Author/s:
Demazy, Antonin Pierre Béatrice

Title:
Risk Management Frameworks and Methodologies for Modern and Resilient Power Systems Planning Using Machine Learning Techniques

Date:
2020

Risk Management Frameworks and Methodologies for Modern and Resilient Power Systems Planning

Using Machine Learning Techniques

Ir Antonin Demazy

Supervisor: Prof. T. Alpcan

Prof. I. Mareels

Department of Electrical and Electronic Engineering

The University of Melbourne

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

July 2020

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

The content included in Part II of this dissertation: "Extreme weather events risk and energy community resilience method in Bangladesh" has been developed as part of a contractual arrangement between the University of Melbourne and the Asian Development Bank (namely Contract Number 145572-S53313 - "Community-Energy Systems and clean energy resilience in Bangladesh") that grants the resulting Intellectual Property Rights to the Asian Development Bank. In order to protect that IP, the content of this dissertation will be embargoed.

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 100,000 words, exclusive of tables, maps, bibliographies and appendices.

Ir Antonin Demazy

July 2020

Acknowledgements

I would like to express my deep gratitude to Professor Tansu Alpcan and Professor Iven Mareels, my main research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work all along the way. Additionally, I specifically thank them for having enabled me to perform a portion of my research under a contractual arrangement between the University of Melbourne (UoM) and the Asian Development Bank (ADB), namely Contract Number 145572-S53313 - "Community Energy Systems and Clean Energy Resilience in Bangladesh".

I, therefore, extend my deep gratitude to Dr. Reihana Mohideen (UoM), Professor Pierluigi Mancarella (UoM), Mr. Francesco Tornieri (ADB) and Dr. Priyantha Wijayatunga (ADB), my research supervisors and project coordinators under the above-mentioned contract for the immense opportunity given to me to perform cutting edge socio-technical research applied to power system resilience in developing countries.

My thanks are also extended to the personnel and academics of the Melbourne School of Engineering as well as my peer students and close ones for their valuable support, advice and encouragement.

Preface

The research activities and outcomes are listed in this section. The content of this research, and the writing of this dissertation has been executed predominantly by the student, but credit must be given to the supervisors who provided technical comments and guidance.

Financial support provided by the Australian Government Research Training Program Scholarship, in particular the fee offset scholarship is gratefully acknowledged as well as the partial stipend granted by the University of Melbourne.

The content of this research that is illustrated in Part II of this dissertation has been developed as part of a broader contractual agreement between the University of Melbourne and the Asian Development Bank for the project: "Community Energy Systems and Clean Energy Resilience in Bangladesh", which final report will be published by the Asian Development Bank. The related content has been developed the student, but credit must be given to the project supervisors Dr. R. Mohideen and Prof. P. Mancarella who provided technical comments and guidance.

During the research, the following deliverables have been produced:

- Conference Paper: A.Demazy, T.Alpcan, I.Mareels and S.Saha 2017, **Assessment of voltage stability risks under stochastic net loads using scalable SVM classification**, published by IEEE 2017 Australasian Universities Power Engineering Conference (AUPEC) on Nov 2017.

The contribution of each author is as follows: First author: Developing the network model, bifurcation point calculation and support vector machine model in Matlab, performing simulations, writing the paper, and presenting in the conference. Second, Third and Fourth authors: Supervision, proofreading, and providing technical comments.

- Industry project and conference Paper: A.Demazy, A, Kalloniatis and T.Alpcan 2018, **A Game-Theoretic Analysis of the Adversarial Boyd-Kuramoto Model**, published by GameSec 2018: Decision and Game Theory for Security on Oct 2018.

  This paper has been produced as the final outcome of a contractual agreement between the University of Melbourne and the Australian Defence Force. The content developed during this project is not included in this dissertation being off-topic but the skills developed by the student during the course of this project (such as Python coding) has been highly beneficial for his main research activities. The contribution of each author is as follows: First author: Developing the Boyd-Kuramoto game-theoretic approach model in Python, performing and interpreting the simulations, writing the paper. Second, Third authors: Supervision, proofreading, providing technical comments and presenting in the conference.

- Industry project: A.Demazy, R. Mohideen and P. Mancarella 2019, **Inclusive Community Energy Resilience in Bangladesh**.

  This project, contracted between the University of Melbourne and the Asian Development Bank, attempted to answer the following key question: "How to design a clean Power System extension in Bangladesh that is resilient to major weather events, affordable and accessible by the greatest majority of local communities, including the most vulnerable, that enables a social and economic empowerment and respect land, livelihood and environment?". To answer this question, two major components where studied and developed: firstly a Power System Socio-Technical Planning Framework informed by the IEEE Smart Grid domains framework that is Gender Equality and Social Inclusion oriented; secondly a methodology for community and power system resilience assessment that is linked to the power system planning framework. The second component is described in part II of this dissertation. The contribution of each author is as follows: First author: Developing the Power System Socio-Technical Planning Framework and the Resilience Methodology, writing the report. Second, Third authors: Supervision, proofreading, providing technical comments.

