A

Ph.D. THESIS

on

Predictive Analytics and Machine Learning for Mitigating Unintentional Islanding in Solar PV Distributed Generation Systems

submitted for partial fulfillment for the degree of

Doctor of Philosophy in Engineering

Academic Session

(2014-2018)

Supervisors:
Dr. Rajesh Kumar
Dr. Rajesh Kavasseri

Submitted By:
Shashank Vyas
(2014REN9053)

Centre for Energy and Environment

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY, JAIPUR (RAJ.)

Predictive Analytics and Machine Learning for Mitigating Unintentional Islanding in Solar PV Distributed Generation Systems

Ph.D. Thesis

Shashank Vyas
(Registration Number 2014REN9053)

CENTRE FOR ENERGY AND ENVIRONMENT
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR
JANUARY 2018

CENTRE FOR ENERGY AND ENVIRONMENT
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

CERTIFICATE

This is to certify that the thesis entitled, **"Predictive Analytics and Machine Learning for Mitigating Unintentional Islanding in Solar PV Distributed Generation Systems"** submitted by Shashank Vyas (Registration No. 2014REN9053) to Malaviya National Institute of Technology, Jaipur for the award of the degree of ***Doctor of Philosophy*** in Renewable Energy is a bonafide record of original research work carried out by him under our supervision.

It is further certified that:

1. The results contained in this thesis have not been submitted in part or in full to any other university or institute for the award of any degree.
2. Mr. Shashank Vyas has fulfilled the requirements for the submission of this thesis.

Dr. Rajesh Kumar
(Supervisor)
Adjunct Associate Professor
Centre for Energy and Environment
MNIT Jaipur
India

Dr. Rajesh Kavasseri
(External Joint-Supervisor)
Associate Professor
Electrical and Computer Engineering
North Dakota State University
Fargo, USA

Date:

CENTRE FOR ENERGY AND ENVIRONMENT
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

CANDIDATE'S DECLARATION

I hereby declare that the thesis titled,**"Predictive Analytics and Machine Learning for Mitigating Unintentional Islanding in Solar PV Distributed Generation Systems"** is my own work conducted under the supervision of Dr. Rajesh Kumar and Dr. Rajesh Kavasseri. I confirm that:

- This work is done towards the partial fulfillment of the degree of "*Doctor of Philosophy*" at MNIT, Jaipur.
- To the best of my knowledge, the results of this work have not been published anywhere else.
- Wherever I have consulted the published work of others, this has been clearly mentioned with proper citation.
- I have acknowledged all main sources of help.

Shashank Vyas
(2014REN9053)

Date:

Place: MNIT Jaipur

ACKNOWLEDGMENTS

It gives me immense pleasure to express gratitude and regards to all those people who supported me during the course of this doctoral research work at MNIT Jaipur. I acknowledge the involvement and contribution of each one of them.

I deeply owe a debt of gratitude to my supervisor Dr. Rajesh Kumar for all the support and motivation that he showered upon me throughout the doctoral program. He kept the faith in me and always backed my research proposal and ideas. I attribute the successful realization of this research work to his sincere and consistent guidance, versatile and unconventional thinking and his penchant for trying something new and taking risks.

I am deeply indebted to Dr. Rajesh Kavasseri, Associate Professor at the Department of Electrical and Computer Engineering, North Dakota State University, USA for having agreed to jointly supervise my thesis work. I thank him for his invaluable guidance, critical comments and for investing his precious time for my work. His kind cooperation and suggestions throughout the work guided me with an impetus to work and made its completion possible. His modesty and simplicity in thinking provided me with new insights into conducting research.

I express my sincere gratitude and respect to Prof. Jyotirmay Mathur, (DPGC Convenor, Centre for Energy and Environment, Malaviya National Institute of Technology, Jaipur) for engaging in rounds of beneficial discussion and letting me avail all the facilities to pursue this work. His critical questions and constructive arguments helped in developing my research acumen. I am also grateful to the members of my DREC: Dr. Sanjay Mathur, Head of the Centre for Energy and Environment at our institute and Dr. Sandeep Shrivastava for their constant support and encouragement.

This thesis work was carried under the National Renewable Energy Fellowship scheme. I acknowledge the financial support provided by the Ministry of New and Renewable Energy, Government of India, under this flagship scheme.

I convey special regards to Dr. Kapil Pareek, Dr. Vivekanand Vivekanand, Dr. Amartya Chowdhury and Dr. Rohit Bhakar, Assistant Professors at the Centre for Energy and Environment (CEE), for their ever-present inspiration and encouragement.

I also wish to record my gratitude to the team of office staff at the CEE, led by Mrs. Sunita Sharma, for being always willing to provide logistics-related help.

This thesis work involved interactions and deliberations over various aspects to assess the relevance of the research problem and practical applicability of the proposed solution. Therefore, I wish to acknowledge the help provided and support extended by various experts whom I consulted and approached. My deepest gratitude to Mr. Fransisco José Pazos from Iberdrola, Spain for sharing important field-test results from real islanding events. Special thanks to Mr. Gabriel Mafra from Petrobras, Brazil for rectifying my Simulink model of the IEEE feeder. My heartfelt gratitude to Prof. Srete Nikolovski from the University of Osijek, Croatia who constantly provided me relevant material on islanding and involved me in solving a real islanding problem. Special thanks to Dr. R.K. Purohit, Jodhpur for all the constructive criticism and discussions.

Mr. Alekhya Datta and Mr. Arun Joshy from the Smart Controllers Laboratory at The Energy and Resources Institute, New Delhi deserve a special mention for extending support to access their experimental facilities that provided me with an important breakthrough. My sincere gratitude to Dr. Deepak Fulwani and Mr. Vinod Kumar from IIT Jodhpur for allowing me to access their RTDS facility and training me on the same.

Special regards to colleagues and friends at RAMAN Lab, MNIT namely Chandra Prakash sir, Sujil sir, Vaiju sir, Venu, Shalini Ma'am, Srikanth, Lokesh, Akanksha, Om Ji sir, Harish, Pulkit and Vishu for their constant support, warmth and affection. Thanks to Mr. Rahul Singhal for helping me in embedded hardware. Thanks to all members of the robotics group ZINE for being a source of inspiration. Special thanks to Prateek, Umesh Ji, Gajendra, Mayand and Amit for always being there. Deepest regards to Ajay, Priyam, Falti, Boola, Pulak, Manaswi and Namrata for having taken care of me and having cheered me up always. Thanks to all the M.Tech. scholars and my friends in the CEE, and the Departments of Electrical, Computer Science and Electronics and Communication Engineering for their encouraging words.

As regards to my family, I express my deepest gratitude (inexpressible through mere words) to my father, mother, sister and grandmother for having encouraged me and having kept me updated on latest in smart grid technologies. Special regards to my uncles Mr. Durgadutt Bohra, Mr. Surendra Vyas and Mr. Narottam Vyas for making my stay at Jaipur enjoyable.

Contents

Abstract

High penetration of solar Photovoltaic (PV) generation on distribution feeders is anticipated in view of current policy changes favouring renewable energy adoption in the country, and elsewhere across the world. The rising penetration of PV on distribution networks will increase the possibility of power balance between the PV inverter and the loads, a fact well acknowledged in international studies. The load behavior is also changing from the traditional constant power model and hence the interaction of static inverters and different types of load models can produce an altogether new class of disturbances. These interactions, in the presence of grid-side disturbances, may lead to anomalous situations that can impact the integrity of the distribution network. Unintentional islanding of a feeder section containing both solar PV inverter and loads is deemed to be an immediate consequence of such events that can trigger the protective devices at the Point of Common Coupling (PCC). To protect such active distribution networks, it will be important and essential to investigate and analyze such situations, deemed to cause unintentional islanding. This framework is in contrast to the existing anti-islanding measures present inside commercial grid-connected PV inverters that 'react' once an island is formed. The proposed framework is built on the pillars of continuous system monitoring (via available methods), data collection and use of machine learning in real-time to detect any anomalies that can accidentally island a feeder section. The collected data consists of instances from well-known power system transients like short-circuit faults, switching surges, load switching etc. and from recently discovered situations that have been proved to lead to anomalously high currents that can island a feeder or its section. The strategy is to enable the PV inverter to learn how to differentiate these from other power system transients so as to pre-emptively detect and avoid formation of an unintentional island or activate timely change in operation mode and operate an island with acceptable power quality. Use of online classification algorithms executed in portable, dedicated computing platforms fit inside the inverter have reported prediction (pre-detection) accuracies and speeds that show promising applications in small-scale distributed generation systems like rooftop net-metering on residential or institutional establishments.

List of Tables

List of Figures

List of Abbreviations

ADC Analog to Digital Converter

AFD Active Frequency Drift

AI Artificial Intelligence

ANN Artificial Neural Network

BFS Backward Feature Selection

BMU Best Matching Unit

CART Classification and Regression Trees

CI Computational Intelligence

CIGRE Conseil International des Grands Réseaux Électriques

CSI Current Source Inverter

DC Direct Current

DFT Discrete Fourier Transform

DG Distributed Generator

DOCR Directional Over Current Relay

DSP Digital Signal Processor

DT Decision Tree

DWT Discrete Wavelet Transform

EMTDC Electro Magnetic Transients including DC

FDD Fault Detection and Diagnosis

FFS Forward Feature Selection

FFT Fast Fourier Transform

FNN Fuzzy Neural Network

GA Genetic Algorithm

GIS Geographical Information Systems

GPIO General Purpose Input Output

HC Harmonic Content

ID3 Iterative Dichotomizer-3

IEC International Electrotechnical Commission

IEEE Institution of Electrical and Electronics Engineers

INR Indian National Rupee

KLD Kullback-Leibler Divergence

KLI Kullback-Leibler Information

KNN k-Nearest Neighbors

LAN Local Area Network

LDA Linear Discriminant Analysis

LLLG Line to Line to Line to Ground

LVRT Low Voltage Ride Through

MLP Multi Layer Perceptron

MPPT Maximum Power Point Transfer

NB Naive Bayesian

NDZ Non-Detection Zone

P-R Proportional-Resonant

PC Principal Component

PCA Principal Component Analysis

PCC Point of Common Coupling

PDF Probability Density Function

PLL Phase Locked Loop

PMU Phasor Measurement Unit

PNN Probabilistic Neural Network

PSCAD Power System Computer Aided Designing

PV Photovoltaic

PWM Pulse Width Modulation

RAM Random Access Memory

RBF Radial Basis Function

RFC Random Forest Classifier

RoCoF Rate of Change of Frequency

RoCoP Rate of Change of Power

RTDS Real Time Digital Simulator

SAS Statistical Analysis System

SEL Schweitzer Engineering Laboratories

SOM Self Organizing Map

SPM Salford Predictive Modeler

SPSS Statistical Package for the Social Sciences

SQL Structured Query Language

STC Standard Test Conditions

SVD Singular Value Decomposition

SVM Support Vector Machine

T & D Transmission and Distribution

THD Total Harmonic Distribution

UCL Upper Control Limit

UDP User Datagram Protocol

UL Underwriters Laboratory

V2G Vehicle to Grid

VRT Voltage Regulating Transformer

VSI Voltage Source Inverter

WEKA Waikato Environment for Knowledge Analysis

List of Symbols

R Resistance in Ohms
L Inductance in Henry
C Capacitance in Micro-Farad
P Active Power in kilo Watts (kW)
Q Reactive Power in KvAr
ω Frequency in radians per second
Q A statistical measure of PCA
Q_f Quality factor
μ Statistical mean
σ^2 Variance
κ Cohen's Kapppa statistic
χ^2 Statistical chi-square probability distribution

Chapter 1

Introduction

1.1 The Distributed Generation Paradigm

The electrical power system is traditionally built on the model of centralized-generation, transmission and distribution which has been in existence for over a century now with inherent operational inefficiencies. However, the incorporation of renewable energy based distributed sources has brought a paradigm shift towards the concept of localized-generation that reduces losses and saves additional generation capacity. The introduction of such generators downstream of a distribution network also brings dynamic changes and modifies the configuration. The difference between such active networks and the traditional ones and the challenges and issues in operating them have been discussed in the following sub-sections respectively.

1.1.1 Traditional and Active Power Distribution Networks

Electricity is a critical public utility besides water and gas that needs to be supplied to the end users through an efficient distribution network. The electricity distribution infrastructure consists of a systematized arrangement of conductors and cables which can be laid in an overhead or an underground configuration. Generally, for reasons of economy and ease in operation and maintenance, the overhead system is preferred. However, in areas where erection of poles for overhead conductor mounting is infeasible or restricted by right of way issues, the underground system is adopted wherein cables are routed through conduits and

trenches under the surface of streets and sidewalks. Such a configuration has its own advantages in the form of long-term reliability, public safety and aesthetics.

A distribution system represents the consumer-side terminating portion of a large and complex network called the power grid that segments the power system into its generation, transmission and distribution levels. It is the electrical sub-system between the transmission substation and consumers' meters. Physically, an overhead distribution network consists of a hierarchical topology of three types of conductors: feeders, distributors and service mains. The radial system is the most common topology which is employed for power generation at low voltage-level and when the substation is sited near the load centre. Due to its simplicity, confined reach and economical operation, it is popular in rural areas. Figure 1.1 shows a single line diagram of a radial distribution system. Three feeders can be observed to be radiating from the substation. A feeder is the main conductor that connects the substation to the distribution area and from which no tappings are taken since it is designed for continuous current carrying capacity. In the figure 1.1, SL is a feeder that connects to the distributor IB at node I. The tappings for supplying loads are taken at different points along IB. A distributor must maintain the voltage profile along its length since point-wise tappings lead to a variable current passing through it. It must be ensured that the voltage variations are within the nominal limit of $\pm 5\%$ of the value rated to be at the consumers' terminals [1]. Service mains are the cables that connect the consumer premises to the distributor tappings. A ring-main system is a more common topology in urban areas that have a dense conglomeration of different types of customers. If multiple distributors are introduced in figure 1.1 and the primaries of their distribution transformers are connected to form a loop such that it closes on the substation, then a topology as shown in figure 1.2 is obtained. The closed feeder SLNAOPQR has several sections that can operate as independent feeders in case a fault occurs or maintenance and repair is being done on any part of the ring. Therefore, such a topology ensures reliability of supply to consumers. In addition to this, the voltage fluctuations at the customers' terminals are less than those experienced in a radial system. If more than two generating stations or substations are added to the feeder ring, an interconnected system is formed which is generally considered as the third topology of distribution networks. The interconnected system topology adds a dimension of efficiency through pooling of generating resources that facilitate load sharing during peak hours and reduce reserve power capacity.

The traditional distribution systems are a part of the conventional electricity supply model that follows the centralized hierarchy: a large plant generates power

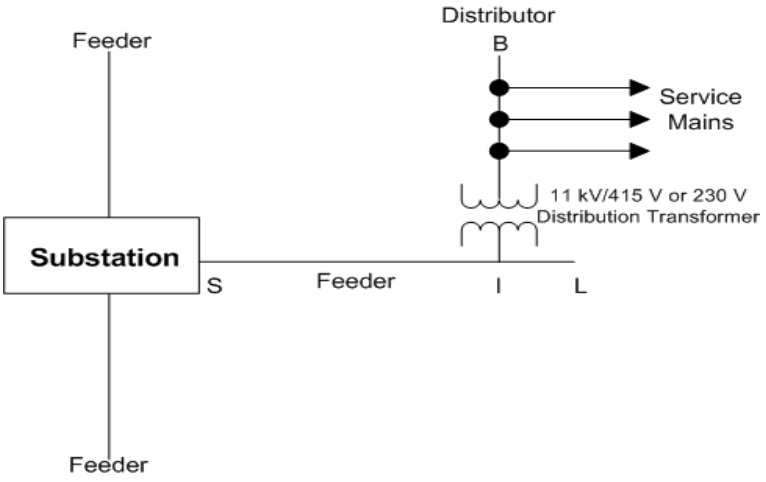

Figure 1.1: Single line diagram of a radial distribution system

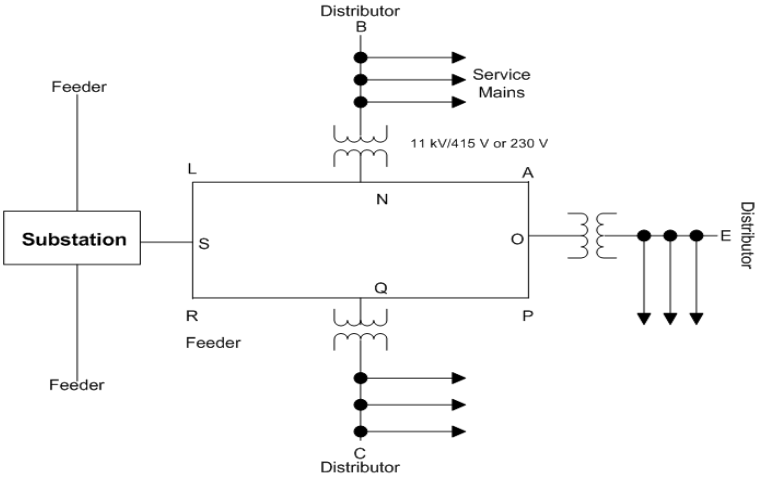

Figure 1.2: Single line diagram of a ring-main distribution system

which is transmitted over high voltages and then distributed to the consumers at voltage levels determined by the end-use type i.e. residential, commercial or industrial. These systems experience the highest burden among the three power system components and work under high operational stress. They also suffer from inefficiencies in power delivery and consumption. Moreover, the generating sources consume fossil-fuels for the thermodynamic energy-conversion process to produce an alternating current. Due to inherent conductor-resistance and thermal limits of power transfer, the centralized delivery model largely remains in-efficient because of huge losses in the transmission and distribution lines. Each unit of electrical energy lost in the form of unrecoverable heat during transmission or due to inefficient consumption at the utilization end has a quantum of carbon emission released in the atmosphere. Such attributes have made the power industry a major source of environmental imbalance amidst rising concerns of global warming. The

de-carbonization of the power system was thus envisaged alongside improvements in the delivery and consumption efficiency of the traditional power grid. Radical changes were then conceptualized and implemented with strong focus on optimizing the distribution system operation. The 'smart grid' paradigm thus evolved and most of its technologies were targeted for the distribution network that represents the back-bone of the power supply system. One element was identified to be an integral part supporting the idea of a smart grid: incorporation of generating sources at the consumption-end with the ability to operate in isolation from the main-grid during contingencies and extreme events. This, however, came as a major challenge since the traditional distribution feeders are mostly radial in nature and integration of a generator in the service stream would transform them from passive to active networks. Thus the centralized model, in existence since the later half of the 19^{th} century was being gradually replaced by a distributed framework. The transformation has slowly picked up pace but the introduction of distributed generation facilitated some disruptive innovations to happen at the distribution level: customer-side generation or the idea of 'prosumers' (producing-consumers), Vehicle to Grid (V2G), distributed storage and many more.

The concept of electricity generators operating in parallel with the electrical distribution network has now been in practice since the last few years. The integration of such generators, distributed in nature, space and availability has become an essential feature of the smart grid framework. Generation sources based on renewable energy, due to their green-attributes, have given impetus to the same. Technologies like solar PV inverters, wind-electric generators, small hydro turbines, microturbines, biomass generators etc. are clean sources of electrical power whose output can be synchronized with the utility grid. However the addition of these sources brings some operational complexities [2] which need to be addressed and planned about. However, the scheme has gained commercial weight whose popularity can be adjudged by the various names that have found use in practice:

1. Distributed Generation
2. On-site Generation
3. Dispersed Generation
4. Embedded Generation
5. Distributed Energy Resources.

On a comparative front, distributed generation offers the following advantages over conventional centralized generation:

1. Generation near the site of consumption; no additional transmission infrastructure required

2. Cleaner sources like renewable energy-based generators can be used more effectively with added value to their generation
3. Around 20% of Transmission and Distribution (T & D) losses, as found in the centralized system, can be reportedly avoided
4. Design can be customized to be made reliable by becoming compatible to local requirements
5. Multiple modes of generation like co-generation, tri-generation that provide more value for energy conversion and utilization can be used
6. Can be used to power load independent of the grid; micro-grid configuration is possible.

The major reason for the growth and adoption of distributed-generation has been the inherent losses prevalent in the existing supply model. The centralized generation based power delivery model encounters power dissipation in conductors over long distances, voltage drop, extreme weather events and other disturbances that affect its reliability. On the other hand, distributed-generation located close to the points of consumption and integrated with the power grid supports power flow in the network besides enhancing the reliability of supply to consumers in the event of main-grid failure, brown-out, black-out or shut-down. Rising interest in distributed-generation can also be attributed to other important reasons like:
1. Deregulation of power utilities
2. Emerging power markets that involve customer participation
3. Provision of ancillary services by distributed generators like reactive power support and stand-by capacity etc.
4. Savings in investment of additional generation , transmission and distribution capacity by utilities.

Since the power grid was not designed for the inclusion of sources in its distribution-pathway, the integration of distributed energy resources is a technically complex issue. When the sources are renewable-energy based, their stochastic nature and the variability in availability of generation induces extra randomness in the system operation. Hence such active distribution systems face a lot of issues that have been covered in detail in the subsequent section. The aim of this section is to build up on these issues and take forward the issue of unintentional islanding in solar PV systems connected with the distribution grid which is the main focus of this research. Figure 1.3 shows the conventional form of power delivery based on large, fossil-fuel powered centralized generators. The different levels of the system are shown with the direction of power flows. The distribution system is radial in nature, meaning that power can only flow from the substation towards the loads.

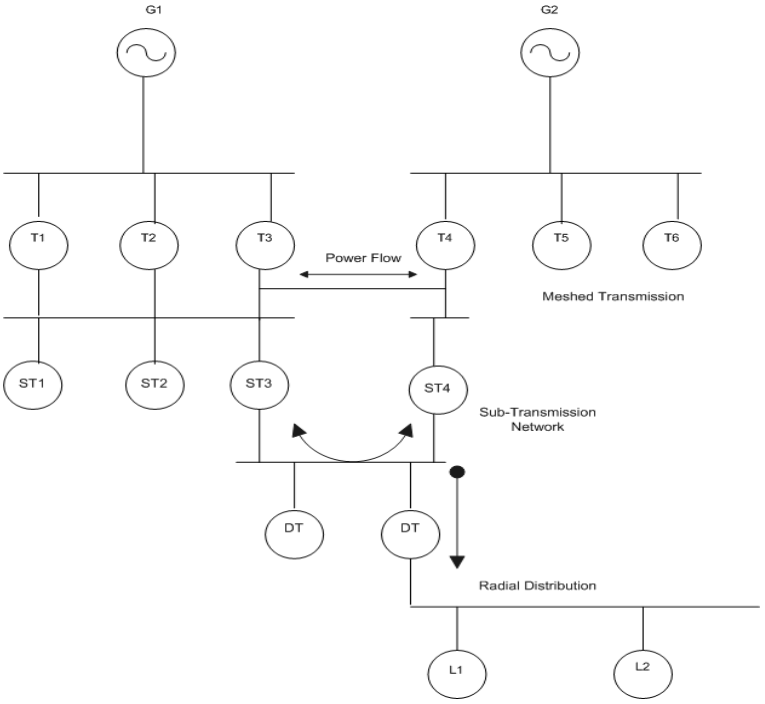

Figure 1.3: Conventional power delivery model

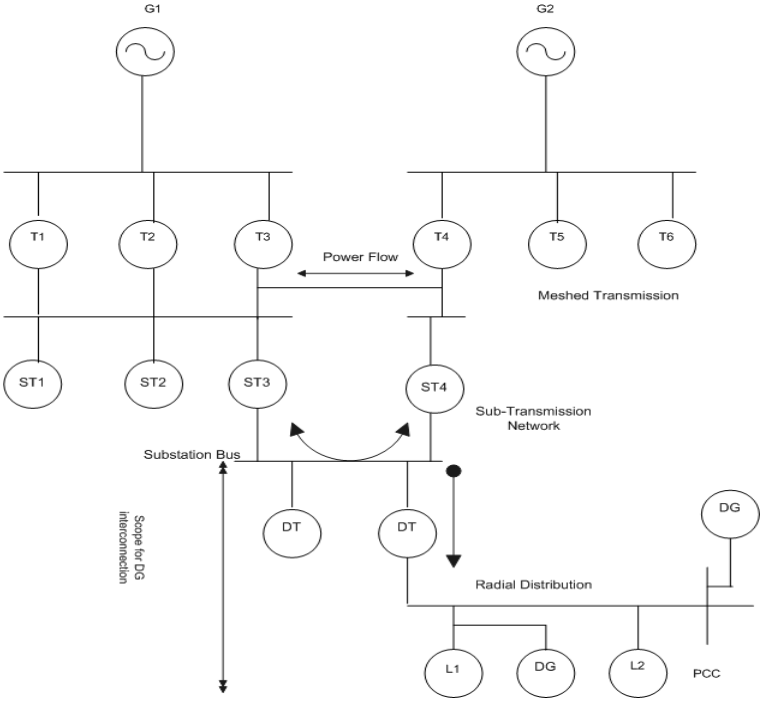

Figure 1.4: Active distribution network due to DG interconnection

Figure 1.4 shows the topology of the same network, as above, when a Distributed Generator (DG) is connected at a node on the distribution feeder which is called the Point of Common Coupling (PCC). Another DG is directly connected to a load bus on the feeder. There is now a possibility of power flow in reverse direction

due to the DG output exceeding load demand. Such a situation is one of the consequences of DG interconnection.

1.1.2 Challenges and Issues with DG Integration

A smart grid can accommodate rising distributed generation. Most of the integration takes place at the distribution and utilization level, as discussed previously. However there are a lot of challenges and issues associated with connection of distributed resources on the grid. The biggest challenge is to connect the DGs to the existing distribution system without making major changes or modifications to the intrinsic or inherent nature of the distribution network.

The prevalent issues can be categorized as:-
1. Technical Issues
a. Power Quality Issues
b. Utility Interconnection Issues
c. Protection Issues
2. Business and Financial Issues.

Technical Issues:-

a. Power Quality Issues:

DG interconnection causes problems like voltage rise, sag, flicker and harmonics mainly due to the nature of sources (like solar PV), the power converter interface or due to improper synchronization with the utility grid. Figure 1.5 displays voltage disturbances in a section of a small feeder that are typically caused due to integration of a PV system. Such random fluctuations in voltage waveforms affect the quality of power being delivered to the connected loads and put stress on the reserve units for grid-support to provide quick inertial response.

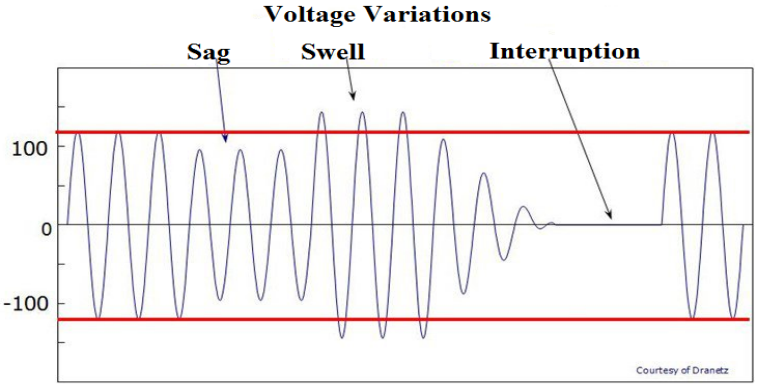

Figure 1.5: Typical deviations in voltage due to a solar PV array

b. Utility Interconnection Issues:
The distributed generator that wants to connect to the utility grid must be synchronized in phase, frequency and voltage with the grid. Different types of DGs have different methods of interconnecting with the grid. Unsuccessful interconnection has implications on power flow in the line including circulating currents and power quality reduction. To maintain the discipline of the grid, many standards exist worldwide for utility interconnection of distributed generators. An example is the IEEE 1547 Standard for Interconnecting Distributed Resources with Electric Power Systems, 2004 [3].

c. Protection Issues:
Since the DG is a source in addition to the centralized generator(s) on the grid, its integration is expected to disturb the protection coordination of the utility system. The DG infeed adds to the fault current level however the direction of this contribution is such that the grid-side I_{fault} is reduced causing a PCC over-current relay to under-reach. Thus a high penetration of DGs, solar PV inverters in particular, can even lead to de-sensitization of protective relays during short-circuit faults[4]. Apart from this phenomenon, DGs can cause several other protection-related issues that affect the integrity of the distribution system. Some of the common protection issues can be noted as:

1. Over current protection sensitivity
2. Instantaneous Reclosing
3. Unintentional Islanding
4. Ferroresonance
5. Grounding Issues

Business and Financial Issues:
Schemes like net-metering allow customers to install DGs like solar PV inverters on their rooftops and sell the excess power back to the utility grid. There is a fear among utilities that rising DG penetration will reduce their revenues. Solar PV generated electricity reduces the amount of money utilities need to pay for fuel to generate power, but those savings are offset by "revenue erosion" from lower sales and from deferring investment in poles, wires, and other infrastructure, found a study by the Lawrence Berkeley National Laboratory [5]. According to the study's results, if distributed solar PV adoption rose to the equivalent of 2.5% of utility retail sales in USA, it would cut shareholder earnings by 4%. However, the impact on electricity rates for consumers would be minimal - only 0.1 or 0.2% increase. These figures are reported from a set of results generated for different penetration

scenarios for two prototypical utilities in the USA. Despite reported resistance to DG integration from distribution utilities worldwide, the penetration level has been increasing, and rapidly in developing economies. The long-term impacts of DG integration will culminate into monetary benefits over the years for the utility owners.

Key Points:

1. Distributed generation is an integral aspect of smart grid framework
2. The integration of DGs with the grid introduces operational aspects
3. Renewable energy based generators like solar PV induce extra problems due to stochastic generation
4. Proper planning is required to integrate DGs at the distribution level and mitigate operational threats
5. Unintentional islanding is identified as a protection issue which will be the focus of the next section.

1.2 Research Context

The general description of advantages and impacts of grid-connected distributed generation systems was provided in the previous section. This section focuses on the particular case of solar PV interconnection with the distribution network. The concept of islanding, the associated problems and the scope of the reported research work (including a brief note on the relevance of research in this domain) have been discussed in this section.

1.2.1 Grid-Integration of Solar PV

Solar energy is a clean and inexhaustible source that can be transformed into electrical energy by PV cells stacked into modules and arrays, with some efficiency limits. Advancements in inverter technology have ramped up the pace of grid-interfaced solar PV systems and more number of installations are taking place at the distribution level. However, there is still a lot to explore regarding how such inverter-interfaced DG technology can affect the centralized grid in different scenarios. Solar PV generators can supply power to the electric grid through a grid-interactive or grid-connected solar PV system. The connection of solar PV can take place at two levels in the power system: Transmission level and distribution

level. Large-size PV systems of multi-MW capacity feed three-phase power to the high-voltage transmission grid and are known as utility-scale PV power plants. On the other hand, kW or sub-MW scale PV systems that usually feed one-phase power to the distribution network are popularly called as distributed generation systems.

Penetration level is an important term associated with the amount of PV based generation feeding the distribution network. From the distribution system point of view, PV penetration level is defined as : $\frac{\text{PV Capacity}}{\text{Peak load of line section or feeder}}$. This is the most commonly used definition for the penetration level. In other definitions, minimum system load or transformer/ station rating become the denominator respectively. At low penetration levels, PV does not affect the system much apart from small issues like sudden ramp-up of generation from scheduled generators [6] or a dynamic battery response due to variable solar insolation. At high penetration scenarios, many issues are expected to arise from PV integration, namely as noted in [6]:

1. Over voltages at the customer node depending on the PV location; if PV is located away from the substation, the voltage profile will increase towards the load end thus causing voltage regulation problems
2. Possibility of reverse power flow
3. Distribution system losses may increase as penetration increases above 5%
4. More harmonics, voltage disturbances and flicker
5. Increased chances of unintentional islanding.

Since a PV array is a current-source, most of the grid-connected inverters are Current Source Inverters (CSIs) or current-controlled Voltage Source Inverters (VSIs) that employ pulse-width modulation (PWM) based switching to fabricate a sinusoidal current wave. As grid synchronization requires PV system output to match the utility phase and frequency, the inverters need a continuous reference to achieve the same. Figure 1.6 shows a schematic block diagram of the basic architecture of a grid-connected solar PV system. The early CSIs were line-commutated which means they used the utility line voltage as a reference trigger for commutation of switches in the power electronic bridge to create a sine wave output that matched the utility requirements. The latest CSIs and VSIs are self-commutated and can sustain operation based on their internal reference to trigger the switching accordingly. Consequently they can also operate in absence of the utility voltage until the inbuilt protection trips them off-line. This ability to 'island' is not unique to self-commutated CSIs as per popular perception [7] but this phenomenon of operating independently of the grid without changing the operating mode from

constant power control to V/f regulation mode [8] has many implications as discussed in next sub-section. Figure 1.7 shows the fundamental control structure of grid-connected PV inverters. The current-controlled or constant PQ control strategy is generally employed in the grid-connected mode of operation. There are two loops: the outer voltage control loop and inner current control loop. Proportional Integral (PI) controllers are most commonly used for error-compensation however the use of a Proportional-Resonant (PR) controller for the current loop has also been proposed to reduce harmonics [9]. Since the DC link voltage has to be maintained constant at a value $> V_{\text{grid RMS}}$, the outer loop regulates the same by taking this value as the reference $V_{\text{ref. DC link}}$. The P-I error compensator of the outer loop outputs the magnitude of the grid-reference current which is multiplied by the phase information extracted by the Phase-Locked Loop (PLL) resulting into $I_{\text{AC ref.}}$. The inner loop controls the output current to be aligned with the phase of the utility voltage and this exhibits the constant power control philosophy also known as constant PQ control. In the grid-connected operational mode, the inverter is not allowed to regulate the voltage or frequency of the utility grid [10], unlike in the V/f regulatory mode. Most of the interconnection regulations require PV inverters to operate at unity power factor and not to contribute any support during voltage and frequency disturbances. Hence they cannot ride-through any small disturbance and have to trip even during such transients leading

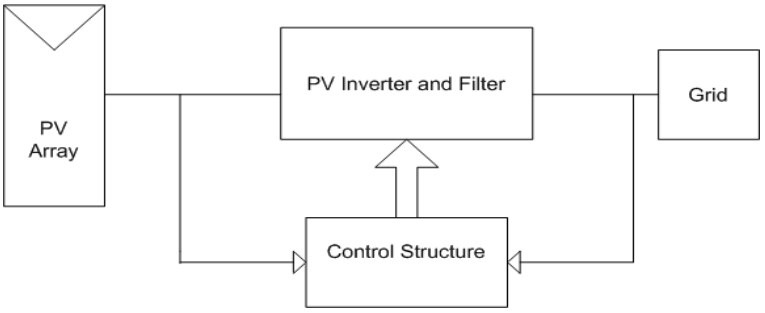

Figure 1.6: Block diagram of a grid-connected solar PV system

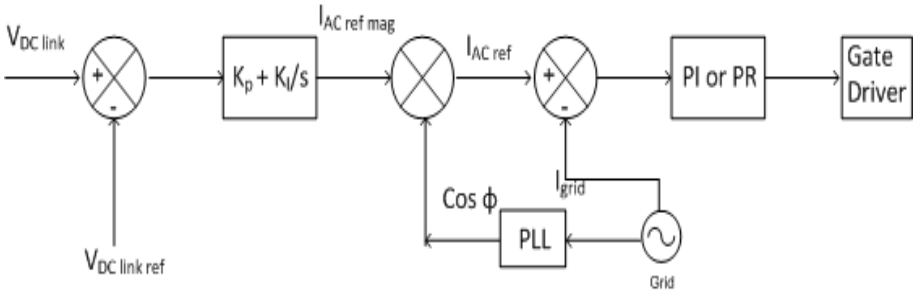

Figure 1.7: Typical control structure of a grid-connected solar PV inverter

to wastage of generation capacity. Rising penetration will lead to inverters going off-line and coming back repeatedly which may trigger a sustained disturbance on the line-section leading to a kind of 'yo-yo' effect that was observed in Germany [11].

It is understood that distributed solar PV connected to the central grid will be a key component of the smart distribution network for reasons stated previously. It is also acknowledged that higher penetration levels may lead to operational problems that need to be addressed. One of the problems, identified as unintentional islanding will be the focus of this research and more details related to it are provided in the next sub-section.

1.2.2 Islanding in Grid-connected PV systems

Since a grid-connected PV system operates in parallel with the synchronously running power network, the coherent operation of both the systems is absolutely necessary for maintaining power quality and reliability. The vulnerability of the existing power grid to frequent outages and interruptions gets magnified at the distribution level because it is the most stressed out portion of the network. Interconnecting a solar PV system thus becomes challenging but careful operational planning has ensured no hassles so far in operation of such systems world-wide. However the continued feeding of loads in vicinity of the PV inverter when the main supply suddenly goes off is a situation that must be avoided.

