTP models). For example, transmission line impedances are often lumped or approximated based on standards or "base cases". It is also extremely difficult to get an accurate measure of line impedance from existing equipment. These assumptions will lead to further uncertainty components.

This generative process would allow for creating more representative and robust datasets to address the effects on machine learning applications, as presented in Chapter 5. A training set that contains a more diverse set of data scenarios will yield better prediction results with lower uncertainty.

Bibliography

[1] T. A. Short. *Electric Power Distribution Handbook*. 2004. 1

[2] University of Tennessee. FNET/GridEye Web Display. 3, 55

[3] Ieee standard for performance and test requirements for instrument transformers of a nominal system voltage of 115 kv and above - redline. *IEEE Std C57.13.5-2019 (Revision of IEEE Std C57.13.5-2009) - Redline*, pages 1–81, 2020. 5

[4] Ariana Hargrave, Michael J. Thompson, and Brad Heilman. Beyond the knee point: A practical guide to ct saturation. In *2018 71st Annual Conference for Protective Relay Engineers (CPRE)*, pages 1–23, 2018. 6, 7

[5] IEEE Recommended Practice for Monitoring Electric Power Quality. *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)*, pages 1–98, 2019. 8, 15

[6] Lou van der Sluis. *Transients in Power Systems*. John Wiley & Sons Ltd, 2001. 11, 12, 13, 14

[7] Edmund O. Schweitzer, Armando Guzmn, Mangapathirao V. Mynam, Veselin Skendzic, Bogdan Kasztenny, and Stephen Marx. Locating faults by the traveling waves they launch. In *2014 67th Annual Conference for Protective Relay Engineers*, pages 95–110, 2014. 15

[8] Instrument transformers - Part 2: Additional requirements for current transformers. *IEC 61869-2*, 2012. 21

[9] Additional requirements for inductive voltage transformers. *IEC 61869-3*, 2011. 21

[10] Instrument transformers - Part 8: Electronic current transformers. *IEC 60044-8*, 2002. 21, 26

[11] G. Crotti, A. Delle Femine, D. Gallo, D. Giordano, C. Landi, P.S. Letizia, and M. Luiso. Traceable characterization of low power voltage instrument transformers for pq and pmu applications. In *2020 Conference on Precision Electromagnetic Measurements (CPEM)*, pages 1–2, 2020. 21

[12] A. Cataliotti, V. Cosentino, G. Crotti, A. D. Femine, D. Di Cara, D. Gallo, D. Giordano, C. Landi, M. Luiso, M. Modarres, and G. Tin. Compensation of nonlinearity of voltage and current instrument transformers. *IEEE Transactions on Instrumentation and Measurement*, 68(5):1322–1332, 2019. 21, 22

[13] A. Cataliotti, V. Cosentino, G. Crotti, D. Giordano, M. Modarres, D. Di Cara, G. Tin, D. Gallo, C. Landi, and M. Luiso. Metrological performances of voltage and current instrument transformers in harmonics measurements. In *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1–6, 2018. 21

[14] A. Cataliotti, D. Di Cara, A. E. Emanuel, and S. Nuccio. Characterization of clamp-on current transformers under nonsinusoidal conditions. *IEEE Transactions on Power Delivery*, 24(1):373–380, 2009. 21

[15] M. Faifer, C. Laurano, R. Ottoboni, S. Toscani, and M. Zanoni. Characterization of voltage instrument transformers under nonsinusoidal conditions based on the best linear approximation. *IEEE Transactions on Instrumentation and Measurement*, 67(10):2392–2400, 2018. 21, 22

[16] Oppenheim, A. V. and Schafer, R. W. Pearson, 3 edition, 2010. 22

[17] A. Riepnieks and H. Kirkham. An introduction to goodness of fit for pmu parameter estimation. *IEEE Transactions on Power Delivery*, 32(5):2238–2245, 2017. 24, 26

[18] Ahmad, Tabia and Senroy, Nilanjan. Statistical Characterization of PMU Error for Robust WAMS Based Analytics. *IEEE Transactions on Power Systems*, 35(2):920–928, 2020. 35

[19] Wang, Shaobu and Zhao, Junbo and Huang, Zhenyu and Diao, Ruisheng. Assessing Gaussian Assumption of PMU Measurement Error Using Field Data. *IEEE Transactions on Power Delivery*, 33(6):3233–3236, 2018. 36

[20] Zhao, Jiecheng and Tan, Jin and Wu, Ling and Zhan, Lingwei and Yao, Wenxuan and Liu, Yilu. Impact of the Measurement Errors on Synchrophasor-Based WAMS Applications. *IEEE Access*, 7:143960–143972, 2019. 36

[21] Cheng, Gang and Lin, Yuzhang and Chen, Yanbo and Bi, Tianshu. Adaptive State Estimation for Power Systems Measured by PMUs With Unknown and Time-Varying Error Statistics. *IEEE Transactions on Power Systems*, 36(5):4482–4491, 2021. 36

[22] Nugroho, Sebastian A. and Taha, Ahmad F. and Qi, Junjian. Robust Dynamic State Estimation of Synchronous Machines With Asymptotic State Estimation Error Performance Guarantees. *IEEE Transactions on Power Systems*, 35(3):1923–1935, 2020. 36

[23] Ahmad, Tabia and Senroy, Nilanjan. An Information Theoretic Approach to Power-Substation Level Dynamic State Estimation With Non-Gaussian Noise. *IEEE Transactions on Power Systems*, 35(2):1642–1645, 2020. 36

[24] Chakrabarti, Saikat and Kyriakides, Elias. PMU Measurement Uncertainty Considerations in WLS State Estimation. *IEEE Transactions on Power Systems*, 24(2):1062–1071, 2009. 36

[25] Chen, Lei and Zhao, Wei and Zhao, Dongfang and Huang, Songling. Frequency-Domain Sampling Theorem-based Harmonic Phasor Estimator. In *2020 Conference on Precision Electromagnetic Measurements (CPEM)*, pages 1–2, 2020. 36

[26] Chen, Lei and Zhao, Wei and Wang, Fuping and Huang, Songling. Harmonic Phasor Estimator for P-Class Phasor Measurement Units. *IEEE Transactions on Instrumentation and Measurement*, 69(4):1556–1565, 2020. 36