- Conference and consultations: A.Demazy, R. Mohideen and P. Mancarella 2019, **Community Energy Systems and Clean Energy Resilience in Bangladesh**, Third Lateral Learning Program on Smart Grid Technologies and Implications for Inclusive Development, Melbourne, Australia [1].

  Presentation of the findings and draft of the final report by the first author to the audience composed of Asian Development Bank and South-Asian countries officials.

- Start-up launch: A.Demazy, J.Fowler and P.Maasoumi 2019, **Dr Dingo Pty LtD**, Second winner of the Energy Hack 2019 competition organised by the Melbourne Energy Institute.

  Dr Dingo is an innovative educational tools creator, seeking to utilise games to educate people about the complexity of the energy mix, the disruption of emerging technologies and how it's operated to produce electricity. By spreading awareness about these topics, it intends to empower people to make more sustainable energy decisions, better understand the energy-related discussion in the media and know-how their choices impact the rest of the community. The authors are co-founder and manager of the start-up company that will soon release a power system management game.

- Conference Paper: A.Demazy, T.Alpcan and I.Mareels 2020, **A Quantitative Risk Framework for DER-rich Power System Planning and Decision Making**, published by IFAC: 21st IFAC World Congress, July 2020

  The contribution of each author is as follows: First author: Developing the quantitative risk framework model in Matlab, performing simulations, writing the paper, and presenting in the conference (tbc). Second and Third authors: Supervision, proofreading, and providing technical comments.

- Journal Paper: A.Demazy, T.Alpcan and I.Mareels 2020, **A Probabilistic Reverse Power Flows Scenario Analysis Framework**, submitted for publication to IEEE Open Access Journal of Power and Energy on June 2020 .

  The contribution of each author is as follows: First author: Developing the AI-based scenario analysis framework based on Python/OpenDSS/Keras, performing simulations and writing the paper. Second and Third authors: Supervision, proofreading, and providing technical comments.

- ADB South Asia Working Paper: A.Demazy and R. Mohideen 2020, **Inclusive Community Energy Resilience in Bangladesh**, submitted for publication to ADB on July 2020.

  First author: Developing the Power System Socio-Technical Planning Framework and the Resilience Methodology, writing the working paper. Second author: Supervision, proofreading, providing technical comment

Abstract

Renewable energy technologies, customer behaviour, and new regulations are key factors contributing to a change in power generation paradigm that is becoming increasingly decentralized and embedded in the distribution network. The new paradigm, together with strong opportunities, is bringing challenges for power networks that must be adequately anticipated and planned to maintain the security and reliability of the power supply. This research addresses two key challenges for developed power networks and one challenge for developing networks located in countries vulnerable to extreme weather events. For developed power networks, this research formulated risk assessment models based on Artificial Intelligence techniques that enable power system planners to analyse vast numbers of scenarios and assess the impact of voltage excursions and reverse power flows as a result of elevated penetration of distributed energy resources. The novelty of the work is derived from the scalability of the proposed models and its end-to-end approach that includes financial modelling of the impacts. For the developing network, this research developed one risk-based methodology to assess resilience to extreme weather events that is linked to power system planning. The novelty of the proposed methodology is derived from the problem formulation that explicitly considers both the technical power system resilience and the social community energy resilience in quantifiable terms that are linked to power system planning via an optimization problem.

Table of contents

List of figures

List of tables

Chapter 1

Background and motivation

1.1 An overview of the Australian power network and market

Australia has one main two other electricity grids that were centrally designed for the reliable delivery of electricity to households and businesses across the different states of the country:

- the main grid, or the "National Electricity Market" (NEM), serves Queensland, New South Wales, Victoria, South Australia and Tasmania,
- the South West Interconnected System (SWIS) serves the southern part of West Australia,
- the North West Interconnected System (NWIS) serves the northern part of West Australia,
- the Darwin-Katherine Electricity Network serves part of the Northern Territory.

Additionally, multiple ’island’ grids are in place in remote areas of the country.

The NEM is an arrangement in the electricity sector for the connection of the synchronous electricity transmission grids of the eastern and southern Australia states and territories to create a cross-state wholesale electricity market. In total, there is over 850,000 km of distribution grid and 45,000 km of transmission grid in operation across Australia, with the eastern and south-eastern states boasting the National Electricity Market – one of the longest interconnected electricity markets in the world with an end-to-end distance of more than 5000 kilometres, and 40,000 kilometres of high voltage transmissions line (see Fig.1.1).