Most of the DGs including PV inverters operate in constant PQ control mode or constant power (active and reactive) control mode. This means they are commanded to give a fixed output power in synchronism with the grid by providing them fixed power set-points. Thus the operational strategy is to control the output current of the inverter such that it is in phase with the utility voltage. The PV inverter is generally operated at unity power factor because this is a strategy of maximizing the energy yield from the array through Maximum Power Point Tracking (MPPT). The inverter thus cannot adjust its active or reactive power output according to the system frequency and voltage since it is programmed to only inject maximum active power in accordance with the utility requirements. However if the utility grid suddenly disappears, I_{grid} is lost but the error signal is still passed onto the inner current control loop. This signal is further compensated and the output activates the gate-driver that provides the triggering pulses for the switches. However, the fabricated output wave is not sinusoidal but the inverter

continues operating. A sudden disappearance of the mains and continued feeding of loads by the inverter will thus lead to an isolated network continuing to operate in the same grid-connected mode which is not required. This system configuration in this condition is analogous to an island where a section of live conductors exist in a large network having no power from the main source.

Islanding is a phenomenon in which a power network gets separated into controllable/uncontrollable sections containing both sources and loads and operating independently of the main power grid. For grid-mode distributed generation systems, islanding is said to occur when the DG continues supplying power to the distribution grid even after a portion of the network, that includes the PCC, gets disconnected from the main power system. This is clear from a single line diagram representation shown in figure 1.8. The inverter will continue to operate and feed the loads until the voltage and frequency in the island cross the inverter operating limits and the inbuilt under/over voltage/frequency relays operate. The V_{island} and f_{island} are controlled by $\triangle P$ and $\triangle Q$ between the islanded loads and the inverter respectively. The Non-Detection Zone in which the inverter cannot detect the islanding condition is shown in figure 1.9. For any values of V_{island} and f_{island} at the inverter terminals, the island continues to operate until the degree of $\triangle P$ and $\triangle Q$ increase to a level severe enough to trigger the relays. Therefore, the NDZ represents the region where the load - generation balance is such that a PV inverter employing only voltage and frequency based detection can continue to operate in the sudden absence of the main utility.

The dynamics of an islanded system are governed by the level of $\triangle P$ and $\triangle Q$ between the islanded loads and the inverter, as discussed above. The behaviour of the isolated system will depend on the values of $\triangle P$ and $\triangle Q$ just at the instant of opening of the PCC breaker, recloser or lateral fuse to form the island [12]. In

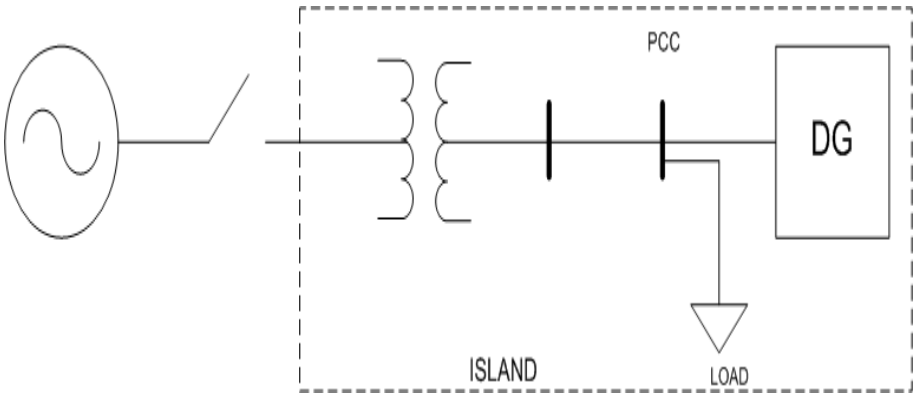

Figure 1.8: Creation of a Power Island

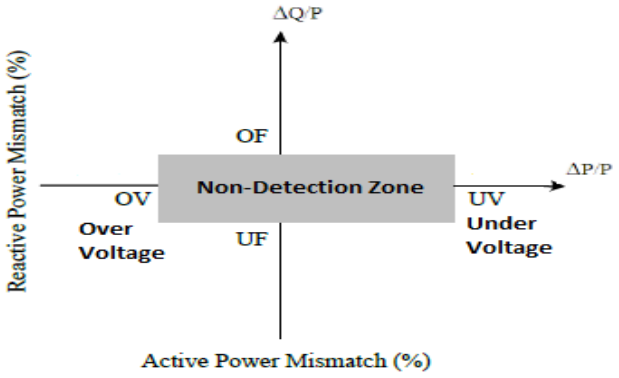

Figure 1.9: Non-detection zone

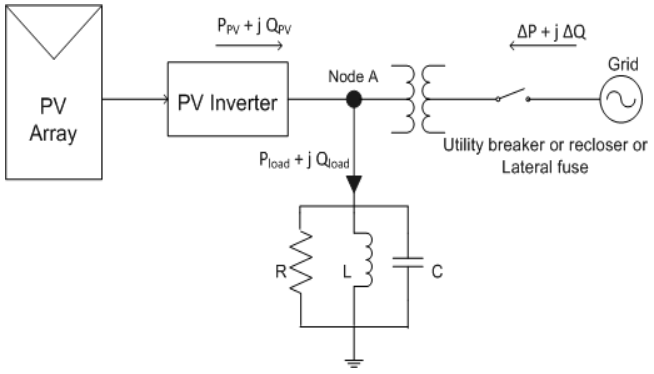

Figure 1.10: Power flows in a grid-connected PV system

order to understand the correspondence between V_{island} and $\triangle P$ and f_{island} and $\triangle Q$, it is important to understand the expressions for P_{load} and Q_{load}. Consider a simple configuration of a solar PV inverter connected to a feeder at node A which is the PCC as shown in figure 1.10. The power flows have been defined with their directions indicated. The load is single phase and has been taken to be of the parallel RLC type which is the most common configuration adopted in many test standards for islanding detection. Accordingly, the expressions for P_{load} and Q_{load} can be written as:

$$P_{load} = V_A \times \frac{V_A}{R} \tag{1.1}$$

$$Q_{load} = V_A \left[\frac{V_A}{\omega L} - \frac{V_A}{\frac{1}{\omega L}} \right] \tag{1.2}$$

It is visible from equations 1.1 and 1.2 that voltage and frequency at the inverter terminals have a relation with the active power and reactive power respectively. When the switch shown in figure 1.10 opens, $\triangle P = 0$ and $\triangle Q = 0$. The values of the resultant V_{island} and f_{island} are determined by the values of $\triangle P$ and $\triangle Q$ just before the opening of the switch. The power balance for this circuit is expressed

as:

$$P_{load} = P_{PV} + \triangle P \tag{1.3}$$

$$Q_{load} = Q_{PV} + \triangle Q \tag{1.4}$$

The equations 1.3 and 1.4 show that the LHS values are controlled by the PV power outputs once an island is formed. For example, if $\triangle P_{\text{before island}} > 0$, then $P_{load} > P_{PV}$ and when the switch opens, P_{load} has to decrease since $\triangle P = 0$. Therefore the $V_{island} = V_A$ decreases in relation with the $\triangle P$. If the value of V_{island} decreases to be $< V_{\text{inv. lower limit}}$, then the under-voltage relay inside the inverter will trip and the island will collapse after it is detected. Similarly, for the case of $\triangle Q_{\text{before island}} > 0$, there will be an impact on f_{island}. As $Q_{PV} = 0$ for most grid-connection schemes, Q_{load} will become zero following equation 1.4. Hence the term inside the square brackets in equation 1.2 will approach zero for which the inductive part must fall and the capacitive part must increase. Consequently, the ω will rise and thus the f_{island} in the resulting network will increase. If the value of f_{island} rises to a value $> f_{\text{inv. upper limit}}$, then the over-frequency relay of the inverter will trip and detect the island formation. The concept of under/over voltage/frequency and its relation with $\triangle P$ and $\triangle Q$ can now be well understood with the NDZ shown in figure 1.9.

Based on the reasons for occurrence, islanding can be categorized into - intentional or unintentional. Accidental isolation of a portion having PV and loads from the main grid due to reasons like outages/faults that cause disruption in upstream supply is called unintentional islanding. Intentional islanding is adopted as a part of the utility strategy to split the grid into controllable sections for supporting peak demand and/or mitigating network congestion.

Islanding, if caused by an unplanned event, is harmful to the power system. The problems due to unintentional islanding can be enumerated as :
1. The frequency and voltage of the island remain unregulated as it operates as an autonomous entity and hence degrades the power quality. This can upset the load or devices connected to the island.
2. It disturbs the co-ordination of the utility protection devices as the automatic reclosers can complete the broken circuit even during power flow in the branch/island.
3. Out-of-phase reclosure can cause circulating currents to flow when grid supply resumes as both the grid and island may be out of synchronism after reconnection, giving rise to large electro-mechanical forces that can damage the connected

Distributed Generators (DGs)
4. There remains a potential safety hazard to the utility personnel maintaining or repairing the apparently de-energized distribution system
5. Transient over voltages that affect system reliability may build up
6. Inadequate grounding of the islanded system by the DG interconnection.
On the other hand, intentional islanding, if practiced strategically by utilities, has the potential for ensuring smoother real-time operation of the distribution network.

1.2.3 Relevance of Research in the Domain

Islanding as a concern is still discussed today. However the ill-effects of such an event remain a matter of debate among the utilities and PV system operators. In many situations the utilities refuse interconnection to their networks because of the fear of accidental islanding. Although the internationally documented experiences of such events are limited, actual records of unintentional islanding events have been reported in [13] and [14].

Important reports like [15] have published results of sustained efforts in studying the phenomenon which concludes that the possibilities of its occurrence are rare however, there is lack of clarity and consensus on the validity of results for the scenario of rising PV penetration. A recent survey in [16] has clearly brought out the thinking of the distribution system operators for whom unintentional islanding is the biggest concern among others that can result due to DG interconnection as discussed in section 1. Figure 1.11 is reproduced from that published survey and shows the response of a group of different types of utilities based on investment holding.

The inverter technology is changing swiftly and the nature of loads is also changing from linear to non-linear (with more complex static and dynamic models) thus the load-PV interaction seems an important feature for islanding-related studies. Also the risks of islanding-protection performance in the case of multiple inverters of different make and protection strategies connected to the same feeder have not been clearly outlined. Furthermore, it is more critical now to keep track of the power system occurrences at the distribution level for high PV penetrations. Thus it is felt that starting from the basic study and analysis on dynamic interactions between loads and PV inverters for different grid conditions and operating situations may provide some new outlook to the problem of unintentional islanding. It

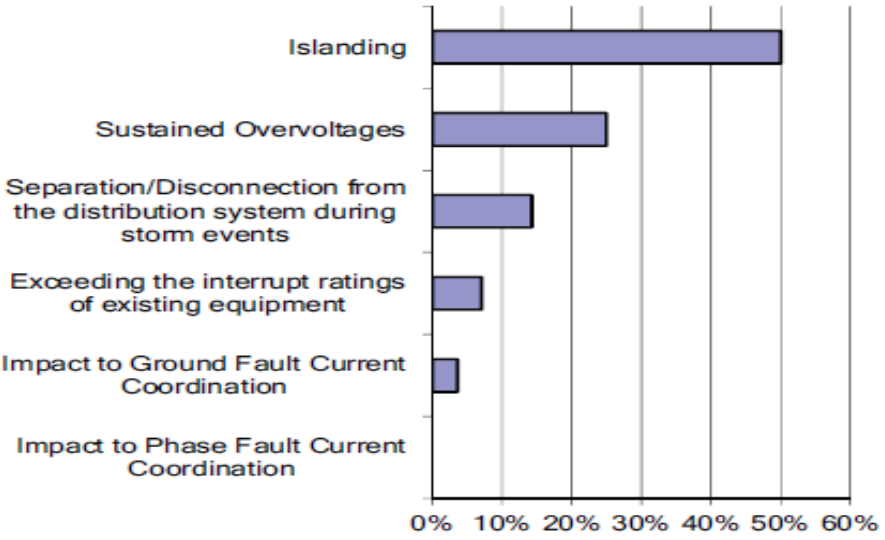

Figure 1.11: Concerns Among Distribution Utilities from DG Integration as published in [16]

is important to look at the issue of PV inverter islanding from the perspective of a new self-isolating, parallel-running energy systems paradigm. Rather than treating it solely as a protection issue, it should be tackled comprehensively given the challenging scenarios ahead. Since solar PV systems involve a lot of initial investment, the value of the energy they deliver to the grid has a premium. Therefore, adopting a reactive approach of removing the inverter as an anti-islanding measure leads to loss of valuable power, even for a few seconds, which multiplies to units of energy lost or undelivered. Thus, it is felt that the pre-hand actionable intelligence about suspicious triggers to islanding is essential. Therein lies the scope of the work that proposes a pro-active islanding-management strategy in response to unique islanding-initiators. Detecting such triggers, before the feeder protection acts, can alert PV inverters to make a decision regarding their operational mode in the lead time obtained.

The following section will outline the theme of the proposed research wherein initial attempts have been made at exploring and discovering some new situations that may cause unintentional islanding and enabling intelligence inside inverters to preemptively detect them. The research problem, the specific objectives and the problem formulation has been discussed in the portion that follows.

1.3 Research Objective

In this study, some anomalous occurrences arising out of impacts of high localized-penetration of solar PV generation on radial feeders have been investigated. One of the two discovered phenomena has been identified as a potential initiator of unintentional islanding on the modeled feeder. The aim is to detect this event and signal a mode-change trigger to the solar PV inverter before a feeder protective device isolates the section. The preemptive mode-change triggering logic is the major contribution of this work. It aims to alert the inverter to change its operational mode from constant PQ control to V/f regulatory mode so that once an island is formed, the inverter is ready to maintain and sustain the loads until load-generation imbalance occurs or the utility resumes. Elements of machine learning have been used to discover patterns from the data points belonging to the explored events and other known power system transients. Various learning techniques have been tested offline and realized in a dedicated microcomputer for online classification of signals that provides actionable intelligence to the inverter. An effective feature extraction methodology for real-time applications has also been proposed to enable online identification of such patterns. This research work is thus in contrast to the existing work being done on improving the accuracy of detecting an island formation.

The formal research objectives pertaining to this thesis work can be stated as follows :

1. To discover and create different situations and analyze various behaviours occurring in grid-connected solar PV systems that can lead to unintentional islanding and creating a bank of such precursor signatures.
2. To create a machine-learning based framework for enabling a predictive strategy for predicting the possibility of island formation in such networks based on knowledge of predicting features/signals/symptoms/signatures.
3. To create a preemptive detection based mode change triggering logic for solar PV inverters in the situation of an imminent islanding condition on a distribution network with high penetration.

The problem, keeping in view the objectives mentioned above, can be framed as a mathematical expression in the following form:

Let there be two sets Set $S_1 \in \{E_1\}$ and $S_2 \in \{E_2\}$ where E_1 = [events that cannot cause islanding] and E_2 = [discovered islanding initiator]. A set of class labels $C = \{0, 1\}$ exists such that $\forall x \in S_1, y = 0$ and $\forall x \in S_2, y = 1$ where x is

the feature extracted from event data. The problem is to design a learning system that can execute the following task:
Given a training data set $D = S_1 \cup S_2$, learn a model $G = f(x, y)$ such that for any $X \notin D$, G should return a $y \in C$ accurately and quickly. Also, for any $x \in S_2$ that comes from a different sampling rate, $y = 1$. Additionally, when only a set of features x is available, the system must group the events as belonging to either S_1 or S_2.
In this context, 'quickly' is explained by the condition $t_{\text{prediction}} < t_{\text{feeder protection}}$.

1.4 Thesis Organization

The remaining thesis is organized in the following manner. Chapter 2 presents a review of the state of the art in islanding detection based on intelligent techniques. Apart from covering various types of tools and techniques reported in the literature, the chapter also includes a section on reported efforts towards proactive islanding mitigation and then discusses broader prospects for CI techniques in tackling islanding comprehensively. It also cites some examples used in other applications to set the theme for the work carried in this research. Chapter 3 describes the exploration and investigation of the unique phenomena observed on different radial feeder models. The Real Time Digital Simulator (RTDS) and emulator-network model verification of the island-initiating event have been explained. Collection of data pertaining to the events discovered and to other power system transients simulated along with its pre-processing is explained in Chapter 4. The chapter mainly details the offline feature extraction and event detection by multivariate statistical methods and supervised learning models. Offline detection by unsupervised learning is also described. Chapter 5 describes online feature extraction and event detection. The physical realization of the mode-changing trigger is also described. Chapter 6 concludes the study. Figure 1.12 provides a flow of the thesis in the form of a block diagram highlighting the major portions covered in each chapter.

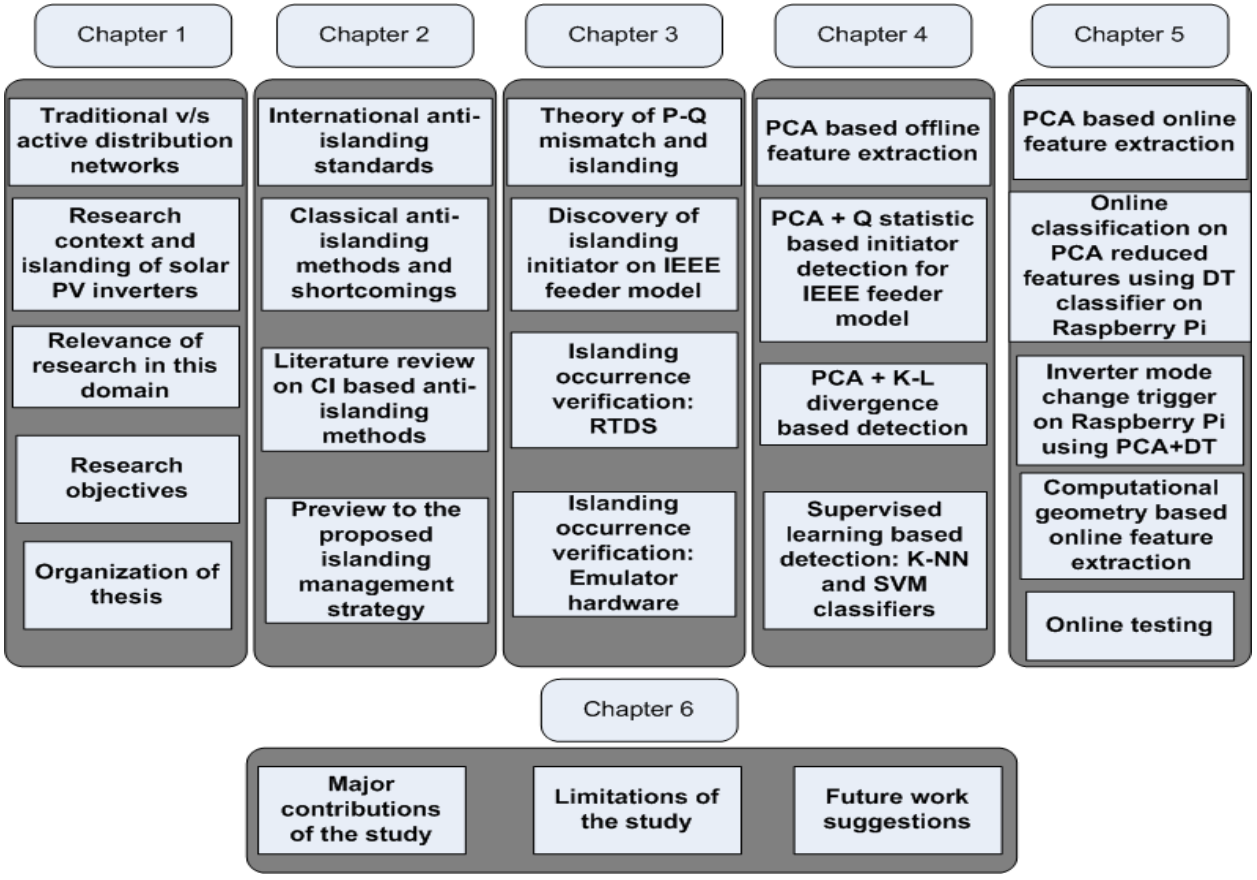

Figure 1.12: Block diagram of thesis organization

Chapter 2

Literature Review

2.1 Introduction

The fundamental concepts related to islanding of solar PV interfaced distributed generation were discussed in the previous chapter. This chapter presents a survey of the state of the art in the domain of Computational Intelligence (CI) based islanding detection and then shifts focus towards islanding-prediction that has been attempted in the literature. A summary of recent work reported in the literature has been provided through a comparative analysis of the different CI techniques used. The chapter begins with a description of the classical anti-islanding techniques and their shortcomings. For the purpose of understanding the practical significance of anti-islanding in grid-connected DG systems, a dedicated section on various international standards has been included. The section that follows provides a thorough review of the recently published works and briefly explains the underlying methods and techniques used in the cited works. This also includes a sub-section on the various software platforms used to implement the same. The limited amount of references to efforts towards predicting an islanding situation, as found in the literature, have been focused upon in the last section. Apart from detecting an islanding condition, CI based techniques have a lot more to offer for a comprehensive management of such a condition. These interesting prospects have also been discussed with real examples (in other applications) and build the theme for this research apart from highlighting the importance of the chosen problem. The following paragraphs discuss islanding, its impacts and detection methods alongside examples of actual events.

For grid-connected distributed generation systems, islanding is said to occur when

the DG continues supplying power to the loads on a section at permissible voltage and frequency [17] even after that portion of the network gets disconnected from the main power system. This can happen either accidentally (unintentional islanding) due to various reasons like outages, faults and breakdowns that interrupt the upstream supply. Recently, some new cause of unintentional islanding was reported in [18] wherein the inverter anti-islanding method itself caused a voltage imbalance that initiated the isolation of the PCC section. Islanding can be intentional if it is executed as a part of the utility strategy to split the network into controllable sections for supporting peak demand and for managing congestion. Although internationally documented experiences of such events are limited [13], actual records of unintentional islanding events found in some works have highlighted the profound impacts on the network and the loads, as observed in the field. Intentional islanding, on the contrary, can ensure a seamless operation of the distribution network in real-time contingencies. Non-availability of DG power to the loads during islanding situations carries a financial loss. Although cost of power generated by many DGs is approaching grid-parity , the loss of available power even for a few seconds leads to value-reduction. In the UK, the latest feed-in tariff is 4.32 pounds/kWh for a 10 kW solar PV system. This example proves that the operational cost involved with such DGs is high relative to centralized thermal generators and hence their unintentional islanded operation involves huge financial implications due to misuse of valuable power. Thus efficient and cost-effective islanding-management techniques have become a necessity.

Accidental islanding disrupts the integrity of the power system as discussed in the previous chapter. Unregulated V_{island} and f_{island} were observed in a real islanding event on a 60 kV network in Portugal [19]. The island was formed due to a phase to ground fault and the voltage supplied to the consumers was below the level prescribed by the European standard EN50160. The issue of the utility protection co-ordination getting disturbed presents a serious concern as the automatic reclosers can complete the broken circuit even during power flow in the island. In all the islanding events described in [13], the islands just managed to successfully collapse before the reclosing of the feeder. However, there always remains a risk of circulating current flow due to automatic reclosure during islanding operation. Transient over voltages affecting system reliability is a serious outcome as experienced during an islanding event caused during maintenance work on a MV feeder in Spain [20]. Voltages as high as 258% of nominal value were observed due to islanding behavior of the solar PV inverters. Damages to the PV plant and customers connected to the feeder were also confirmed.

As a part of the compliance to follow the regional grid-code and maintain grid-discipline [10], the DGs have to avoid islanding by detecting the condition within a stipulated time interval. The anti-islanding measures differ in operation for different types of generators feeding the grid. A solar PV inverter employs a technique that might not work well for a biogas-run alternator[21] owing to distinct characteristics like input-output relation or large inertia. Furthermore, the performance of these techniques is also affected by the type of load connected to the system [22]. In principle, the loss of utility supply can be detected in two ways. The first one involves monitoring the changes in grid parameters until they cross the values defined for healthy system operation. The second way is to inject some disturbances artificially and analyze the network's response to get an indication of deviation from normal state resulting from the loss of upstream supply. These two ways represent the classical islanding-detection methods and have been named as passive and active respectively. An amalgamation of the two results in hybrid techniques, and these three together cover the category of local techniques under the classical methods. There is also a second broad category of classical techniques that rely on transmitting signals over power lines to alert the DG in the wake of an islanding event. These remote telecommunication based techniques are too costly for use in interconnected distributed generation systems. The classical techniques have matured at the commercial level and mostly passive and active methods are embedded inside PV inverters. The spectrum of anti-islanding methods also includes the CI based techniques that have not yet found commercial breakthrough. They are being actively pursued in research and will be the main focus of this chapter. This approach of islanding-detection incorporates the elements of data mining, machine learning and optimization to learn the patterns prominent in different events and distinguish them from the islanding event. However, the detection is built upon the basics of classical techniques. Accordingly, they have been termed as 'passive based' and 'active based' in the general classification of anti-islanding techniques shown in figure 2.1. 'Passive based' methods learn the patterns from measurements of local parameters to identify power system events and classify them as islanding and non-islanding. The 'active based' ones are characteristically of two types. Those belonging to the first type generate optimal disturbances to help system parameters exceed the threshold in an islanding condition. The aim is to enable a quick detection with least impact on the system. The second type includes those that can optimize the response of the DG to a given disturbance to improve the detection of an islanding condition. The classical methods and their shortcomings and different CI based techniques reported in recent works have been described in separate sections later in the chapter.

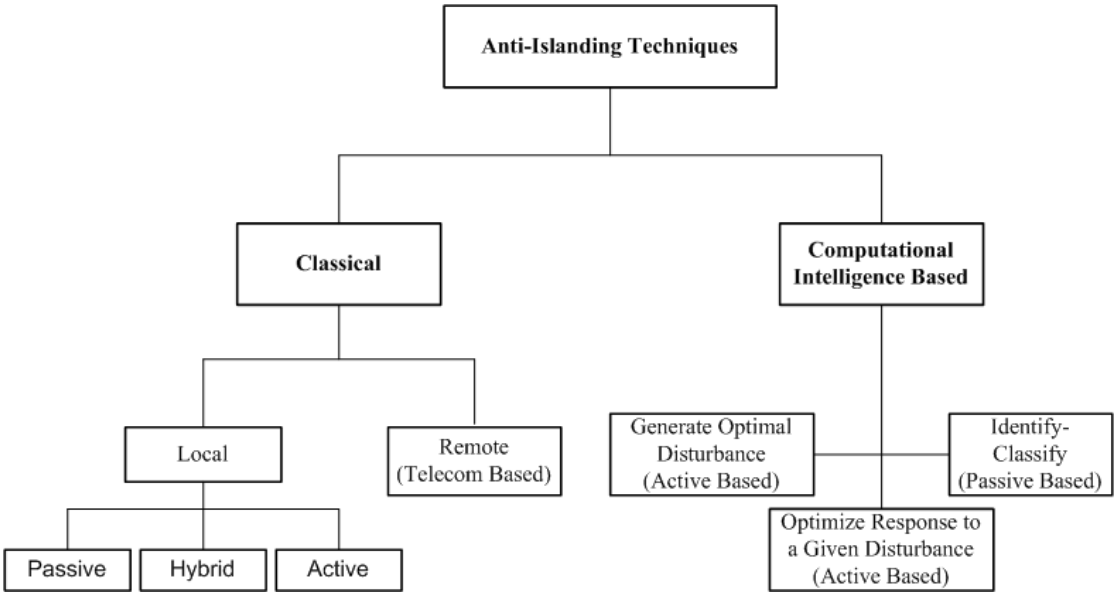

Figure 2.1: General classification of anti-islanding techniques

2.2 International Standards for Anti-Islanding

Many global standards exist for interconnecting distributed energy resources to the centralized utility power grid. These are documented guidelines and recommendations containing sections that specify certain anti-islanding requirements for the generators. The most basic requirement is expressed in terms of minimum islanding detection time for which quality factor of the load or Q_f is an important determinant. The generator must detect the formation of an island and disconnect itself from the grid within the stipulated time period. The Q_f at the time of island formation is an important factor that assesses the capability of an islanding-detection technique. It defines the strength of the resonance for the parallel RLC load in the island. This load model is frequently used to represent feeder loads to test the performance of an islanding-detection method. In the test procedures specified in these standards, the quality factor is adjusted in such a way that the loads resonate at grid frequency. The condition of power balance between the DG and the load alongside resonance in the load present the toughest case for island detection. A high Q_f at resonance tends to stabilize the island at grid frequency making it more difficult to detect. Besides these, the standards also mention the lower and upper limits of the operating range of V_{Grid} and f_{Grid} for the DG after interconnection. This range also determines the extent of the NDZ since a wide range will lead to a broader area covered in this zone.

Some of the prominent and most referenced standards that are followed across

Table 2.1: Various anti-islanding standards

Standard	**Q-factor**	**Detection Time**	**Frequency Range**	**Voltage Range**
IEEE 929-2000	2.5	Within 2 s	59.3 Hz to 60.5 Hz	0.88 p.u. to 1.1 p.u.
IEEE 1547	1	Within 2 s	59.3 Hz to 60 Hz	0.88 p.u. to 1.1 p.u.
IEC 62116	1	Within 2 s	59.3 Hz to 60 Hz	0.88 p.u. to 1.15 p.u.
UL 1741	1	Within 2 s	59.3 Hz to 60.5 Hz	0.88 p.u. to 1.1 p.u.
Japanese JIS C 8692:1997	0	Within 2 s (active) Between 0.5 s to 1 s (passive)	Setting Value	Setting Value
Korean	1	Within 0.5 s	59.3 Hz to 60 Hz	0.88 p.u. to 1.1 p.u.
German VDE 0126-1-1	2	Within 0.2 s	47.5 Hz to 50.2 Hz	0.88 p.u. to 1.15 p.u.
Australian AS 47777.3-2005	1	Within 2 s	Setting Value	Setting Value
Canadian C22.2 No. 107-01	2.5	Within 2 s	59.5 Hz to 60.5 Hz	0.88 p.u. to 1.06 p.u.
French ERDF-NOI-RES_ 13E	2	Instantaneously	49.5 Hz to 50.5 Hz	0.85 p.u. to 1.15 p.u.
UK G83/2 (DGs upto 16 A/phase)	> 0.5	within 0.5 s	47.5 Hz to 51.5 Hz (stage 1) 47 Hz to 52 Hz (stage 2)	0.87 p.u. to 1.14 p.u. (stage 1) 0.8 p.u. to 1.19 p.u. (stage 2)
UK G59/3 (17 kW/phase or 50 kW 3 phase)	> 0.5	within 0.5 s	47.5 Hz to 51.5 Hz (stage 1) 47 Hz to 52 Hz (stage 2)	0.87 p.u. to 1.14 p.u. (stage 1) 0.8 p.u. to 1.19 p.u. (stage 2)

the globe are the ones given by the IEEE, IEC and UL. The major examples are IEEE 929-2000 [23], IEEE 1547 [3], IEC 62116 [24] and the UL-1741 [25]. Besides these internationally accepted standards, various countries have come up with their own set of rules and guidelines to formulate standardizing operating procedures for grid-connected DG systems. The summary of the various standards and the information they convey is given in table 2.1 where per unit (p.u.) representation of electrical quantities has been used. Many country-specific standards have proposed some good recommendations alongside the internationally accepted standards from major organizations. It can be seen that the Canadian, Korean and German standards impose the most strict anti-islanding requirements for the DGs. Many developing countries who plan to increase the DG penetration on their distribution networks can study these standards as ready reference for developing their own set of protocols based on the operating conditions they desire.

2.3 Classical Techniques and their Shortcomings

The local techniques among the classical ones are prevalent in commercial practice as described previously. However several shortcomings seem to render them slightly incompetent for the foreseen scenario of rising DG penetration. The passive methods are named so because they act by monitoring the various electrical quantities in the PCC region around the DG, without actively manipulating them. The values of these quantities during the normal grid-connected operation of the DGs are used to define the islanding protection thresholds. An islanding situation is detected on the violation of the thresholds caused due to disconnection from the utility supply. Voltage and frequency and their time-derivatives, power, Rate of Change of Power (RoCoP), Rate of Change of Frequency (RoCoF), change in power factor, current and voltage THD, change in phase angle among others are some of the system parameters in the region local to the DG that are monitored

in passive detection [26]. Under-voltage, over-voltage and under-frequency, over-frequency relays are the most basic protection systems residing inside the DG. To explain their working, the example of a solar PV inverter connected to a distribution network in the UK and compliant with the G83/2 standard can be taken. An unintentional islanding situation may reduce the frequency of the section around the PCC below 47.5 Hz due to reactive power mismatch. This crosses the lower frequency threshold value defined in the standard and hence the under-frequency relay will immediately detect the island formation. However, in the case of a close balance between the Q_{load} and the inverter output, the island may persist if the frequency remains within the limits of the operational range defined in the standard. In such a situation, the RoCoF based passive detection may be effective since the Hz/s variation in the islanded mode may be different from that in the grid-connected operation.

Speed and accuracy are generally the two performance indicators for any islanding-detection method. Passive methods, as discussed above, can quickly detect changes in local parameters but they are not always 100% accurate in detecting islands. Islanding-detection accuracy is affected by the extent of the NDZs that represent the region of $\triangle P$ and $\triangle Q$ between the load and DG for which the V_{island} and f_{island} remain within the safe operating limits, even in the absence of grid. This has been depicted earlier in figure 1.9. An anti-islanding technique fails to detect the islanding condition when the power mis-match levels lie in the NDZ. Both passive and active techniques possess a NDZ. The variation of load with voltage and frequency (static and dynamic characteristics) and the control functions of the DG, particularly inverter based, also impact the area covered by the NDZ for a passive technique.

The passive methods somehow fail to capture the dynamics of the multi-dimensional variations in different parameters taking place on a distribution network. A passive islanding protection relay is effectively able to monitor and process only one parameter at a time. For multiple parameters, setting the threshold becomes a problem. The thresholds should be such that they are not too low to cause false detection and not too high that island-formation goes undetected. For the set up as shown in figure 1.10, $P_{load} = P_{inv}+P_{grid}$ and $Q_{load} = Q_{inv}+Q_{grid}$. However, under power balance condition, P_{grid} and Q_{grid} approach zero and this fact guides the theory behind the relation between V_{island}, f_{island} and $\triangle P$ and $\triangle Q$ between the loads and PV inverter. For a parallel RLC load, the resonant frequency is given by $\omega_0 = \sqrt{\frac{1}{LC}}$ and the quality-factor is expressed as $Q_f = R\sqrt{\frac{C}{L}}$. When an inverter-based DG like solar PV that operates in a constant PQ control mode gets

accidentally disconnected from the utility grid[27], the voltage and frequency of the island vary with the active and reactive power output that are controlled by the solar insolation. Noticeably the NDZ for such a situation, as given by equations 2.1 and 2.2 and discussed in [28], conveys the association between V_{island}, f_{island} and $\triangle P$ and $\triangle Q$ and thus highlights the importance of the term 'P-Q mismatch level' for islanding-related studies.

$$Q_f \left(1 - \left(\frac{f}{f_{min}}\right)^2\right) \leq \frac{\Delta Q}{P_{load}} \leq Q_f \left(1 - \left(\frac{f}{f_{max}}\right)^2\right) \tag{2.1}$$

$$\left(\frac{V}{V_{max}}\right)^2 - 1 \leq \frac{\Delta P}{P_{load}} \leq \left(\frac{V}{V_{min}}\right)^2 - 1 \tag{2.2}$$

These fractions are the defining parameters of a NDZ. Based on the V_{lower} and V_{upper} and f_{lower} and f_{upper} threshold settings in utility relays and the load quality factor as defined in a standard, the NDZ for a passive technique can be determined. Since the approach in such threshold based techniques is limited to comparing the value of the monitored parameter with the design threshold, they often overlook the fact that a true islanding condition is defined by the status of multiple parameters. The passive islanding detection techniques are identified and named according to the various quantities that they monitor. A few examples are under/over voltage method, under/over frequency method, change of source impedance, voltage unbalance and THD of DG current method, RoCoP, RoCoF and vector surge relays (similar to RoCoF).