[27] Duda, Krzysztof and Zieliǹski, Tomasz P. and Bieǹ, Andrzej and Barczentewicz, Szymon H. Harmonic Phasor Estimation With Flat-Top FIR Filter. *IEEE Transactions on Instrumentation and Measurement*, 69(5):2039–2047, 2020. 36

[28] Jain, Sachin K. and Jain, Preeti and Singh, Sri Niwas. A Fast Harmonic Phasor Measurement Method for Smart Grid Applications. *IEEE Transactions on Smart Grid*, 8(1):493–502, 2017. 36

[29] Jain, Sachin K. and Singh, S. N. Exact Model Order ESPRIT Technique for Harmonics and Interharmonics Estimation. *IEEE Transactions on Instrumentation and Measurement*, 61(7):1915–1923, 2012. 36

[30] IEEE Standard Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Nonsinusoidal, Balanced, or Unbalanced Conditions - Redline. *IEEE Std 1459-2010 (Revision of IEEE Std 1459-2000) - Redline*, pages 1–52, 2010. 36

[31] Wilson, Aaron J. and Warmack, Bruce R.J. and Kerekes, Ryan A. and Brukiewa, Patrick D. Comparison of Power System Current Sensors via Playback of Electrical Disturbances. In *2022 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–5, 2022. 39

[32] T. W. Anderson and D. A. Darling . A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268):765–769, 1954. 40

[33] M. A. Stephens. EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, 69(347):730–737, 1974. 40

[34] Type 4 wind turbine generators. 42, 58

[35] G. T. Alencar, R. C. Santos, and A. O. Neves. Fault Location Using Wavelet Transform and Independent Component Analysis. In *Proc. IEEE CHILECON CHILEAN Conf.*

on Electrical, Electronics Engineering, Inf. and Commun. Tech., pages 1–7, Valparaiso, Chile, Nov 2019. 55

[36] F. B. Costa. Fault-Induced Transient Detection Based on Real-Time Analysis of the Wavelet Coefficient Energy. *IEEE Trans. Power Del.*, PWRD–29(1):140–153, Feb. 2014. 55

[37] F. Costa and B. A. Souza. Fault-Induced Transient Analysis for Real-Time Fault Detection and Location in Transmission Line. In *Proc. IPST Int. Conf. on Power Syst. Transients*, pages 1–6, Delft, Netherlands, Jun. 2011. 55

[38] S. Devi, N. K. Swarnkar, S. R. Ola, and O. P. Mahela. Detection of Transmission Line Faults Using Discrete Wavelet Transform. In *Proc. IEEE CASP Conf. on Adv. in Signal Process.*, pages 133–138, Pune, India, Jun. 2016. 55

[39] M. S. Pranav, C. Karthik, D. Kavitha, K. Vishal, J. Tarun, and V. Vanitha. Fault Detection and Classification in Three Phase Transmission Lines using Signal Processing. In *Proc. IEEE RTEICT 3rd Int. Conf. on Recent Trends in Electronics, Inf. & Commun. Tech.*, pages 347–350, Bangalore, India, May 2018. 55

[40] G. Feng, M. Chen, P. Wang, M. Xu, J. Yuan, and Z. Jiao. Faulted Line Detection Using Time-Frequency Information for Distribution System. In *Proc. IEEE/IAS ICPS Asia Ind. and Commercial Power Syst. Asia*, pages 976–981, Weihai, China, Jul. 2020. 55

[41] Youzhen Liang, Xiping Ma, Fengzhan Zhao, Shuai Hao, Songhuai Du, Juan Su, Yu Zhang, and Tingting Zhao. A High Accuracy Detection Method of Voltage Flicker Signal Based on Time-Frequency Transform. In *Proc. IEEE ICPES 9th Int. Conf. on Power and Energy Syst.*, pages 1–5, Perth, WA, Australia, Dec. 2019. 55

[42] Atif Maqsood, Nick Rossi, Yue Ma, Keith Corzine, Leila Parsa, and Damian Oslebo. STFT-Based Event Detection and Classification for a DC Pulsed Load. In *Proc. IEEE ESTS Electric Ship Tech. Symp.*, pages 492–498, Washington, DC, USA, Aug. 2019. 55

[43] I. Paraskevas, K. Prekas, S. M. Potirakis, and M. Rangoussi. On the Use of Time-Frequency Distributions for the Power Quality Problem of Harmonics. In *Proc.*

IEEE/IET MedPower Mediterranean Conf. and Exhibition on Power Generation, Transmission, Distribution and Energy Conversion, pages 1–5, Agia Napa, Greek Republic of Southern Cyprus, Nov. 2010. 55

[44] I.J. Rivera, A.B. Ramirez, and D. Rodriguez. A Time-Frequency Signal Analysis System for Power Quality Assessment. In *Proc. IEEE 48th Midwest Symp. on Circuits and Syst.*, pages 1677–1680, Covington, KY, USA, Aug. 2005. 55

[45] Dilan Senaratne, Jinsub Kim, and Eduardo Cotilla-Sanchez. Spatio-Temporal Frequency Domain Analysis of PMU Data for Unsupervised Event Detection. In *Proc. IEEE ISGT Power & Energy Society Innovative Smart Grid Tech. Conf.*, pages 1–5, Washington, DC, USA, Feb. 2021. 55

[46] D. Gabor. Theory of Communication. *IET J. Inst. of Electrical Eng.*, 93(26):429–457, Nov. 1946. 55

[47] Lingwei Zhan, Bailu Xiao, Fuhua Li, He Yin, Wenxuan Yao, Zhi Li, and Yilu Liu. Fault–tolerant Grid Frequency Measurement Algorithm During Transients. *IET Energy Syst. Integration*, 2(3):173–178, Sep. 2020. 55, 61

[48] Harry Urkowitz. Energy Detection of Unknown Deterministic Signals. *Proc. IEEE*, 55(4):523–531, Apr. 1967. 55, 57

[49] A. R. Ekti, A. J. Wilson, J. Olatt, J. Holliman II, S. Yarkan, and P. L. Fuhr. Simple and Accurate Transient Waveform Detection for Smart Grids: RealWorld Field Data Performance. submitted to MDPI. 55