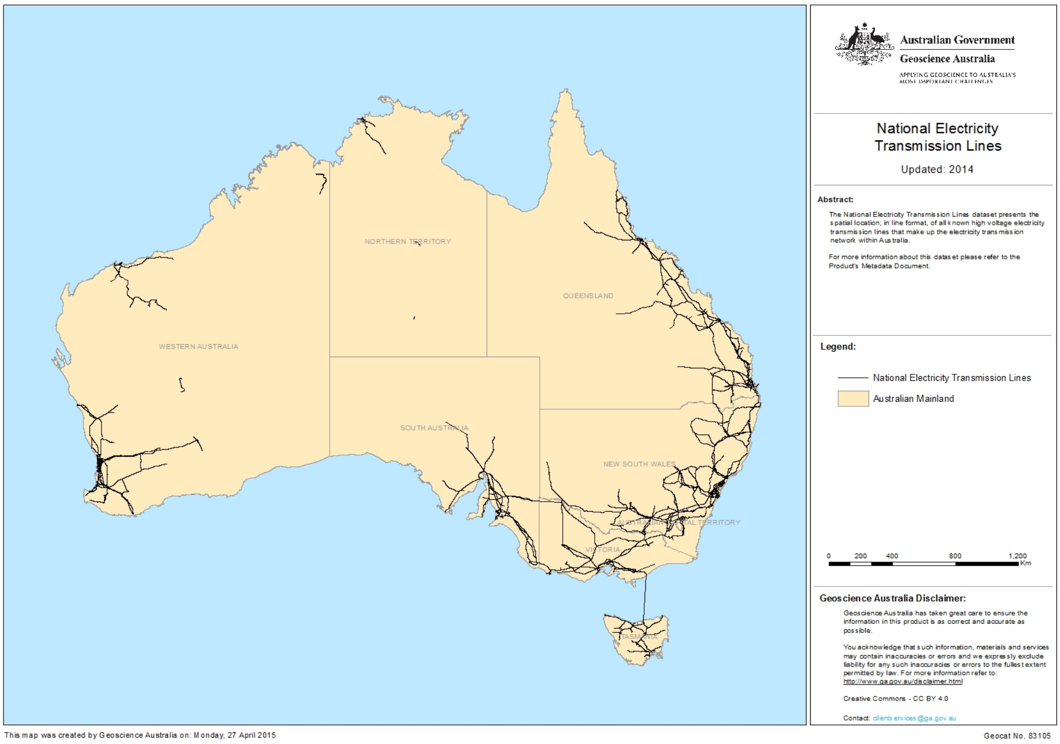


Fig. 1.1 Australian transmission line network (Source: Geoscience Australia)

NEW has a total electricity generating capacity of almost 54,421 MW (as at December 2017). During the 2017-18 financial year, 203 TWh were generated serving 9.7 million end-use consumers. Approximately 40% of NEM generation is consumed in New South Wales, while Victoria and Queensland consume approximately 25% each.

NEM leverages a traditional supply chain structure, with centralized power generation, transmission lines, distribution through poles and wires, and delivery of supply to end consumers through retail businesses (Fig.1.2).