The response of the PCC region of a feeder to disturbances during healthy system conditions and in an islanded condition can be different. This difference is utilized in active methods to detect islanding. The active methods try to bring the local parameters, used in passive detection, out of the NDZ by injecting a disturbance in the DG output. If the main supply is present, such a disturbance will not make any impact but during grid-disconnection, the response of the system to the disturbance gets magnified that forces the local parameters to cross the threshold. This deviation is finally detected by the underlying passive-detection based relay that identifies the island formation. The active methods have the advantages of a smaller NDZ but they tend to degrade the power quality of the network. They also take longer to detect islanding as compared to the passive methods. Current injection methods, active frequency drift (AFD), Sandia frequency shift and Sandia voltage shift [29], adaptive logic phase shift method are some of the examples of active islanding detection methods. Considering the example of the PV inverter cited above, if both the under/over frequency and RoCoF based passive methods

fail, then the AFD method can ensure island detection. This method will force the frequency of the V_{island} to drift above or below the limits defined in the standard G83/2 which advocates the use of an active method for a reliable loss of mains detection. Suitable for PV inverters, the basic AFD slightly distorts the current waveform injected by the inverter into the utility grid so that the natural frequency of the V_{PCC} drifts outside the safe limits. The shift is overpowered by the utility supply frequency regulation but once it is removed, the shift gets amplified. This is requirement for accurate detection and it must be met specially during the condition when $Q_{\text{Load}} = V^2[(\omega L)^{-1} - (\omega C)] = Q_{\text{PV}}$, which is when the passive methods fail. Advanced variants of this method use positive feedback to increase the deviation intensity as reported in [30] or propose changes in the injected current waveforms, like in [31], where a sudden change over half cycle of the fundamental produces the necessary phase shift. This modifies the disturbance so as to have less impact on the grid, mainly in terms of reduced THD.

The hybrid methods leverage the speed of passive methods and the accuracy of active methods combined together but the detection time is increased. They fare better in some respects when compared individually to each of the two methods since they are not an implementation of sequential operation of active and passive techniques but a single technique incorporating the strengths of both. However, these set of techniques also do not seem to provide a comprehensive, dynamic and robust solution to the issue since they enforce the same reactive strategy of islanding detection. As the number of renewable energy based DG installations on the centralized power grid will increase, the challenges and complexity are expected to increase and new triggers to section islanding will emerge. In such a scenario, the local methods may not suffice to provide a reliable and foolproof solution to ensure network security and real-time smooth operation.

2.4 Computational Intelligence based Techniques for Anti-Islanding

The classical methods do not possess embedded intelligence to learn from events and decisions. They perform event detection by executing a conditional rule-based logic for given thresholds. An islanding event is defined when a set of conditions is satisfied, requiring sequential execution of if-then-else rules in a span of few cycles to detect its occurrence. Moreover as DG penetration and their advanced control functions increase, the thresholds will also keep changing dynamically.

These challenging scenarios exist alongside the fact that new type of phenomena keep occurring on active distribution feeders giving rise to potentially malicious events. Given such complexities, more robust islanding-protection can be achieved with the use of CI based techniques. These set of techniques can process multiple quantities simultaneously and compute inferential parameters to build predictive models. The DG's control structure can then be embedded with a 'trained model' enabling it to identify different events correctly in real time.

Designing a computational model for islanding-detection comprises a few steps. First, raw data belonging to different events is collected from which distinct event-specific features are extracted. These features are then archived into a database and inherent patterns are discovered to extract useful knowledge that develops intelligence in the DG to recognize and classify events and take decisions accordingly. A detailed comparison of the CI techniques reported in literature for anti-islanding has been presented in this section. The CI techniques can solve highly non-linear problems including those that cannot be easily translated into an algorithm. In this respect, they slightly differ from AI based methods [32] under whose paradigm data mining and machine learning have been generally put. However there is a very thin boundary between the two classes of intelligence and thus the broad classification has not been made while reporting the specific techniques in this section. Processing event data-extracted features to build models that identify and classify (passive based) is not the only approach towards anti-islanding. CI based techniques have also been used in generating optimal disturbance signals as well as in optimizing the response to a given disturbance signal in detecting islanding and this approach has been termed as 'active based'. The following subsections report the various feature extraction techniques, the CI techniques and special software packages used in the anti-islanding studies mentioned in recent literature by briefly explaining each one of them.

2.4.1 Feature Extraction Techniques

Features are processed forms of power grid signals like current, voltage, frequency and others that carry inherent information about the physical events that generate them. Various signal processing methods have been applied in extracting important features from different signals as explored from the literature survey. The islanding-detection technique reported in [33] has used a software model that executes a certain signal processing algorithm for extracting features from the voltage and current signals. The parameter deviations discussed in [34] have been directly

obtained as the features for training the learning model. Twenty-one suitable features were extracted from the voltage and current waveforms in [35].

A feature extraction window is an interval of time in which an instantaneous signal is processed to derive features. The window can be stationary in time or moving known as sliding window. The literature contains references to both kinds of windows used for feature extraction at different sampling rates for continuous signals. In reference [36], the governor input signal has been sampled at 175 Hz in a 200 ms window to extract 35 features for training the learning model used. Discrete Wavelet Transform (DWT) is a discrete sampling technique that decomposes a signal into different frequency scales having different temporal resolutions[37]. The mother wavelet is decomposed into detail wavelets known as levels. The DWT of the transient current signal has been used to extract feature vectors at a sampling rate of 20 kHz for the classifier modeled in[38]. The study used Daubechies-4 (DB-4) wavelets with 6 decomposition levels for analyzing the signal in a 0.01 sec window. DWT of 166 samples from a 60 Hz voltage waveform has been used to extract energy contents of 7 detail wavelets from a sliding mode window at a sampling rate of 10 kHz in [39] that used Daubechies-1 (DB-1) as the mother wavelet. The DWT of transient voltage and current signals at a sampling frequency of 10 kHz has been utilized to derive the input feature vectors for the classification learner reported in [40]. The study used Mother wavelets belonging to the Daubechies family and signals in each mother wavelet were decomposed upto 9 levels to train the classifier exhaustively. However, the study resulted in a finding that using only the voltage signal, DB-4 wavelet and the third level gave the best results.

Discrete Fourier Transform (DFT) is another popular method of feature extraction that has many different variants. It is the equivalent of the continuous Fourier transform and is generally applicable to stationary sequences. For a sequence made up of equally-spaced samples of a function, the DFT gives a sequence of equivalent length containing samples of the discrete time FT that are equally spaced and represent a function of frequency that is complex-valued. Many variants of this method have been reported in the literature for deriving the features for use in different approaches for islanding detection. Goertzel algorithm has been used in the active based method reported in [41]. This method is an algorithmic approach to generate a kind of DFT that directly extracts the magnitude and frequency of a particular harmonic. In the cited study, the Goertzel algorithm has been used for the 9^{th} harmonic of the current wave to detect islanding. The 2^{nd} harmonic component of the PCC voltage waveform was utilized in [42] that proposed an active islanding detection method using the Goertzel algorithm. For quick islanding

detection, the use of Fast Fourier Transform (FFT) has been aptly found in the literature. This technique reduces the time complexity of computing the DFT of a signal by using a sparse matrices product to express the DFT matrix. FFT has been used in [43] to calculate the Harmonic Content (HC) of the equivalent reactance seen at the DG terminals X_{DG} as one of the input features for the learner model applied.

Auto-regressive signal modeling is a parametric spectral analysis technique and finds use in deriving important features from data points. It has been utilized in [44] to extract various features from the current and voltage signals for the classification model used. In reference [45], voltage and current were measured and 21 possible features were extracted by sampling them at 1 kHz. Feature selection algorithms like Forward Feature Selection (FFS) and Backward Feature Selection (BFS) were applied to select four main features for island detection.

2.4.2 Computational Intelligence Techniques

1. *Decision Tree*: A Decision Tree (DT) is a non-parametric predictor model that can be used for both classification and regression. It essentially employs an optimized attribute-search throughout the complete data set of features and labels. Concepts of information theory are used to find the best attribute to classify the data at each sub-tree's root node. As a classifier, it is of high dimensions for accurate modeling of non-linear decision boundaries, it has to balance dimensionality [46]. The root node contains the initial classification problem which splits into internal nodes that test other attributes for respective thresholds or settings to further branch out. Each branch of a node represents a decision function or logic. If the conditions for the logical test are matched, the leaf nodes contain the predicted class label.

A DT classifier trained using 11 features obtained from 54 events generated in simulation has been used in [33] to classify events into islanding and non-islanding ones. Reference [34] also uses a DT classifier that is based on the ID3 algorithm and trained from the events generated from the same test-circuit model as used in [33]. Current and power factor were selected by the tree as the best features for learning the classification model used. DT was used as one of the classifiers in a study in [38] that employed transient signals for loss of mains detection. DT is also used as one of the classifiers in [45] to detect islanding after being trained and tested by k-fold cross validation on two different sets of features. A DT was used to

initialize the classification boundary of the fuzzy inference system in [47] and[48] by using the three most significant features selected by the tree. A Classification and Regression Tree (CART) used in [43] was trained from the input features derived from the reactance across the synchronous DG terminals. The DT used in the study reported an accuracy of 100% during both training and testing stages with an appreciable time of detection. An efficient anti-islanding relay proposed in [40] employed a DT as the classification model for identifying islanding events from the non-islanding ones. The classifier achieved an average test accuracy of 98% after multiple-folds of cross-validation.

2. *Random Forest Classifier*: A Random Forest Classifier (RFC) is a ensemble of a different DTs. The RF algorithm follows a voting method among the trees to reduce the bias in classifying unknown instances. A RFC made up of C4.5 DTs [49] was used as the main classifier in [45]. Both entropy and information gain functions were used to split the training data at each sub-tree to construct the trees.

3. *Naive Bayesian classifier*: It is a classification model based on the Bayesian paradigm of probability. From the training data set, it basically learns the conditional probability $P(C = c|X = x) = \frac{P(X=x|C=c)P(C=c)}{P(X=x)}$ for each feature $x_i \in X$ with the predicted class label value $c \in$ set of class labels C. The expression thus estimates the probability of a test data point belonging to a class c conditioned on its set of features x_i. The assumption that all the data sets are independent of each other helps in calculating the probabilities from a large set of training data points. This assumption is also the reason behind calling it Naive. The study reported in [35] used a Naive Bayesian (NB) classifier whose classification accuracy was validated with a Support Vector Machine (SVM) classifier and 4-fold cross-validation. Four-fold cross validation for training and testing a NB classifier was reported in [45]. The classifier was trained with all possible 21 features and 4 important features, in addition, to identify and classify various events simulated for all three scenarios described in the work.

4. *Support Vector Machine*: A Support Vector Machine (SVM) is a prediction model based on the statistical decision-theory. For discrete-valued data, it represents a parametric classifier that finds an optimally separating hyperplane to divide the training data into different classes. The optimal hyperplane can then categorize unseen data points into one of the possible classes. For non-linearly separable training data, a SVM classifier uses a kernel function K(x_i, x_j) that relates a subset of training data vectors x_i to the testing data vectors x_j. This is a

non-linear transformation comprised of some basis functions and projects the input data onto a high-dimensional feature space. The kernel for linearly-separable data is simply $\mathrm{K}(x_i) = x$. Building a SVM classifier involves solving a constrained optimization problem of finding the best separating hyperplane with minimum training error. In the study reported in [38], SVM has been used as one of the classification models. A kernel using the Radial Basis Function (RBF) has been used in the SVM classifier to validate the classification results of the main classifier in[35]. The use of a SVM classifier has also been been reported in [44] wherein it was trained using the derived feature vectors. RBF kernel was used where the kernel parameter λ and the term that controls the optimization known as error-penalty C were varied in a range. The best combination was then selected by using five-fold cross-validation. SVM was one of the classifiers used in [45]. It was trained and tested for two different sets of features using k-fold cross validation.

5. *Artificial Neural Network*: It is a technique that creates a set of neurons similar to those found in a biological neural network. It has great capability to learn and recognize from patterns. The input data vectors activate the neurons. A transfer function weighs and transforms them and other neurons receive these activations. The process is repeated until an output neuron is finally activated and it recognizes the particular data set which was to be identified. In [50], the learning capabilities of a Neural Network (NN) have been utilized in a wavelet-fuzzy neural network. An ANN based classifier has been used in[39] to build an event-classifier trained with the used feature data. The unsupervised learning paradigm was tried in [36] that used a self Organizing Map (SOM) type neural network to detect islanded operation of a synchronous generator. The network used 7 neurons in a hexagonal topology. The learning parameters of the SOM NN were set in the form of a Gaussian neighborhood function with $h_0 = 1$ and two learning rates $n_1 = 0.9$ and $n_2 = 0.02$ divided among 650 cycles with all the initial neuron weight vectors set as 0.5.

A NN clasifier has been used as one of the benchmark classifiers in [45] to validate the accuracy of the main classifier. A Multi-Layer Perceptron (MLP) type ANN with 4 layers has been designed in [51] as the main learning network to identify and classify islanding events. The MLP was built with 8 perceptrons in the first layer, 4 in the second and 2 and 1 in the third and output layers respectively. $V_{\mathrm{DG\ output}}$ sampled at 64 and 128 samples per second was used as the input feature vector.

6. *Probabilistic Neural Network*: A Probabilistic Neural Network (PNN) is a

feed-forward neural network with four layers. There is communication between the neurons in each layer. When a test vector is given to the network, the first layer computes the distance of this vector from the training-data input vectors. The result is a vector whose elements indicate the nearness of the test input to the training input. The second layer sums the contribution for each class of inputs and produces a vector of probabilities as its net output. Finally, a complete transfer function on the output of the second layer picks the maximum value among these probabilities. It produces a '1' (positive identification) for that class else gives a '0' (negative identification). A PNN has been used as one of the classifiers in [38] that used 4-fold cross validation for performance optimization. The study reported in [52] used phase-space analysis to extract features. These were then used to train a PNN based classifier to implement an effective anti-islanding strategy. $V_{\text{DG terminalk}}$ was used as a parameter to extract five input features in the high dimensional phase space.

7. *Fuzzy Logic Control*: Fuzzy logic extracts fuzzy variables from crisp, real-world input variables. This is based on a reasoning logic that is multi-valued and can express beyond true and false. The literature mentions the use of fuzzy logic based techniques mostly in active-islanding strategies since they are best at handling and processing uncertainties. A fuzzy inference system whose classification boundaries were initialized by a DT built using the three most significant features has been reported in [47] and [48]. The classification boundaries given by the DT were used to assign membership functions to the generated fuzzy variables. The fuzzy rule base contained 4 rules and they were used to identify and classify events into islanding and non-islanding. Reference [53] utilized fuzzy logic in an active islanding detection strategy. In the active based technique used in the work, the reference values of I_d and I_q were modified by their respective fuzzy output disturbances. The rules in the fuzzy inference system were then used to detect islanding. V_d and ω_{PLL} were used as input signals and triangular and trapezoidal membership functions were assigned to them respectively to create the disturbances. A Wavelet fuzzy neural network based controller was used in [50] to regulate the error between Q_{inverter} and the Q_{command}. The drift in f_{PCC} from the NDZ value was detected to confirm islanding after the injection of the disturbance signal in the form of a direct-axis current. The uncertainty-handling was done by the fuzzy logic portion in this active based method.

8. *Wavelet Fuzzy Neural Network*: It is a combination of Fuzzy Neural Network (FNN) and Wavelet Neural Network (WNN). Wavelets, as discussed above, are powerful processed signals that retain the time and frequency information of each

harmonic of the original waveforms. FNNs on the other hand combine the advantages of robustness of fuzzy reasoning in real-life situations and the learning capabilities of an ANN. In [50], the conventional Proportional-Integral (PI) controller was replaced by a WFNN based controller. Uncertainty was handled by the fuzzy portion and learning capabilities of the ANN were used to adjust the interaction among controller levels to minimize the energy function V based on Q_{command} - Q_{inverter}.

2.4.3 Special Software Packages used in Anti-Islanding Studies

A distributed energy resource can have many possible locations for integrating it with a given distribution network. The network-portion between the PCC and the utility protection on the DG side becomes the region of interest for testing the anti-islanding capabilities of the connected DG. Accordingly, the test set-ups described in the anti-islanding standards are realized or modeled in software to test and analyze the performance and results of a particular anti-islanding technique. Certain benchmark feeder models are also used as test systems since placing DGs on a standard network having defined parameters is a justifiable practice. The IEEE 34 bus system is a suitable test-bed for analyzing the impact of integrating renewable energy based DGs [54], including islanding in particular, as reported in literature[55, 54, 56, 57]. The IEEE 13 node feeder is another test-system that represents a small yet highly un-balanced distribution network [58] and provides a challenging test-framework for anti-islanding studies that can be then scaled to larger networks. The same network, with some modifications, has been used in this research work. Apart from those given by the IEEE, many other standard distribution systems exist that can be modeled to study the impact of DG integration and consequently test the anti-islanding techniques applied to them [59]. The CIGRE benchmark medium voltage system [60] is another popular test-bench for modeling DG integration with the power grid and testing for anti-islanding performance.

The test systems that are based on standard radial feeders have been simulated in Matlab-Simulink in [35, 45] and in PS-CAD/EMTDC in [38, 44]. PSIM simulator has been used in [41] for creating the test bench specified in the IEEE-1547 standard. Digsilent Power Factory has been used to model the CIGRE MV system in [40] wherein various islanding and non-islanding conditions have been simulated. Digsilent Power Factory was also used to model the IEEE 33 bus system in [52].

Studying the performance of a given anti-islanding technique of a feeder-integrated DG in simulation requires the models of some basic elements to be available inside the software package being used. These fundamental components are the grid substation, the feeder, DGs of various types and the loads. Models of the required basic components like transmission lines, transformers and circuit breakers are available in almost all the modeling and simulation softwares. The over-current relay model and the Voltage Regulating Transformer (VRT) model is available in all the software tools discussed except in Simulink. Different software tools discussed above have been compared for their capabilities in designing test systems for studying anti-islanding techniques in table 2.1. The ability to model a feeder based on its geographic coordinates available from the GIS data makes any tool convenient to use and the same has been discussed in the table.

MATLAB provides special toolboxes and library packages for implementation of various CI based techniques. Fuzzy logic toolbox is generally used for creating fuzzy systems and its use has been reported in [47, 48, 53]. The neural network toolbox of Matlab has been used to create the ANN in [45] and the PNN in [38]. The SVM classifier modeled in MATLAB in [38] was created using the LibSVM package. MATLAB has the capability to implement various CI techniques through interactive toolboxes however the source-code of the underlying algorithms are proprietary and not easily visible. User-side modifications or adjustments to the algorithms thus cannot be made.

CART data mining package as used in [33, 38] and Insightful Miner package in [47] are other DT building softwares reported in the literature. The CART® data mining package is owned by Salford Systems and is available under the Salford Predictive Modeler (SPM) software suite. CART stands for Classification and Regression Tree which is a popular variant of the DT algorithm. The package provides robust performance and built-in automation courtesy the strong original proprietary code. It also offers ease of use with a user-friendly interface. Insightful Miner is an enterprise-wide data mining tool that can be integrated with many data-handling applications. It provides tools for cleaning and manipulating data for analytical modeling on pre-processed data and can perform all types of statistical visualizations. It can fit a variety of statistical models including linear and logistic regression and DTs. It comes with an in-built support for popular analytical platforms like SPSS and SAS.

WEKA is another powerful open-source data mining and machine-learning package whose use was reported in training the Naive-Bayesian classifier modeled in [35]. Waikato Environment for Knowledge Analysis (WEKA) is a freely available

Table 2.2: Comparison of software tools reported for test system modeling

Software Tool	**GIS data input**	**DG model available**	**VRT model**	**Residential load model**
MATLAB-Simulink	Yes	PV + wind	No	No
PS-CAD	Yes	PV + wind	No	Yes
PSIM	No	PV + wind	No	No
Power Factory	Yes	PV + wind	Yes	Yes

software that offers portability across any computing platform as it is programmed in Java. It supports standard data mining operations like classification, clustering, regression including data pre-processing and visualization. WEKA offers the capability to provide connectivity to Structured Query Language (SQL) databases and it can also process the result returned by a database query. WEKA was used to train the DT used inside the anti-islanding relay proposed in [40]. It was also used in [45] for training and testing the Random-Forest classifier along with MATLAB.

A comparative evaluation of the different computational intelligence based anti-islanding techniques reported has been presented in table 2.3 and table 2.4. Goertzel algorithm has also been included as a computational technique however its brief explanation had been covered in section 2.4.1.

Table 2.3: Comparison on simulation parameters

CI Technique	Events Simulated	Test System	DG Type
Decision Tree [33]	54 prescribed events: faults and switching actions	Own system model	Synchronous Generator
Decision Tree [34]	15 types of faults	Own system model	Synchronous Generator
Decision Tree [38]	154 islanding cases: Opening of bus and feeder Circuit Breakers (CBs) 121 non-islanding cases: Temporary faults, load and DG switching	CIGRE MV system	Induction (DG 1) Synchronous (DG 2)
Decision Tree [43]	72 scenarios of 7 event types: Faults and DG, capacitor and load switching	WSCC 9-bus system	Synchronous generator
Decision Tree [40]	568 islanding and non-islanding events: Faults, DG trips and capacitor and load switching	CIGRE MV system	1 Synchronous generator 1 Induction generator
NB classifier [35]	P and Q mismatches, individually and all together load and capacitor switching	IEEE 34 bus system; DG at node 848	Inverter based
Goertzel algorithm [41]	Grid disconnection and power-balance condition	IEEE 1547 test system	1 phase PV inverter
Goertzel algorithm [42]	Grid disconnection and voltage disturbances	Own system model	1 phase PV inverter
SVM Classifier [38]	154 islanding cases: Opening of bus and feeder CBs 121 non-islanding cases: Temporary faults, load and DG switching	CIGRE MV system	Induction (DG 1) Synchronous (DG 2)
SVM classifier [44]	350 islanding events: $\pm$ 40% P mismatch and $\pm$ 5% Q mismatch 350 non-islanding cases: Faults and load switching	IEEE 13 bus system	1 phase PV inverter
PNN [38]	154 islanding cases: Opening of bus and feeder CBs 121 non-islanding cases: Temporary faults, load and DG switching	CIGRE MV system	Induction (DG 1) Synchronous (DG 2)
PNN [52]	1012 islanding and non-islanding events: 3 types of faults and capacitor and load switching	IEEE 33 bus system	4 synchronous generators
Fuzzy logic [47, 48]	6 different islanding and non-islanding conditions: 3 ϕ faults, CB and DG trips, sudden load decrease	Own system model	2 Synchronous generators
Fuzzy logic [53]	Power balance condition at high Q values	Own system model	1 phase inverter
ANN [36]	Islanding events: Opening of 3 CBs for DG loading from 30% to 400% Non-islanding events: Load switching at various buses (0.25 p.u. to 2 p.u.) Capacitor switching (0.2 p.u. to 1.6 p.u.) 3 phase and 1 phase faults	Own system model	Synchronous generator
ANN [39]	Total 360 islanding and non-islanding events Islanding events: Opening of utility and DG side CBs (1 cycle) Non-islanding events: Faults, loads/capacitor bank switching	Own system model	Two 3 phase PV inverters
ANN [51]	2000 simulations created: Switch opening, load switching and abnormal frequency	Own system model	Synchronous generator
WFNN [50]	Disconnection of the grid	UL-1741 test system	Inverter based
RF classifier [45]	Total 200 events Islanding events: P mismatch upto $\pm$30%, Q mismatch upto $\pm$5% Both of the above mismatch conditions together Non-islanding cases: Load, capacitor and motor switching at different buses faults at different buses	IEEE 34 bus system	Synchronous generator and inverter

Table 2.4: Comparison on performance parameters

CI Technique	Features Extracted	Time Taken in Detection	Accuracy of Detection
Decision Tree [33]	11 grid parameters	45 to 50 ms	100% (non-islanding) 83.33% (islanding)
Decision Tree [34]	Deviations in I ,V , pf , f , P and Q	Not reported	Not quantified
Decision Tree [38]	Signal energy of different wavelets at each of the 6 decomposition levels	0.032 s + processing time	90% average
Decision Tree [45]	21 different grid parameters extracted: V, f, Z_1 and ϕ selected	Not reported	95% (inverter DG) 97% (synchronous DG) 96% (Both DGs)
Decision Tree [43]	Max.(df/dt) and $HC_{odd}(X_{DG})$ /$HC_{even}(X_{DG})$	300 ms	100% (Training and testing)
Decision Tree [40]	DWT energy coefficients of transient voltage	0.01 s	98% (both DGs)
NB classifier [35]	21 different grid parameters	within 120 ms	96.66% average
NB classifier [45]	21 different grid parameters extracted V, f, Z_1 and ϕ selected	Not reported	56% (inverter DG) 100% (synchronous DG) 100% (Both DG types)
Goertzel algorithm [41]	Magnitude and frequency of: 9^{th} harmonic component of V_{PCC}	1 cycle of V_{PCC}	Active method
Goertzel algorithm [42]	2^{nd} harmonic component of PCC voltage	Less than 120 ms	Active method
SVM Classifier [38]	Signal energy of different wavelets at each of the 6 decomposition levels	0.032 s + processing time	74% average
SVM Classifier [35]	21 different grid parameters	Not reported	88.33% average
SVM Classifier [44]	62 features from voltage and current signals	50 ms	98.94% average
SVM Classifier [45]	21 different grid parameters extracted: V, f, Z_1 and ϕ selected	Not reported	94% (inverter DG) 100% (synchronous DG) 100% (Both DGs)
PNN [38]	Signal energy of different wavelets at each of the 6 decomposition levels	0.032 s + processing time	85% average
PNN [52]	5 phase-space features of DG terminal voltage	0.24 s	100% (islanding/ non-islanding)
Fuzzy Logic [47]	11 grid parameters extracted; DT selects $\frac{\Delta P}{\Delta T}$, $\frac{\Delta f}{\Delta T}$ and Δf as 3 most significant features	Not reported	100%
Fuzzy Logic [48]	11 grid parameters extracted; $\frac{\Delta P}{\Delta T}$, $\frac{\Delta f}{\Delta T}$ and Δf selected	Not reported	100 % for 36 training cases (with and without 30 dB noise)
Fuzzy Logic [53]	Features monitored: V_d and ω_{PLL}	Not reported	Active based method
ANN [36]	240 input data vectors: 35 × 1	200 ms	97.9% average
ANN [39]	Energy contents of: 7 detail wavelets (levels) for each phase voltage signal	Not reported	97.22% average (DG 1) 97.77% average (DG 2)
ANN [45]	21 different grid parameters extracted: V, f, Z_1 and ϕ selected	0.30 s (inverter DG) 0.20 s (synchronous DG) 0.26 s (Both DGs)	97% (inverter DG) 100% (synchronous DG) 100% (Both DGs)
ANN [51]	DG output voltage sampled at: 64 and 128 samples/sec	> 2.75 s (64) 2.32 s (128)	99.28% average
WFNN [50]	Feature monitored: $\omega - \omega_{grid}$ Input feature: Q - Q_{inv} and its derivative	0.68 s	Active based method
RF Classifier [45]	21 different grid parameters extracted: V, f, Z_1 and ϕ selected	0.18 s (inverter DG) 0.18 s (synchronous DG) 0.18 s (Both DGs)	98% (inverter DG) 100% (synchronous DG) 100% (Both DGs)

2.5 Pro-active Islanding Management: Classical Attempts and Prospects for CI Techniques

The classical and intelligent approaches to islanding mitigation discussed so far essentially follow a reactive strategy. Despite commercial advancements in controlling this condition, islanding still remains an issue that needs to be comprehensively addressed in practice. It is one of those protection-related issues whose effects are expected to be influenced by DG penetration levels. As has been mentioned previously, the practice of disconnecting a DG upon island detection leads to loss of valuable power carrying financial implications. Although it is essential for maintaining system integrity, such an approach is not expected to be 'smart' enough to tackle this situation in the foreseen scenario of rising penetration. This is more pertinent to solar PV based DGs that remain the focus of this work. Given the rise in advanced PV inverter functions that now allow supporting grid voltage and frequency during short-duration contingencies, it makes sense to have the inverter connected to the network to supply the islanded loads by adjusting its operational mode.

The probability of occurrence of unintentional islanding was estimated to be practically zero by [15]. However the findings of this study, done in 2000, stand challenged presently when the penetration level is rising steadily. A study done for a PV interfaced feeder in [61] calculated the number of hours of load-generation balance to estimate the risk of accidental islanding. If this probability is a finite number P, then the cost of PV power undelivered can be expressed as a representative value $N \times P \times W_n \times Y$ INR/year where N is the total number of hours the PV system is in operation annually at a nominal power output of W_n Watts sold at a levelized tariff of Y INR/kWh. Thus the revenue made by the PV operator reduces to $N \times W_n \times Y(1-P)$ INR/year due to the inverter tripping offline following the prevalent reactive islanding management strategy. This can increase the payback period that may render the project financially unviable. It is thus important to have a pro-active management strategy in place to tackle unintentional islanding. By proactive, it is meant that the inverter can identify an imminent islanding situation and adjust its operational mode so as to run the loads in the island. It must also be responsive in the same manner to new triggers to section islanding that emerge in the high-penetration scenario, and one has been described in this thesis. The classical methods that attempted this approach have been listed and similar prospects for CI methods have been discussed in this section.

The PV generation estimate, the load profile and the typical nature of the yesteryear's grid were leveraged in [62] to predict unintentional islanding events in distribution networks. This was one of the initial attempts towards a predictive approach. It was a simulation study that developed simplified analytical models of the self-commutated inverters available at that time. Various conditions of $\triangle P$ and $\triangle Q$ were simulated to analyze the run-on time to estimate the duration of a possible islanding condition based on deviations in the control-circuit responses of the inverters. A contemporary work of that period, reported in [63], estimated the parameters for which a power balance condition could exist for possible formation of an island. It used the concepts of limit-cycle behaviour, small-signal stability and describing function methods for predicting the fundamental frequency and its multiples at which islanding can occur. In [64], a similar task was performed by analysis of the various inverter control functions modeled for the same.

The power and capabilities that the CI based techniques can offer are being undermined by the way they are currently being used for islanding mitigation. Fitting a learning model inside a DG to distinguish the occurrence of islanding from other events can be oriented towards incorporating predictive capabilities regarding the possibility of islanding. This can lead to robust and foolproof real-time operation for which data analytics and CI techniques possesses the necessary capabilities. Supporting this statement is the fact that prediction and forecasting techniques are used on a day to day and term to term basis in running the power system. Load forecasting is the most visible example of using historical data to predict future values. Therefore, monitoring, recording and processing data distinct to different types of events can be effectively used in incorporating learning into the DG to develop event predictive capabilities. Suitable optimization techniques can then be used to minimize the prediction error or tune the parameters of a real-time event identifier and classifier to enhance the accuracy and computational efficiency.

An encouraging example from the field is the application of data mining in preventing future blackouts on feeders in the USA. The commercial technology reported in [65] uses predictive-grid analytics to learn a model from the patterns of historical outages. The abnormalities or anomalies preceding an outage train the predictive model that identifies them in real time and takes corrective action before they magnify. Similar to the use of signature data of anticipated events, sensitivity-analysis of signature data has been used in detecting the occurrence of malicious events in [66]. However, characterizing the signatures of impending islanding events from other events data is possible after a thorough study of the distribution network and its response to the dynamics of the connected renewable energy DG [67]. Current

research reported on islanding-prediction has mostly focused on using actual and simulated records from PMU measurements to predict island formations in bulk interconnected transmission networks having centralized thermal generators feeding the islands formed. The study in [68] identified pre-event signatures from the PMU records of an actual islanding event while simulated real-time PMU measurements were used in [69]. However, PMU data is not always available for distribution grids [70] where much of the risk of unintentional islanding prevails. Thus alternative approaches are required to build better performing predictive models. Few examples in other applications as noted in the literature are: real-time prediction of extreme events in underground power distribution system using historical events data [71] and use of learning-to-rank algorithms for real-time failure-susceptibility rankings of primary feeders of an underground system in [72].

2.6 Conclusion

This chapter discussed the classical techniques of islanding detection that are prevalent in use. With the help of some examples, their functioning and shortcomings were also explained. Various CI based techniques reported in the literature for islanding protection of renewable energy based DGs were also discussed. A comparative analysis of these techniques in terms of speed, accuracy and other performance parameters was presented for a clear understanding. The available software tools and packages used in the reported techniques, with their relative merits and capabilities highlighted, were also mentioned to create awareness for the same. The need for proactive islanding management strategies was stressed upon that relates to the importance of the problem chosen in this research. Following this, some innovative prospects for the application of CI techniques in islanding detection were discussed. The rise in clean energy based DG penetration on distribution feeders poses new challenges for system operation. Anomalous occurrences due to combined effect of uncontrollable generation, evolving nature of loads and rising network complexities may lead to section islanding triggers. The next chapter describes one such trigger that has been discovered in this study. Given a challenging scenario ahead, it is felt that the CI based techniques should be oriented towards predictive detection approaches to enable zero wastage of power during an island forming situation.

Chapter 3

Discovered Initiators of Unintentional Islanding

3.1 P-Q Mismatch and Islanding

The degree of active power mismatch $\triangle P$ and reactive power mismatch $\triangle Q$ between the loads and the PV inverter on a feeder section has a close relation with islanding of the section. The behaviour of the islanded portion of a network is determined by the level of $\triangle P$ and $\triangle Q$ as described in section 1.2.2. The voltage magnitude and frequency of the supply in the island are governed by the interaction between the loads and the PV inverter. If the mismatch rises to a level such that the inverter operating limits of voltage and frequency are exceeded, the island collapses. If the mismatch level is such that the resulting V_{island} and f_{island} are within the Non-Detection Zone (NDZ), as defined in section 1.2.2, the island continues to operate and the inverter in such a situation is said to be in 'run-on' mode. The value of P-Q mismatch at the instant of isolation of the feeder section actually determines whether the isolated section will form an island or not. Once an island is formed, the level of $\triangle P$ and $\triangle Q$ determines whether it will sustain itself or collapse and in how much expected duration.

The various international anti-islanding standards described in chapter 2 also mention of regulating the levels of P-Q mismatch as an important step in the standard tests prescribed by them. Subjecting the PV inverter to different incremental levels of $\triangle P$ and $\triangle Q$ is a part of the procedure for testing the performance of the inverter's anti-islanding method in different conditions. Many simulation, laboratory based and field studies reported in the literature have tested the various

anti-islanding approaches proposed by them in the worst conditions of P-Q mismatch. The closely matched power situation of $\triangle P \rightarrow 0$ and $\triangle Q \rightarrow 0$ represents the toughest condition for an inverter to detect an islanding state. The field tests in [14] resulted into an interesting finding. The study concluded that the dynamic nature of loads could itself force them into a P-Q match condition with the inverter, on the basis of the resulting voltage and frequency, just after section isolation, leading to sustenance of an unintended island.

Few of the practical studies reported in [73], [74] and [75] have characterized PV inverters' islanding behavior for different levels of P-Q mismatch on feeder sections. However in the expected scenario of high PV penetration levels on distribution feeders, complete P-Q match can not be ruled-out as a remote possibility. Internationally accepted reports like [15] have also acknowledged that P-Q balance between PV inverter and loads is a possibility that can be quantified. Field tests in [14] and laboratory tests in [76] have studied and assessed the inverters' anti-islanding capabilities for the case of $\triangle P = 0$ and $\triangle Q = 0$. In a further study done for a spot type distribution network in India in [61], the risk of unintentional islanding was estimated by keeping an account of the number of hours for which this condition occurs. These documented practices have highlighted the significance of power mismatch levels for the condition of islanding, but this correspondence has been studied for the situation after the occurrence of the event. This study investigates the impact of a complete P-Q match case, coupled with some other conditions, on the possibility of creation of an imminent islanding situation.

It is clear from the above-mentioned points that P-Q mismatch plays an important role both before a feeder section's isolation and during the islanding condition. However, the P-Q interaction in high PV penetration levels can also lead to situations that may cause section islanding. Combined with the model of loads' power consumption variation with voltage and frequency, some levels of P-Q mismatch may lead to anomalous power flows during grid-end disturbances. Such power flows can transform into over-current spikes capable of activating the utility side protection at the PCC. The same condition was observed on the different radial feeder models used in this study. The next section describes the modeling of a modified IEEE feeder model. The modifications were mostly made in the original model to enforce different levels of P-Q mismatch during event simulation to observe the impacts on section islanding. The resulting anomalous over-currents have been described alongside verification of one them as an island-initiating anomaly.

3.2 Simulation Verification of the Observed Islanding Initiator

High localized penetration of solar PV generation on radial feeders can lead to anomalous power flows. Such occurrences get magnified in the presence of system disturbances and many other component-level parameters also affect their intensity. The PCC protection is vulnerable to any resulting over-current and this can trigger section islanding. Two of such over-current anomalies observed in simulation on a modified IEEE feeder model have been described in this section. The feeder modeling, observed over-currents and verification of one of the over-currents as an islanding-initiator is described in the following subsections.

3.2.1 Modified IEEE 13 Node Feeder Simulink Model

A distribution network model was required to be created for modeling the impacts of high PV penetration and to observe the resulting power flows. A simulation based study was thus appropriate as a preliminary step towards exploring the same. Accordingly, the IEEE 13 node feeder was modeled in Simulink platform of MATLAB with some modifications that seemed necessary for carrying the exploratory study. This feeder model was particularly chosen because it represents a complex network for carrying out studies relating to impacts of high PV penetrations on small feeders. Although it is small, it is highly unbalanced and provides a challenging test-bed for islanding-related studies and favourable results obtained for this feeder can help in scaling the study to larger networks.