[50] North American Electric Reliability Council. Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn? https://www.nerc.com/docs/docs/blackout/NERC_Final_Blackout_Report_07_13_04.pdf, 2004. [Online; accessed January, 2022]. 67

[51] North American Electric Reliability Council. Odessa Disturbance Texas Events: May 9, 2021 and June 26, 2021. https://www.nerc.com/pa/rrm/ea/Documents/Odessa_Disturbance_Report.pdf. [Online; accessed July, 2022]. 67

[52] A.C. Parsons, W.M. Grady, E.J. Powers, and J.C. Soward. A direction finder for power quality disturbances based upon disturbance power and energy. *IEEE Transactions on Power Delivery*, 15(3):1081–1086, 2000. 68

[53] Yong-June Shin, Edward J. Powers, W. Mack Grady, and Ari Arapostathis. Signal processing-based direction finder for transient capacitor switching disturbances. *IEEE Transactions on Power Delivery*, 23(4):2555–2562, 2008. 69

[54] H. Khani, M. Moallem, and S. Sadri. On tracking and finding the location of switched capacitor banks in distribution systems. In *2009 Transmission Distribution Conference Exposition: Asia and Pacific*, pages 1–4, 2009. 69

[55] H. Khani, M. Moallem, and S. Sadri. A novel algorithm for determining the exact location of switched capacitor banks in distribution systems. In *2009 Transmission Distribution Conference Exposition: Asia and Pacific*, pages 1–4, 2009. 69

[56] H. Khani, M. Moallem, M. Dolatshahi, A. Davari, and S. Sadri. A new index for online tracking of the switched capacitor bank location in distribution systems. In *2010 18th Iranian Conference on Electrical Engineering*, pages 856–860, 2010. 69

[57] C. N. Bhende. S-transform technique for identifying the location of a switched capacitor. In *2011 10th International Conference on Environment and Electrical Engineering*, pages 1–4, 2011. 69

[58] G.W. Chang, M.H. Shih, S.Y. Chu, and R. Thallam. An efficient approach for tracking transients generated by utility shunt capacitor switching. *IEEE Transactions on Power Delivery*, 21(1):510–512, 2006. 69

[59] Gary W. Chang, Ju-Peng Chao, Hunter M. Huang, Cheng-I Chen, and Shou-Yung Chu. On tracking the source location of voltage sags and utility shunt capacitor switching transients. *IEEE Transactions on Power Delivery*, 23(4):2124–2131, 2008. 69

[60] David J. C. MacKay. A practical bayesian framework for backpropagation networks. *Neural Comput.*, 4(3):448472, May 1992. 70

[61] Matthew D. Hoffman, David M. Blei, Chong Wang, and John Paisley. Stochastic variational inference. *Journal of Machine Learning Research*, 14(4):1303–1347, 2013. 70

[62] Alexander Amini, Wilko Schwarting, Ava Soleimany, and Daniela Rus. Deep evidential regression. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 14927–14937. Curran Associates, Inc., 2020. 70, 75

[63] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 64056416, Red Hook, NY, USA, 2017. Curran Associates Inc. 70, 75

[64] Tim Pearce, Felix Leibfried, and Alexandra Brintrup. Uncertainty in neural networks: Approximately bayesian ensembling. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 234–244. PMLR, 26–28 Aug 2020. 70

[65] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1050–1059, New York, New York, USA, 20–22 Jun 2016. PMLR. 70

[66] Pei Zhang, Siyan Liu, Dan Lu, Guannan Zhang, and Ramanan Sankaran. A prediction interval method for uncertainty quantification of regression models. Technical report, Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 2021. 70

[67] Richard D. De VlEAUX, Jennifer Schumi, Jason Schweinsberg, and Lyle H. Ungar. Prediction intervals for neural networks via nonlinear regression. *Technometrics*, 40(4):273–282, 1998. 70

[68] J. T. Gene Hwang and A. Adam Ding. Prediction intervals for artificial neural networks. *Journal of the American Statistical Association*, 92(438):748–757, 1997. 70

[69] J.G. Carney, P. Cunningham, and U. Bhagwan. Confidence and prediction intervals for neural network ensembles. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, volume 2, pages 1215–1218 vol.2, 1999. 71

[70] Tom Heskes. Practical confidence and prediction intervals. In *Proceedings of the 9th International Conference on Neural Information Processing Systems*, NIPS'96, page 176182, Cambridge, MA, USA, 1996. MIT Press. 71

[71] Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. 71

[72] Roger Koenker and Kevin F. Hallock. Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156, December 2001. 71

[73] Abbas Khosravi, Saeid Nahavandi, Doug Creighton, and Amir F. Atiya. Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Transactions on Neural Networks*, 22(3):337–346, 2011. 71

[74] Tim Pearce, Alexandra Brintrup, Mohamed Zaki, and Andy Neely. High-quality prediction intervals for deep learning: A distribution-free, ensembled approach. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4075–4084. PMLR, 10–15 Jul 2018. 71

[75] Eli Simhayev, Gilad Katz, and Lior Rokach. Piven: A deep neural network for prediction intervals with specific value prediction, 2021. 71

[76] Tárik S. Salem, Helge Langseth, and Heri Ramampiaro. Prediction intervals: Split normal mixture from quality-driven deep ensembles. In Jonas Peters and David Sontag, editors, *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*

(UAI), volume 124 of *Proceedings of Machine Learning Research*, pages 1179–1187. PMLR, 03–06 Aug 2020. 71, 75

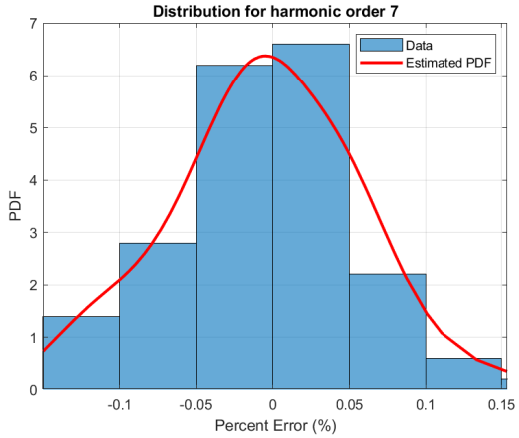
[77] Siyan Liu, Pei Zhang, Dan Lu, and Guannan Zhang. PI3NN: Out-of-distribution-aware prediction intervals from three neural networks. In *International Conference on Learning Representations*, 2022. 71, 72, 73, 77