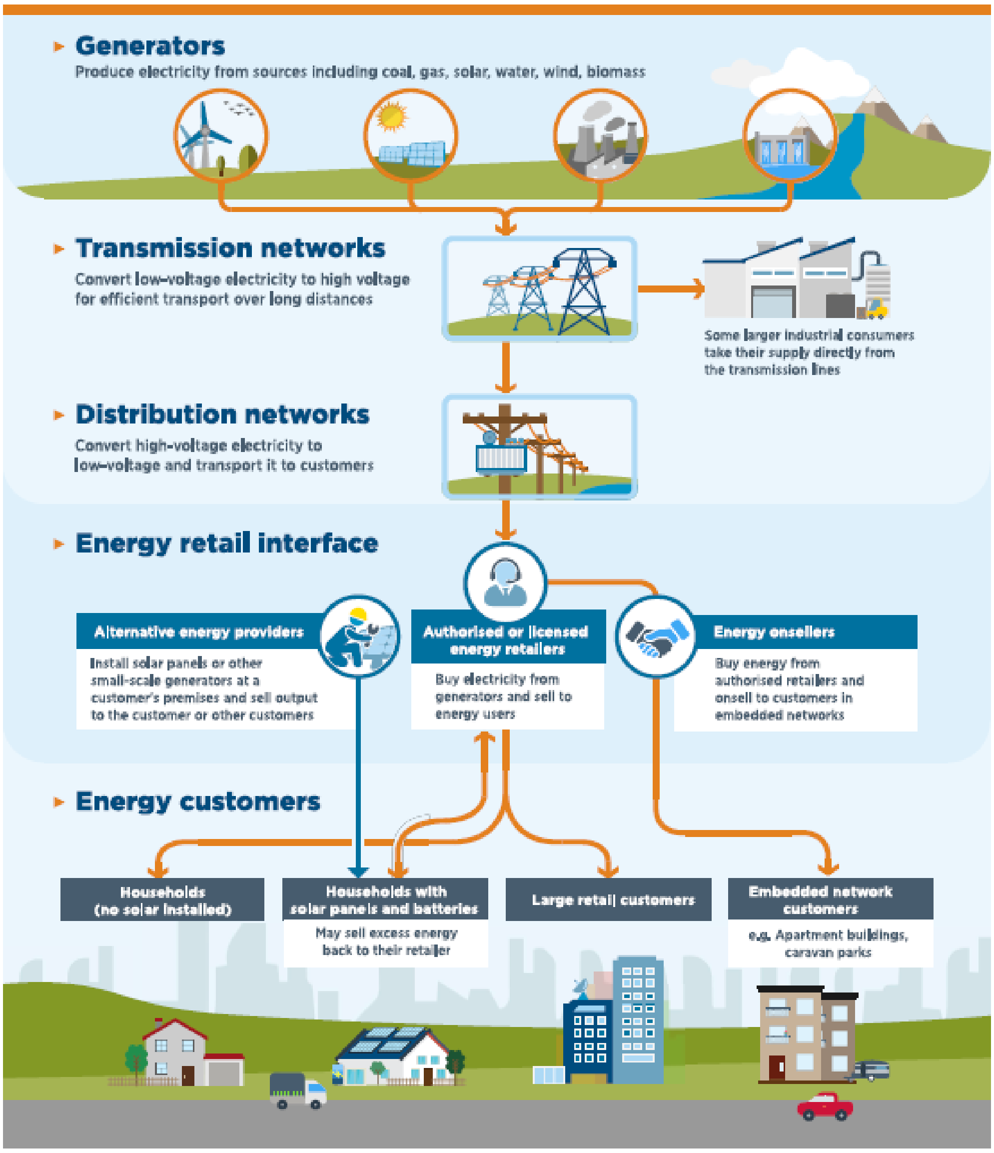


Fig. 1.2 Electricity supply chain infographic (Source: T&O Energy Consultant)

There are three main electricity markets and four types of participants as illustrated in Fig.1.3:

- The NEM is a wholesale market involving exchanges between electricity producers (generators) and retailers (companies that purchase electricity from generators and then sell it to homes and businesses).
- The Retail Market involves electricity retailers selling the energy they have purchased wholesale (via the NEM) to homes and businesses.
- The Financial Market describes various contracts set up between electricity producers, retailers and investors, which act as insurance policies by reducing the significant risk of financial exposure faced by market participants due to electricity price volatility that can occur. These financial contracts may lock in a firm price for electricity that will be produced or consumed at a given time in the future.

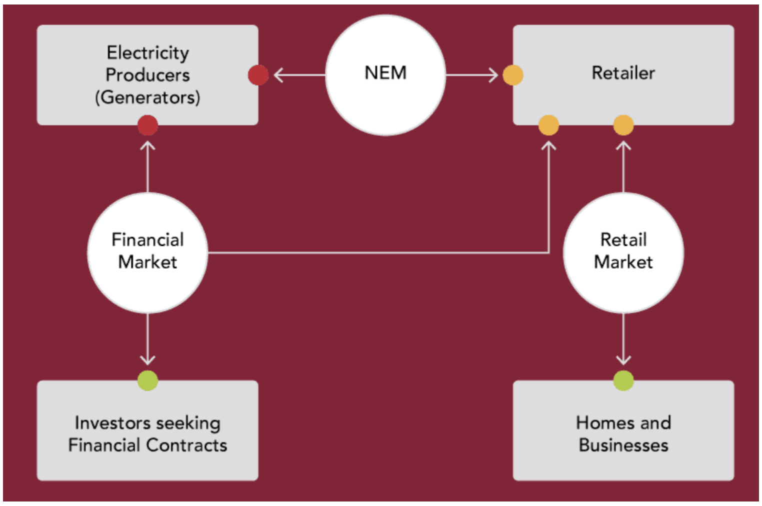


Fig. 1.3 Electricity markets and participants (Source: AEMO)

1.2 The emergence of a new power system paradigm and the need for new planning tools

In recent years, the classical paradigm of power systems has been disrupted primarily because of the incorporation of intermittent renewable energy sources.