Among the foremost additions to this feeder model, a 100.7 kWp solar PV array was integrated at node 692 through a three-phase inverter. This capacity of the solar PV based DG was chosen because it is a typical value used in small-sized commercial projects. Based on this value, the load ratings were adjusted keeping in view the intention of enforcing near 100% penetration on the section 671-692. Specifically for node 692, $P_{\text{load}} = P_{\text{PV}}$ and $Q_{\text{load}} = Q_{\text{671-675 Cap. bank}} + Q_{\text{inv. filter ckt.}}$. Node 692 was chosen as the PCC because of two reasons. First, the results of the Simulink power flow study agreed with the benchmark results that this node has one of the lowest voltages, although the values were not found to be same in both results. Therefore to improve the voltage profile by active power injection, the PV system was added. Secondly, the section 692-675 has a switch whose location conveniently facilitates the creation of an islanding situation in simulation.

The automatic voltage regulator connected at the substation node 650 in the original feeder was not modeled. This was done intentionally so as to avoid any possible dynamic interaction of the PV system with the tap-changing controls of the AVR. This is a phenomenon often practically observed in the field [6] and modeling it could affect the recording of certain patterns related to islanding on the feeder. To support the change's justification, it was found in the technical documentation that usually no On-Load Tap Changer (OLTC) action is seen in medium or low voltage transformers on distribution feeders. Also, when no OLTC exists, the voltage rise effect limits the maximum PV penetration [77] and thus the assumption was considered to be safe. A recent study assessing PV induced over voltage on distribution feeders, reported in [78], has also presented its findings on the IEEE 13 node feeder without considering the AVR in its model. The single line diagram of the modified IEEE 13 node feeder is shown in figure 3.1.

The rest of the changes were made in the power consumption of the loads on section 692-675. The constant-current load at node 675 was removed and the P and Q demands of the constant-current load at node 692 were adjusted to observe the desired load-PV interactions at high penetrations. The modeled PV inverter operates on unity power factor and thus to attain the P-Q balance, the loads had to be scaled according to the chosen PV capacity and the fixed feeder capacitor bank size. In the inverter model, the DC-AC conversion stage is preceded by a DC-DC converter that regulates the DC link voltage to a value of 500 V at the inverter input terminals. An incremental conductance based MPPT algorithm having an integral regulator method has been used to track the maximum power point of the PV array. A detailed schematic is shown in figure 3.2. The three phase inverter is

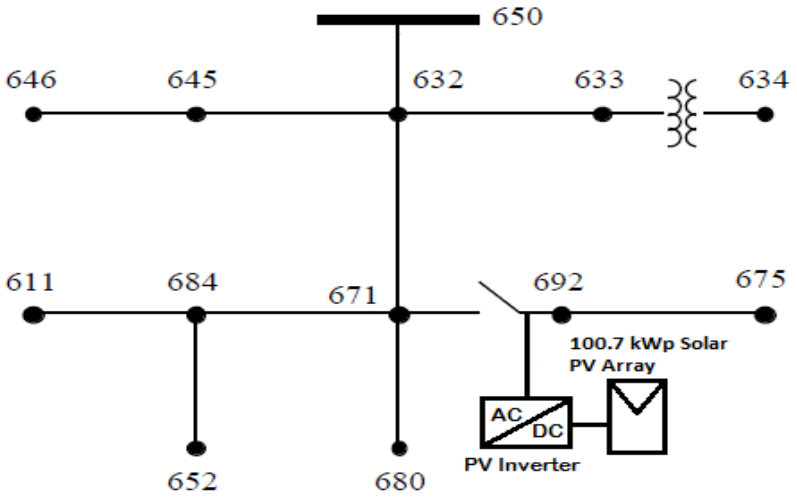

Figure 3.1: The modified IEEE 13 node feeder

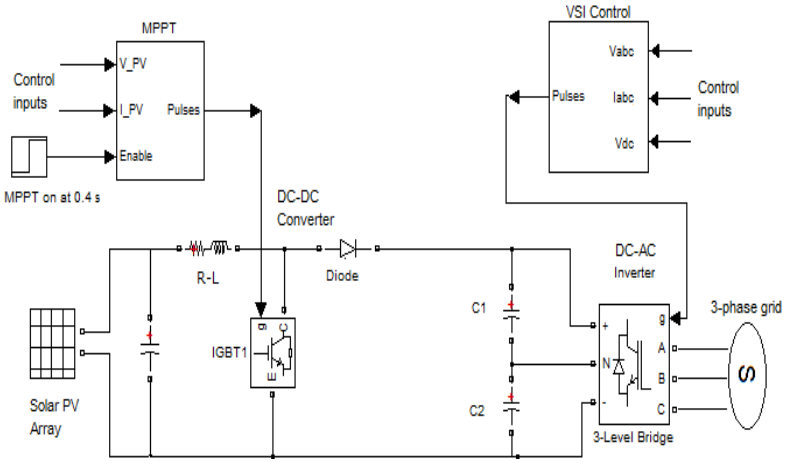

Figure 3.2: Schematic diagram of the modeled PV inverter

a voltage control based current controlled CSI interfaced to the feeder via a 100 kVA coupling transformer. The section 692-675 in the modified feeder model now contains the three-phase inverter, a single-phase constant-current load and a 600 kVAr capacitor bank as shown in figure 3.3. The capacitor inside the filter circuit of the inverter plays an important role in enforcing P-Q balance as discussed in the next paragraph.

To verify the modeling approach, the operation of the system was checked in islanded mode to see whether the parameter variation is in accordance with the theory. The section 671-692 was isolated with the PV inverter and the resulting voltage and frequency were checked for a P-Q mismatch condition. For simulating the same, the values of P_{Load} and Q_{Load} and those corresponding to the inverter (along with a R-C filter circuit) and the capacitor bank in the section were set as follows: $P_{\text{load}} = 90$ kW, $Q_{\text{load}} = 151$ kVAr; $P_{\text{PV}} = 100$ kW AC at peak radiation,$Q_{\text{PV}} = 0$; $Q_{\text{capacitive}} = Q_{\text{Cap. bank}} + Q_{\text{inv. filter ckt.}} = 600$ kVAr + 10 kVAr = 610 kVAr. It must be recalled that the load is single-phase while the inverter

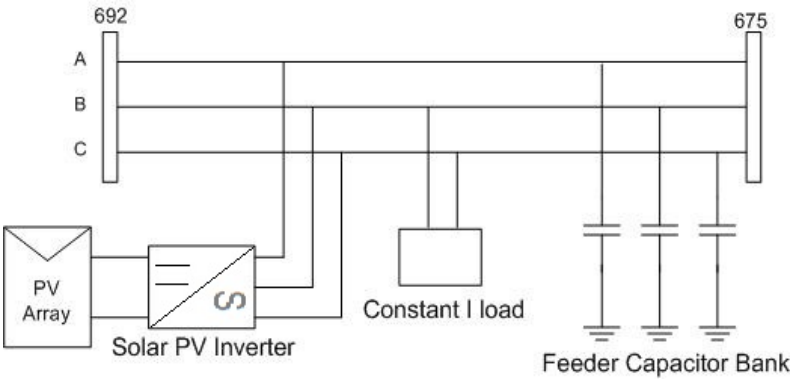

Figure 3.3: The section 692-675 of the modified feeder

and the capacitor bank are three-phase devices. The islanding switch was opened for a period of 30 ms starting at t = 0.45 seconds and reclosed at t = 0.48 seconds, from start of $t_{\text{simulation}}$. The solar irradiance in all the events simulated in this study has been kept fixed at 1000 $\frac{W}{m^2}$. This was done to ensure a static frame of study in which the PV system operates at its full capacity rated at the Standard Test Conditions (STC). For all the simulations discussed in this paper, the MPPT of the PV array is switched on at t = 0.4 seconds and after a few transients, reaches the peak operating power point of the PV array. The resulting islanding operation has been shown in figure 3.4. It can be observed that the isolated network's voltage and frequency pertain to the values explained in the basic theory of islanding of solar PV systems. The resulting under-voltage and under-frequency agrees with the theoretical trends applicable for the case of P-Q mismatch during islanding. The highly unbalanced nature of the feeder is also evident. The scale of the figure apparently suggests of no possible impacts of non-inclusion of an AVR since the observed three phase voltage remains at its nominal value. Also, since the major objective of this exercise was to observe the trends and the behaviour via the waveform, slight changes in the expected magnitude shall not affect the trend-related information carried in the signal's shape. The harmonics observed in the three phase voltages and currents are natural when a PV system is integrated and the $I_{\text{grid side}}$ harmonics averaged to be inside allowable range. The simulation modeled a discrete circuit by sampling the voltages and currents at a rate of 1 MHz. Although such a high sampling rate is not practically seen in power systems, this was used to closely approximate a continuous system in the discrete time-step simulation. Similar kind of wave shapes were observed in field tests car-

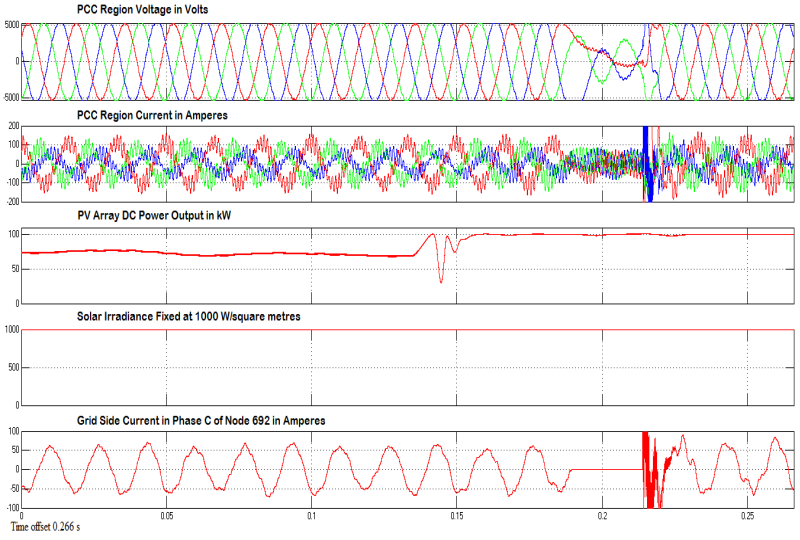

Figure 3.4: Islanding operation in the P-Q mismatch case

ried on a MV feeder in Spain [20]. The island reported over-voltages are visible in the field-test results reproduced with permission in figure 3.5. The time-scale of observing voltages and currents in the field tests is 10 ms/division. Although, comparing simulation results with field-results may not be completely appropriate, the similarity in the general trend of the obtained system parameters does provide confidence about the system modeling approach.

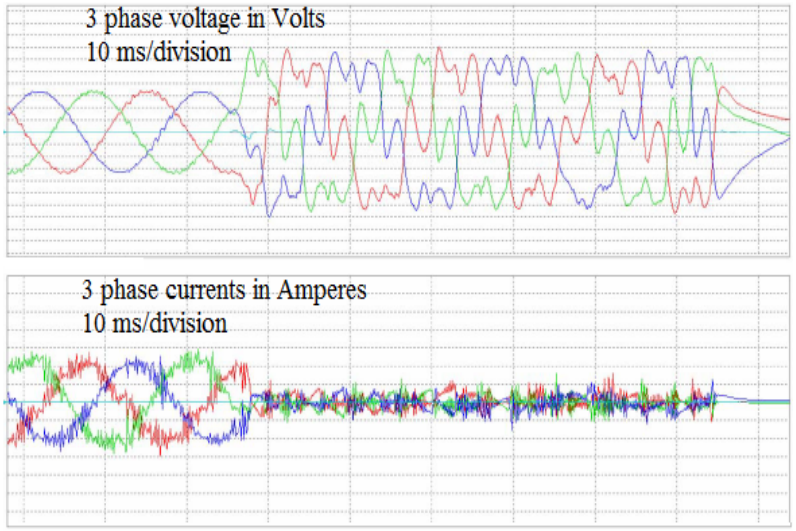

Figure 3.5: Field test results as in [20]

3.2.2 Anomalous current liable to cause section islanding

Two types of dynamics are commonly observed on any distribution feeder: disturbances from the grid substation side and variation of P_{load} and Q_{load} with the bus voltage and frequency. The injection of relatively large quantities of solar PV generation will add to the dynamic interactions between the inverter and the loads. Such occurrences alongside different grid conditions, coupled with load-model dynamics, can cause abnormal power flows that may trip the protective devices at the PCC. To investigate the impact of such events on the modified feeder model, two types of disturbances were simulated from the utility substation side - under voltage and over voltage, in concurrence with exact P-Q balance. Reportedly, these two types of disturbances are commonly used in practice for carrying out islanding related studies [79]. In either case, the following values were set to enforce complete P-Q match between the 1 phase load and the PV inverter and $Q_{\text{capacitive}}$ on the feeder section 671-692: $P_{\text{Load}} = 23.33$ kW and $Q_{\text{Load}} = 203.33$ kVAr. A 10 kW resistor in parallel with this load model creates exact P match as the P and Q

demands of the inverter filter circuit are negligible. The values of P_{PV} and Q_{supply} were kept the same as used in the P-Q mismatch case earlier.

The modeled solar PV inverter was given a lower value of minimum operating voltage which is slightly below the nominal floor values given in standards (0.8 p.u. to 1.15 p.u.). This is however comfortably outside the Low Voltage Ride Through (LVRT) operating zone as given in [80]. The intention was to provide flexibility in the operating limits in view of the recent developments in adoption of LVRT for PV inverters to attain uninterrupted distribution system operation. However, for the purpose of this study, the magnitude in the simulated under-voltage event was kept within the operating range of the inverter (0.7 p.u. to 1.3 p.u.) so that no LVRT functionality, if incorporated in the future, is activated. The modeled inverter cannot support the grid-voltage however even if this advanced functionality is ever incorporated, the time lapsed $t_{\text{voltage restoration}}$ is suspected to be $> t_{\text{identifying anomaly}}$. The island-initiating anomaly was induced by the combined effect of the grid-end low-voltage disturbance and the load model's relation of power consumption with V_{bus}.

The simulated under-voltage disturbance was 0.7 p.u. of the nominal voltage amplitude and was programmed to occur from the substation end for 30 ms from t = 0.45 s to t = 0.48 s with $t_{\text{simulation}} = 0$ s as reference. The low voltage event simulated in this study is not attributed to loading on the feeder. Such kind of voltage sag can be caused by practical reasons like generation-end short-circuit faults and brownouts on a feeder-lateral. Recently, a new phenomenon was reported in the field that caused voltage fall due to the action of inverter's anti-islanding protection [18]. All the disturbances in this work have been simulated for a period of 30 ms considering the inverse time-current characteristics of a commercial digital over-current relay model SEL-751 having directional over-current protection. This relay's catalog specifications show a $t_{\text{max. pickup}}$ of 8 ms. Accordingly the 30 ms nominal period was considered as a reference so that any over-current instance has a duration that is of the order of this value. Also, the $t_{\text{sampling}} < t_{\text{max. pickup}}$ which ensures that if the previous condition is met, the current can trip the relay. Since the P-Q match condition had already been configured on the feeder and the under-voltage was simulated alongside, this event has been named as UV+P-Q match and will be referred to as the same throughout the thesis. The resulting grid-side current flowing in phase C of section 671-692 due to this event's simulation run is shown shown in figure 3.6. The current has many anomalous peaks during the period of voltage disturbance and after the disturbance ends. The islanding-initiator is observed in the UV+P-Q after disturbance stage. It must also be noted that

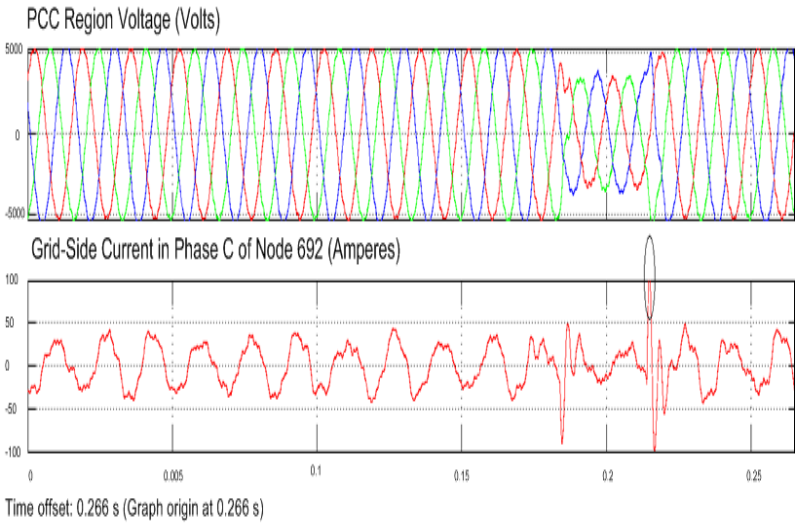

Figure 3.6: The discovered islanding-initiator

the instantaneous L-G voltages in phase C in the period from 0.48 s to 0.512 s are of the order of 0.85 p.u. which is well within the inverter's operating range.

The single phase load on the section 671-692 has a constant-current model that follows the relation $P \propto V_{\text{bus}}^{n}$ where $n = 1$. Because of the substation-side under-voltage disturbance, the low voltage propagates at the PCC bus and the P-Q balance is broken. There is now an excess of power from the PV inverter that begins flowing in reverse, towards the lower potential at the substation. Once the condition of $||V_{\text{bus}}|| = 1$ p.u. is satisfied, the superposition of the phasors $V_{\text{grid}}(t) + V_{\text{PV reverse}}(t)$ gets magnified. This magnification causes an instantaneous spike in the phase C voltage resulting into the spike in the corresponding current as shown in the figure by a circle. The obtained anomaly is consistent with the reported observation that reverse power flow leads to over voltages in high-PV penetration feeders even when the AVRs might not be present or not working properly [77]. Considering the magnitude and duration of this over-current, how this spike can become an island-initiating trigger has been explained next.

For the feeder model used in this study, an event will be called island-initiating if its intensity and duration can trip the utility protection near 692. For radial feeders, generally over-current relays and lateral fuses are used as protective devices [81]. However due to PV integration effects, the use of Directional Over-Current Relay (DOCR) at the PCC is recommended in practice. Since the observed anomaly is an over-current spike, its RMS value and duration must match with the trip-settings and time-current characteristics of the PCC protection. However, the feeder circuit has no actual protective device modeled. In a static frame, the comparison of the

over-current's magnitude with a calculated threshold has been followed to check whether the spike would trip the device, if it were kept at the PCC. To verify the severity of this current, the possible threshold for a protective device in this section was estimated. The addition of a solar PV inverter on a radial feeder brings an additional source of I_{fault} in the system. This impacts the magnitude and direction of fault currents flowing in the network. The radial nature of the feeder is lost since fault current from PV can flow in reverse for a short-circuit fault at an upstream node. The traditional system protection is not designed for such reverse flows and thus their coordination is disturbed.

The magnitude and direction of the maximum possible fault current at node 692 will decide the setting of the over-current relay to be placed there. Since the over-current anomaly is caused due to reverse power flow, a DOCR must be preferably chosen as the PCC protection. Accordingly, a three phase to ground short-circuit fault was created at the substation bus. The fault current was recorded in the section 671-692. This was done following the observations made in a short-circuit study performed for the IEEE 13 node feeder in [82]. The study found that the maximum fault current at the PCC was recorded when the fault occurred at a node farthest from it. For our network, node 650 is farthest from node 692 and hence for a L-L-L-G fault at 650, the RMS value of $I_{\text{fault 692}}$ was found to be less than that for the over-current spike which is 0.1 kA. Therefore, the current in the UV+P-Q after-disturbance event exceeds the threshold estimated for a DOCR at node 692. Also, the magnitude of this spike is more than the maximum continuous current $I_{\text{max. 671-692}} = V_{\text{L-N peak}}/Z_{671-675} \sim 60$ A. This is more than 1.3 times $I_{\text{max. load}}$ which defines the $I_{\text{pick up}}$ and hence the severity of the discovered over-current spike is again confirmed. Although $t_{\text{sampling}} < t_{\text{max. pickup}}$, the duration of the over-current spike is 0.6 ms based on the used sampling rate. Such a small duration spike cannot trip a DOCR as it does not last long enough to match the relay's time-current characteristics. However, this spike can blow a lateral fuse or trip a recloser placed before it for the fuse-saving scheme. Since a lateral fuse is also commonly used in radial feeders, the status of the discovered anomaly as an islanding-initiator is confirmed in simulation.

With an objective of investigating more of such anomalous occurrences due to multiple dynamics in active distribution networks, an over-voltage + P-Q match event was simulated. This was done to observe the behaviour of the modeled inverter, having a liberal operating range, at the other extreme of the wide operating voltage range. A voltage-rise, symmetrically 1.3 p.u. of the nominal voltage amplitude, was programmed from the substation-end for a period of 30 ms from

t = 0.45 s to 0.48 s. The power settings were kept the same as in previous cases so as to enforce P-Q balance alongside the disturbance occurrence. The resulting grid-side current is shown in figure 3.7. There are characteristic sharp peaks and valleys in this current during the period of over-voltage disturbance. Its magnitude however, could not cross the previously calculated threshold. Hence, based on the same static fault study approach, this over-current was not found to be severe enough to trigger section islanding. Traditionally, voltage-swells on distribution feeders occur due to reasons like Ferranti-effect. Modeling such a high voltage-rise was motivated by the fact that over-voltages due to PV integration effects are also becoming a common phenomena and their magnitudes have not been characterized.

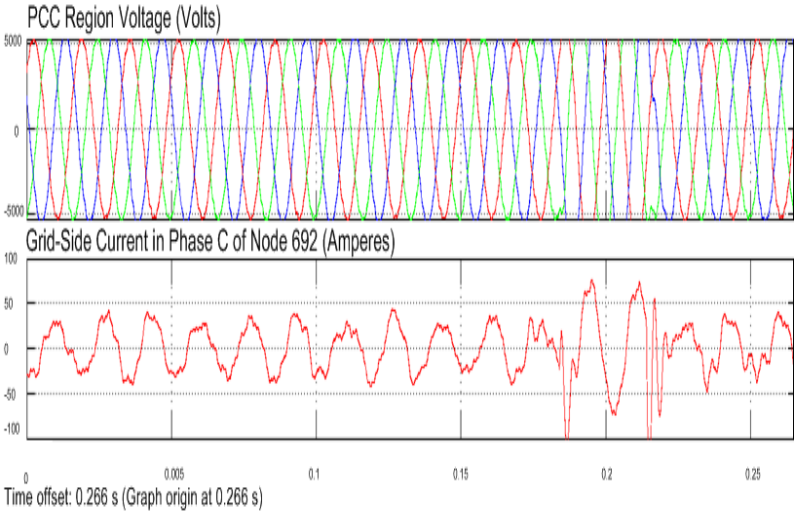

Figure 3.7: OV+P-Q match

3.2.3 RTDS Verification of the Observed Islanding Initiator

Since a simulation based study provides the control over system parameters, it allows convenient investigation of events facilitating extensive collection of data. Accordingly, the simulation based study was carried-forward in a realistic framework by running the model in real time. The complete system involving all the components was re-modeled and executed in a Real Time Digital Simulator (RTDS). Both the events discussed above were simulated. The Simulink model was re-structured in a master-slave configuration with inclusion of block programs necessary to interface the Simulink system with the RTDS mainframe. A RTDS provides both hardware-in-loop and processor-in-loop capabilities to execute models designed in

Simulink and various other simulation softwares. Rapid control prototyping in real time is also supported. An RTDS contains a collection of powerful processors arranged in a rack. Each processor is a multi-core central processing unit usually running at 3.2 GHz and above with upto 32 cores.

An OPAL-RT make RTDS (model OP-5600) available at the Department of Electrical Engineering, Indian Institute of Technology Jodhpur was used in the study. The RTDS rack has been shown in figure 3.8. The system with an oscilloscope attached was used to observe the real-time waveforms and has been shown in figure 3.9. The system comes with its own modeling software known as RT-LAB. The Simulink model was re-designed in this software where the model blocks that interface the Simulink model with the RTDS processor were added. Since real-time simulations cannot adjust sampling times dynamically, the Simulink models have to be designed with a fixed time-step of simulation. Accordingly, the model configuration was changed to fixed time-step discrete and it was run with a sample time of 10 μs. This is the lowest resolution of sampling time possible on the RTDS used for the study.

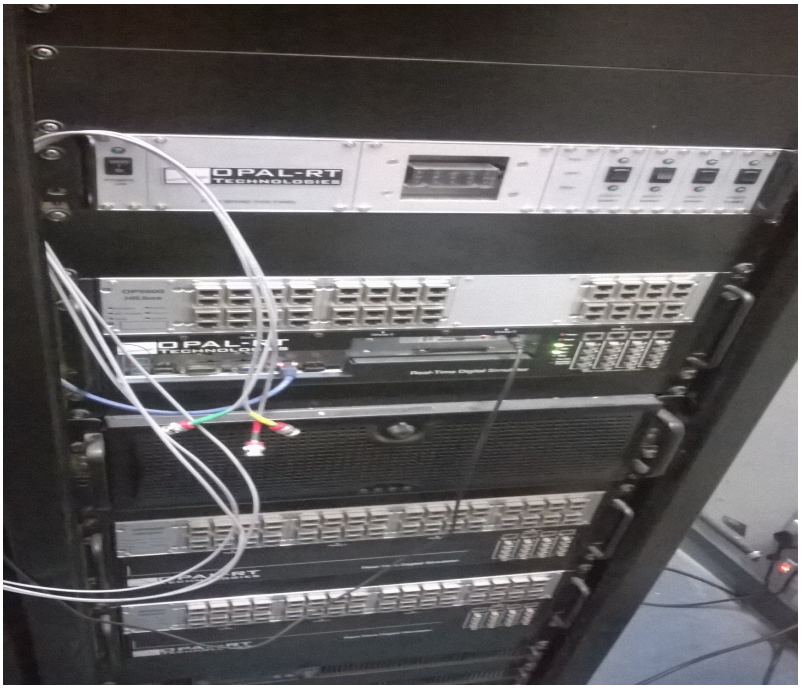

Figure 3.8: The RTDS rack

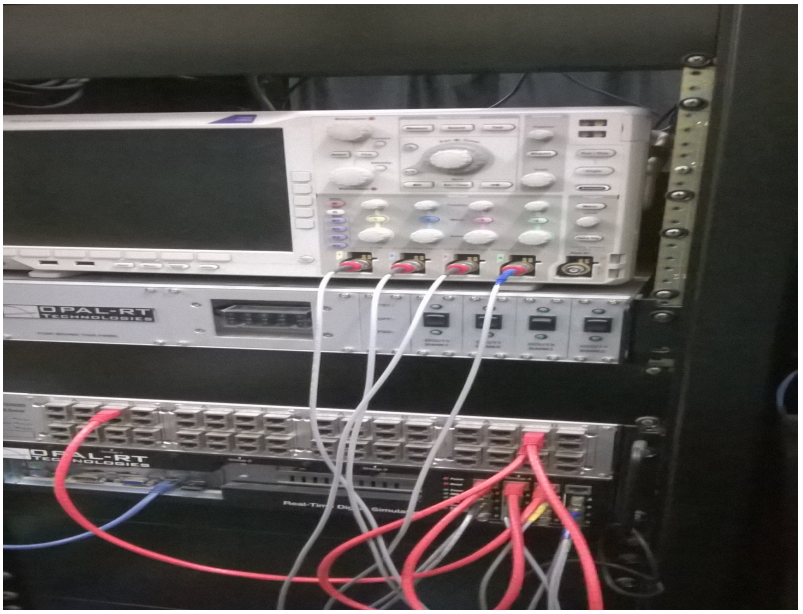

Figure 3.9: The oscilloscope connected to OPAL-RT OP-5600

The system was configured to have a stop-time that makes the duration of a run as 0.5 s. Both the UV+P-Q match and OV+P-Q match cases were simulated for 0.5 s in real time. The model was compiled in Simulink on the personal workstation and the execution of the compiled C code was done in the RTDS. The results generated by the RTDS were communicated to the personal workstation and were saved in the MATLAB workspace from which the waveshapes were plotted as shown in figures 3.10 and 3.11 respectively. The spikes in the initial phase are due to the PV system integration transients.

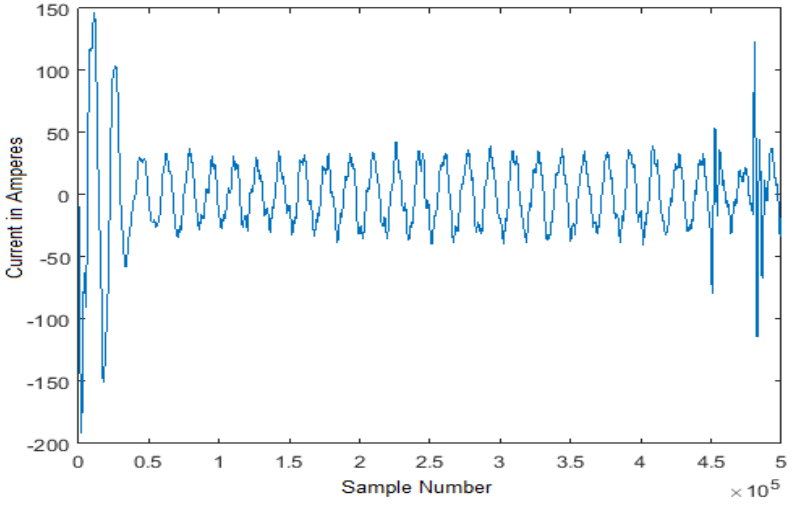

Figure 3.10: RTDS result for the UV+P-Q match case

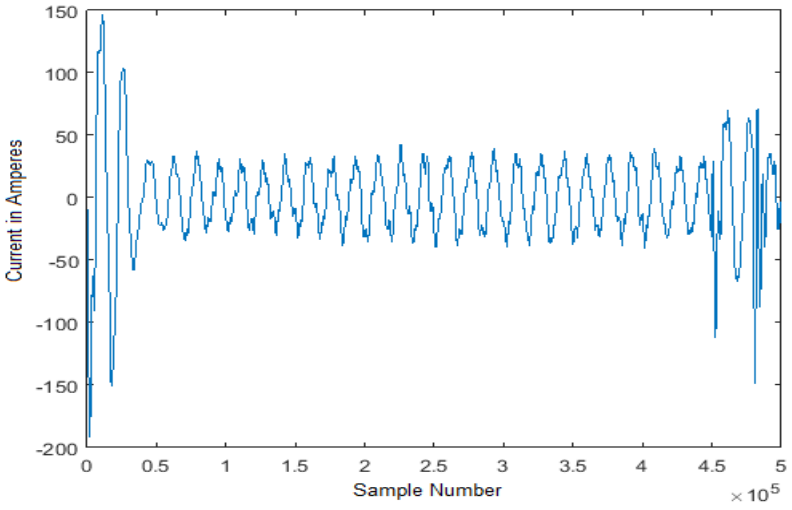

Figure 3.11: RTDS result for the OV+P-Q match case

3.3 Emulator Network Verification

A discovered anomalous phenomenon must be network independent. Thus the over-current spike observed for the UV+P-Q match event on the modified IEEE 13 node feeder must be replicable. It must be observed if the same set of conditions that cause it occur on a different radial feeder. Accordingly, a one-phase, one-bus network was created using emulators and the associated hardware in the Smart Controllers laboratory of the Centre for Distributed Generation at The Energy and Resources Institute, New Delhi, India to verify the spike's occurrence. There were some differences from the settings configured for the simulation study. Firstly, there was no source of reactive power so only P-match condition could be realized. Secondly, the inverter used was a commercial model (SMA Sunny Central) whose operating voltage range was 0.8 p.u. to 1.15 p.u. and which worked on unity power factor. The power grid was realized using a grid-simulator which could operate only in 1-phase mode due to some technical issues. The system frequency was 50 Hz unlike the simulation study. The solar PV array was replaced by a 30 kW solar PV array simulator whose radiation was also kept fixed at 1000 W/m^2. A programmable load or the load emulator provided a constant impedance type load model however only a purely resistive load was implemented due to the availability of an active power supply only. The feeder model realized in hardware is shown in figure 3.12.

The grid-simulator used was a 30 kVA programmable voltage source from which all kinds of voltage disturbances can be programmed. The 5 kVA inverter functioned as a typical CSI with a pre-configured operating margin. The radial feeder

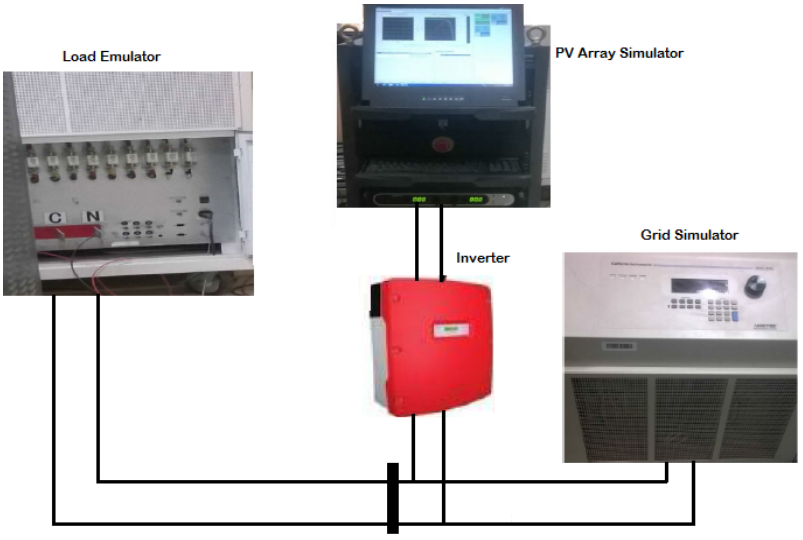

Figure 3.12: The laboratory emulator network

was created using 2.5 mm^2, three core single phase cables with effective resistance of 5 Ω and negligible inductance. The load values were set on a workstation and communicated to the load emulator through RS-485 protocol. The PV inverter was supplied with the solar PV emulator output corresponding to the STC programmed from the interactive dashboard. The loads were fine tuned to a value such that their consumption matched the inverter's output. An under-voltage disturbance of 80% of the nominal one-phase grid-voltage of 230 V was programmed on the dashboard of the grid-simulator. The resulting current as observed on the digital oscilloscope is shown in figure 3.13.

A similar peak in the grid-side current was observed however it occurred during the

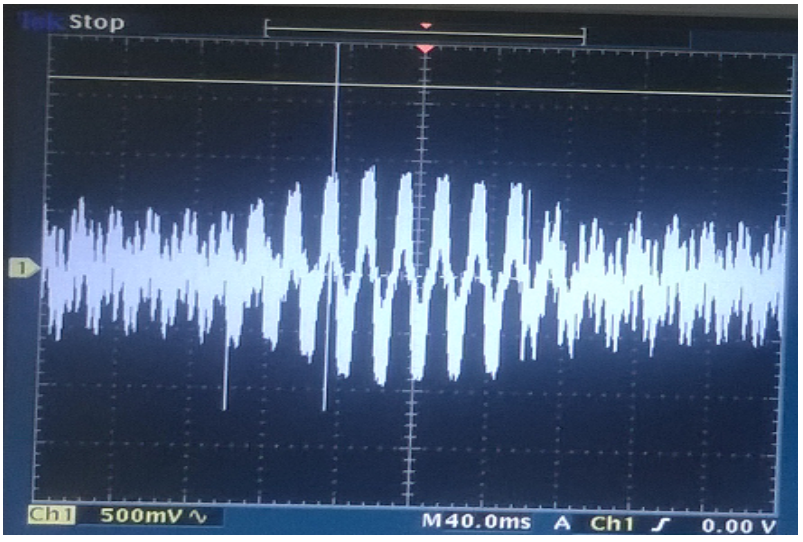

Figure 3.13: The over-current observed on the emulator network

period of under-voltage disturbance. This was because the load model in the laboratory was that of the constant-impedance type whose $P \propto V_{\text{bus}}^2$ led to an abrupt rise in excess reverse power from the PV inverter, inside the disturbance duration, thus causing the early occurrence of the peak. Hence $P_{\text{net flow}} = P_{\text{grid}} + P_{\text{PV reverse}}$ caused net flow in reverse that caused the sudden rise in the current. A replica of the emulators-based network was designed in Simulink for cross-verification of the emulator findings. Since Simulink does not contain a model block for a programmable single phase voltage source, a three phase network had to be created. The relevant parameters were scaled accordingly to replicate the laboratory network in the three phase domain.