[78] W. C. Duesterhoeft, Max W. Schulz, and Edith Clarke. Determination of instantaneous currents and voltages by means of alpha, beta, and zero components. *Transactions of the American Institute of Electrical Engineers*, 70(2):1248–1255, 1951. 79

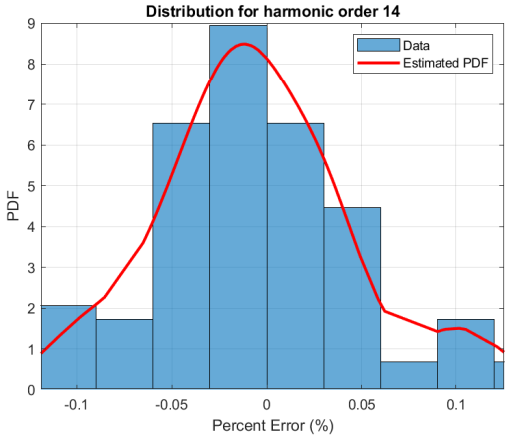
Appendices

A Harmonic Event Distribution KDE Plots

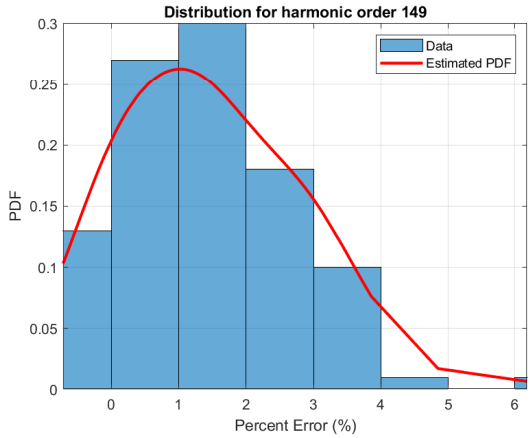
The distributions shown in Figs. A.1-A.5 represent the percent error distributions, both empirical (blue, histogram plots) and estimated via KDE (red, solid lines) for the wind fault event as depicted in Chapter 3.