Traditionally, the security, resilience and efficiency of power systems are guaranteed by large-scale centralised and predominantly thermal generation units, and, highly redundant transmission/distribution network. The generation units constantly adapt to follow in real-time the power demand at the edge of the network. In industrialised countries, where the network has evolved over more than a century, the transmission and distribution infrastructures are well suited and appropriately sized to be able to transfer the energy required at peak-time from large scale centralised generators to customers. In that paradigm, the power flow is uni-directional, from the generators to the load (consumers) and the amount of power generated at any instant is following the instantaneous aggregated load. Limited ancillary services ensure that the network is managed within its physical constraints, whilst maintaining the required quality of service: a constant frequency and regulated voltages at all the nodes.

In the industrialised countries, this paradigm is slowly evolving with the emergence of distributed energy resources (DER) within the established power system, mainly from a variable renewable source that is challenging the unidirectionality and predictability of power transfer. Distributed energy resource refers to power generation at the point of consumption. When the distributed source of generation is intermittent, meaning that the source is not continuously available for conversion into electricity and outside direct control (eg. Solar PV, Wind), the net-demand at point of consumption can be negative for some time and with fast intermittency. When the net-demand is negative, if not curtailed or stored, power will start flowing from the consumption point to the grid, therefore activating a "bi-directional" power flow within the grid which is a totally new phenomenon that was inexistent in the old centralised generation paradigm. The significance and impact of the phenomenon will depend on the intensity of each imbalance, their variability, duration and concentration within the entire power system.

Until the adoption of DER is contained to marginal values within the power network, the effects on the existing power system infrastructure are negligible and can be ignored. Conversely, if the scale of adoption and individual capacity of each DER was to increase to significant values, their impact would cease to be negligible and would need to be appropriately forecasted and assessed to plan for adequate adaptation measures in the power network.

The traditional tools and techniques used to size and configure a power network were not constructed to identify and assess the impact of intermittent decentralised generation within the network, and therefore there is a need for new or modified approaches to assess that specific impact.

The driving forces behind the emergence of a new power system paradigm is a conflux of multiple disrupting factors, of technical, social and political dimensions. The next sections propose a deeper analysis of the disruption factors and their impact on the current power system.

1.3 Main disruption factors

In recent years, driven by technical progress, customer behaviour and new regulatory landscape, the old paradigm of power systems is slowly evolving with the emergence of distributed generation within the system, mainly from a variable renewable source that is challenging the unidirectionality and predictability of power transfer. The main disruption factors can be summarised as follows:

1.3.1 Technological progress

Generally speaking, the progress in technology over time usually grows at a fast pace and while the number and level of sophistication in technological capabilities is growing, their size and cost are typically decreasing. Related to power systems, concurrent progress in multiple areas are contributing to potentially disrupt the old centralised paradigm for power system:

- **Renewable Generation Technology** - Improvements in efficiency and affordability of renewable generation is putting them "on the map" at different scales. The technologies that are significantly increasing their penetration nowadays are the PV (Photo-Voltaic) panels and Wind Turbines. Those technologies are dependant on the availability of intermittent energy source (sun or wind) to produce electricity and are typically installed within the distribution network at medium (Wind Farm) or low voltage level (Solar Rooftop). Therefore, due to intermittency, their penetration is increasing the variability of net demand (local load minus local generation) and occasionally creating reverse power flow within the system when the local generation exceeds the local demand. Increasing penetration of decentralised and intermittent generation is an important factor of changes for the old centralised power system paradigm. It is to mention that further renewable generation technologies are emerging and being economically viable (such as Maritime, Geothermal, Concentrated Solar,..) and will potentially all play a role in evolving the power system towards a new paradigm.

- **Storage Technology** - Similarly to renewable generation, energy storage technologies are emerging in terms of affordability and are expected to play a growing role in actively managing the variability of the net load within the system. In its broader definition, energy storage can be a device that stores energy produced in excess (eg. batteries, flying wheels, pumped water, heated molten salt,..) or a managed deferral of energy consumption at certain time (eg. controlled Electric Vehicle charge, controlled air-conditioner, controlled water heater,..). The concept of energy storage is providing flexibility in the system allowing for some buffers.

- **Connected SMART electric/electronic devices** - current and growing affordability and compactness of microprocessors, controllers, and sensors are enabling their use in a wide range of electronic devices or electric appliances and therefore include a certain level of logic and sophistication in the operation of those so-called "smart" electric/electronic devices. Besides, the affordable connection of those to a wider network is enabling collaboration amongst them and/or collaboration with a centralised control unit. The progress and innovation in connected smart devices impact are twofold for the power system. First, smart devices can be inserted and used for active management of the power network itself, in that case, the smart devices or sensors are the network assets. Second, smart devices can be the end-users of electric power. In both cases their capability to be "programmed" and to communicate and collaborate enable a level of flexibility in the power and energy management of the power system.