The nominal system voltage was changed to 400 V rms L-L and the network resistance was changed to 15 Ω but a small 1 mH inductance was added. The models of the PV array and the three phase inverter used for the modified IEEE 13 node feeder model were used in designing this network also. However the operating range of this inverter model was different from that pre-configured for the inverter used in the laboratory. A 100 kW resistive load, as used in the laboratory study, was added having a constant impedance model characteristics. The Simulink model was run as a discrete-time step circuit for 0.5 seconds with a sampling rate of 1 MHz for reasons stated above. In the simulation run, the MPPT was switched on at t = 0.4 s and after a few transients, the P-Q balance was achieved at t = 0.42 s. To simulate the action taken in the laboratory study, an under voltage disturbance of 0.7 p.u. was introduced at 0.42 seconds, immediately after achieving the power match condition. The resulting patterns in the grid-side current are shown in figure 3.14.

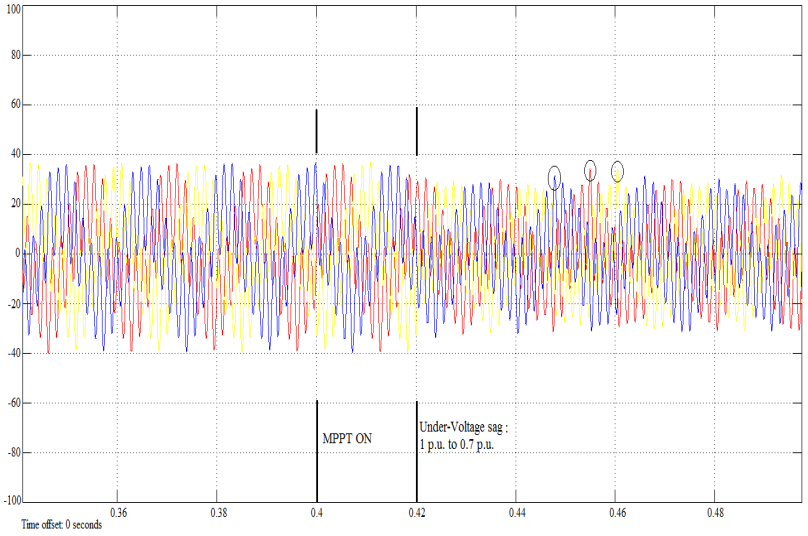

Figure 3.14: Results for the Simulink study of the emulator network

Each of the individual phase currents resemble the wave-shape obtained in the hardware study. The spikes occur inside the duration of the under-voltage disturbance. This occurrence is attributed to the constant-impedance characteristics of the load that resulted into the reverse flow of the excess PV power during the voltage-sag period. Since the simulation was run for 0.5 seconds only, the progressive wave forms could not be observed. It is intuited that more of such peaks will repeat in a sequence. To find whether the obtained current can trigger isolation of the section of this radial feeder model, the same static fault study approach was adopted. A three phase to ground short-circuit fault was created at the point of interconnection of the PV inverter with the network. Since the fault could be created at any instant after t = 0.42 s, it was programmed to occur at t = 0.45 s persisting for a period of 30 ms till t = 0.48 s. The process was repeated for the case when the PV inverter was removed from the model. The fault current obtained in both the cases were sinusoidal. The peak value of the sinusoidal fault current for the case when PV inverter was connected came to be less than the value obtained in the case of no PV interconnection. This is in agreement with the results of the static fault study done for the modified IEEE 13 node feeder model.

The RMS value of the over-current spike due to the UV+P-Q match event was comparable with the fault current RMS value of 0.02 kA. Thus, in a manner similar to the verification done for the case of the modified IEEE feeder, the island-initiating status of the emulator study anomaly was assessed.

3.4 Conclusion

This chapter described the investigation of the impacts of high localized-penetration of solar PV generation on a modeled radial feeder. The IEEE 13 node feeder was modeled in Simulink with some modifications to carry out the investigative study as desired. Anomalous instances of over-currents were found for two cases: P-Q match between loads and PV inverter alongside under-voltage and over-voltage disturbances from the grid-side. The over-current observed in the after-disturbance period of the UV+P-Q match event was found to island the section 671-692 on the feeder. A static fault study was performed to verify the status of this over-current as an islanding-initiator for the modeled feeder. The discovered phenomenon's occurrence was attributed to the reverse flow of excess PV power that lead to an instantaneous spike in phase C current of the section. The constant-current

model of the 1-phase load in the section caused a sudden reduction in active power consumption due to the voltage-sag. This disrupted the P-Q balance and the magnitude of superposition of $V_{\text{PV excess}} + V_{grid}$ peaked after the disturbance was removed. The post-disturbance over-current observed for the OV+P-Q match event was not found to be an islanding-initiator for the section 671-692 based on the same static fault study.

To have a realistic feel of the simulation, both the events were simulated on a RTDS. The same nature of waveforms were observed for both the UV+P-Q and OV+P-Q match cases confirming the occurrence of the anomalous current spikes in real-time simulation. For the same set of conditions corresponding to a UV+P-Q match event, the over-current spike was also observed on a 1-ϕ, 1-bus network created using emulators in a laboratory setup. However, due to the constant-Z model of the resistive loads, the spike was observed during the voltage-sag period, verifying the phenomenon's occurrence and displaying the impacts of various load models.

The next study, presented in the successive chapter, initiates the use of learning models for event detection. It describes how the data points belonging to the two events discovered on the IEEE feeder were used to extract features to fit different learning models. The different learning models implemented are based on multivariate statistical methods and supervised and unsupervised machine learning. Their common objective is to identify the island-initiating current from other transient events. However, the detection is offline as the data is saved first and then the prediction models are applied.

Chapter 4

Offline Feature Extraction and Event Detection

4.1 PCA Based Feature Extraction

The system designed on the modified IEEE 13 bus feeder was configured as a discrete time-step model in Simulink. The sampling time was kept as 1μs and different event based simulations including under-voltage + P-Q match and over-voltage + P-Q match cases, described in previous section, were run for a time duration of 0.5 seconds. The voltages and currents were thus sampled at 1 MHz for reasons given in the previous section. These simulation runs generated a huge amount of data points. Based on the sampling rate and the actual time at which the simulations were stopped, an order of magnitude of 50,000 data points were generated for each run. These are instantaneous values of current and voltage of phase C recorded at section 671-692. Thus a $N \times 2$ dimensional data matrix was saved for each event simulation run where N typically remained of the order of 50,000. The entries under both the columns showed a strong correlation as also exists practically between voltage and current and the phase difference present between them was also visible. Since in each event simulation run, the MPPT was switched on at t = 0.42 s, the active-power contribution from the PV system could be stabilized only after that. Each event simulation had two portions in general - the 'normal' system operation and the disturbance portion. The 'normal' system operation corresponds to the duration when no external disturbances exist in the system apart from the harmonics introduced by the PV inverter. For the UV+P-Q match and OV+P-Q match cases, both the 'during disturbance' and 'after-

disturbance' periods were used to collect data for building predictive models.

The 'normal' system operation portion was common to all the different types of event simulations. It was thus treated as a separate event in itself. Any predictive model built for this study carries the objective of identifying the island-initiating over-current from the UV+P-Q match case from other different kinds of power system transients - known ones and those discovered in this study. Hence a large amount of data points must be collected to form a composite matrix (that may or may not be having final outcome values) to fit a predictor model whose final outcome is the class or group to which an event has been identified to belong to. For this study, there are two possible classes where events can be categorized or grouped into :- island-initiating or not island-initiating. An appropriate step before building a predictive model is to reduce the dimensionality of the data. Since the data matrix for each event-simulation is quite large, the collection of such matrices will compound the number of instances. It is thus important to extract only meaningful information and discard data redundancy and keep only those attributes or dimensions that contribute to the information content. Accordingly, Principal Component Analysis (PCA) was thought to be an appropriate technique to extract features from the $N \times 2$ dimensional data since current and voltage are correlated quantities.

PCA is a linear-algebra based projection technique that maps highly correlated features into a low dimensional space. The resultant feature vectors in this sub-space are non-correlated and orthogonal to each other. Mathematically, it represents an orthonormal basis transformation wherein the input data vectors are projected onto new basis vectors of an m dimensional feature space, where m is the number of features in the original data. The best way to represent a given dataset is to select those features:
1. Along which the variance of data points is maximum
2. Whose covariance with other features is minimum.

These two conditions also translate into the desired properties of low noise and reduced redundancy. The basic mathematics of PCA can be explained as follows: Let $X \in R^m$ be a real-valued matrix having n samples or observations where each sample vector is a column. Hence X is arranged as a $m \times n$ data matrix which has to be represented by Y, of the same dimensions. Let P be a transformation matrix that converts X into a representation matrix Y such that the features of Y follow the two conditions mentioned above. Accordingly, P must be a $m \times m$

matrix that transforms X as in equation 4.1:

$$PX = Y \tag{4.1}$$

The covariance matrix of Y is given by equation 4.2:

$$Cov(Y) = \frac{1}{n-1} YY^T \tag{4.2}$$

The diagonal elements of $Cov(Y)$ represent the variance of each dimension while the off-diagonal elements are the covariances between different dimensions. To satisfy the two conditions given above, the values of diagonal elements must be high and those of the off-diagonal terms must be least. The best possible way to achieve this is to simply diagonalize this matrix.

If P can be represented as collection of row vectors $\begin{bmatrix} p_1 & p_2 ... & p_m \end{bmatrix}^T$

Then, Y can be expressed as in equation 4.3:

$$Y = \begin{bmatrix} p_1x_1 & p_1x_2 ... p_1x_n \\ p_2x_1 & p_2x_2 ... p_2x_n \\ . & . & . \\ . & . & . \\ p_nx_1 & p_nx_2 ... p_nx_n \end{bmatrix} \tag{4.3}$$

When, X is a collection of column vectors, $\begin{bmatrix} x_1 & x_2 ... & x_n \end{bmatrix}$.
It is evident from equation 4.3 that the vectors of X are projected onto the vectors of P which now act as new basis vectors. Since basis vectors of a Euclidean vector space are always orthogonal, they must now be orthonormal so that the projection represents only a rotation, not a stretch or change in length. Therefore, now the task is to find P such that $Cov(Y)$ becomes a diagonal matrix. Also, the essential requirement becomes that P must be a matrix of orthonormal vectors.

Using equation 4.1, equation 4.2 can be expanded into equation 4.4 as:

$$\begin{aligned} Cov(Y) &= \frac{1}{n-1} PX(PX)^T \\ Cov(Y) &= \frac{1}{n-1} PXX^TP^T \\ Cov(Y) &= \frac{1}{n-1} PCov(X)P^T \end{aligned} \tag{4.4}$$

To diagonalize $Cov(Y)$, the basic PCA algorithm takes advantage of the fact that

Eigen Value Decomposition (EVD) of a symmetric matrix results into orthogonal eigen vectors that can be normalized to unit length. $Cov(X)$ is an $m \times m$ symmetric matrix that can be expressed as in equation 4.5 based on the standard EVD:

$$Cov(X) = EDE^{-1} \tag{4.5}$$

Here, E is the matrix of eigen vectors of $Cov(X)$ and D is a diagonal matrix of its eigen values. The EVD based PCA suggests the choice of P as $P = E^T$. To make the eigen vectors of $Cov(X)$ orthonormal, the matrix is mean-centered along the dimensions so as to normalize the length to 1. This makes P an orthonormal matrix for whom $P^T = P^{-1}$. Using equation 4.5, equation 4.4 can be further expanded to result into equation 4.6 as:

$$\begin{aligned}
Cov(Y) &= \frac{1}{n-1} P(EDE^{-1})P^T \\
Cov(Y) &= \frac{1}{n-1} P(EDE^{-1})P^{-1} \\
Cov(Y) &= \frac{1}{n-1} P(P^T D P^{T-1})P^T \\
Cov(Y) &= \frac{1}{n-1} P(P^T D P^{-1-1})P^T \\
Cov(Y) &= \frac{1}{n-1} P(P^T DP)P^T \\
Cov(Y) &= \frac{1}{n-1} (PP^T)D(PP^T) \\
Cov(Y) &= \frac{1}{n-1} (PP^{-1})D(PP^{-1}) \\
Cov(Y) &= \frac{1}{n-1} D
\end{aligned} \tag{4.6}$$

Thus the selection of P as E^T made $Cov(Y)$ a diagonal matrix. There is another version of PCA in which the Singular Value Decomposition (SVD) of the input data matrix is used to diagonalize the required matrix. In this approach, the data matrix X is decomposed into two orthonormal and one diagonal matrix as in equation 4.7:

$$X = U\Gamma V^T \tag{4.7}$$

Here, U is a $n \times m$ matrix with orthonormal columns, V is a $m \times m$ orthonormal matrix and Γ is a $m \times m$ diagonal matrix of singular values. The columns of U are orthonormal eigen vectors of XX^T, columns of V are orthonormal eigenvectors of $X^T X$ and singular values contained in Γ are the square roots of eigen values from U or V in a descending order.

The basic PCA algorithm based on the EVD of covariance matrix can now be formulated. The same has been described in Algorithm 1.

Algorithm 1 PCA

1: Arrange data sample vectors into a data matrix X = samples $\times$ observations
2: Mean centering of $X \rightarrow$ mean centre X by subtracting column means from each observation along the columns to have zero means along the columns. This results into matrix A
3: Compute the covariance matrix of the mean-centered data matrix A by $\frac{A^T \times A}{n-1}$. It is a square symmetric matrix of $p \times p$ dimension
4: Find the eigen vectors of the covariance matrix of A
5: Sort the eigen vectors in decreasing order of eigen values
6: Project the sample vectors of A onto these eigen vectors to find their feature vectors for building prediction models

To re-iterate the basic vectors involved in this application of PCA:
Eigen vectors = principal component axes
Score vectors = projection of mean centered data samples onto the Principal Components (PCs)
Eigen values = variance of each score vector = latent values

The 'normal' system operation portion was taken as a reference case. Linear PCA technique was applied on the data set corresponding to the 'normal' case to create the reference PCA model. PCA based on the standard SVD approach was implemented in MATLAB to find the principal component matrix. The two-dimensional data belonging to the 'normal' case resulted in two principal components (PCs) as explained in equation 4.8 $\forall n = 1$ to N = |Normal|.

$$\begin{bmatrix} V(n) & I(n) \end{bmatrix}_{N \times 2} \rightarrow \begin{bmatrix} PC_1 & PC_2 \end{bmatrix}_{2 \times 2} \tag{4.8}$$

Latent analysis was used to select a component. A latent matrix contains the values of variances of projections of data points onto the PCs known as scores. Based on the variance of the projections onto the two PCs, the 1^{st} PC was retained for all analyses. The latent matrix is shown in table 4.1 that explains the choice of retaining PC_1. Thus the projection of mean-centered data points onto the 1^{st} principal component of the reference PCA model, $F_1(n)$, was taken as the feature attribute of data points belonging to different events as explained in equation 4.9 where each of the two PCs is a 2×1 vector.

$$\begin{bmatrix} V(n) & I(n) \end{bmatrix}_{N \times 2} \times \begin{bmatrix} PC_1 & PC_2 \end{bmatrix}_{2 \times 2} = \begin{bmatrix} F_1(n) & F_2(n) \end{bmatrix}_{N \times 2} \tag{4.9}$$

Table 4.1: Latent values for the reference PCA model

On PC_1	On PC_2
5.37 $\times 10^6$	486.6

Many kernel functions also exist that can transform the original data into feature vectors in the principal component sub-space non linearly. For N number of data sample vectors $x \in R^m$ stacked above one another to form a data matrix $X^{N \times m}$, application of PCA on X leads to a $p \times p$ coefficient matrix P. If $r \leq m$ number of PCs are retained based on latent values, then X can be resolved into a PCA model and a residual model as $X = X_{PCA} + X_{res}$. The projection onto the PC or loading matrix leads to formation of a score matrix $T^{N \times m}$. The original data matrix X can be reconstructed using score matrices T_{PCA} and T_{res} and loadings P_{PCA} and P_{res} as $X = T_{PCA}P_{PCA}^T + T_{res}P_{res}^T$ where P_{res} and T_{res} are of $N \times m - r$ dimension [83].

The application of PCA on the 'normal' system conditions dataset to form the reference PCA model resulted into two things - reduction of reference model dataset into a uni-dimensional matrix and a methodology to extract features from any data set for training and testing different predictive models. The next section discusses two multivariate statistical models that were used in conjunction with PCA for event detection purposes.

4.2 Multivariate Statistics based Event Detection

The PCA technique used above was applied in building models that can detect any abnormal instance or occurrence using statistical process control strategy. This approach is well suited for multivariate statistical models that can sense the statistical significance of deviation in a dataset. Although PCA itself is a multivariate technique, it has to be used with some statistical parameter that can define control limits for a process variable and sense its violation. Therefore, the PCA feature extraction methodology was used to implement two different types of multivariate statistics based event detection models discussed below. The data belonging to different events simulated was pre-processed and projected onto the 1^{st} PC of the reference PCA model. The aim of the underlying statistical process control strategy is to differentiate a condition that can cause unintentional

islanding on the modeled feeder from those like faults and other transients that appear similar to an islanding-initiator but actually are not. These events are thus tricky to detect and identify correctly. As highlighted in chapter 2, the techniques described in the literature detect an islanding condition among other transients like faults, load and capacitor switching surges, and transients after the island has been formed. This study explores the possible practical causes of the event, as described in chapter 3, and then applies models for detecting such conditions from the ones that appear close enough to fool the inverter. For the offline detection study discussed in this chapter, the four cases resulting from the two grid-side disturbances (described in chapter 3) and a three-phase short-circuit fault case have been simulated.

4.2.1 PCA + Q Statistics based Model

The different event based cases simulated for this offline study and the number of samples generated are given in table 4.2. The 'normal' system condition is an event in itself as mentioned above. A three phase to ground short-circuit fault was introduced at the PCC (node 692) with a fault resistance $R_f = 0.01\ \Omega$. This event was simulated for a 30 ms interval and only the disturbance period data points were collected. Thus $N \sim 30,000$ for this case. Similarly, for the 'during-disturbance' and 'after-disturbance' scenarios of the UV+P-Q match and OV+P-Q match cases, number of samples $N \sim 30,000$ were recorded respectively since both the scenarios were equally long having a 30 ms duration.

The statistical process control method is widely used in industrial engineering for quality control purposes. It has found other applications in many domains for outlier detection to check if the process variables are in control. Any process variable is indicated as out of control when a certain statistic related to it crosses its upper limit. PCA has two multivariate statistics associated with it: the Hotelling's T^2 statistic and the Q statistic. Both have an Upper Control Limit (UCL) defined in different ways and when these limits for both of them are crossed by the corresponding statistics of data points in a data set, it indicates an anomalous and abnormal behavior.

The Hotelling's T^2 statistic is a multivariate distance of a set of data points from a target value. It indicates the variance of data points inside the PCA model. If f is a mean-centered (scaled) sample data vector then $t_{PCA} = fP_{PCA}$ is a score vector. The T^2 statistic for f is defined as $T^2 = t' \wedge t$ where $\wedge$ is a diagonal matrix having

Table 4.2: Cases Simulated for the First Offline Study

Case	No. of Samples	Event
Case 1	30484	Normal
Case 2	30584	UV+P-Q match (during disturbance)
Case 3	30622	UV+P-Q match (after disturbance)
Case 4	30585	L-L-L-G fault at PCC
Case 5	30656	OV+P-Q match (during disturbance)
Case 6	30644	OV+P-Q match (after disturbance)

r Eigen values of data matrix $F^{N\times m}$ for $r \leq m$ number of retained PCs. The UCL for the statistic is denoted by T^2_α. If all data points are linear and normally distributed, the T^2_α follows an F distribution and is given as in equation 4.10

$$T^2_\alpha = r\frac{(N^2-1)}{N(N-r)}F_{r,N-r} \tag{4.10}$$

at a given level of confidence α. Here r and $N-r$ are the two degrees of freedom of the F statistical distribution.

The Q statistic is a measure of deviation of the original data points from the projection onto the PC axes. Hence it measures variance among data points inside the residual subspace. The Q statistic is calculated using residuals and for a residual vector e of a scaled sample vector f, the Q statistic is given as $Q = e^Te = f^T(I - P_{pca}P^T_{pca})$ where I is an identity matrix. For normally distributed and linearly co-related data points, the Q statistic follows a central χ^2 distribution and its UCL is given by equation 4.11

$$Q_\alpha = \frac{\sigma^2}{2\mu} \times \chi^2\left(\frac{2\mu^2}{\sigma^2}\right) \tag{4.11}$$

at given level of confidence α. Here μ and σ^2 are the mean and variance of the Q statistic respectively.

Recently, PCA based process control strategy has been applied for detecting the occurrence of an islanding condition and distinguishing it from several non-islanding events. PMU recordings of frequency measurements on 6 different sites in the UK power grid were used as reference data for implementing T^2 and Q statistic based islanding detection in [84]. The occurrence of an islanding situation was evident

only when the Q_α was crossed in addition to the crossing of T^2_α by the corresponding multivariate statistics for a test event data set. Since the power system is a dynamically varying system, the system variables used for creating the reference PCA model also change dynamically causing temporal variations in the model also. To tackle this issue, a recursive PCA algorithm was developed in [85] for the same UK power system case. The reference PCA model was updated in every iteration and the detection results for abnormal transients verified its effectiveness over the static PCA approach. This research thesis reports a study that has made use of the usual SVD for creating the reference PCA model since the reference data does not change from one event to other. This is so because the simulation has been performed for fixed settings to observe some unique changes that occur in fixed-time windows as described previously.

The application of the PCA + Q statistics model involves the calculation of Q value for each of the 5 events except for the one represented by case 1. Each data set $X^{samples \times 2} \in$ different test cases (2 to 6) underwent scaling to make the mean along the columns as zero. The mean-centered data set X_{mc} was then projected onto the 1^{st} PC of the reference PCA model as $X_{mc} \times P_{pca}$. Correspondingly, the T^2 and Q statistics were calculated. Since crossing of the T^2_α limit for the reference case by the T^2 statistic of any test case indicates only a faulty or out of control event, the Q statistic was used as the only parameter for detection. Q statistic measures deviation inside the residual subspace and hence is a strong indicator of any abnormal or anomalous condition.

Following the same, the 5 test cases were subjected to mean-centering as before and were projected onto the 1^{st} PC of the reference PCA model. The Q statistics for the projected data matrix for each case was found and compared with the UCL Q_α of the reference case score. The Q_α at 98% confidence level was calculated to be $= 3.846 \times 10^7$. For finding this value, the Q matrix was first found as $Q = \text{residuals}^T \times \text{residuals}$ where residuals is a $N \times 1$ vector containing values of the residuals of the reference PCA model. The mean and variance of the first column of the resulting matrix were used in equation 4.11 to calculate the same. The results of this multivariate statistics based detection are given in table 4.3.

As seen in table 4.3, this approach identifies the anomalous case (case 3) of the island-initiating over-current correctly. It also correctly identifies that the disturbance event in case 2 is not an anomaly that can island the system. However the cases 4, 5 and 6 are incorrectly identified. It must be noted that the L-L-L-G short-circuit fault at PCC (case 4) is not an islanding-initiator. This is because for

Table 4.3: Event detection results using Q statistic

Case	Q	$Q > Q_\alpha$	Islanding initiator
2	1.57×10^7	NO	NO
3	4.05×10^7	YES	YES
4	1.23×10^9	YES	NO
5	8.57×10^7	YES	NO
6	4.32×10^7	YES	NO

a given short-circuit capacity of the feeder, a short-circuit fault at the PCC will cause very low voltages which can cross the $V_{\text{inv. lower limit}}$. Thus the PV inverter will shut-down itself following the tripping action of the under-voltage relay and hence no PV generation will remain even if the section is isolated during the fault. Although the Q value for the data set belonging to this case is statistically the highest indicating out-of-control process, it is not an island-initiating event. This shows that the Q statistic based statistical process control approach is not completely reliable for detecting any anomalous currents liable to island a section on the system. To improve upon the false detection rate, the Kullback-Leibler (K-L) divergence based approach using the PCA model is presented in the next section.

4.2.2 PCA + K-L Divergence based Model

K-L divergence is an important statistical measure that belongs to the domain of information theory. It is also known as relative entropy. It has shown a great potential for application in Fault Detection and Diagnosis (FDD) and has been aptly applied to incipient fault detection in mechanical and electrical systems as reported in [86]. It has also been widely used in multimedia security and neuroscience. However, the application of K-L divergence in islanding detection could not be confirmed in the literature. This section details the use of K-L divergence in conjunction with the PCA model for improved accuracy of event detection.

K-L divergence is basically a measure of dissimilarity between two probability distributions. If two data samples are drawn from two populations having the same distribution, the K-L divergence for the two samples will be zero. For two continuous Probability Density Functions (PDFs) $f(x)$ and $g(x)$ of a random variable x, the K-L Information (KLI) is defined as in equation 4.12:

$$I(f||g) = \int f(x) \log \frac{f(x)}{g(x)} dx \tag{4.12}$$

The K-L divergence is then given as a symmetric operation of the KLI in the form of KLD(f,g) = $I(f||g) + I(g||f)$. For discrete distributions, K-L divergence is defined as the mean value of the log-likelihood ratio of the two distributions.

For an anomalous behavior or a sudden change in a process, the PDF of the corresponding data points changes from that for the reference scenario and if it goes beyond a safe threshold ϵ, it can be statistically detected. This approach is also illustrated graphically in figure 4.1. The left hand side plot shows the original probability distributions while the plot on the right is of the K-L divergence between the two distributions with a threshold value ϵ shown for illustration purposes. For two Normal (Gaussian) probability densities f and g having means and variances as μ_1, μ_2 and σ_1^2, σ_2^2 respectively, the K-L divergence between them can be given by a simple expression of equation 4.13.

$$KLD = \frac{1}{2}\left[\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} + (\mu_1 - \mu_2)^2\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) - 2\right] \tag{4.13}$$

For this study, the divergence between distributions of projections of two types of data sets onto the 1^{st} PC of the reference PCA model was found. The data set belonging to case 1 (also called reference PCA score when projected) denotes the reference while the test data sets come from different event based cases. The probability distribution of the reference case (projection of the 'normal' case data points onto PC_1) and those of each of the test cases were not Gaussian. Accordingly, non-parametric kernel density estimation was used to approximate each of the two distributions as normal distributions. For a real-valued function $f(x)$ of a continuous real variable x, the kernel density estimate of $f(x)$ at $x = x_0$ is given

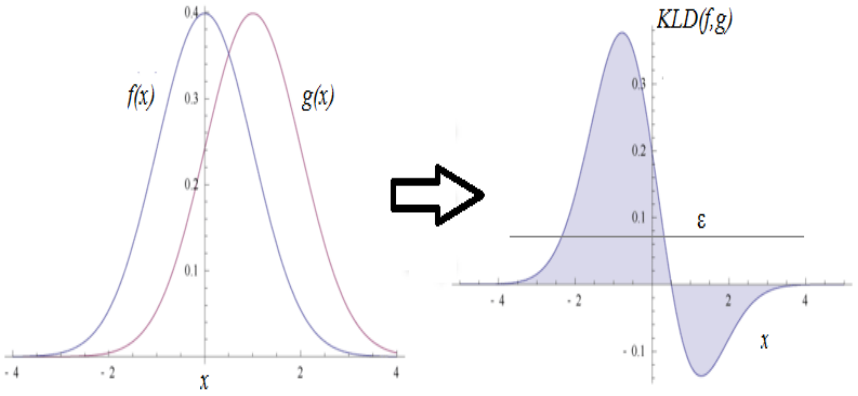

Figure 4.1: A sample illustration of K-L divergence between two Gaussian distributions

by equation 4.14:

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{x_i - x_0}{h}) \tag{4.14}$$

This is also known as a Parzen window density estimator and represents a simple non-parametric density model. Here, $K()$ is a kernel function that gives more weight to points that are closer to x_0. This function is symmetric about zero and integrates to one. For some x_0, $K(x_0) = 0$ if $|x_0| \geq 0$ or $K(x) \to 0$ as $x \to \infty$. Common kernel functions are Gaussian, triangular, rectangular, Epanechnikov and cosine. All these, except the Gaussian, have a cut-off point beyond which $K() = 0$. The bandwidth h is an important parameter that determines how quickly the cut-off point is reached. A small bandwidth will cause the kernel density estimate to depend only on values close to the point of evaluation, while a larger bandwidth will include more of the values in the vicinity of the point, yielding a smoother estimate. For the Gaussian kernel density estimate of the reference case score, the value of bandwidth was taken as $h = 476.783$. An advantage of the non-parametric approach over a parametric model is that no model fitting is required. However, tuning of the bandwidth is required which is usually done by cross-validation

Since mean-centering of data samples is an intrinsic step of the applied PCA algorithm, the means of the projections of both - the test cases and the reference case are zero. The PC scores have been assumed to be fairly normally distributed since they are linear combinations of the original data samples [87]. Taking this assumption, the following formula to calculate K-L divergence between a test case and the reference case was used:

$$KLD = \frac{1}{2} \left[\frac{\sigma^2_{\text{test case}}}{\sigma^2_{\text{ref.}}} + \frac{\sigma^2_{\text{ref.}}}{\sigma^2_{\text{test case}}} - 2 \right] \tag{4.15}$$

Here $(\mu_{\text{ref.}} - \mu_{\text{test case}})^2 = 0$ and $\sigma^2_{\text{ref.}}$ is nothing but the variance of projection on the 1^{st} PC which equals 5.736×10^6, as given in table 4.1. The other variances are those of the projections of the different test cases onto the 1^{st} PC of the reference PCA model.

Using equation 4.15, the test cases 2 to 6 were used as the second distribution and the case 1 was taken as the reference distribution. The values of K-L divergence calculated for different cases are given in table 4.4. The results from table 4.4 show an important picture. All those cases which had $Q > Q_\alpha$ and were wrongly detected seem to have been differentiated by their K-L divergence values. It can be clearly seen that cases 4,5 and 6 do not fall in the same category as they had been previously clubbed by the Q statistics based approach. The extremely large

Table 4.4: Event detection results using K-L Divergence

Case	σ^2	KLD	Islanding initiator
2	1.95×10^6	0.554	NO
3	4.92×10^6	0.004	YES
4	5.53×10^3	485.55	NO
5	8.42×10^6	0.1023	NO
6	5.12×10^6	0.0012	NO

and small values of case 4 and 6 respectively segregate them into different category of events however the similar order of values for case 3 and case 6 do not give a clear boundary. The L-L-L-G fault case has the least variance among the sample data points and hence it has the largest K-L divergence among all cases. However after looking at the divergence values of cases 3 and case 6, setting the correct ϵ for an event to be identified as the anomalous islanding trigger seems to be the problem with this approach although the false alarm detection rate has reduced to 1/5 from 3/5 in the previous section. To tackle this issue of threshold selection, a machine-learning based approach to detect anomalous events correctly has been presented in the next section.

The kernel-density estimated Gaussian PDFs for cases 3, 4 and 6 and their divergence from that of case 1 are shown in figures 4.2, 4.3 and 4.4 respectively. The higher the value of the K-L Divergence, larger is the gap between the two densities. Although the numerical values of the respective K-L divergences for each case have not been plotted, the figures show the asymmetry between the two distributions in each case. This asymmetry is an indicator of the statistical divergence between the test case data and the reference case data.

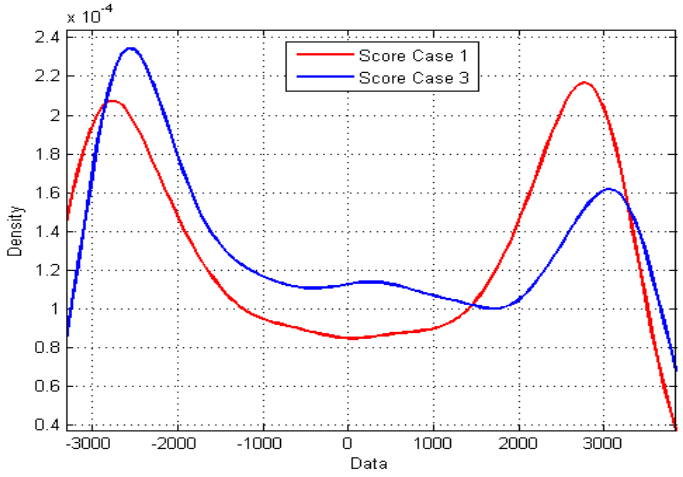

Figure 4.2: K-L Divergence: Case 3 v/s Case 1

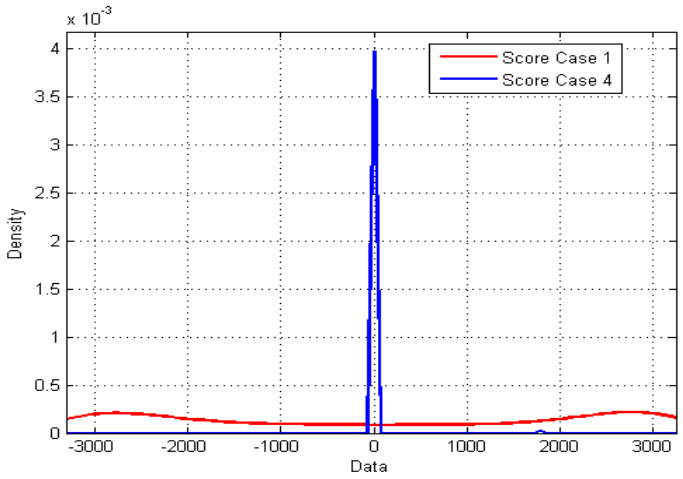

Figure 4.3: K-L Divergence: Case 4 v/s Case 1

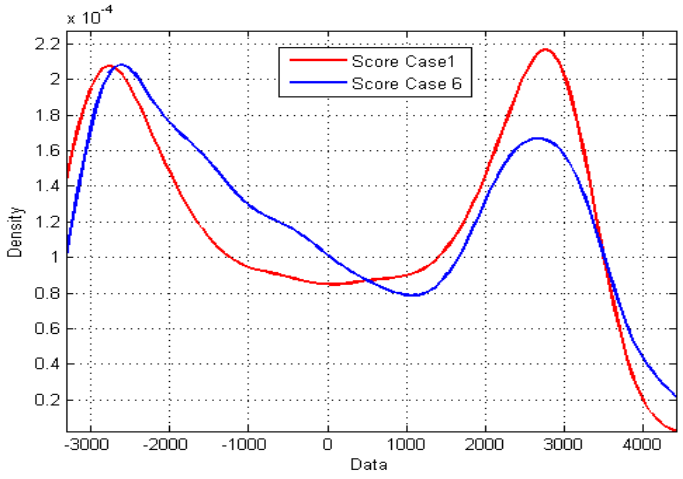

Figure 4.4: K-L Divergence: Case 6 v/s Case 1

4.3 Supervised Learning Based Event Detection

The statistical process control methods used alongside PCA extracted features have a fundamental limitation. They are threshold based approaches that identify an abnormal event or an anomaly only when its deviation from the normal system conditions is statistically significant so that the upper control limit of the statistic is crossed. The changes taking place on an active distribution feeder with high levels of PV generation are so dynamic that the thresholds change adaptively. For example the adoption of LVRT in a PV inverter will stretch the voltage operating limits that will affect the thresholds for an islanding condition. For an event to become an island-initiator, generally the following conditions must be met for over-current triggers:

$I > I_{\text{max. PCC protection}}$ and $V_{\text{inv. lower limit}} < V_{\text{inv. terminal}} < V_{\text{inv. upper limit}}$. Even if the thresholds for a particular network remain fixed for a given inverter, there exist many events that can satisfy the thresholds defining an islanding initiator, but are not actually liable to island a section. An example is the case of a motor starting transient. This event causes high inrush currents and a sharp dip in voltage that can cross the respective current and voltage thresholds but no distribution feeder relay is configured to operate for a current surge due to motor starting. Hence such an event cannot island a section and therefore a learning model is required to identify all such possible events as different from an islanding-initiator.

A threshold based detection approach will also involve a large number of sequential if-then-else operations. A pattern recognition model fit using training data consisting of samples and their actual outcomes can perform faster operations on unseen test data. The supervised learning approach can cover all the possible events on a high PV penetration distribution feeder along with their class labels in the knowledge base. This will enable the learner to recognize the narrow dissimilarities between the initiators and non-initiators of accidental islanding. Considering the model $G(x, y)$ mentioned in section 1.3, the set of class labels $C = \{0, 1\}$ defines the discrete values for binomial classification required in this case. Class label '0' defines an event that cannot island a section on the modeled feeder while class label '1' defines the opposite. The next two subsections detail the classifier models applied.

4.3.1 K-NN Classifier

The data points collected from the event simulations discussed above were analyzed for a preliminary inspection. It was observed that the anomalous instances in the UV+P-Q match (after disturbance) case were instantaneous values of positive half-cycles of current and voltage corresponding to the over-current spike. They were 555 in number since the event occurred for a short duration at the given simulation sampling rate. Based on the limited number of such instances, all the 555 values were included in a data-set. Since a supervised learner must recognize this anomaly as distinct from the healthy system conditions, a part of the 'normal' system operation data-set was included in the same set. The model must also recognize the patterns of the event just preceding this over-current spike and hence all the data points in the UV+P-Q match (during disturbance) case were also included. Thus a composite data-set was created and it was subjected to multiplication with the 2×2 PC matrix. The projections onto PC_1 were retained as features (by

retaining only the 1^{st} column of the resulting 51139×2 matrix) and from these the training data-set was created. This composite data matrix was 51139×1 dimensional in which a second column of class labels was added. For labeling, the condition of $I > 100$ A was used on the original un-transformed composite data matrix to label the rows corresponding to the islanding-initiator in the training dataset as '1' with rest labeled as '0'. The scatter plot of the original dataset of instantaneous data points in the 51139×2 matrix is shown in figure 4.5.