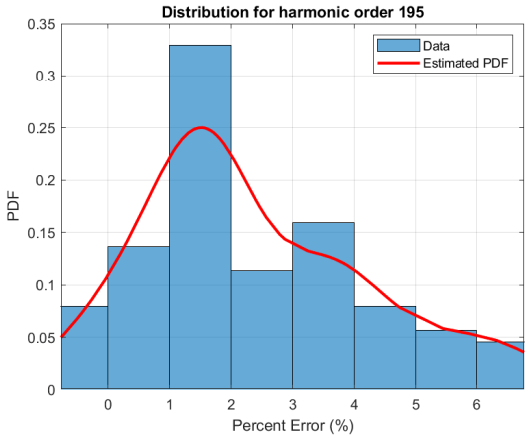
(a) Harmonic order 7



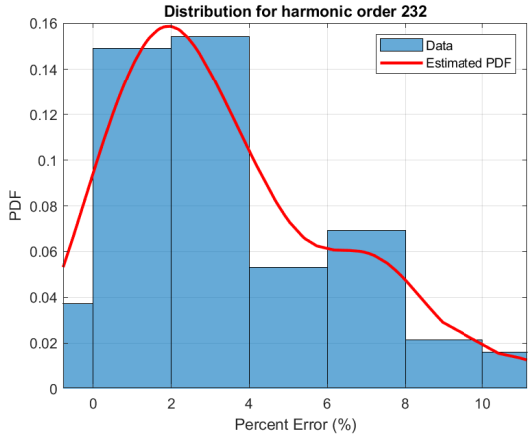
(b) Harmonic order 14



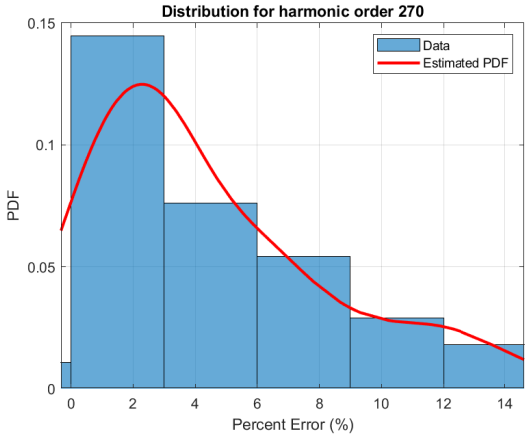
(c) Harmonic order 149



(d) Harmonic order 195

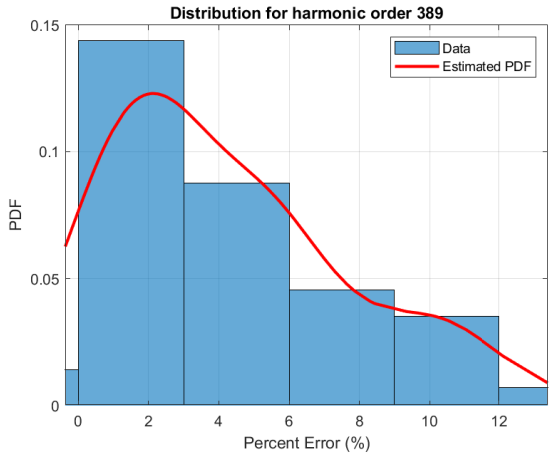


(e) Harmonic order 232

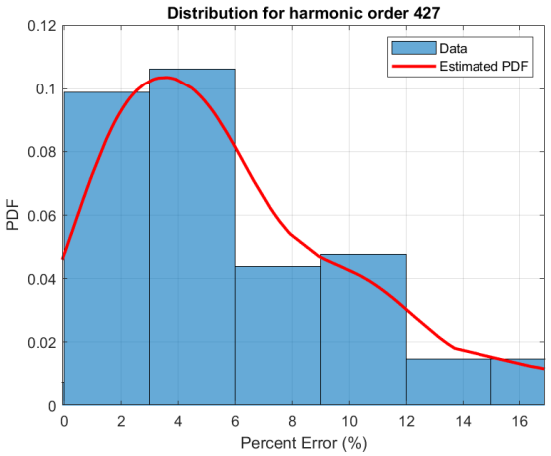


(f) Harmonic order 270

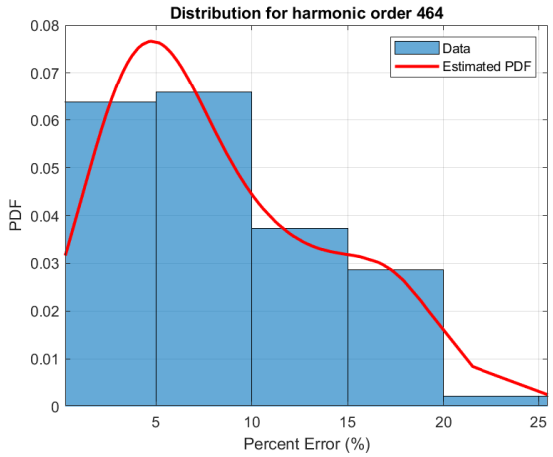
Figure A.1: Wind fault G&W error statistics for harmonic indices 1-6 (from Table 3.7)



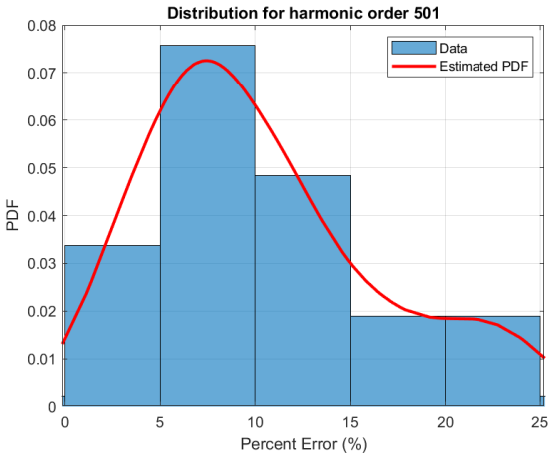
(a) Harmonic order 389



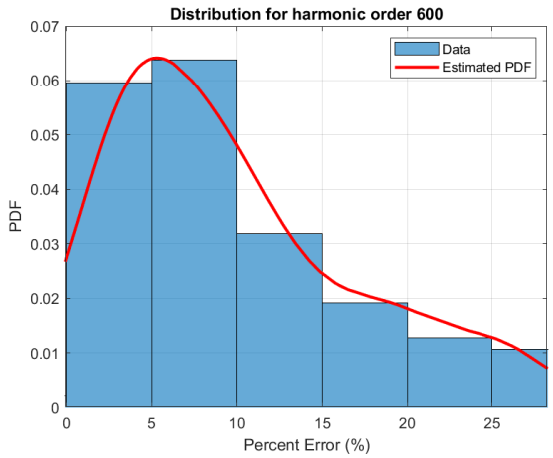
(b) Harmonic order 427



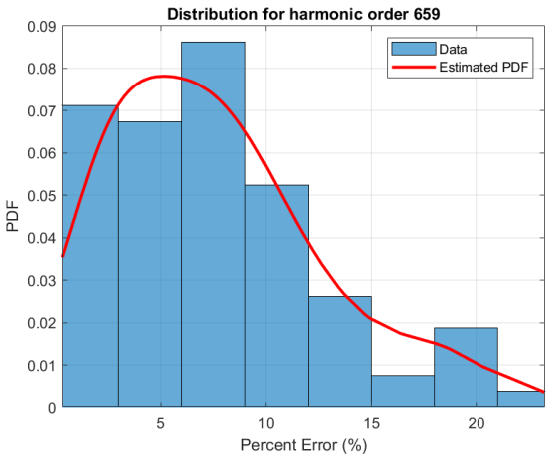
(c) Harmonic order 464



(d) Harmonic order 501

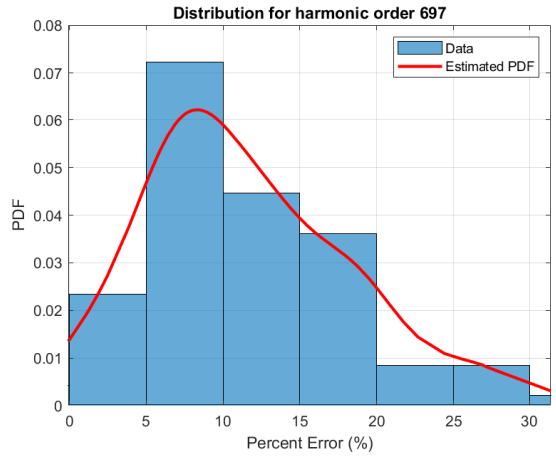


(e) Harmonic order 600

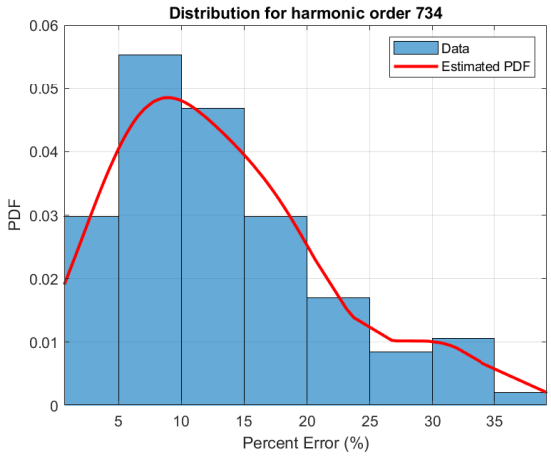


(f) Harmonic order 659

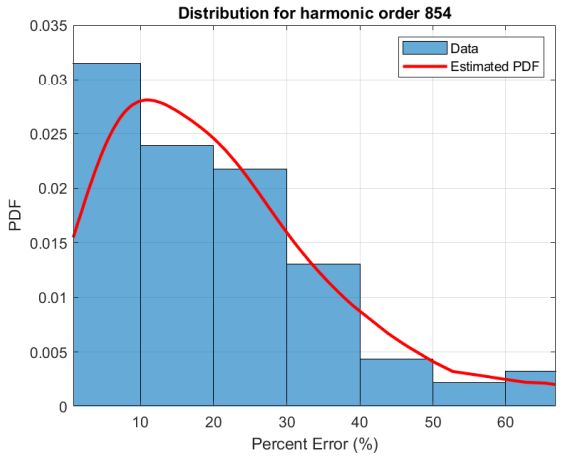
Figure A.2: Wind fault G&W error statistics for harmonic indices 7-12 (from Table 3.7)