- **Power Electronic / Inverters** - Similar to SMART devices, affordability and technological advancement in power electronics contribute to reinforce the control capabilities and potentially the reliability of integrated power network solutions where decentralised generation and storage play a significant role.

- **Communication Technology** - mobile and internet communication technology and deployed infrastructure in addition to standard communication protocols are enabling a wide array of interconnection between devices which can be the backbone to centralised or distributed control algorithms.

- **Information Technology** - affordability, availability and scalability of cloud computing is a strong enabling force for the deployment of advanced and real-time network management algorithm and active demand management systems.

The parallel improvement and advancement of technology in the above-mentioned dimensions and their mutual combination is an important factor disrupting the old centralised power system paradigm. Nevertheless, the business case for change is strictly speaking not compelling yet, but additional factors, sometimes less tangible such as the sociological aspects are additional forces to change.

1.3.2 Customer behaviour

Interestingly, in parallel to the technical progress, people beliefs, mindset and motivation are also evolving while new economical and business models are emerging. Those changes, usually enabled by technology, are deeply modifying customer behaviours to the point of disrupting past and well-established industries.

Relevant to the power industry, interesting trends are emerging in the customer behaviours that play a crucial role in disrupting the old paradigm by significantly increasing the penetration of decentralised renewable generation within the power system. The scale and complexity of those decentralised generations can vary from individual household solar PV to communities hybrid-microgrids that can run on islanded mode.

It is currently extremely complex to estimate with sufficient level of certainty the future level and pace of this penetration. Individual business case and reasonable payback period will surely influence the level and pace of penetration but there are some additional tangible or intangible factors that can influence it. From a mindset or sentiment point of view, individual customers are becoming increasingly sensitive to environmental issues and sustainably in general and this is an important intangible factor that can influence the choice to equip a house with rooftop solar PV and potentially batteries even if the rational business case is not hugely compelling. This will mainly apply to the wealthier part of the population who can afford those type of equipment. It is to mention that the same mindset might also influence the political choice of the population and increase the relative power of "green" political parties that will naturally foster and incentivise the penetration of renewable power generation.

Another potential influencing factor is the rise of the "share" or "peer-to-peer" economy where users can freely trade products or services on a distributed platform. Start-ups are timidly emerging in multiple countries with the intent to offer an energy trading platform for individual consumers based on distributed ledger (blockchain) to record the transactions the availability of such platforms would enable the prosumers (consumers who are also locally producing power) to freely and seamlessly trade their excess energy and/or modulate their consumption based on real-time spot energy price. The trading capability could entice more customers to invest in local generation capacity and take advantage of the high volatility of energy spot price. This could impact individual customers as well as small or medium communities or commercial estates with growing complexity and capacity of the installed assets (generation, storage, controllers). The potential success and uptake of the peer-to-peer trading platform are uncertain and hard to predict but as already observed in other industries (eg. Transport - UBER, Hospitality - AIRNB,..) the rate of adoption can rise very sharply.

1.3.3 Regulation

Network and Electricity Market Regulation

In Australia, the traditional supply-chain model of electricity includes:

- Generators that make electricity from a primary energy source which then flows into the high-voltage transmission network

- The transmission network that transports the electricity to the distribution network
- The distribution network that transports the electricity to residential and commercial buildings
- The end-users of electricity

The model includes the retailers who arrange the delivery and billing of electricity to customers and customers who pay the electricity to the retailers. With the emergence of new technology and in particular the DER, the Australian electricity system is less linear as the line between generators, consumers and retailers become less clearly defined.

The flow of electricity through the system is controlled by the National Energy Market (NEM) that includes a competitive wholesale generation sector, monopoly network businesses and competitive retail sector. The physical electricity system and the electricity market are regulated by three bodies: Australian Energy Market Commission (AEMC) which makes the rules for the energy system and market; the Australian Energy Market Operator (AEMO) which operates the energy systems and markets; the Australian Energy Regulator (AER) which monitors and regulates. The regulation structure was tailored to orchestrate unidirectional and centralised power generation but is less adequate to regulate embedded generation at the end-user point.

Decarbonisation incentives

The Paris Climate Agreement entered in effect the 6th November 2016 with the aim of "enhancing the implementation" of the UNFCCC through: (a) Holding the increase in the global average temperature to well below 2 °C above pre-industrial levels and to pursue efforts to limit the temperature increase to 1.5 °C above pre-industrial levels, recognizing that this would significantly reduce the risks and impacts of climate change; (b) Increasing the ability to adapt to the adverse impacts of climate change and foster climate resilience and low greenhouse gas emissions development, in a manner that does not threaten food production; (c) Making finance flows consistent with a pathway towards low greenhouse gas emissions and climate-resilient development.