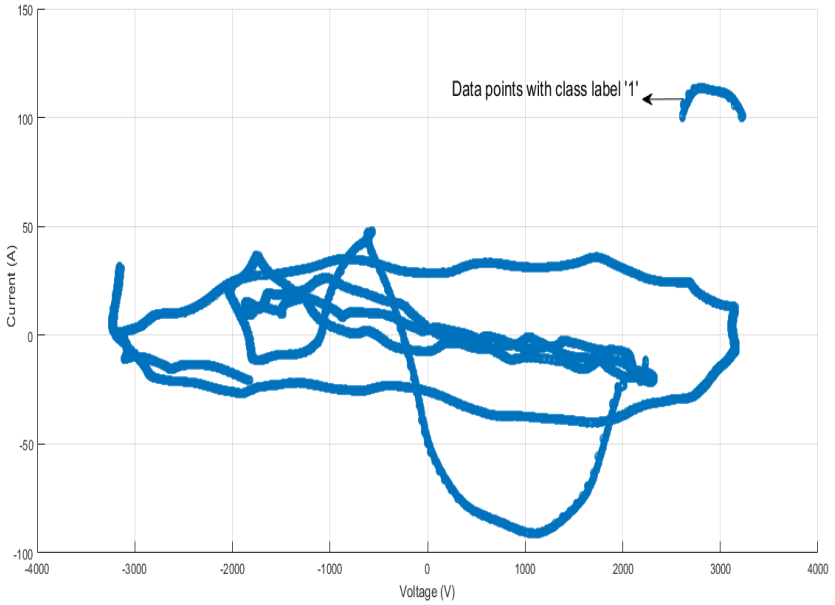

Figure 4.5: Scatter plot for the raw 51139×2 dataset

It can be seen that the data points with class label '1' are separated from those having class label '0'. However a horizontal straight line or even a slanting line in the current-voltage plane cannot create the decision boundary for learning a classifier to classify test data points. The reason for this, as mentioned above, is that many events can cross the thresholds defined by straight line boundaries and lie in class '1' but they actually do not belong to that class. A decision boundary that is not rigid but rather more smooth like a parabola visibly seems more appropriate for such a scatter. Hence a complex non-linear decision boundary is required to learn a classifier such that the events are assigned the correct classes. Exploiting the fact that the class label '1' data points in the training set lie close to each other, the simplest classifier based on a K-Nearest Neighbours (K-NN) approach was used.

The K-NN method employs an instance based learning paradigm that does not specifically fit a rigorous mathematical model for recognizing patterns in a set of

data points. It simply classifies an unknown instance based on its similarity with the labeled training data points in its vicinity. A K-NN classifier is non-parametric and its performance depends on the selection of k that defines the number of training data points in the neighbourhood of a test point. Therefore it does not actually create a model for classification of test data. It rather calculates the class of a point as that encoded variable which has the highest frequency from the K-most similar instances. Since the classification occurs only when a new instance comes up, it is also called as a lazy-learning technique. A distance measure is used to find the nearness of a test instance to its neighboring points. The most basic explanation of this approach can be given mathematically as follows:
Assuming all instances correspond to points in R^n, the nearest neighbours of an instance is defined by the Euclidean distance. Let an arbitrary instance x be described by the feature vector

$$\langle a_1(x), a_2(x), ...a_n(x)\rangle$$

Here, $a_r(x)$ is the value of the r^{th} attribute of x. The distance between two instances x_i and x_j based on the Euclidean measure is given by equation 4.16

$$d(x_i, x_j) \equiv \sqrt{\sum_{r=1}^{n}(a_r(x_i) - a_r(x_j))^2} \tag{4.16}$$

A class label is then assigned to the point usually based on a majority vote among the training points in its neighborhood [88]. The algorithm used here is described in Algorithm 2 given below. A distance-weighed K-NN approach can also be incorporated in the basic algorithm. Thus, for any unknown query point x_q, the

Algorithm 2 K-NN Classifier

1: Add each training example $(x, f(x))$ to the list of training examples. Here $f(x)$ is a binary encoded indicator response variable storing the class label for x.
2: **for** a query point x^q to be classified **do**
3: Let $x_1, x_2.....x_k$ be the k training instances nearest to x_q based on a distance measure
4: Return: $\hat{f}(x_q) \leftarrow \frac{\sum_{i=1}^{k} f(x_i)}{k}$
5: **end for**

estimate of the encoded indicator response variable is given by equation 4.17:

$$\begin{aligned} \hat{f}(x_q) &\leftarrow \frac{\sum_{i=1}^{k} w_i f(x_i)}{\sum_{i=1}^{k} w_i} \\ w_i &\equiv \frac{1}{d(x_q, x_i)^2} \end{aligned} \tag{4.17}$$

The number of neighbours k is practically determined by two methods. The first one is by running cross-validation to find the value that is computationally the most appropriate. The other way is to set $k = \sqrt{N}$ where N is the number of sample observations. The training data set used was the same 51139×2 matrix that has been described above. Since the PCA based feature extraction maps the original two-dimensional data onto a one-dimensional space, the value of k was varied between 1 to 10 since the class '1' data points are more close to each other and overlap with class '0' at less number of instances. The classifier was trained with the one-dimensional dataset where the choice of selecting the constituent event datasets was justified in the previous paragraph. Since the ultimate aim is to identify the island-initiator correctly and distinguish other 'similar looking' events from it, the testing was done on one known transient that appears similar and one anomalous over-current event. Accordingly, the following test sets (as projections onto PC_1) were used for the same reasons as given previously:
Test set I : Last 10,484 points of case 1
Test set II : Data points $\in$ case 4 (L-L-L-G fault at PCC)
Test set III : Data points $\in$ case 6 (UV+P-Q match, after disturbance).

Based on the accuracy and time taken in predicting a class label, k was set = 5. The euclidean measure of distance was used. The average cross-validated classification error or loss for 10-folds on the training data was 0.0070 indicating high training accuracy. The classifier performance was tested for each of the test cases using the 'majority vote with nearest point tie-break' rule. For test set II, the classifier identified all data points to $\in$ Class '0' indicating a 100% accuracy for this case. The reason for class '0' status of a L-L-L-G fault event at the PCC has already been given above. This confirms the correctness of the classifier in assigning label '0' to this case. The classifier accuracy for test sets I and III was found to be 97.42% and 90.12% respectively after multiple runs. The confusion matrices for these test sets are shown in table 4.5 and table 4.6 respectively. Case 6 comes very close to the case of actual islanding initiator discovered in case 3 and hence a large number of data points were assigned label 1. The average classifier accuracy can be reported as 95.75%. The classifier took an average time of 294 ms

in classifying a new test data point. As the k was reduced till 1, the time taken remained same but the accuracy improved even for testing on the third training data set 3. Clearly in this approach also, the three-phase short-circuit fault case is identified to be different from all other cases with 100% accuracy.

Table 4.5: K-NN: Confusion Matrix for test set I

	Class 0	class 1
Class 0	10214	270
Class 1	0	0

Table 4.6: K-NN: Confusion Matrix for test set III

	Class 0	class 1
Class 0	27618	3026
Class 1	0	0

4.3.2 SVM Classifier

The distribution of the binomial-class data points or feature vectors in the one-dimensional PC subspace is shown in figure 4.6. It can be observed that there is a fair degree of overlap between the two classes. Visibly, another transformation of data vectors towards high-dimensional feature space looked an alternative option to build a classification model. Accordingly, a Support Vector Machine (SVM) based classification model was thought of. The SVM approach is based on the statistical learning theory and can be applied to the case of both separable and non-separable training data vectors. For the separable case, the SVM outputs an optimally separating hyperplane defined by $f(x) = \beta_0 + x^T\beta$ with $y_i f(x_i)\ \forall i$. Here (x_i, y_i) is a part of the training data set consisting of a p dimensional sample vector $x_i \in R^p$ and $y_i \in \{-1, 1\}$, the set of binary class labels. The objective is to maximize the distance of each class's data points from the hyperplane defined by $M = \frac{1}{||\beta||}$ as shown in figure 4.7. This is expressed as an optimization problem in equation 4.18:

$$\begin{gathered}\min_{\beta,\beta_0} ||\beta|| \\ \text{subject to } y_i(x_i^T\beta + \beta_0) \geq 1, i = 1.,,,,N.\end{gathered} \tag{4.18}$$

Once the optimal values for β and β_0 are obtained, the decision function or classification rule is framed as $\hat{G}(x) = sign[\beta_0 + x^T\beta]$.

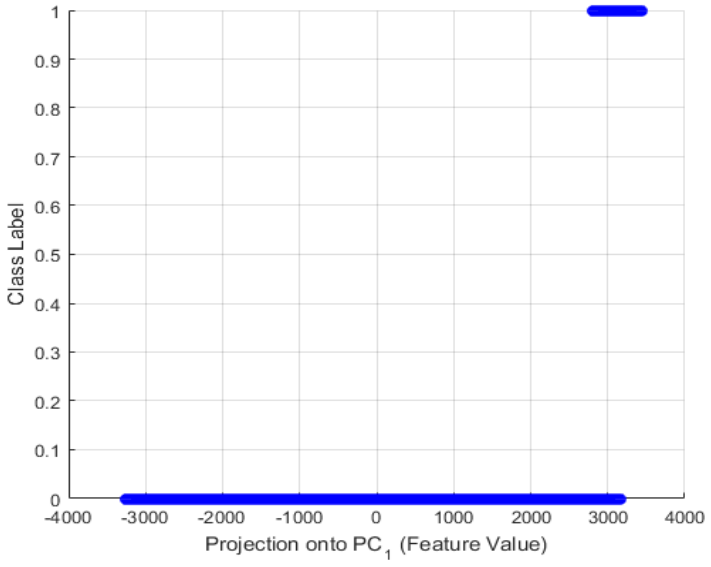

Figure 4.6: One-dimensional scatter of features in PC subspace

Consider the non-separable case shown in figure 4.8 where some points labeled ξ_i are on the wrong side of the margin by a factor $M\xi$. The objective remains the same however with one of the additional constraints that $\sum \xi_i <$ constant. Accordingly the optimization problem changes to equation 4.19:

$$\begin{gathered} \min_{\beta,\beta_0} \frac{1}{2}||\beta||^2 + C\sum_{i=1}^{N} \xi_i \\ \text{subject to } \xi_i > 0,\ y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i \forall i. \end{gathered} \tag{4.19}$$

Here C is the 'cost' or tuning parameter. Its value is estimated by cross-validation. However for the non-separable case, SVM follows the kernel trick wherein a basis function transformation $h(x)$ maps the original feature space onto a high-dimensional space where the feature vectors become separable. Thus the input space is grown in dimensions and the inputs of the model $\hat{G}$ get replaced by the basis functions used. The inverse mapping then translates into a non-linear decision function. The kernel function is based on the inner-product of the basis transformation functions and the distance between two points given as $K(x, x') = \langle h(x), h(x') \rangle$. Now, the classifier model changes to $\hat{G}(x) = sign[\hat{f}(x)] = sign[\sum_{i=1}^{N} \hat{\alpha}_i y_i K(x, x_i) + \hat{\beta}_0]$ where $\hat{\alpha}_i$ is a vector of non-negative Lagrangian multipliers, coming from the primal of equation 4.19 and whose upper limit is C.

The training dataset for this case has one-dimensional features. Thus a separating hyperplane in the original input space will be just a point. Therefore, a

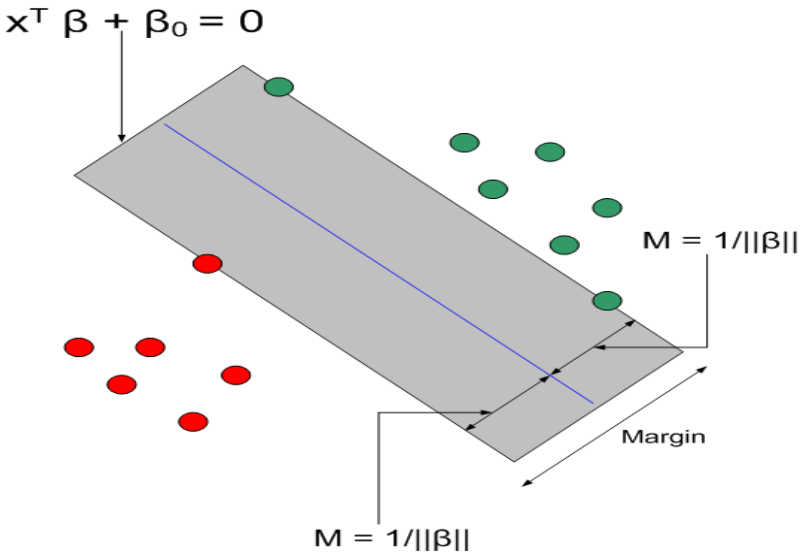

Figure 4.7: Separating hyperplane for the separable case

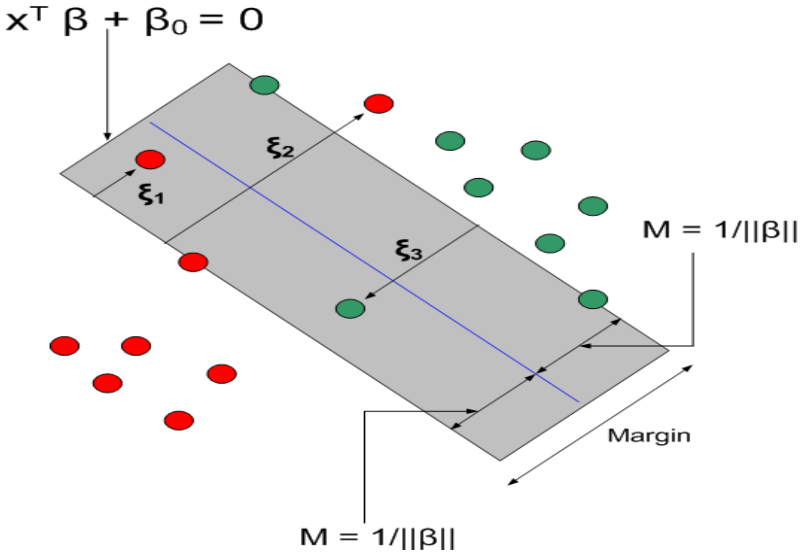

Figure 4.8: Separating hyperplane for the non-separable case

Radial Basis Function (RBF) kernel was applied in the SVM model fit for this data taking the assumption of normally distributed features due to the PCA operation involved, as mentioned above. A kernel using the RBF is defined as $K(x, x^{'}) = \exp(-\gamma||x - x^{'}||^2)$ where $\gamma = \frac{1}{2\sigma^2}$. The values of the tuning and kernel parameters were taken as $C = 1$ and $\gamma = 0.5$ respectively. The classifier was trained and then tested with the same test data sets (as used in the K-NN case):
Test set I : Last 10,484 points of case 1
Test set II : Data points $\in$ case 4
Test set III : Data points $\in$ case 6.

The classifier performance was tested for each of the independent test cases given

a promising 10-fold cross validation accuracy during the training. For test set II, the classifier identified all data points to $\in$ Class 0. This pertains to a 100% accuracy for this case. The variance of data points in case of the 3-ϕ fault is the least among all cases. Fault is thus detected naturally and, due to reason given above, cannot be labeled as an islanding initiator. This confirms the correctness of the classifier in assigning label '0' to this case. The classifier accuracy for test sets I and III was found to be 67.5% and 83% respectively after multiple runs. The time taken in classifying test cases is important as it should be less than or equal to the delay of the utility relay to trip a breaker so that the PV inverter trips well in time and this is also noted in table 4.7, although not required in an offline study.

Case 6 comes very close to the case of actual islanding initiator discovered in chapter 3 and hence a large number of data points were assigned label '1'. Clearly in this approach also, the three-phase short-circuit fault case is identified to be different from all other cases with 100% accuracy.

Table 4.7: SVM Classifier Performance for the 3 Testing Data Sets

Test Set	Accuracy	Time Taken (Including training time)
I	67.5%	437 ms
II	100%	63 ms
III	83%	432 ms

4.4 Unsupervised Learning Based Event Detection

Revisiting the generalized predictor model G described in section 1.3, one of the objectives of such a model is also to identify events as belonging to S_1 or S_2 when it has not learnt the class labels. Obtaining a labeled dataset for training of predictor models is not always practical or feasible. Thus $G(x)$ must be able to cluster events into 'can cause islanding' and 'cannot cause islanding' when only their feature vectors x_i are available. Accordingly, as a further exercise, unsupervised learning was explored with the motivation to create self-learning and acting systems. Such kind of systems will find more ground applications in real-time scenarios. However, initial attempts for offline testing were made to begin with. This section presents preliminary results of application of the same where a Self-Organizing Map (SOM) neural network was created for preemptive detection of unintentional islanding by classifying the discovered islanding initiator

from other power system events. Classification of a L-L-L-G short-circuit fault at the PCC was found to be invariant to input feature reduction. However, the other two test cases investigated produced contrasting results when the number of input features were changed from 1 to 2.

SOM type neural networks are based on the concept of competitive learning. Also known as Kohonen maps, they are a type of feed-forward networks that learn without any supervision. The weights are learned by the network on the basis of similarity among input features. There exists a winning neuron, also known as the best matching unit, whose weight vector is nearest to the input vector. Utilizing this 'winner takes it all strategy', a SOM network also learns the input space topology. A SOM is used to represent multidimensional data in low dimensions, usually 1 or 2. Each neuron is a functional unit arranged in a topological space. Since a SOM NN arranges neurons or nodes on the basis of similarity with input-data features, it clusters the data into similar groups. Unlike traditional ANNs, neurons in a SOM are not laterally connected to each other. Also, the meaning of 'weight' is different in their context. A weight vector signifies the relative degree of similarity of each node with the features of the input vectors. Each node is characterized by two properties: its position in the reduced-dimension space and a vector of weights. The weight vector is of the same dimensions as each input data vector. It has no relation with the spatial position of a node. To illustrate the basic philosophy behind a SOM NN, consider the schematic diagram given in figure 4.9. Here the neurons are arranged in a 4×4 grid or lattice in a two dimensional plane. Two-dimensional vectors constitute an input data matrix that has been used to train the network. Based on the similarity of the weights with the input data features, the nodes or neurons organize themselves into regions of similar features. Any new, previously unseen input vectors presented to the network will stimulate nodes in the zone with similar weight vectors.

The training process of the network is explained as follows. Each node is randomly assigned weight vectors. For a given input vector presented to the network, the distance between weights of each neuron and the vector is calculated by equation 4.20:

$$d = \sqrt{\sum_{i=0}^{n} (V_i - W_i)^2} \tag{4.20}$$

Here, V_i and W_i are the current input vector and the i^{th} node's weight vector. The node with the lowest value of d becomes the Best Matching Unit (BMU) for that particular vector. The neighbourhood of this BMU is then defined and all the

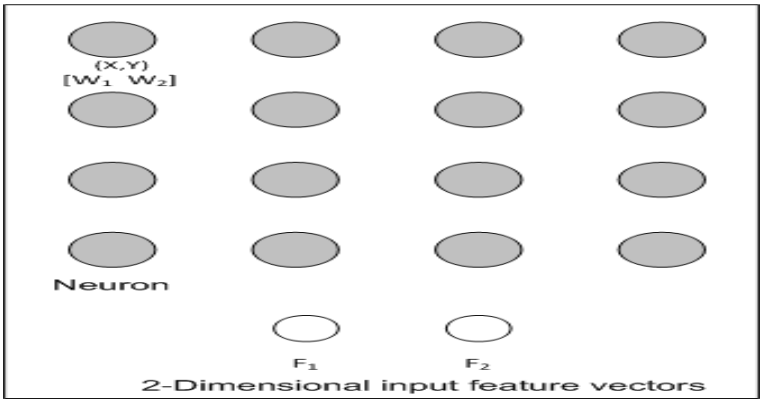

Figure 4.9: A SOM network with neurons in a 4×4 lattice in a two-dimensional space

neurons inside this radius make a similar zone. However this radius shrinks with time and ultimately the neighbourhood consists of the BMU (for that particular vector) itself. The radius is given by equation 4.21:

$$\sigma(t) = \sigma_0 \exp(\frac{-t}{\lambda}) \tag{4.21}$$

Where σ_0 = lattice width at t_0 and λ = time constant.

Once the radius is known, all the nodes are iterated through in the lattice to determine if they are within the radius or not. If a node is found to be within the neighbourhood then its weight vector is adjusted. For every node within the BMU's neighbourhood (including the BMU), the weight vector is adjusted according to equation 4.22:

$$\begin{aligned} W(t+1) &= W(t) + L(t)[V(t) - W(t)] \\ L(t) &= L_0 \exp\left(\frac{-t}{\lambda}\right) \end{aligned} \tag{4.22}$$

Here $L(t)$ is the learning rate that exponentially decays with time. The amount of influence a node's distance from the BMU has on its learning also affects the weight update process. Accordingly, a factor F has been added in the weight adjustment rule now expressed in equation 4.23:

$$\begin{aligned} W(t+1) &= W(t) + FL(t)[V(t) - W(t)] \\ F &= \exp\left(\frac{d^2}{2\sigma(t)^2}\right) \end{aligned} \tag{4.23}$$

This process keeps on repeating for a specified number of iterations and the neurons organize themselves in different regions of similarity. For training the network,

either each data vector is passed one by one or all the vectors are given together as a whole (batch training). Any unseen input vector when passed to the network is then assigned a cluster based on the process mentioned above.

In general, for the winner neuron (BMU), $||X - V_W|| = min_{k=1.,,,.N_n}||X - V_k||$ where V_W and V_k are the weight vectors of the winner neuron and the k^{th} neuron respectively, X is a training example and N_n is the number of neurons. The neurons are generally arranged in a two-dimensional map space and thus the high dimensional input space can be represented in lower dimensions by the grid of neurons. Such networks are widely used in clustering applications and hold a lot of promise in power system events grouping. In the domain of islanding detection, only one such study has been reported in [89] where input signals to the governor of a synchronous generator based DG were grouped into two classes to detect islanding conditions from the non-islanding ones.

The SOM network was created with two neurons based on the fact that the data points from different test cases have to be categorized into either islanding initiator or non-islanding initiator. The network was initially created using the above used 2 dimensional matrix of 51139 samples during training, after removing the class labels column. The different cases simulated and the total number of data points for the given sampling rate are the same as given in table 4.2. The created network is shown in figure 4.10.

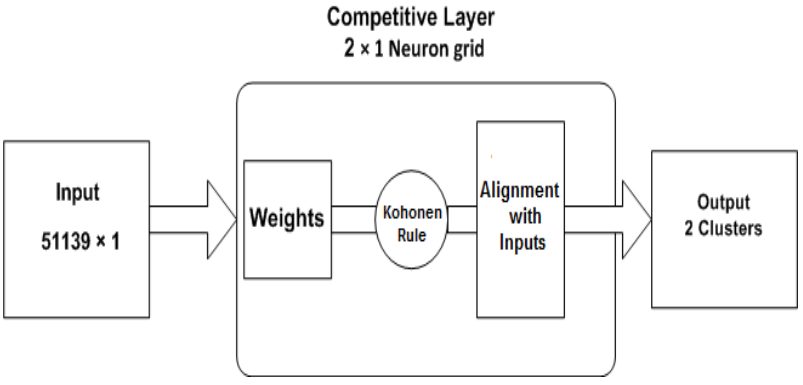

Figure 4.10: The trained SOM network

Three data sets were clubbed together to create a composite set for training of the network (projections of event data points onto reference PC_1):
Initial 20,000 instances $\in$ case 1; Complete data points from case 2 and 555 anomaly data points from case 3. The batch training algorithm was used for training the network modeled in MATLAB. The network was trained with no class labels corresponding to each input feature and it was tested on three test

data sets (projections of event data points onto reference PC_1):
Test set I : Last 10,484 instances of case 1
Test set II : Data points $\in$ case 4
Test set III : Data points $\in$ case 6.
The network took around 7 seconds to train and based on the similarities of the features, the input data vectors were clubbed into two clusters represented by the two neurons arranged in a 2×1 grid in the map space with a hexagonal topology. The number of training data points assigned to each cluster is shown in figure 4.11.

As part of the preliminary efforts, the network showed promising results for the Test sets I and II. For test set II, the network based classifier assigned class label '0' to all the data points with a mis-classification rate of around 0.2%. The identification of a three-phase short-circuit fault at the PCC as not an islanding initiator is physically attributed to the same fact mentioned previously. Accordingly, grouping of almost all the data points into the non-islanding initiator cluster is a correct action taken by the network. The classification accuracy for test sets I and III was found to be 81.09% and 54.35% respectively. The similarity between the anomalies represented by case 3 and case 6 seemed to have confused the network in its training and a little bit of assistance provided by the corresponding labels could have strengthened the network accuracy for test set III. Hence it is felt that a semi-supervised learning based approach might work very well in this case of preemptive detection of islanding in solar PV inverter interfaced distribution network.

As mentioned above, case 6 is in close proximity to case 3 in which the actual islanding precursors were discovered. Thus many of the data instances were wrongly

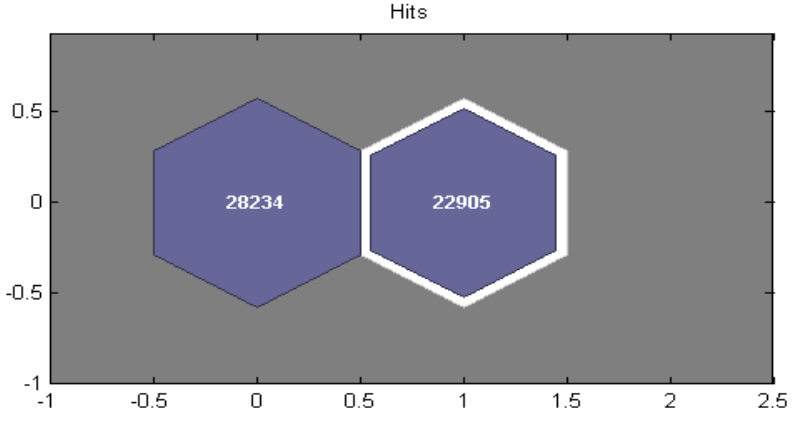

Figure 4.11: Number of input vectors assigned to each cluster after training

given class label '1'. The performance of the classification can be reported as an average 78.42% accuracy. Table 4.8 gives the testing results. However the highlight of the approach, which is also in agreement with the results obtained by supervised learning methods used above, has been that the case of L-L-L-G fault at the PCC was identified to be distinct from all the other cases with near 100% accuracy which was exactly 100% in the supervised learning cases. Since the applicability of the technique discussed in this section has to be evaluated for a real-time online identification scenario, discussed in the next chapter, the SOM network was also trained on two-dimensional input data. This was done to acknowledge the fact that real feeder data sampled at any practically possible rate will not be relayed in the form of a mean-centered data set. This is in contrast to the feature scaling step automatically performed by the SVD based PCA used in the initial training. Thus the original two input features of the simulated data were used for training a 2×1 network. The network took 14 seconds to train on the 51139×2 data set. Surprisingly, for test set II, the classification rate was found to be the same as obtained for the one dimensional data set. However the accuracy improved very little to rise to 56% for test set III. The worst performance in identification was observed for test set I where only 36% of the input data vectors were grouped into the correct cluster. Clearly, one thing is understood from this exercise that the scaling and processing of features is very much important for unsupervised learning techniques and hence considering the physical limitations of feature-scaled data transmission, appropriate unsupervised learning methods must be chosen and carefully trained to obtain an acceptable accuracy. A semi-supervised paradigm can be a recommendable approach to begin with for establishment of self-learning islanding predictive solar PV inverters.

Table 4.8: SOM Classification Accuracy for the Testing Data

Test Set	Accuracy
I	81.09%
II	99.83%
III	54.35%

The event detection models based on multivariate statistical methods were tested on all the event-based cases except for case 1 (the 'normal' operation case). The machine-learning based models, both supervised and unsupervised, were testing on different independent test data sets. It would be good to combine the performance of all these models for the two broad test data sets for a comparative evaluation. Accordingly, the different learning models and their event detection results have been summarized in table 4.9. The results for the two multivariate statistics based

models have been presented in terms of a single overall detection accuracy. For the machine-learning based models, detection accuracies have been reported test set-wise where Test Set II and III correspond to cases 4 and 6 respectively and Test Set I contains last 10,484 points of case 1.

It can be inferred that the K-NN based model performs in the way as desired for this particular application. For an offline-based event detection scheme, the nearest-neighbours approach suits based on the scatter of data points belonging to different events, to be classified into two broad categories. However, labeling of data points may be an intensive exercise but the experimentation with SOM NN has opened the widows for a semi-supervised based solution since the network's accuracy was high in some particular test cases.

Table 4.9: Performance comparison of different learning models used for event-based test sets

Learning Model	Test Set I	Test Set II	Test Set III
PCA + Q statistic	40%		
PCA + K-L Divergence	80%		
K-NN	97.42%	100%	90.12%
SVM	67.5%	100%	83%
SOM	81.09%	99.83%	54.35%

4.5 Conclusion

This chapter gave the first-hand account of the use of event-simulation data for building various learning models for event detection. The preliminary step of extracting important features from the raw input data was described first. Exploiting the existence of a strong correlation between voltage and current in the modeled power system, PCA was used as the feature extraction technique on instantaneous values of these two quantities. The inherent pre-processing step executed during the application of the PCA algorithm used was also explained. Projection onto the 1^{st} PC of the reference PCA model corresponding to the 'normal' operation of the modified IEEE 13 node feeder was taken as the uni-dimensional feature for creation of different event-detection models. Apart from the events described in chapter 3, a L-L-L-G fault at PCC was simulated. The selection of data points for creating the training and testing datasets for different learning models was also justified.

A statistical process control strategy using the PCA Q statistic was employed as the first multivariate statistics based event detection model. However the PCA + Q statistic model had a high false-alarm rate and hence a PCA + K-L divergence model was applied on the same data. This was a unique application of the divergence measure for islanding-related studies, previously not reported in the literature. The detection accuracy improved however, for a dynamically changing active distribution network scenario, this multivariate statistical model might not be reliable since it is fundamentally threshold based, finding whose value is a challenge. Accordingly, machine learning based models were applied on a composite training dataset. The use of the simplest model of a K-NN classifier on independent test data sets reported much higher accuracy than the multivariate statistical models. A SVM classifier using the RBF kernel was also applied to create separable features in a high-dimensional space from the original 1-dimensional input PC subspace. However the classification accuracy was less than that for K-NN and the time taken to predict the class label was also high. The different machine-learning models were tested offline, after storing the simulation generated data, and it was found that both the models give a 100% accuracy in classifying the PCC fault case correctly. Finally, considering the practical challenges in obtaining labeled data, an unsupervised learning method based on SOM was applied on the same training dataset - with and without PCA derived features. The network :

- Performed with near 100% accuracy for the PCC fault case regardless of the number of features
- However for the other two test cases, the accuracy reduced when features increased from 1 to 2
- Had a very poor accuracy of grouping the events correctly. The performance could have improved follwoing semi-supervised learning approach.

Although the use of this unsupervised technique did not give favourable results of event detection, the considerations made in its application were a prelude to the online feature extraction and event detection case described in the next chapter.

Chapter 5

Online Feature Extraction and Event Detection

5.1 Introduction

The different types of learning models for offline event detection were discussed in chapter 4. The PCA based feature extraction methodology was also discussed. Both, the multivariate statistical and machine learning based predictive models were implemented. It was found that the supervised machine learning based models provided much better accuracy in correctly classifying events than the multivariate statistical models. With the practical considerations relating to availability of labeled data and feature scaling and dimensionality reduction in real-time applications, an unsupervised learning technique was also attempted in the offline mode. This represented the initial set of steps towards realization of the preemptive islanding detection strategy as described in the research objectives section. However, the application of the implemented techniques was done in offline mode on a standard workstation (personal computer) running on a proprietary operating system. Since the ultimate objective is to design and realize a strategy for detecting islanding-initiators in real-time, an online event detection approach needs to be implemented in a dedicated, portable computing platform.

The techniques described in this chapter and their implementation differ from that mentioned in the previous chapter in the following ways:
1. Only one supervised learning technique has been sticked to in this chapter unlike both supervised and unsupervised learning explored in the previous chapter
2. Two methods for feature extraction have been used: PCA and a proposed

computational geometry based approach
3. All the techniques are executed in a dedicated microcomputer that can control the mode-transfer for a solar PV inverter
4. The testing data points are not stored as offline matrices but are transmitted in real-time from event simulations run on the personal computer
5. The predicted class labels are also translated into digital voltages to provide the mode-change trigger.

Section 5.2 carries the PCA based methodology forward, in the online testing mode. The projections onto the 1^{st} PC of the reference PCA model are used as features to train a decision tree classifier. The testing database is expanded to include a lot of different events $\in$ class '0' for generalizing the classification accuracy. The online testing results on the microcomputer show better accuracy than those obtained in the offline mode on the personal computer. Lastly, in this section, the set of class labels predicted by the online classifier module during the event-simulation runs was realized into a digital high or low voltage corresponding to the respective binomial class labels. Section 5.3 highlights certain shortcomings in the PCA based feature extraction approach and describes the proposed computational geometry based feature extraction methodology for online implementations. The online testing results compare well with those obtained from the application of the PCA based methodology however in terms of time taken, both the approaches were at par with each other. The time taken to predict the class label of the islanding-initiator was found to be less than that its duration. Naturally, this is less than the time taken by an over-current relay to trip the breaker at PCC or that by a lateral fuse at PCC to blow up. This lead time can be used for changing the mode of the PV inverter based on the trigger provided. Thus a preliminary implementation of classification of data points transmitted from event-simulation on a modeled feeder has been made in this study.

5.2 PCA based Online Feature Extraction and Classification

The reference PCA model corresponding to the 'normal' system operation described in chapter 4 was used to test the robustness of this feature extraction approach for online classification. The reference PCA matrix was stored in the microcomputer's classifier program and the data samples transmitted from the workstation simulations were multiplied with it to find the projections onto PC_1.

These were used as test-data features to predict the class labels of the corresponding events. The offline training of the classifier and the online feature extraction and classification is described in the next subsection.

5.2.1 Decision Tree Learning

The need for machine learning based models for identifying the discovered islanding-initiator from other power system transients was justified in the previous chapter. The appropriateness of adopting a supervised learning strategy, at this stage, was also observed. Completely relying on an unsupervised learning approach for real-time applications might not be justifiable at the preliminary level of such a study although self-learning and acting systems must be the idealized target to achieve absolutely robust systems. However, as exhibited by the example of predictive grid analytics in [65], the historical data of past events collected over a time frame supports labeled-data training, subject to its availability. Examining the scatter of original features in the training dataset used in the offline study and the scatter in the PC subspace, the use of K-NN and RBF-SVM classifiers was suggested. However for the implementation of an online classification strategy, the use of a Decision Tree (DT) was found to be more appropriate. This is attributed to the following reasons pertinent for the online testing application for this study:
1. The training features are not linearly separated and require complex decision boundaries.
2. Decision trees are invariant to feature scaling since they depend on ranking of attributes and not their absolute values.
3. Training data based on PCA is mean-centered while real-time testing data is non-scaled. Only use of DT gave acceptable change in accuracy (detailed later).
4. Only a DT is robust to a high imbalance in class-distribution and can learn the inherent pattern effectively.

As observed in chapter 2, the use of DT classifiers in islanding-detection has relatively higher number of references in the literature than others. Its use has given better accuracy values than other types of classification models. For this study also, it gave better results than the other models, in the offline testing mode. Analyzing the scatter plot shown in figure 4.5, the required non-linear decision boundary can also be learnt by a decision tree, that too very quickly. The fact that decision trees are also robust to class-imbalance in the training data is particularly a strong justification for the present case. The reasons for using a DT

classifier are more pertinent for the online testing case and become visible in the next subsection.

A decision tree is a non-parametric learning model. It does not assume any specific form for the mapping function that is to be approximated. A DT recursively partitions the input training data to create a hierarchical tree structure following a top-down approach, to classify data instances. Given the features or attributes and class labels in a training data set, a DT finds a root node and leaf nodes that are connected by branches. The root node is the starting node of the tree that descends to other nodes based on different values of the attribute that is chosen for the root node position. The number of training data examples having this attribute value then decide the nature of the descendant node on the basis of the class-label distribution. If all the examples belong to one class, the descendant node becomes a leaf otherwise it becomes a root for a sub-tree that grows further. Thus an attribute and a threshold value are chosen at every terminal node to grow the tree. Statistical measures are used to find the root node and assess the quality of the partition created by the chosen attribute. It is important to understand two important terms regarding the same: entropy and information gain. These two measures are generally used by different DT algorithms to measure the data 'splitting' ability of an attribute and to find root nodes respectively. An attribute may be discrete-valued or continuous-valued. The concept of root, leaf and sub-tree growth is illustrated in figure 5.1 for discrete-valued and continuous-valued attributes.