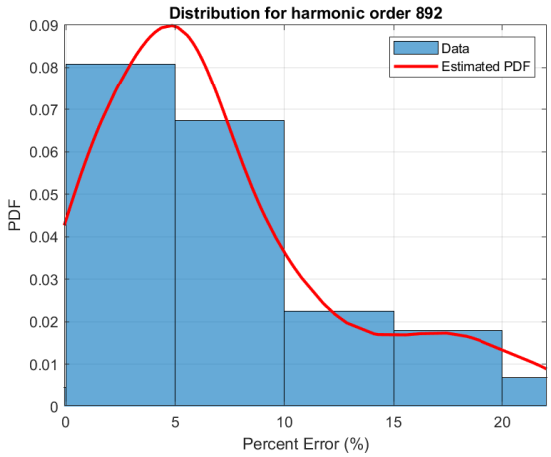
(a) Harmonic order 697



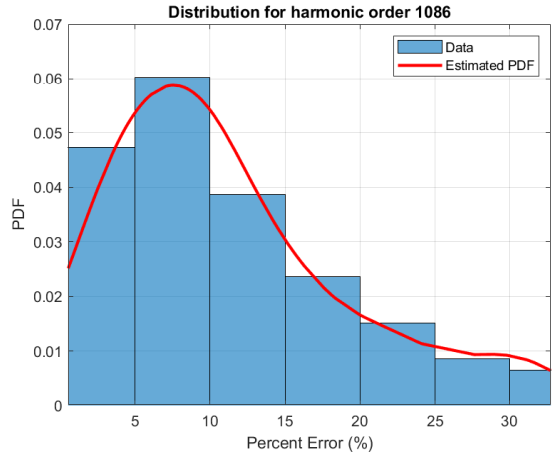
(b) Harmonic order 734



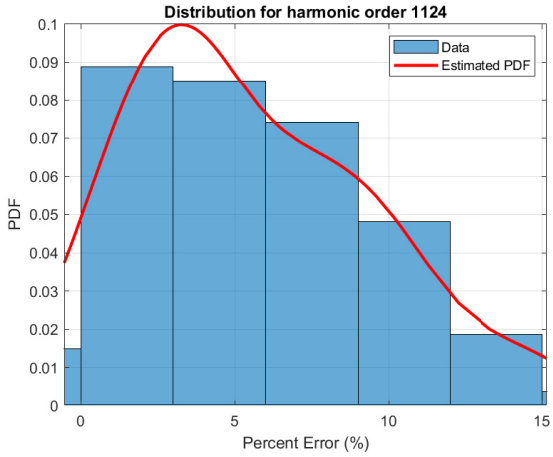
(c) Harmonic order 854



(d) Harmonic order 892

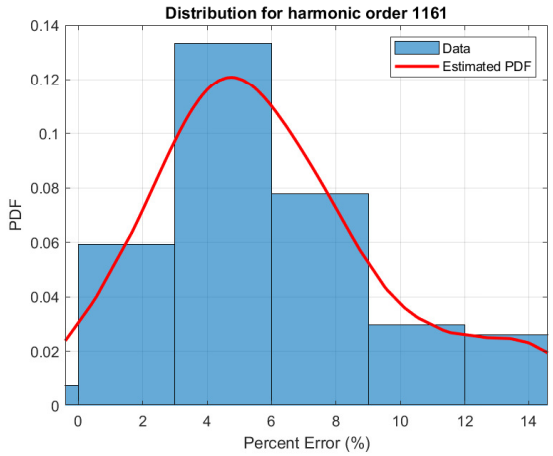


(e) Harmonic order 1086

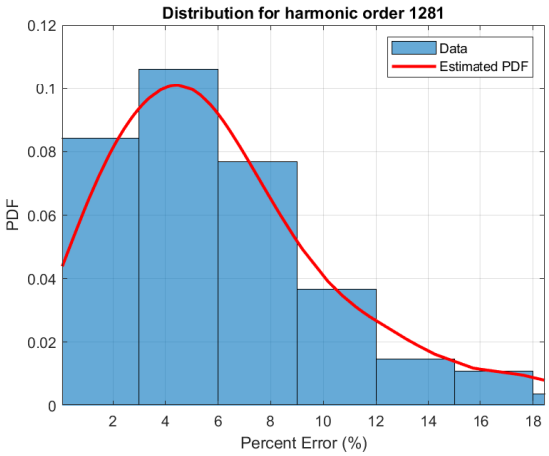


(f) Harmonic order 1124

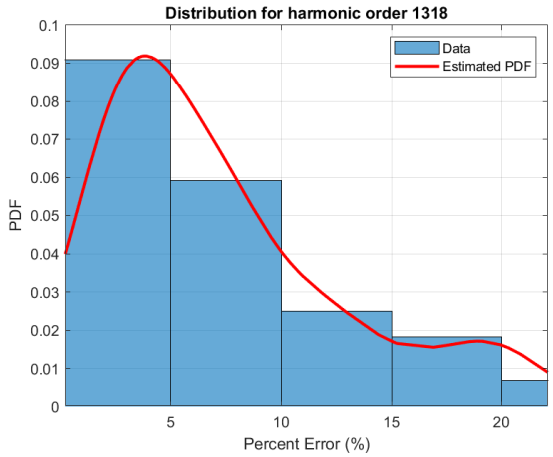
Figure A.3: Wind fault G&W error statistics for harmonic indices 13-18 (from Table 3.7)