The accord sets carbon emission objective targets to each country, which have to plan for partial and progressive de-carbonisation of their economy. One of the industrial sectors that are currently highly carbonised is the electric power generation together with heating and transport even to a greater extent. Depending on the ruling political party, it is observed that in multiple countries incentive schemes have been introduced to encourage individuals to install rooftop solar PV or generation companies to install wind farms in parallel with carbon emission trading schemes or taxes.

Subsidised incentives are most certainly a factor that encourages an increased penetration of decentralised renewable generation within the power system, disrupting the old centralised paradigm in developed countries.

Interestingly in developing countries, most power networks are still under development to reach the vast majority of the population. This allows for those countries to leapfrog the technology gap and adopt modern renewable technologies in the original design of the overall power network.

1.4 Key Impacts on Power Systems

1.4.1 In developed countries

The previous section is describing some key factors that are disrupting the old centralised unidirectional power system paradigm well established in the developed countries, and especially in Australia which is the major developed country's case study for this research. A major disruption comes from increased penetration of renewable generation within the medium and low voltage distribution network. The high variability and the uncontrollable nature of those power sources are impacting the reliability and security of the entire power system, and when there is an excess local generation, the power flows from the consumer to the network (bi-directional power flow). The main impacts can be briefly summarised as follows:

- **Capacity planning and dispatch** - The long-term and short-term generation adequacy, being the capability to meet demand with a certain level of reliability of supply, are impacted by the mostly unpredictable nature of the renewable power generation. The high variability of renewable sources requires the availability and dispatch of more reserves in the system to maintain reliability. The long term planning faces a trilemma by contemporaneously minimising the carbon emission, maintaining the security of energy supply and minimising the energy cost.

- **Security and reliability of supply** - The fast variation of intermittent generation and the misalignment between peaks of renewable generation and peaks of load increase the risk of unbalance between demand and supply within the system. A point in time power unbalance will manifest through frequency and voltage shift. In the old paradigm, the primary frequency response (for non-catastrophic) event is guaranteed by the inertia of the spinning rotors of generators. The inclusion of some renewable technologies that do not have intrinsic inertia (typically the solar PV or some type of wind turbines) is weakening the resilience of the entire system and some alternative frequency response solutions will need to be activated if the penetration of those technologies become significant. Similarly, an unbalance between power demand and supply can cause voltage fluctuation within the system and causing a voltage excursion risk. Unlike frequency which is unique for the entire system, the voltage values are different at every bus and nodes of the system and their fluctuation is local and dependent on the network topology. The inclusion of decentralised generation and the potential excess energy generated in a non-peak period that is injected in the system can cause the voltages to exceed statutory limits in portions of the distribution network that were typically sized and installed in a "fit-and-forget" mode without considering a bi-directional flow of energy.

- **Security - thermal limits** - A significant penetration of local renewable generation deep into the low-voltage distribution network is limited by the rating and thermal limit of the local infrastructure in place, mainly due to the bi-directionality of power within the system that was not taken into account in the original sizing of the infrastructure. As per the voltage fluctuations, the risk of exceeding thermal limits within the system is localised and depending on the topology of the network.

- **Security - fault protection** - Traditional fault protection assets such as circuit breakers are located and configured to sense and interrupt abnormal current intensities that would be caused by a fault within the power system. Traditional design and calibration of those assets didn't take into consideration eventual reverse power flows that might exist in the system due to DER and influence their operation.

The impacts are broad and deep, mostly due to the variability and uncontrollability of some renewable sources of energy that are commonly used in the decentralised paradigm. Long term investment in capacity and reserve planning are disrupted as there is an increasing uncertainty on the return of investment for traditional generation technologies. This is mainly due to a decrease in traditional capacity utilisation factor when renewable generation is available at virtually null marginal cost. But on the other hand, the variability and unpredictability of renewable source dictate the requirement to increase the level of the spinning reserve at any time to maintain a certain level of security within the network. We can see the difficult trilemma in action: trying to balance carbon emission, level of security and cost of energy.

Additionally, from a network planning perspective, the reverse power flows originated by embedded generation are potentially disrupting the network configuration and might necessitate network re-enforcements or re-configuration. From an operational point of view, the inclusion of decentralised renewable is disrupting the traditional mechanisms in place to guarantee the stability and resilience of the network at any time. Balancing at all time the demand and supply is made more difficult by an intrinsic increase of variability and a loss of inertia within the system, while additionally, the protection mechanisms have to cope with bi-directional power flow in the system.

1.4.2 In developing countries

In developing countries, the emergence of affordable renewable decentralised generation technology presents an opportunity to provide modern energy to most remote and or vulnerable communities where it would have been uneconomical to extend the main traditional power grid. Nevertheless, some of those countries are also the most vulnerable to extreme weather events and while planning for their power network extension, the resilience of the power network to those events must be taken into consideration. Bangladesh, which is one the most vulnerable country to extreme weather event is the main developing country's case study for this research.