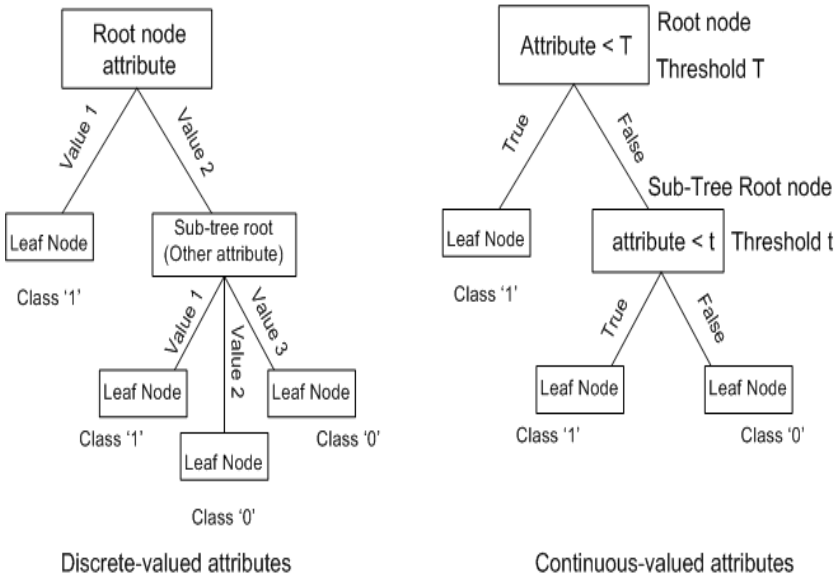

Figure 5.1: A basic tree structure for a binomial classification problem for discrete and continuous valued attributes

Consider a dataset of positive and negative training examples. Entropy measures the heterogeneity of such a dataset based on the distribution of class labels. In general, the entropy of a dataset S containing examples having 'c' classes is given as in equation 5.1:

$$E(S) = -\sum_{i=1}^{c} P_i \log_2 P_i \tag{5.1}$$

Here, P_i is the number of examples having class i divided by the total number of examples in S. Since entropy is a concept belonging to the domain of information theory, it specifies the minimum number of bits of information required to encode the classification of any instance $\in S$. This justifies the use of base 2 in the logarithm. An entropy value of 1 shows a highly impure dataset signifying highest randomness in estimating the class label of an arbitrarily drawn example. Zero entropy means that all the examples belong to one class and the dataset is pure. A descendant node, as shown in figure 5.1, becomes a leaf node when the entropy of the set of data points belonging to that partition is zero. Information gain, on the other hand, is the reduction in entropy of a dataset when a particular attribute alone is used to partition the training data examples according to its values. For the same dataset S, if A is a an attribute having $V(A)$ as the set of possible values, then the information gain of attribute A relative to a collection of examples S is defined in equation 5.2:

$$Gain(S, A) = E(S) - \sum_{v \in V(A)} \frac{|S_V|}{|S|} E(S_v) \tag{5.2}$$

Here, S_v is a set of values of S for which A has a value $v \in V(A)$. $E(S)$ is the entropy of the original dataset S and $E(S_v)$ is the entropy of the set S_v. The attribute having the highest information gain is selected as the root node of the tree and this heuristic is used by the most basic variant of the DT algorithm known as Iterative Dichotomiser-3 (ID3). Misclassification rate and Gini-index are measures other than entropy that are used to judge the quality of a 'split'. Gain ratio is another alternative to information gain to select attributes for the root nodes.

The major DT algorithms are ID3, C4.5 and Classification and Regression Trees (CART). All these algorithms follow the same approach of a top-down greedy-search over the input space for building the tree. However, they differ in terms of controlling the growth of tree, incorporating continuous attributes, selecting the attributes and handling missing values. C4.5 and CART are an extended version of the ID3 that work best for discrete-valued attributes. In all these algorithms,

the process of selecting an attribute and the threshold value for partitioning the data is repeated for each descendant node. Only the training examples that pass the test for the attribute in the parent node are used. This process is continued for each new leaf node until either of the following conditions is met:

1. Every attribute has been included along the concerned branch at least once
2. The entropy of the set of training examples associated with the leaf node becomes zero.

The basic ID3 algorithm for a binomial classification problem involving discrete-valued attributes is given in algorithm 3. Here, $Examples$ are the training examples and $Target$ is the class label. A test is performed for each attribute selected for a root node in order to branch out to terminal nodes for partitioning the input space. For discrete-valued attributes, the test is $A = v_i$ and S_{v_i} is the subset of $Examples$ which pass this test, in the present case. The 'best-attribute' is the one having the highest information gain.

Algorithm 3 Basic ID3: ID3 ($Examples$, $Target$, $Attributes$)

Create a *Root* node for the tree
If all $Examples$ are positive
Return: A single-node tree with *Root* having label = '1'
if all $Examples$ are negative **then**
    Return: A single-node tree with *Root* having label = '0'
end if
If $Attributes$ is empty
Return: A single-node tree with *Root* having label = most common value of $Target$ in $Examples$
Otherwise Begin
$A \longleftarrow$ the best attribute
The decision attribute for $Root \longleftarrow A$
for Each possible value v_i of A **do**
    Add a new branch below *Root* corresponding to the test $A = v_i$
    if S_{v_i} is empty **then**
        Add a leaf node with label = most common value of $Target$ in $Examples$
    else
        Below this new branch add the subtree ID3 ($Examples$, $Target$, $Attributes - A$)
    end if
end for
return *Root*

The CART algorithm is fundamentally based on the same philosophy as that of ID3. However, it has several additional features. It can handle continuous attributes, there is provision of avoiding over-fitting by 'pruning' the tree and it

can handle missing values in a much better way. Also, a CART based tree can use other measures like Gini-index or Gain-ratio for splitting a node or for selecting attributes. For continuous-valued attributes, the test performed at each node for branching-out is usually of the form $A < t$ where t is a threshold value that is to be decided at each step during the growth of the tree. The most basic heuristic is to find those pairs of attribute values where the class-labels change in the training dataset. The threshold must lie in-between these two extreme points and hence t is usually taken as the average of these values. This leads to different thresholds for the attribute A tested at different nodes. Effectively, this leads to the creation of a discrete boolean attribute from A which will hold only two values: True when $A < t$ is satisfied otherwise False. This is also depicted in figure 5.1.

Since the training features for this binomial classification problem are continuous values, CART which is an advanced version of the C4.5, was used instead of the ID3 algorithm. Accordingly, a DT based on the CART algorithm was programmed in Python 2.7 on a Raspberry Pi model 2 B using scikit-learn [90]. Raspberry Pi is a small, portable microcomputer that runs on a freely available customizable distribution of the Linux operating system. It has Python inbuilt so that programming scripts can be dedicatedly run on its computing platform. The model 2 B has a 900 MHz Broadcom quad-core ARM v7 processor with 1 GB RAM and is pocket-sized so that it can fit conveniently inside a solar PV inverter. More details can be referred to in Appendix II. The tree was not pruned considering the low dimensional input feature space and the Gini index was used as the split criteria in place of entropy. The Gini index is defined in equation 5.3.

$$H(X, A_i) = -\sum_{i=1}^{i=c} P_{m,k}(1 - P_{m,k}) \tag{5.3}$$

In equation 5.3, $P_{m,k}$ stands for proportion of training examples in the m^{th} tree node that belong to k^{th} class among 'c' class labels in that partitioned training subset. The built tree had a very wide horizontal growth and hence cannot be shown here due to space constraints. The training data consisted of the uni-dimensional features for the same set of events as in the composite training dataset used in chapter 4. For the reasons given previously, the test data sets were kept the same as in chapter 4 (projections onto reference PC_1 of following):

Test Set I : Last 10,484 data points of case 1
Test Set II : Data points belonging to case 4
Test Set III : Data points belonging to case 6.

The SVD procedure automatically performs mean-centering on the reference data for creation of the reference PCA model. Accordingly, the training and testing data points had to be mean centered before projecting them onto the 1^{st} PC of the reference PCA model. The classifier modeled in the Raspberry Pi was tested in the offline mode on feature-reduced data (prepared on the personal computer) belonging to the different test cases simulated. The performance of the Raspberry Pi classifier tested in offline mode is shown in table 5.1. The mean 10-fold cross-validation accuracy was 98.873%. An accuracy of 100% for Test Set II is justified in the same manner as for previously used classifiers. Since the classifier performed perfectly for the unseen test set belonging to case 6 whose data points were not included in the training set, it was decided to use the same set for the online testing described in the next subsection.

Table 5.1: Offline DT Classifier Testing Results on Raspberry Pi

Test Set	Accuracy	Time Taken to Predict Class Label
Test Set I	95.16%	0.005 s
Test Set II	100%	0.0082 s
Test Set III	89.33%	0.011 s

5.2.2 Online Testing on Raspberry Pi

Different event simulations for the modified IEEE 13 node feeder Simuink model were run on a 2.3 GHz, 4 GB RAM personal computer running on Windows environment that supports real-time Simulink data transfer to Raspberry Pi. The interface between the Simulink model and the DT classifier programmed in Raspberry Pi was made via a Local Area Network (LAN) based connectivity between the two computers. A secure shell connection to the Raspberry Pi was established in order to remotely access its graphical user interface through the personal computer screen. The online testing scheme requires two things: extraction of features from data samples transmitted from event simulation and prediction of class labels of these features, as they arrive. In the case of the PCA based approach, feature extraction was executed by multiplying individual samples with the PC matrix. The instantaneous values of current and voltage at the PCC of the feeder model were sent to the classifier programmed in the Raspberry Pi over the LAN.

Since for online testing, the test data is to be transmitted to the classifier in the form of individual voltage and current samples arranged as single 1×2 arrays, mean-centering of each of such individual entries was not possible. It requires

a dynamic array that takes up memory. Hence the test data sets for the online testing case were simply projected onto the 1^{st} PC, as they arrived, and then given to the classifier. This is only possible with a DT based classifier because it is not affected by feature scaling [91] due to use of information-theory measures like entropy, dependent only on the relative proportion of classes in the dataset, for classifying a test instance. Since the classifier is trained with projections of mean-centered data sets (belonging to different event cases), testing it with non mean-centered test instances (due to real-time data generation) looked ambiguous. However the classifier still tested with each one of the offline test data sets having non-mean-centered projections to track any changes in accuracy due to non-scaling of the features in the test data. The accuracy was found to be unchanged for Test Set II. For others, a negligible change for Test Set I (0.6% reduction) and a 5.7% increase for Test Set III was observed. Adjudging these figures and considering the limitations with real-time data pre-processing, the stream of data was projected onto the reference PCA model's first component and fed to the classifier in online testing mode. The related ambiguities are further cleared by looking at the final accuracy of online classification.

The simulation was run with a fixed-step solver due to Simulink real-time hardware requirements and the sampling time was fixed to 10μ seconds based on previous experience with the RTDS simulation. Each event simulation was run for 0.5 s with every disturbance having a 30 ms occurrence. The PCC current and voltage signals were converted from the complex data type to floating point ('double' in the Simulink C compiler system) type. Each one of these double data format signals was further encoded into an unsigned integer 8-bit string. Both these strings were sent as a multiplexed signal to the classifier over the local area network using the User Datagram Protocol (UDP). The data packets as received by the Raspberry Pi via UDP for the anomalous islanding initiator case can be seen in figure 5.2. The sags and swells in the voltage and the corresponding anomalous current are clearly visible.

The 16-bit multiplexed string was received by the classifier module and unpacked into floating data point numbers. These unpacked numbers were made to arrange into 1×2 numerical python (numpy) arrays. As mentioned in the previous section, these data points, sampled at every 10 μ seconds were multiplied by the 1^{st} PC of the reference PCA model, stored in the Python program and thus transformed into 1×1 numpy arrays. The same set of conditions will produce the same magnitude of values and hence the trend does not change with the sampling rate. However the duration of spike in the UV+P-Q match after disturbance case is

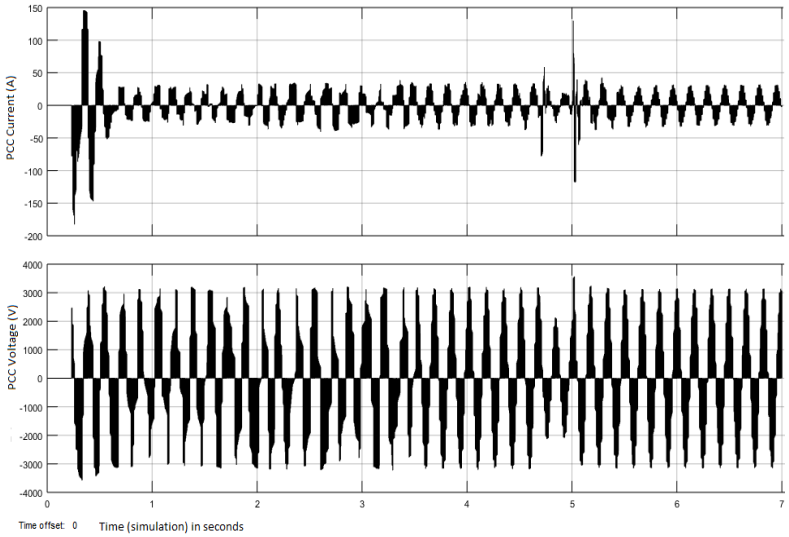

Figure 5.2: Data packets received over UDP by the Raspberry Pi for the UV+P-Q match case

of 0.79 ms which is more than that from the old sampling rate. The different sampling rate and a repeated simulation run will surely produce different data points and hence the island initiator case has been used as a testing data set also. This stream of data being transmitted from the simulation run was thus input as feature processed data to the DT classifier. Testing on case 5 above does not make sense as a 1.3 p.u. over-voltage is expected to occur rarely while case 6 has been tested offline. Cases 2 and 3 were included in the test set since these anomalies have to be accurately identified in real time by the online classifier model. Based on the classifier's offline performance, case 3 was decided to be kept as one of the testing sets for the online case also. Testing a DT on one of the training data sub-set is often required to check for over-fitting on the training data. However a low accuracy of 89.33% for case 3 does not indicate a high degree of overfit. The online classification was done for the following test cases:

Test Case A: Under-voltage + P-Q match (30 ms window during the disturbance)

Test Case B: Under-voltage + P-Q match (30 ms window after the disturbance ends)

Test Case C: Three phase short-circuit fault at PCC

Test Case D: Single line to ground fault at PCC

Test Case E: Capacitor bank switching at node 675

Test Case F: Single phase load switching at PCC

The online accuracy and the average time taken to predict the class labels for each test case is summarized in table 5.2. Table 5.3 shows the consolidated confusion matrix for Test Cases A, B, E and F. The classifier took 4 seconds to learn the pattern (python serial code interpretation) and at the 5^{th} second, it started printing

Table 5.2: Online DT Classifier Testing Results on Raspberry Pi

Test Set	Accuracy	Time Taken to Predict Class Label
Test Case A	98.498%	0.00032 s
Test Case B	97.064%	0.00037 s
Test Case C	100%	0.000363 s
Test Case D	100%	0.000367 s
Test Case E	96.24%	0.000317 s
Test Case F	97.20%	0.000336 s

Table 5.3: DT Confusion Matrix for Test Cases A, B, E and F

Test Case	A	B	E	F
Actual 0s into 0	3279	2986	2787	2780
Misclassified Points	50	0	109	80
Actual 1s into 1	0	122	0	0
Misclassified Points	0	94	0	0

the predicted class label for each of the input arrays, in the Python console. Figure 5.3 shows a screen shot of the process of online printing of predicted class labels in Raspberry Pi's Python console. The Simulink simulation model running on the workstation is shown in the inset. 100% accuracy is repeated for online testing on Test Case C and the same reason is applicable to the same result obtained for Test Case D.

Re-assessing the performance of the DT classifier, the distribution of the binomial-class data is important. Since the class distribution in the training data is highly non-uniform and skewed, classification accuracy is not a comprehensive and in-

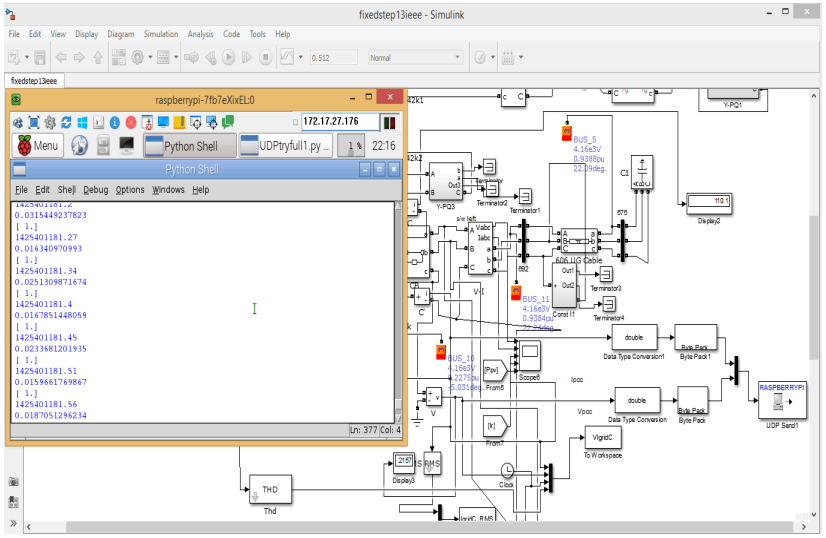

Figure 5.3: Online prediction of class labels

dicative measure of its performance. Since the islanding-initiator event being investigated is rare and occurs for a very short duration, the number of instances in this class '1' event are expected to be less in comparison to class '0' events. For such cases, it is relevant to consider the Kappa statistic. It is a measure of classifier accuracy normalized by the imbalance of classes in the data and is defined as in equation 5.4:

$$\kappa = \frac{\eta_{observed} - \eta_{expected}}{1 - \eta_{expected}} \tag{5.4}$$

$\eta_{observed}$ is the accuracy calculated from the confusion matrix while $\eta_{expected}$ is calculated as in equation 5.5:

$$\eta_{expected} = \frac{\text{Product of Marginal Frequencies}}{\text{Total Instances}} \tag{5.5}$$

The marginal frequency of a class defines the relative proportion of instances put in that class by the learner and those actual in that class to the total instances. The value of the Cohen's Kappa statistic for all the online test cases were evaluated. For the test case B, the value of κ was found to be 0.707. For all the other test cases, it came out to be 0. This shows that the classifier performance for the test cases of short-circuit faults at PCC, switching surges and UV+P-Q disturbance equals the accuracy any random classifier would give on their respective confusion matrices. However since these events belong to class '0' whose instances are relatively more in number, it is not expected to get their respective number of classifications correct by chance however a κ value of 0 for the test case B would have been a concern. Its non-zero, high value indicates that the classifier has learnt the pattern correctly.

5.2.3 Inverter Mode-Change Trigger

From the online testing results reported in table 5.2, it can be seen that the DT classifier takes an average of 0.3 ms (about $1/50^{th}$ of a cycle for 60 Hz system) to predict the class label of a data instance with 98.167% average accuracy. The online classification accuracy for Test Case B was found to be 97.064% with time taken for class label prediction to be 0.00037 s. The SEL-751 DOCR model taken as a reference has typical pick-up setting of 8 ms and takes around a minimum specified time of 10 ms - 12 ms to trip a circuit breaker. Referring to the specifications of SEL-751 and considering the physical distance between the PCC circuit breaker and the relay, the $t_{\text{prediction}}$ is found to be less than the time taken by this commercial relay to trip the breaker for any such anomalous spike that is within its time-current characteristics. Also, for a fuse-relay protective scheme with the

lateral fuse placed towards the grid-side, this time period is less than the duration of the spike (0.79 ms based on the circuit-simulation sampling rate used) and hence can quickly alert the inverter before the fuse blows, so that it is ready with the appropriate action once it has blown.

Accordingly, the inverter can be made to trip and avoid the formation of an unintentional island after it has preemptively detected the anomalous conditions liable to island the section, before the utility relay action. Also, this lead-time can be used to change the inverter's mode of operation from constant P-Q control to V/f regulation mode so as to prepare it in advance for maintaining an island, once the PCC protection trips. The PV inverter can be made to work in a V/f regulatory mode before the feeder section isolates itself from the main grid so that the constant P-Q control mode is avoided in the island that will result but will be maintained and sustained as an intentional island so that the connected loads receive acceptable power quality. To implement this mode-change trigger, the predicted class label values were translated into a physical HIGH or LOW voltage on one of the General Purpose Input Output (GPIO) pins of the Raspberry Pi corresponding to 3.3 V and 0 V respectively. This online classification and mode-change triggering scheme are summarized as a schematic diagram shown in figure 5.4.

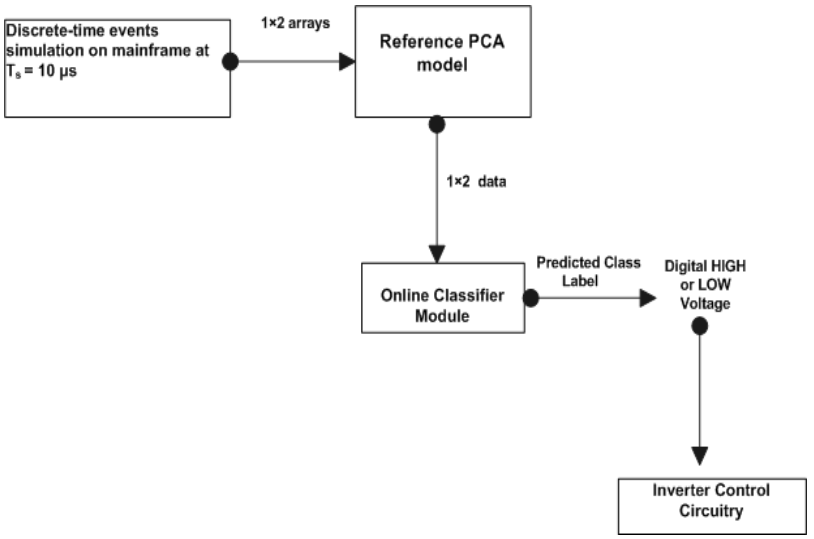

Figure 5.4: Preemptive islanding detection and mode-change trigger strategy

The system responded fairly quickly. However, the change in predicted class labels due to change in test case with the lapse of simulation time could not be observed as fast changing voltages on an ordinary multimeter. The mode change trigger set up is shown in figure 5.5. The output of the GPIO pin is a digital value that

can drive the GPIO pin of a Digital Signal Processor (DSP) that contains the control circuitry for a grid-connected solar PV inverter, as shown in figure 5.6. The figure shows the Raspberry Pi microcomputer on the left hand side while the DSP for the inverter control is shown on the right hand side. Hence there are two separate control circuitry for the grid-connected inverter: the online classifier module implemented on Raspberry Pi for event detection and identification and the switching control that is responsible for grid-synchronization.

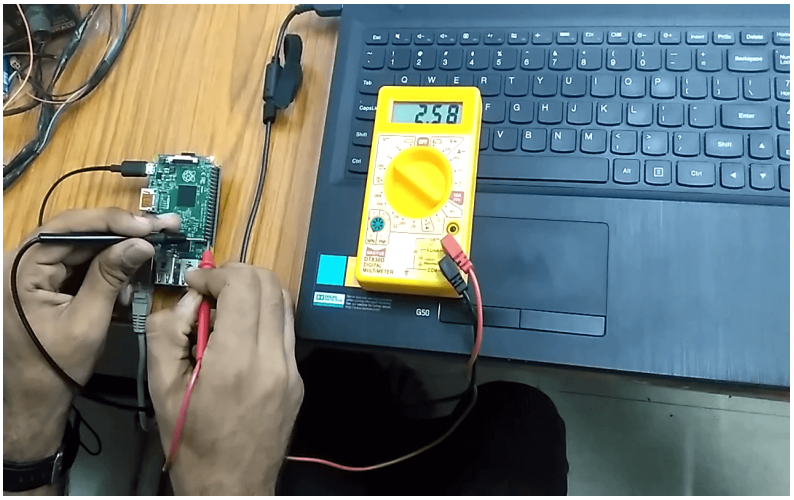

Figure 5.5: GPIO output corresponding to predicted class labels

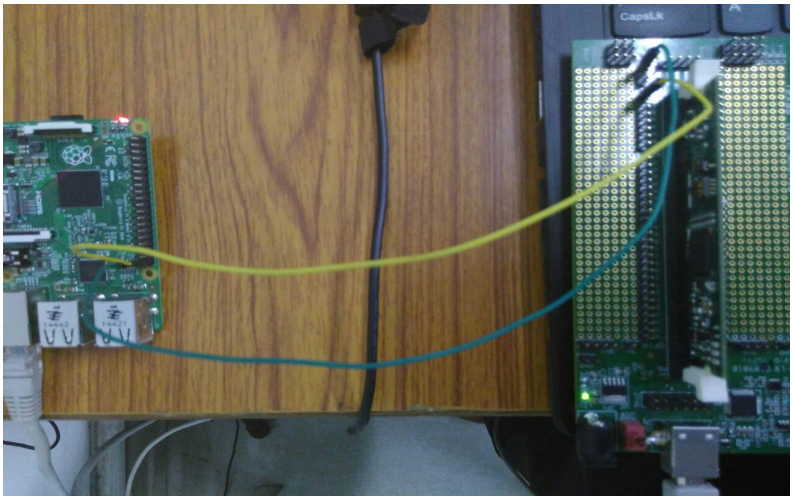

Figure 5.6: Implementation of inverter mode change trigger

Such a preemptive mode-triggering scheme can provide a proactive solution to a comprehensive energy management strategy in imminent islanding situations. However, the actual response of a fuse or an over-current relay for this current spike needs to be ascertained for a more comprehensive understanding. Also, the

time lapsed in changing the inverter's mode of operation need to be determined to check the efficacy of the proposed preemptive detection scheme.

5.3 Proposed Online Feature Extraction Methodology and Classification

The PCA based approach for online feature extraction has a few lacunae. The PC matrix for the reference PCA model contained zeros and ones however it did carry the influence of both V and I raw features. Also, it required the input data samples to be multiplied with the elements of this PC matrix. This study proposes a feature extraction methodology for online classification purposes in which the samples can be input in the form they come, without any transformation. The proposed model uses the concepts of computational geometry and applies an evolutionary optimization technique to create a feature transformation approach for offline training and online testing. For the same reasons as given for its appropriateness for online classification purposes, CART based DT was used as the classifier.

5.3.1 Computational Geometry Based Feature Extraction

The need for the proposed methodology is motivated by he research objective expressed in section 1.3 wherein the task is to learn a model $G = f(x, y)$ with $D = S_1 \cup S_2$ as the training data set and y as a predicted value in C. This represents a binary classification problem with the aim that $\forall x \notin D$, $y \in C$, accurately and quickly. Here 'quickly' stands for the condition: $t_{\text{class prediction}} < t_{\text{response protective device}}$. In contrast to the PCC short-circuit fault case and other transients like lightning and switching surges that cause different levels of changes in voltage and current, the observed anomalous over current as the islanding-initiator is essentially due to a voltage spike and is thus characteristically different. For the feeder model used in this study, any data point becomes an islanding initiator when: $\forall\ x \notin D, y = 1 \Leftrightarrow I > 100A$ and 0.7 p.u. $< V_{\text{Inv. Terminal}} <$ 1.3 p.u. however, there exist many $x \notin D$ that can satisfy the threshold but $y \neq 1$. One example is inrush currents due to motor starting. This event causes high current and a dip in voltage that may satisfy the threshold but no protective relay is made to trip on such transients and hence the corresponding data points can be assigned

$y = 0$. Thus a threshold based approach for detecting potential islanding causes from other system transients will not be reliable when the objective is to detect the initiator quickly and accurately. Accordingly a pattern recognition model based on supervised learning was proposed and realized for online testing. However for effective training of the prediction model and ease in real time testing, a convenient feature extraction methodology is very much important.

A significant amount of data was generated from the simulation of different event based cases. The data points belonging to the 'during-disturbance' for two known power system transients $\in S_1$ and for 'after-disturbance' period for the UV+P-Q match case (that created the islanding initiator) described in the previous section were collected as 'not-normal' case samples for classifier training. Since each simulation was run for a total of 0.5 seconds from the start of the simulation time, the samples belonging to the 'normal' condition, common for all the events, were also collected as the disturbances were programmed to occur at 0.45 seconds from start of the simulation. Among the cases simulated, those considered for offline training and cross-validation of the classifier performance together with the number of data samples generated for each case are summarized in table 5.4.

Table 5.4: Cases Simulated for Training Data Set for the Proposed Feature Extraction Model

Case	No. of Samples	Event
Case 1	8641	Normal
Case 2	501	L-L-L-G Fault at PCC
Case 3	501	Capacitor bank switching
Case 4	555	UV+P-Q match (after disturbance)

For saving on training data and keeping the meaningful information, only positive half cycles for each events were considered for data capture. This is also consistent with the real world data capture where an Analog to Digital Converter (ADC) will sample values for the microcomputer and will not allow negative values to come, if no signal scaler and shifter circuit is used. This will thus save on components and increase the ADC precision. For the normal case, the peaks of both current and voltage were taken in the complete positive half cycles of the two with the phase difference between the current and voltage peaks retained. For the islanding initiator case, only 555 samples were obtained in the simulation run at 1 MHz and all were included in the training data set. To lower the class imbalance in data, only the current-wave peaks were taken with 250 values above and below it for the

L-L-L-G fault at PCC and the capacitor bank switching cases respectively, leaving the voltage peaks in both the cycles due to phase difference.

The training data was split into two matrices by grouping it on the basis of event types. The 3 events: Normal system operation, L-L-L-G fault at PCC and capacitor bank switching at node 675 belong to class label '0'. The islanding initiator case data points belong to class label '1'. Both set of event-based data points were put into two separate matrices. Thus set $S_1 \in \{E_1\}$ and $S_2 \in \{E_2\}$ where $E_1 = \{$normal system operation, 3-phase SC fault at PCC, capacitor bank switching$\}$ and $E_2 = \{$UV+P-Q match after under-voltage case$\}$. Set of class labels $C = \{0, 1\}$; $\forall x \in S_1, y = 0$ and $\forall x \in S_2, y = 1$; here x is a feature vector extracted from event data. The fit classifier $G = f(x, y)$ should also perform in such a way that for any $x \in S_2$ at a different sampling rate, $y = 1$. The next issue is that of selecting or extracting the feature vectors $x_i \forall i = 1$ to $|D|$. Transformations like Fourier transforms require the full cycle to be completed first in order to evaluate and hence are not suitable for online applications. Furthermore, they are not applicable to non-stationary signals while most of the signals in our case are spiky and transient in nature. Instantaneous Discrete Wavelet Transform (DWT) for online transient detection has also been reported in literature [92] however its implementation will be computationally intensive (requiring memory of previous data points) for detecting events at sub-seconds level in our case. Accordingly matrix multiplication based feature transformation was advocated and investigated. The central idea is to bring the transformation onto such a feature space that the resultant dimensions are linearly separated or approach linear separability. Such decision boundaries are simple to learn and prediction of class labels is faster with use of linear classifiers like Linear Discriminant Analysis (LDA) and logistic regression.

The linear separability of training data points makes fitting a classification model easier. However real data is seldom linearly separated and since it also cannot be used in raw form and a few features need to be extracted from the dataset. Although the PC transformation matrix contained mostly ones (with both signs) and zeros, it contained the information of both current and voltage samples and gave good classification results for offline testing as reported in section 5.2. One important observation was the ease and convenience of matrix-transformation based methods for extracting features in real time data transfer. In view of the same, the proposed feature-selection methodology transforms the training data into a feature matrix such that the resultant feature-scatter is more linearly-separable than the original scatter.

The convex hulls of both the data sets were evaluated and drawn using the indices contained in vectors returned by MATLAB. The original scatter plot of the training data is shown in figure 5.7. The data is not linearly separable as the convex hulls of data points $\in S_1$ and S_2 are joint in some common regions. To bring linear separability, the convex hulls of the two event based data sets must be dis-joint. For defining linear separability: two sets X_0 and $X_1 \in R^n$ are said to be linearly separable if there exists n+1 numbers $\in R$ $w_1, w_2, ...w_n$ and K such that $\forall x \in X_0$:

$$\sum_{i=1}^{i=n} w_i x_i > K \tag{5.6}$$

And, $\forall x \in X_1$:

$$\sum_{i=1}^{1=n} w_i x_i < K \tag{5.7}$$

In equation 5.6 and equation 5.7, x_i is the i_{th} component of x. Since this condition can be better realized geometrically, the condition of linear separability can also be stated as:
$C1 \cap C2 = \emptyset$, where $C1$ and $C2$ are convex hulls of the 2 sets respectively. The convex hull of a set of points $x_i \in$ finite set S is defined as in equation 5.8:

$$Conv(S) = \left\{ \sum_{i=1}^{|S|} \alpha_i x_i | (\forall i : \alpha_i \geq 0) \wedge \sum_{i=1}^{|S|} \alpha_i = 1 \right\} \tag{5.8}$$

Positive half cycles of events $\in S_1$ were used while the data points during the 0.6 ms event $\in S_2$ used for creating the training data set D. Original data matrix X

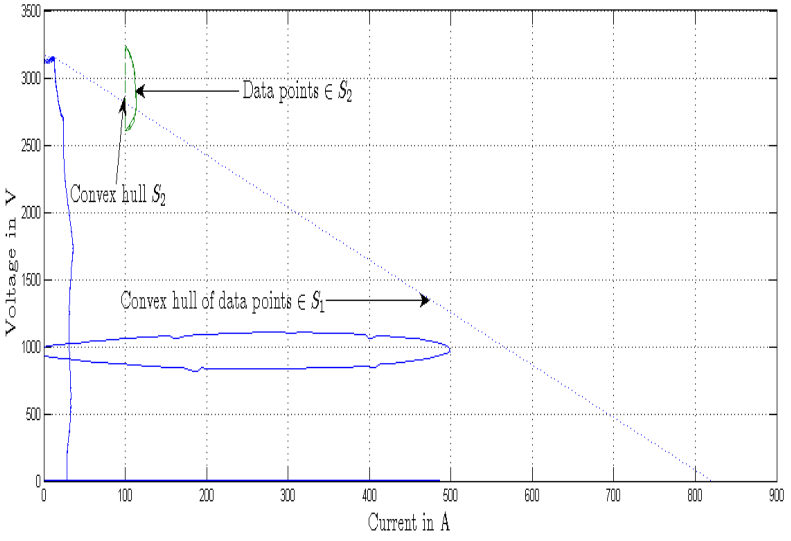

Figure 5.7: Original data scatter plot

is $N \times 2$ where $N = |S_1| + |S_2|$. $|S_1| = |E_1|$ and $|S_2| = |E_2| = 555$. The feature mapping problem can be expressed as in equation 5.9:

$$\begin{bmatrix} V(n) & I(n) \end{bmatrix}_{N\times 2} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} F_1(n) & F_2(n) \end{bmatrix}_{N\times 2} \tag{5.9}$$

5.3.2 Evolutionary Optimization Technique to Determine Feature Transformation Matrix

The problem of finding the 2×2 feature transformation matrix was framed as an optimization problem. The optimization problem can be stated as:
Minimize (length $(k_1 \cap k_2)$); where k_1 = convex hull of dataset S_1 and k_2 = convex hull of dataset S_2. The problem is convex because mathematically, finding the convex hull of a set follows equation 5.6. The matrix transformation is a linear operation and hence is convex. Therefore for any minimization function $f(x, y)$ if $\sum_i f(\alpha_i x_i + \beta_i y_i) \leq \alpha_i x_i + \beta_i y_i$ is followed with $\alpha_i + \beta_i = 1$ then that problem is convex in nature. This proves that the given optimization problem involves steps that are convex-natured. For solving this, derivative based methods are not applicable due to the problem's discrete nature. An iterative approach that uses different sets of candidate solutions and can guarantee convergence is required and accordingly heuristic methods were investigated. Genetic Algorithm (GA) was used in MATLAB running on a 2.3 GHz personal computer having Windows environment to solve the unconstrained optimization problem and find the feature transformation matrix. The new feature matrices will be such that they approach more towards linearly separability.