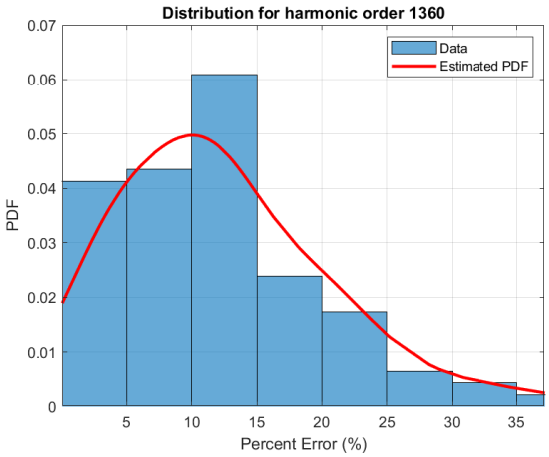
(a) Harmonic order 1161



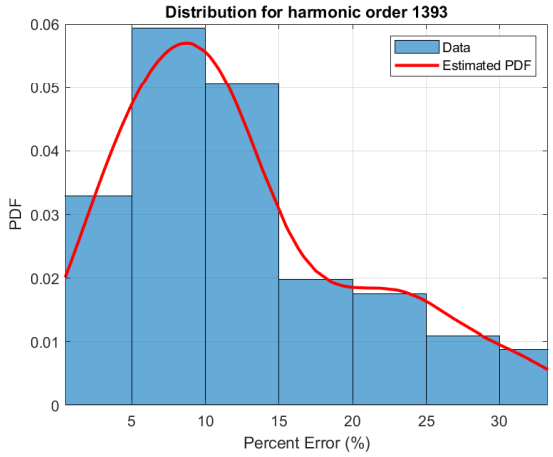
(b) Harmonic order 1281



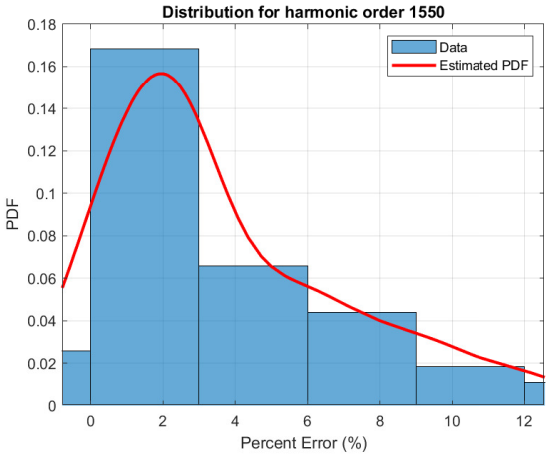
(c) Harmonic order 1318



(d) Harmonic order 1360

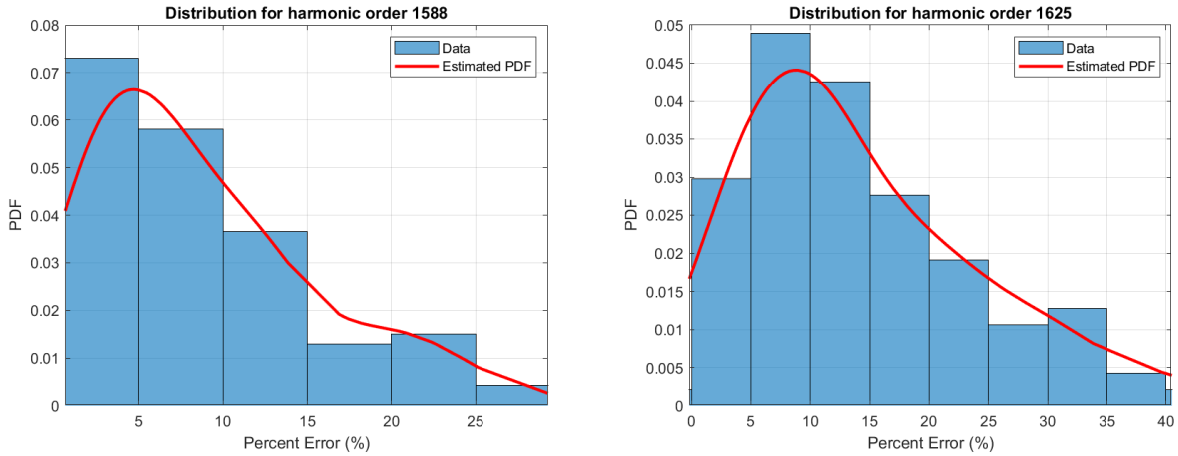


(e) Harmonic order 1393



(f) Harmonic order 1550

Figure A.4: Wind fault G&W error statistics for harmonic indices 19-24 (from Table 3.7)



(a) Harmonic order 1558

(b) Harmonic order 1625

Figure A.5: Wind fault G&W error statistics for harmonic indices 25-26 (from Table 3.7)

B Energy Detector Algorithm: Analytical Derivation of Threshold γ

In Chapter 4, it was mentioned that the presence of a transient may be expressed as a hypothesis test in which the null hypothesis H_0 represents the case in which a transient does not exist in the observation window, and H_1 otherwise. In this appendix, an analytical method for approximating the threshold γ for (4.4) is presented via Maximum Likelihood Estimation.

For the case in which there is no transient (i.e. $H_0: \ p_{Tq}[k] = 0$), the probability density function (PDF) of the energy values $\psi[\cdot]$ is expressed as a chi-squared distribution with N degrees of freedom, which of course assumes that each element of $r[k]$ in the sum of (4.4) is normally-distributed with mean 0 and unit variance:

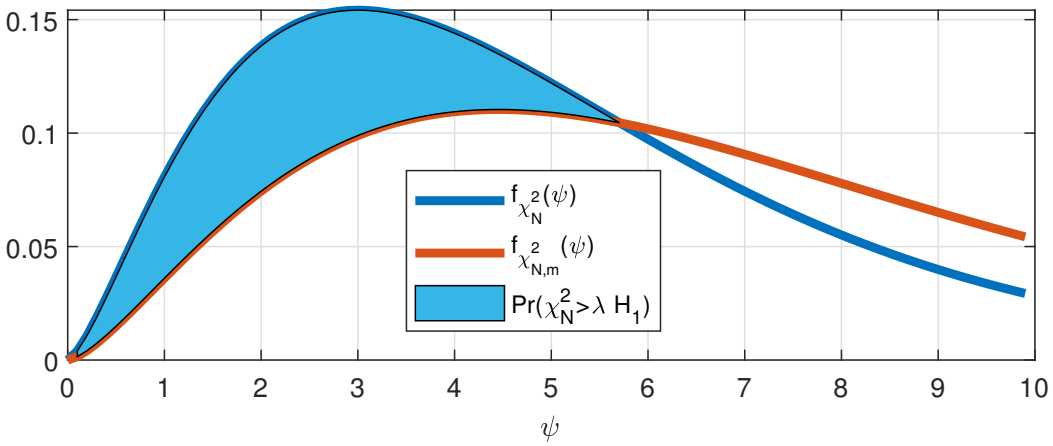
$$f_{\chi^2_N}(\psi) = \frac{\psi^{(N/2)-1} e^{-\frac{\psi}{2}}}{2^{N/2}\Gamma(N/2)} \tag{B.1}$$