After the optimization was applied, the 4 elements of the feature transformation matrix were found. Both the dataset matrices containing elements $\in S1$ and $\in S_2$ were multiplied by this matrix respectively to result into two new matrices S_{T1} and S_{T2} which were combined and then labeled with '0s' and '1s' respectively to be used as a training data set for a classifier. The new scatter plot is shown in figure 5.8. The resulting features are still not completely linearly separable. They are linearly separable for high and medium feature values but a complex non-linear separation boundary exists for low feature values. Also a thing to be noted is that the number of values of F_1 and F_2 that are common to S_{T1} and S_{T2} but having opposite class labels are more in this case than the one plotted in figure 5.7. Therefore, entropy of data split along each individual dimension is not zero all the time thus requiring a tree based decision structure.

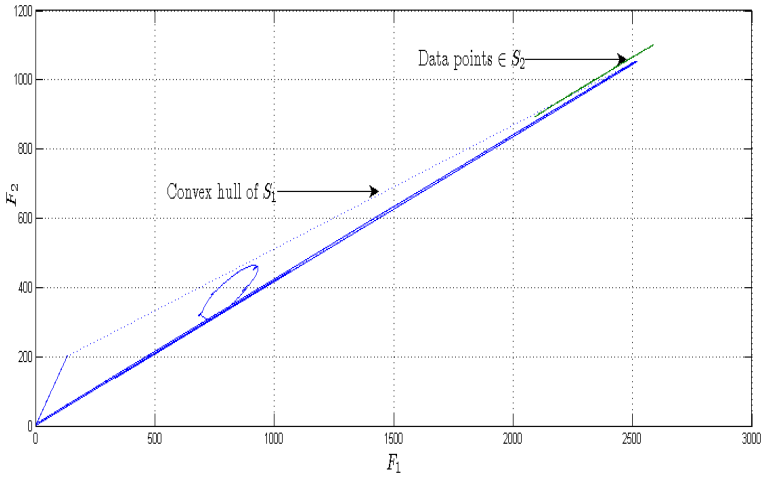

Figure 5.8: Transformed data scatter plot

A decision tree (DT) classifier based on the CART algorithm was programmed in Python 2.7 on the Raspberry Pi, as before. Gini-index was used as the impurity measure. The Gini index was used for the split criteria in place of entropy. The depth of the tree was not controlled by node-pruning as shown in figure 5.9 where the depth of the learned tree is controlled to two levels. For the built DT, $X[0] = F_1$ and $X[1] = F_2$. The insensitivity of DT classifiers to data scaling (a big issue in real-time data transmission) and feature transformation was the main reason for selecting them as classification models. The mean 10-fold cross-validation accuracy was found to be 100% after 3 runs.

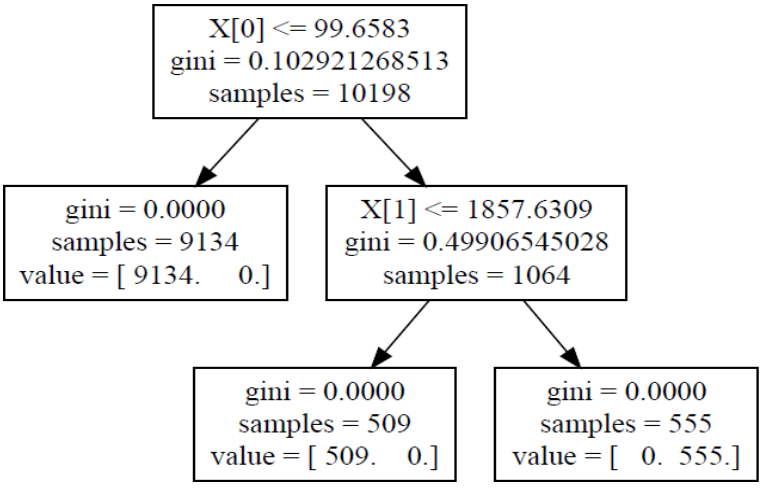

Figure 5.9: The built decision tree

5.3.3 Online Testing

The modified IEEE 13 bus feeder Simulink model was run on the personal computer and the instantaneous values of current and voltage at the PCC were sent to the classifier programmed in the Raspberry Pi. A switching program block that allows only positive half-cycle values to pass through was used to model the effect of an ADC without scaler and shifter circuit. The interface between the Simulink model and the classifier module was made via a local area network based connectivity between the two computers. A secure shell connection to the Raspberry Pi was established in order to remotely access its graphical user interface through the personal computer screen.

The simulation was run with a fixed-step solver due to Simulink real-time hardware requirements and the sampling time was fixed to 10μ seconds, as before, and for purposefully testing $G(x, y)$ from some $x \in S_2 \subset D$ at a different sampling rate. The data-type conversion, multiplexing and de-multiplexing of the voltage and current samples was carried in the same way as described in section 5.2. The 16-bit multiplexed string was received by the classifier module and unpacked into floating data point numbers. These unpacked numbers were arranged into 1×2 numerical python (numpy) arrays. These data points, sampled at every 10 μ seconds, were not multiplied by the 2×2 feature transformation matrix (stored in the Python program) and thus transmitted as raw 1×1 numpy arrays. The online classification was done for the following test cases:
Test Case A: L-G fault at the PCC
Test Case B: Single phase load switching at the node 675
Test Case C: UV+P-Q match case (after disturbance).

The online accuracy and the average time taken to predict the class labels for each test case is summarized in Table 5.5. Table 5.6 shows the consolidated confusion matrix for Test Cases A, B, C. The classifier took 4 seconds to learn the pattern (python serial code interpretation) and at the 5^{th} second, it started printing the predicted class label for each of the input arrays, in the Python console. 100% accuracy is observed for online testing on Test Case C and this is what was desired in real time situation. Table 5.7 shows a comparison of results from the presented approach with those reported earlier from use of other classifiers and PCA.

It can be seen that the DT online classifier module, when trained using the proposed feature extraction methodology, also takes an average of 0.3 ms (about $1/50^{th}$ of a cycle for 60 Hz system) to predict the class label. It does so with an

average accuracy of 98.04%. Also, a point to note from table 5.7 is that the DT used on PCA derived features gives almost same performance (rather a 0.127% higher average accuracy), the individual classification accuracy for the islanding initiator case was not 100% as has been achieved by the use of the computational geometry feature extraction method.

Table 5.5: Online Classifier Testing Results for the Proposed Feature Extraction Methodology

Test Set	Accuracy	Time Taken to Predict Class Label
Test Case A	97.062%	0.00032 s
Test Case B	97.057%	0.00037 s
Test Case C	100%	0.000338 s

Table 5.6: Confusion Matrix for Test Cases A, B and C

Test Case	A	B	C
Actual 0s into 0	3230	3241	2801
Misclassified Points	99	93	0
Actual 1s into 1	0	0	79
Misclassified Points	0	0	0

Table 5.7: Comparison of proposed approach with PCA derived features based classifiers

Classifier	Average Accuracy	Time Taken
K-NN	95.75%	offline testing
SVM (RBF Kernel)	83.5%	offline testing
DT on PCA	98.088%	0.00037 s
DT	98.04%	0.00034 s

The class-imbalance in the training dataset prepared using features derived from the proposed approach was much less than that in the case of the PCA approach. However, to ascertain the classifier's performance on online testing, the Kappa statistic analysis was carried in the same way as before. The value of the Cohen's Kappa statistic for all the online test cases were evaluated. For the test case C, it was found that the value of $\kappa \to \infty$. This was because the $\eta_{observed} = 1$ but $\eta_{expected}$ came out to be 0.946. For all the other test cases, it came out to be 0. This shows that the classifier performance for the test cases A and B equals the accuracy any random classifier would give on their respective confusion matrices. However as these events belong to class '0' whose instances are relatively more in number, their respective number of correct classifications are not expected coincidentally. A κ value of 0 for the test case C however would have been a concern but its high non-zero value indicates that the classifier has learned the pattern correctly.

5.4 Conclusion

This chapter described the implementation of a preemptive detection and inverter mode-change triggering strategy in response to the discovered islanding initiator. The PCA based approach described in chapter 4 was extended to the online feature extraction case. The use of a decision tree classifier for online testing was justified and its implementation with the PCA derived features was explained first. The online feature extraction and event detection was realized on a Raspberry Pi microcomputer. The CART algorithm chosen for the classifier implementation was programmed in Python and the tree was also tested on offline data initially. Simulations were run for the modified IEEE 13 node feeder in the personal computer and the V_{PCC} and I_{PCC} samples were transmitted to the Raspberry Pi over the LAN via UDP socket programming. The data samples were not mean-centered and simply multiplied with the reference PC matrix stored in the Python script. Encouraging results from an offline testing study on the DT trained with mean-centered data and tested with non mean-centered samples verified the practicality of this approach. The accuracy for the three testing sets did not deviate much from the values obtained when the test-samples were mean-centered. This thus proved the insensitivity of DT classifiers to feature scaling.

A convenient feature extraction methodology based on a new feature transformation method for real-time classification was proposed. The proposed methodology utilized concepts from computational geometry to project original training features onto a new feature space such that they approach linear separability. The feature transformation matrix was found by applying an evolutionary optimization technique to the objective function of maximizing the intersection between the convex hulls of the two training subsets. The test data samples came from different combinations of events simulated on the feeder model. It was found that the classifier gave an average accuracy whose value was 0.127% less than that obtained for the PCA based training. However, the testing accuracy for the individual test set pertaining to the islanding-initiator was 100% unlike that for the PCA based feature extraction case. The detection time using both the approaches came out to be the order of 0.3 ms on average. This was found to be less than the spike's duration thus it can be said with confidence that $t_{\text{label prediction}}$ is less than the time taken by a lateral fuse at PCC to blow up. Also, referring to the catalog specifications of the reference SEL-751 DOCR model, this time is less than what the relay will take to trip a PCC breaker.

This lead-time in detecting a potential cause of accidental islanding before the

action of a utility relay-circuit breaker pair can be used to trigger the change in operating mode of the inverter. Accordingly the mode-change triggering strategy was applied on the Raspberry Pi. The predicted class labels were translated into digital HIGH or LOW voltage corresponding to '1' and '0' respectively at the GPIO pins. These pins were further connected to the GPIO pins of the DSP that controls the firing circuitry of 1 phase grid-connected solar PV inverter. In this way, a preliminary realization of the inverter mode-change trigger in response to the discovered islanding-initiators was made. The preemptive islanding detection strategy for a simulated radial feeder was implemented with real-time data transfer from the source of events data (workstation simulation) to the online classifier module (implementation on Raspberry Pi) with encouraging results.

Chapter 6

Conclusion

The assumptions taken in this research work have been described in different chapters, wherever applicable. The reasons for selection of different models and parameter settings like feeder network selection, DG capacity, PCC node and simulation time-settings have been given. They have been aptly justified citing the objective-oriented reason of performing the exploratory study as required. The major contributions of this study are highlighted in section 1 while the next two sections list the limitations of this study and suggestions for further work.

6.1 Major Contributions of this Study

The major contributions of this study can be highlighted as follows:
1. This study has lead to the observation of a potential cause of unintentional islanding on a radial feeder with high solar PV penetration. The discovered over-current's occurrence has been verified in simulation, real-time simulation and on emulator hardware.
2. The use of PCA + K-L Divergence has been shown to be a potentially useful method for offline islanding detection. The use of this statistical measure for islanding-related studies has been reported for the first time.
3. A feature-extraction methodology that may have potential application for online classification of power system signals has been proposed.
4. The preemptive detection and mode-change triggering strategy has presented a possible energy management solution for tackling imminent islanding situations in distributed energy resources based power distribution networks.

6.2 Limitations of this Study

1. The study has been carried for different models of radial feeders, realized in simulation and emulator hardware. However most of the urban networks where grid-connected solar DG systems are coming up are now changing towards mesh based distribution networks.
2. A practical version of the static fault study to verify the status of the discovered over-current spike as an islanding initiator for the feeder models used could not be done.
3. The solar radiation has been kept constant at a value corresponding to the STC of a PV module. Although the assumption has been justified, the occurrences of P-Q match conditions throughout the day for the daily solar profile could not be taken in account.
4. No standard methodology or practice could be adopted for verifying the status of the discovered over-current as liable to cause islanding of a radial feeder section.

6.3 Suggestions for Future Works in this Research Domain

1. Short-circuit studies can be taken up to investigate the impact of solar PV integration on magnitude and direction of fault currents in a radial feeder. This will further strengthen the verification of anomalous over-currents, such as those discovered in this study, as potential causes of islanding
2. Actual protective devices like lateral fuses and DOCRs of small ratings can be used to realize a laboratory scale feeder having PV interconnection. This can help in observing the impact of any such anomalous occurrences on their tripping, when kept at the PCC.
3. To realize a more practical self-acting and self-acting system, semi-supervised learning techniques can be applied. Based on the confidence gained, a gradual shift towards unsupervised learning techniques can be made.
4. Single-phase grid-connected inverter must be actually implemented to test the mode-change trigger. The time taken in changing the operational mode of the inverter in response to an imminent islanding condition is of the most practical relevance.

This preliminary work, initiated on a low-cost hardware-Simulink integration platform, will be continued for signals generated from an emulator based network on

which the same conditions will be created. Once deployed on field, the expected action to be taken is to either trip the PV inverter or change its mode of operation to V/f regulatory mode, before the relay trips. The ability to act appropriately before the utility relay trip is necessitated by the fact that the loads can adjust their power consumption based on their model of dependence of P and Q with voltage and frequency [14]. Thus the preemptive cum predictive action is desired inside the inverter because if allowed to feed the loads even after the grid disconnects, a power balance condition with the loads can be satisfied and such an interaction can ultimately lead to formation and sustenance of an unintended island.

Appendix I

List of Publications

Journal Publications

1. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Multivariate Statistics and Supervised Learning for Predictive Detection of Unintentional Islanding in Grid-Tied Solar PV Systems", Applied Computational Intelligence and Soft Computing, Volume 2016, March 2016, Hindawi Publishing Corporation [SC Indexed; Impact Factor N.A.]

2. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Data Analytics and Computational Methods for Anti-islanding of Renewable Energy Based Distributed Generators in Power Grids", Renewable and Sustainable Energy Reviews, Volume 69, March 2017, pp. 493-502, Elsevier [SC Indexed; Impact Factor 8.050]

3. Shashank Vyas and Rajesh Kumar,"Computational Geometry based Methodology for Identification of Potential Islanding Initiators in High Solar PV Penetration Distribution Feeders", IET Renewable Power Generation, Online, December 2017 [SC Indexed; Impact Factor 2.635]

4. Shashank Vyas and Rajesh Kumar, "Online Classification Module for Identification of Potential Causes of Unintentional Islanding in High Solar PV Penetration Distribution Networks", IEEE Transactions on Emerging Topics in Computational Intelligence [Under Review]

International Conference Publications

1. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Observations from Study of Pre-Islanding Behaviour of Solar PV Systems Connected to a Distribution Network", Proceedings of the 4th International Conference on Advances in Computing, Communication and Informatics (ICACCI), Kochi, August 2015, pp. 645-650

2. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Exploration and Investigation of Potential Precursors to Unintentional Islanding in Grid-Tied Solar Photo Voltaic Systems", Proceedings of the IEEE International Conference on Energy Systems and Applications (ICESA), Pune, October 2015, pp. 577-582

3. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "A Framework for Investigative Analysis of Islanding and Pre-islanding Behaviour of Solar PV Inverters Connected to Distribution Grids", Proceedings of the IEEE IAS Joint Industrial and Commercial Power Systems / Petroleum and Chemical Industry Conference (ICPSPCIC), Hyderabad, November 2015

4. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Statistical Identification and Classification of Potential Islanding Precursors in a Grid-Connected Solar Photo Voltaic System", Proceedings of the 6th IEEE International Conference on Power Systems (ICPS), New Delhi, March 2016, pp. 1-6

5. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Unsupervised Learning in Islanding Studies: Applicability Study for Predictive Detection in High Solar PV Penetration Distribution Feeders", Proceedings of the 3rd IEEE UP Section International Conference on Electrical, Computer & Electronics Engineering (UPCON), Varanasi, December 2016, pp. 361-366

6. Shashank Vyas, Rajesh Kumar and Rajesh Kavasseri, "Preemptive Mode-Change Triggering for Inverters in Response to Unintentional Islanding Situations in High Solar PV Penetration Distribution Networks", Proceedings of the 2016 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), Trivandrum, December 2016, pp. 1-6

7. Shashank Vyas and Rajesh Kumar, "Identification of Potential Islanding Initiators in Radial Feeders with High Solar PV Penetration for Preemptive Preparedness", 12th IEEE International Conference on Industrial and Information Systems (ICIIS), Peradeniya, December 2017

Patents Published

Dr. Rajesh Kumar and Shashank Vyas, "Intelli-Island: A Smart Preemptive Islanding Detection and Mode-Change Triggering Logic for Grid-Connected Solar PV Inverters", 201611015180, Indian Patent Office Journal 23/2016, June 2016.

Appendix II

Raspberry Pi Specifications

Technical Specifications:

- Broadcom BCM2837 Arm7 Quad Core Processor powered Single Board Computer running at 900 MHz
- 1GB RAM
- 40 pin extended GPIO
- 4 USB ports
- 4 pole Stereo output and Composite video port
- Full size high definition multimedia interface
- Camera interface port for connecting the Raspberry Pi camera
- Display interface port for connecting the Raspberry Pi touch screen display
- Micro SD port for loading your operating system and storing data
- Micro USB power source

Bibliography

[1] D.P. Kothari and I.J. Nagrath. *Power system engineering*, chapter Distribution Systems. Tata McGraw-Hill, 2nd edition, 2008.

[2] Roger C. Dugan and Thomas E. McDermott. Distributed generation. *IEEE Ind. Appl. Mag.*, 8(2):19–25, March 2002.

[3] IEEE Std. 1547-2003. *IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems*, 2003.

[4] Benjamin D. Kroposki. *Optimization of Distributed and Renewable Energy Penetration in Electric Power Distribution Systems*, volume 69. 2008.

[5] Andrew Satchwell, Andrew Mills, Galen Barbose, Ryan Wiser, Peter Cappers, and Naim Darghouth. Financial impacts of net-metered PV systems on utilities and rate payers: A scoping study of two prototypical US utilities. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory, september 2014.

[6] M.S. ElNozahy and M.M.A. Salama. Technical impacts of grid-connected photovoltaic systems on electrical networks - A review. *J. Renewable Sustain. Energy*, 5(3), 2013.

[7] John Stevens, Russell Bonn, Jerry Ginn, and Sigifredo Gonzalez. Development and testing of an approach to anti-islanding in utility-interconnected photovoltaic systems. Technical Report SAND 2000-1939, Sandia National Laboratories, August 2000.

[8] Zhayoun Zhang, Wei Chen, and Zhe Zhang. A new seamless transfer control strategy of the microgrid. *The Scientific World Journal*, 2014, 2014.

[9] Mihai Ciobotaru, Remus Teodorescu, and Frede Blaabjerg. Control of single-stage single-phase pv inverter. In *Proceedings of 11th EPE 2005*, pages 1–10. EPE Association, 2005.

[10] Chris Greacen, Richard Engel, and Thomas Quetchenbach. A guidebook on interconnection and islanded operation of mini grid power systems up to 200 kW. Technical report, Lawrence Berkeley National Laboratory, April 2013.

[11] Akhilesh Magal, Tobias Engelmeir, George Mathew, Ashwin Gambhir, Shantanu Dixit, Anil Kulkarni, BG Fernandes, and Ranjit Deshmukh. Grid integration of distributed photovoltaics in India: A review of technical aspects, best practices and the way forward. Technical report, Prayas Energy Group, July 2014.

[12] M.E. Ropp, M. Begovic, and A. Rohatgi. Prevention of islanding in grid-connected photovoltaic systems. *Prog. Photovolt. Res. Appl.*, 7(1):39–59, 1999.

[13] Chun Li, J. Savulak, and R. Reinmuller. Unintentional islanding of distributed generation–operating experiences from naturally occurred events. *IEEE Trans. Power Del.*, 29(1):269–274, 2014.

[14] Fransisco José Pazos. Operational experience and field tests on islanding events caused by large photovoltaic plants. In *Proceedings of 21st International Conference on Electricity Distribution (CIRED)*, Frankfurt, June 2011.

[15] Bas Verhoeven. Probability of islanding in utility neyworks due to grid connected photovoltaic power systems. Technical Report IEA-PVPS T5-07:2002, IEA, September 2002.

[16] Robert F. Arritt and Roger C. Dugan. Review of impacts of distributed generation on distribution protection. In *Proc. 2015 IEEE Rural Electric Power Conference*, pages 69–74, 2015.

[17] M. Karimi, H. Mokhlis, K. Naidu, S. Uddin, and A.H.A. Bakar. Photovoltaic penetration issues and impacts in distribution network A review. *Renewable and Sustainable Energy Reviews*, 53(2016):594–605, 2015.

[18] Fernanda C. L. Trindade, Jose C. M. Vieira, and Walmir Freitas. Potential solutions for minimizing voltage sags caused by DG anti-islanding protection. *IEEE Transactions on Power Delivery*, 30(5):2242–2251, 2015.

[19] Miguel Louro and Carlos Cura. Network islanding A real event. In *proceedings of CIRED Workshop*, pages 1–4, Lisbon, May 2012.

[20] Fransisco José Pazos. Power frequency overvoltages generated by solar plants. In *Proceedings of 20th International Conference on Electricity Distribution (CIRED)*, Prague, June 2009.

[21] Smita Shrivastava, S. Jain, and R.K. Nema. A proposed hybrid method for islanding detection. *International Journal of Engineering Science and Technology*, 2(5):813–817, 2010.

[22] JA Laghari, H. Mokhlis, M. Karimi, A.H.A Bakar, and Hasmaini Mohamad. Computational intelligence based techniques for islanding detection of distributed generation in distribution network: A review. *Energy Conversion and Management*, 88(2014):139–152, August 2014.

[23] IEEE Std. 929-2000. *IEEE Recommended Practice for Utility Interface of Photovoltaic (PV) Systems*, 2000.

[24] IEC standard 62116. *IEC Test procedure of islanding prevention measures for utility-interconnected photovoltaic inverters*, 2008.

[25] UL standard 1741. *Inverters, Converters, Controllers and Interconnection System Equipment for Use With Distributed Energy Resources*, 1999.

[26] Irvin J. Balaguer-lvarez and Eduardo I. Ortiz-Rivera. Survey of distributed generation islanding detection methods. *IEEE Latin America Transactions*, 8(5):565–570, September 2010.

[27] Riccardo Sgarbossa, Stefano Lissandron, Paolo Mattavelli, Roberto Turri, and Alberto Cerretti. Analysis of load-induced unintentional islanding in low voltage grids with PV generators. In *Proceedings of 15th IEEE workshop on Control and Modeling of Power Electronics 2014*, pages 1–8, Santander, June 2014.

[28] Zhihong Ye, Amol Kolwalkar, Yu Zhang, Pengwei Du, and Reigh Walling. Evaluation of anti-islanding schemes based on nondetection zone concept. *IEEE Trans. Power Electron.*, 19(5):1171–1176, September 2004.

[29] Ward Bower and Michael Ropp. Evaluation of islanding detection methods for utility-interactive inverters in photovoltaic systems. Technical Report SAND2002-3591, Sandia National Laboratories, November 2002.

[30] M.E. Ropp, M. Begovic, and A. Rohatgi. Analysis and performance assessment of the active frequency drift method of islanding prevention. *IEEE Trans. Energy Convers.*, 14(3):810–816, September 1999.

[31] Ahmad Yafaoui, Bin Wu, and Samir Kouro. Improved active frequency drift anti-islanding detection method for grid connected photovoltaic systems. *IEEE Trans. Power Electron.*, 27(5):2367–2375, May 2012.

[32] Wodzisaw Duch. *What is Computational Intelligence and what could it become?* Department of Informatics, Nicolaus Copernicus University, 2007.

[33] K. El-Arroudi, G. Joos, I. Kamwa, and D.T. McGillis. Intelligent-based approach to islanding detection in distributed generation. *IEEE Trans. Power Del.*, 22(2):828–835, 2007.

[34] M.S. Thomas and Parveen Poon Terang. Islanding detection using decision tree approach. In *Proceedings of 2010 Joint International Conference on Power Electronics, Drives and Energy Systems (PEDES) and 2010 Power India*, pages 1–6, 2010.

[35] Omar N. Faqhurldin, E.F. El-Saadany, and H.H. Zeineldin. Nave bayesian islanding detection technique for distributed generation in modern disrtibution system. In *Proceedings of 2012 IEEE Electrical Power and Energy Conference (EPEC)*, pages 69–74, 2012.

[36] Ali Moeini Ahmad Darabi and Mohsen Karimi. Distributed generation intelligent islanding detection using governor signal clustering. In *Proceedings of 4th International Power Engineering and Optimization Conference*, pages 345–351, 2010.

[37] Om Prakash Mahela, AbdulGafoor Shaik, and Neeraj Gupta. A critical review of detection and classification of powerquality events. *Renewable and Sustainable Energy Reviews*, 41(2015):495–505, August 2014.

[38] N.W.A. Lidula, N. Perera, and A.D. Rajapakse. Investigation of a fast islanding detection methodology using transient signals. In *Proceedings of IEEE Power and Energy Society General Meeting (PES'09)*, pages 1–6, 2009.

[39] Y. Fayyad and A. Osman. Neuro-wavelet based islanding detection technique. In *Proceedings of 2010 IEEE Electric Power and Energy Conference (EPEC)*, pages 1–6, 2010.

[40] Mehrdad Heidari, Ghodratollah Seifossadat, and Morteza Razaz. Application of decision tree and discrete wavelet transform for an optimized intelligent-based islanding detection method in distributed systems with distributed generations. *Renewable and Sustainable Energy Reviews*, 27(2013):525–532, 2013.

[41] Jae-Hyung Kim, Jun ku Kim, Yong-Chae Jung, Chung-Yuen Won, and Tae-Hoon Kim. A novel islanding detection method using goertzel algorithm in grid-connected system. In *Proceedings of 2010 International Power Electronics Conference (IPEC)*, pages 1994–1999, 2010.

[42] David Velasco, Cesar Trujillo, Gabriel Garcer a, and Emilio Figueres. An active anti-islanding method based on phase-PLL perturbation. *IEEE Trans. Power Electron.*, 26(4):1056–1066, 2011.

[43] M. Vatani, T. Amraee, A.M. Ranjbar, and B. Mozafari. Relay logic for islanding detection in active distribution systems. *IET Generation, Transmission and Distribution*, 9(12):1254–1263, 2015.

[44] Biljana Matic-Cuka and Mladen Kezunovic. Islanding detection for inverter based distributed generation using support vector machine method. *IEEE Trans. Smart Grid*, 5(6):2676–2686, June 2014.

[45] O.N. Faqhruldin, E.F. El-Saadany, and H.H. Zeineldin. A universal islanding detection technique for distributed generation using pattern recognition. *IEEE Trans. Smart Grid*, 5(4):1985–1992, 2014.

[46] Vincent Spruyt. The curse of dimensionality in classification, April 2014.

[47] I. Kumaraswamy, T.K. Sandipamu, and Venkata Prasanth. Analysis of islanding detection in distributed generation using fuzzy logic technique. In *Proceedings of 7th Asia Modelling Symposium (AMS) 2013*, pages 3–7, 2013.

[48] S.R. Samantaray, K. El-Arroudi, G. Joos, and I. Kamwa. A fuzzy rule-based approach for islanding detection in distributed generation. *IEEE Trans. Power Del.*, 25(3):1427–1433, 2010.

[49] John Ross Quinlan. *C4. 5: programs for machine learning*, volume 1. Morgan Kaufmann, 1993.

[50] Faa-Jeng Lin, Kuang-Hsiung Tan, and Jian-Hsing Chiu. Active islanding detection method using wavelet fuzzy neural network. In *Proceedings of 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8, 2012.

[51] V.L. Merlin, R.C. Santos, A.P. Grilo, J.C.M. Vieira, D.V. Coury, and M. Oleskovicz. A new artificial neural network based method for islanding detection of distributed generators. *Electrical Power and Energy Systems*, 75(2016):139–151, 2015.

[52] Aziah Khamis1, Hussain Shareef, and Azah Mohamed. Islanding detection and load shedding scheme for radial distribution systems integrated with dispersed generations. *IET Generation, Transmission and Distribution*, 9(15):2261–2275, 2015.

[53] C.R. Aguiar, R.F. Bastos, R.V.A Neves, G.B. Reis, and R.Q. Machado. Fuzzy positive feedback for islanding mode detection in distributed generation. In *Proceedings of 2013 IEEE Power and Energy Society General Meeting (PES)*, pages 1–5, 2013.

[54] James O. Owuor, Josiah L. Munda, and Adisa A. Jimoh. The IEEE 34 node radial test feeder as a simulation testbench for distributed generation. In *Proceedings of IEEE Africon 2011*, pages 1–6, September 2011.

[55] L. Mihalache, S. Suresh, Xue Yaosuo, and M. Manjrekar. Modeling of a small distribution grid with intermittent energy resources using MATLAB/SIMULINK. In *Proc. 2011 IEEE Power and Energy Society General Meeting*, pages 1–8, 2011.

[56] James A. Silva, Hamed B. Funmilayo, and Karen L. Butler-Purry. Impact of distributed generation on the IEEE 34 node radial test feeder with overcurrent protection. In *Proceedings of 39th North American Power Symposium (NAPS '07)*, pages 49–57, 2007.

[57] Sumit Paudyal, Claudio A. Canizares, and Kankar Bhattacharya. Optimal operation of distribution feeders in smart grids. *IEEE Trans. Ind. Electron.*, 58(10):4495–4503, 2011.

[58] W.H. Kersting. Radial distribution test feeders. In *Proceedings of IEEE 2001 Power Engineering Society Winter Meeting*, volume 2, pages 908–912, 2001.

[59] Project Manager:Chad Abbey. Impact of large scale integration project number TIC706.3. Technical report, CANMET Energy Technology Centre-Varennes, 2007.

[60] CIGRE Task Force C6.04.02. *Benchmark Modeling and Simulation for Analysis, Design and Validation of Distributed Energy Systems*, September 2006.

[61] K.A Joshi and N.M. Pindoriya. Risk assessment of unintentional islanding in a spot network with roof-top photovoltaic system - A case study in India. In *Proc. IEEE Innovative Smart Grid Technologies Asia*, pages 1–6, 2013.

[62] George Vachtsevanos and Hoon Kang. Simulation studies of islanded behaviour of grid-connected photovoltaic systems. *IEEE Trans. Energy Convers.*, 4(2), June 1989.

[63] Satish J. Ranade, Nadipuram R. Prasad, Steve Omick, and Louis F. Kazda. A study of islanding in utility-connected residential photovoltaic systems part I - models and analytical methods. *IEEE Trans. Energy Convers.*, 4(3):436–445, September 1989.

[64] G.A. Smith, P.A. Onions, and D.G. Infield. Predicting islanding operation of grid connected PV inverters. *IEE Proc.-Electr. Power Appl.*, 147(1):1–6, January 2000.

[65] Tollgrade. Bringing our electric grid into the 21st century, 2014.

[66] R.G. Kavasseri, Yinan Cui, and S.M. Brahma. A new approach for event detection based on energy functions. In *Proceedings of 2014 IEEE Power and Energy Sociey (PES) General Meeting*, pages 1–5, 2014.

[67] S. Bhatt. Case studies for preventive maintenance systems: building a knowledge base. Technical Report 1000562, EPRI, Palo Alto, California, December 2000.

[68] Vijay Vittal, Trevor Merho, Mladen Kezunovic, Ce Zheng, Vuk Malbasa, Junshan Zhang, and Miao He. Data mining to characterize signatures of impending system events or performance from PMU measurements. Technical report, Arizona State University and Texas A & M University, August 2013.

[69] Zahra Pakdel. *Intelligent Instability Detection for Islanding Prediction.* PhD thesis, Virginia Polytechnic Institute and State University, May 2011.

[70] Rui Sun. *Wide Area Power System Islanding Detection, Classification and State Evaluation Algorithm.* PhD thesis, Virginia Polytechnic Institute and State University, December 2012.

[71] Cynthia Rudin, D. Waltz, R.N. Anderson, A. Boulanger, A. Salleb-Aouissi, M. Chow, H. Dutta, P.N. Gross, B. Huang, S. Ierome, D.F. Isaac, A. Kressner, R.J. Passoneau, A. Radeva, and L. Wu. Machine learning for the new york city power grid. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(2):328–345, 2012.

[72] Philip Gross, Albert Boulanger, Marta Arias, David L David Waltz, Philip M Long, Charles Lawson, Roger Anderson, Matthew Koenig, Mark Mastrocinquem, William Fairechio, John A. Johnson, Serena Lee, Frank Doherty, and Arthur Kressner. Predicting electricity distribution feeder failures using machine learning susceptibility analysis. In *Proceedings of National Conference on Artificial Intelligence*, volume 21, page 1705, 2006.

[73] A. Woyte, R. Belmans, and J. Nijs. Testing the islanding protection function of photovoltaic inverters. *IEEE Trans. Energy Convers.*, 18(1):157–162, March 2003.

[74] M. Dymond. Practical results from islanding tests on the 20 kW Kalbarri PV system. In *Proc. Solar 97, Sustainable Energy, 35th Annual Conference of the Australian and New Zealand Solar Energy Society*, December 1997.

[75] Michael Ross, C. Abbey, Y. Brissette, and G. Joos. Photovoltaic inverter characterization testing on a physical distribution system. In *Proc. 2012 IEEE Power and Energy Society General Meeting*, pages 1–7, 2012.

[76] R.J. Bravo, S.A. Robles, and E. Muljadi. Assessing solar PV inverters' anti-islanding protection. In *Proc. 40th IEEE Photovoltaic Specialists Conference*, pages 2668–2671, 2014.

[77] Arne Faaborg Povlsen. Impacts of power penetration from photovoltaic power systems in distribution networks. Technical Report IEA-PVPS T5-10:2002, IEA, February 2002.

[78] M. Hasheminamin, V.G. Agelidis, V. Salehi, R. Teodorescu, and B. Hredzak. Index based assessment of voltage rise and reverse power flow phenomena in a distribution feeder under high PV penetration. *IEEE Journal of Photovoltaics*, 5(4):1158–1168, May 2015.

[79] S. Srivastava, K. S. Meera, R.S.S Aradhya, and S. Verma. Performance evaluation of composite islanding and load management system using real time digital simulator. In *Proc. 15th National Power Systems Conference*, pages 242–247, Mumbai, 2008.

[80] Jorge Luis Sosa, Miguel Castilla, Jaume Miret, Jos Matas, and Y. A. Al-Turki. Control strategy to maximize the power capability of PV three-phase inverters during voltage sags. *IEEE Trans. Power Electron.*, 31(4):3314–3323, April 2016.

[81] Ahmed Kamel, M.A. Alaam, Ahmed M. Azmy, and journal = Electrical and Electronics Engineering: An International Journal year = 2013 volume = 2 number = 2 pages = 1-13 A.Y. Abdelaziz, title=Protection coordination of distribution Systems equipped with distributed generation.

[82] Y. Ates, M. Ozunoglu, A. Karakas, and A.R. Boynuegri. The case study based protection analysis for smart distribution grids including distributed generation units. In *Proc. 12th IET International Conference on Developments in Power System Protection*, pages 1–5, Copenhagen, 2014.

[83] Jian Tang, Wen Yu, Tianyou Chai, and Lijie Zhao. On-line principal component analysis with application to process modeling. *Neurocomputing*, 82(2012):167–178, 2012.

[84] X. Liu, D.M. Laverty, R.J. Best, K. Li, D.J. Morrow, and S. McLoone. Principal component analysis of wide-area phasor measurements for islanding detection - A geometric view. *IEEE Trans. Power Del.*, 30(2):976–985, April 2015.

[85] Y. Guo, K. Li, D. Laverty, and Y. Xue. Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches. *IEEE Transactions on Power Delivery*, PP(99), May 2015.

[86] Jinane Harmouche. *Statistical Incipient Fault Detection and Diagnosis with Kullback-Leibler Divergence: From Theory to Applications.* PhD thesis, Supélec, 2014.

[87] Abdulrahman Youssef, Claude Delpha, and Demba Diallo. Performances theoretical model-based optimization for incipient fault detection with KL divergence. In *Proceedings of 22nd European Signal Processing Conference (EUSIPCO)*, pages 466–470. IEEE, September 2014.

[88] Tom M. Mitchell. *Machine Learning.* McGraw Hill, 1997.

[89] Ahmad Darabi, Ali Moeini, and Mohsen Karimi. Distributed generation intelligent islanding detection using governor signal clustering. In *Proc. 4th International Power Engineering and Optimization Conference (PEOCO2010)*, pages 345–351, June 2010.

[90] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos,

D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[91] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*, chapter Additive Models, Trees, and Related Methods. Springer, 2008.

[92] Nima Mahmoudpour, Farhad Haghjoo, and Seyed Mohammad Shahrtash. Implementation of an integrated online instantaneous discrete wavelet transform decomposition toolbox in ATP-EMTP. *Electrical Power and Energy Systems*, 68(2015):373–383, 2015.