where $\Gamma(\cdot)$ denotes the Gamma function. Similarly, for the case in which a transient is present ($H_1: \ p_{Tq} \neq 0$), the PDF may be represented as a noncentral chi-square distribution with parameter m and N degrees of freedom:

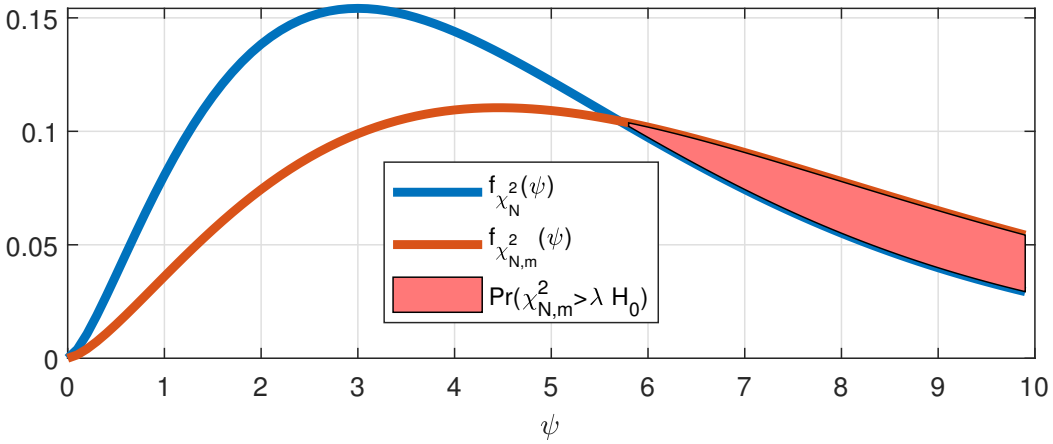
$$f_{\chi^2_{N,m}}(\psi) = \frac{1}{2}\left(\frac{\psi}{m}\right)^{N/4-\frac{1}{2}} e^{-\frac{m+\psi}{2}} J_{(N/2)-1}\left(\sqrt{m\psi}\right) \tag{B.2}$$

where $m = \sum_{i=0}^{N-1} \mu_i^2$ represents a random variable equal to the sum of the squared means of the r_i's, and $J_k(\cdot)$ denotes the k^{th} order modified Bessel function of the first kind.

The probability of detection and probability of false alarm may be analyzed by $\Pr(\chi^2_N > \lambda H_1)$ and $\Pr\left(\chi^2_{N,m} > \lambda H_0\right)$, respectively. Here, $\Pr(\chi^2_N > \lambda H_1) = Q_h\left(\sqrt{2\gamma}, \sqrt{\lambda}\right)$ is the generalized Marcum-Q function, λ represents the decision threshold, and γ represents the instantaneous signal-to-noise ratio (SNR). Thus it is important to find the value of γ that satisfies (4.4).



(a) Probability of detection



(b) Probability of false alarm

Figure B.1: Example probability distributions in blue (B.1) and red (B.2) along with decision probabilities.

B.1 Maximum Likelihood Estimation for Threshold Selection

Given the signal model in (4.2), error residuals $e[k]$ at high SNRs may be tracked at each point k using:

$$e[k] = r[k] - A\cos\left(2\pi f\frac{k}{f_s} + \theta\right) \approx w[k] + q[k]p_{Tq}[k - N_p] + \epsilon_N \tag{B.3}$$

This equation simply represents the portions of the original signal not captured in the model $r[k]$, and the measurement noise $w[k]$. For a 60-Hz signal with additive white Gaussian noise (AWGN) and a high-frequency transient superimposed on top of it, $e[k]$ would simply be the sum of the noise component and the transient. This formulation is suitable for determining the threshold value independent of the original signal's magnitude.

Finding a suitable threshold value is simply a problem of estimating the variance of this quantity. Simple Maximum Likelihood Estimation (MLE) can be used to accomplish this. The MLE solution to this problem, Θ, may be found as:

$$\Theta = \underset{\sigma_e^2}{\text{argmax}}\left\{\mu_e = \frac{1}{N_p}\sum_{k=0}^{N_p-1} e[k],\ \sigma_e^2 = \frac{1}{N_p}\sum_{k=0}^{N_p-1}\left(e[k] - \mu_e\right)^2\right\} \tag{B.4}$$

$$\approx \underset{\sigma_q^2}{\text{argmax}}\left\{\mu_q = \frac{1}{N_p}\sum_{k=0}^{N_p-1} e[k],\ \sigma_q^2 = \frac{1}{N_p}\sum_{k=0}^{N_p-1}\left(e[k] - \mu_q\right)^2\right\} \tag{B.5}$$

where μ_e is the mean of $e[k]$, approximate equivalence between the two expressions is established from (B.3) and the fact that $\mathbb{E}[w[k]] = 0$. Finally, the threshold γ may be found by:

$$\gamma = \sigma_w^2 + \alpha\sigma_q^2 \tag{B.6}$$

where σ_w^2 is the noise power and $0 < \alpha < 1$.

Vita

Aaron Wilson was born in Chattanooga, Tennessee. He received his B.S. and M.S. degrees in Electrical Engineering from The University of Tennessee at Chattanooga in 2017 and 2019, respectively. Currently he is working in the Grid Communications and Security group at Oak Ridge National Laboratory while pursuing the Ph.D degree in Electrical Engineering from the University of Tennessee at Knoxville. His research interests include data analysis for the power grid, evaluation of grid sensor systems, and study of measurements used in grid applications.