1.5 Aggravating Factors

The adoption of renewable energy sources is fast increasing globally. For instance, in its global roadmap REMap to 2050, IRENA [8] predicts an increase of global electricity demand from 20.204 TWh/yr in 2015 to 41.508 TWh/yr in 2050, with a renewable share jumping from 24 % to 85% in the same period and where the solar PV capacity contribution would jump from 223 GW to 7.122 GW. Diverse predictions of the increased share of renewable exist, some more conservative and some more optimistic, but the truth is that no one exactly knows the pace and extend at which renewable energy sources will be penetrating the power networks in the next decades.

In Australia, as of September 2019, more than two million, or 21 per cent, of households had rooftop solar PV, with a combined capacity exceeding 13 GW, of which more than 3 GW where installed in the preceding 12-month [9]. The state of Queensland is leading the pack with nearly 36% of dwellings with PV installation. Such levels of penetration have already caused an unwanted rise of the voltage within the system and consequently and consequently the regulator has increased its voltage upper statutory limit from 230 volts +6% to 230 volts +10%.

Most of the driving forces that are fuelling the increase adoption of household Solar PV systems (eg. Incentives, decreasing price of technology, consumer mindset,..) are under no control of the Distribution Power Network Service Providers (DNSP) who must nevertheless manage the power system and guarantee the security and reliability of the power supply under increasing penetration of DER within their system.

1.6 Use a probabilistic risk-based approach to plan the effect of disruption

The penetration of DER in the power system is steeply increasing in countries like Australia and already reaching levels that cause non-negligible impacts on the reliability and security of the power systems. The pace, level and concentration of penetration are mostly uncontrolled by the DNSP who must react to maintain a secure and reliable supply of electricity to their customer.

With levels of penetration that are now in the double-digit and fast-growing, the DNSPs must proactively equip themselves with tools and methodologies that enable them to forecast and analyse the impact of DER-rich configuration of their network.

The complexity of the task at hand is particularly elevated due to elevated variability and stochasticity that plays at multiple levels. In facts, at macro-level, there is uncertainty regarding the number of DER that will be installed over-time, their pace of installation, their individual nominal capacity and location or concentration in certain part of the network. Additionally, at micro-level, there is uncertainty and high variability of the point-in-time generation values of every single DER due to the intrinsic intermittency of their point-in-time energy source. Rarely one single DER will have any impact on the neighbouring network, but with sufficient concentration, their aggregated effect can impact the security and reliability of the power network locally and globally.

There are certain levels of similarities with financial assets managers who must diversify their portfolio of financial instruments across types of instruments, currency, industry, geography. etc.. to mitigate potential losses and optimise returns. Rarely one single position in an instrument will make the difference in any day, but the aggregated value by highly correlated sectors would make a difference. To manage their portfolio, asset managers employ powerful probabilistic risk management tools and models that would support their decisions in terms of asset allocation within their portfolio. Some of the tools react in real-time to the market values fluctuations (very short-term) and some other analyse trends and opportunities in the long-term.

Similarly, power networks operators must at any time balance the demand and supply of power and react to sudden changes in condition within the power network, but must also plan for long term changes in the load and generation profile and prepare their network accordingly. The variability of DER generation within the network and the vast number of scenarios to investigate are new factors and traditional planning tool are poorly equipped to appropriately manage this new complexity.

1.7 Leverage a data-driven approach

The traditional methods to forecast and analyse the impact of DERs within the network implies the necessity to solve the load-flow equations for the system and for each scenario. This approach is broadly used by power system operators (using professional IT simulation tools) to forecast and analyse the effect of new solar or wind farm to their system, but rarely to assess the impact of individual residential solar PV systems. Solving the load-flow equations for a power system is a relatively complex task that can be computationally intensive. In order to use a probabilistic risk-based approach, a vast amount of scenarios (10^6) must be considered and analysed. Solving the load-flow equations becomes impractical for such amount of scenarios where modern data-driven approach (Machine Learning, Artificial Intelligence) can lead to multiple order of magnitude savings in computational time while retaining acceptable accuracy.

This research focuses on developing data-driven tools and methodologies that can support power network planner to identify and analyse impacts of a vast number of future scenarios of DER penetrations within the network. In particular, we focused the scope of the first part of this research on developing tools and frameworks that can forecast the impact of DER on node voltages and on reverse power flows, both in steady-state, therefore, assuming a stable frequency. This research is intended to support long-term planning of the network in response to future scenarios of DER penetration, and is not intended to address operational issues of the current grid such as real-time frequency and voltage controls.

The second part of this research proposes a risk approach to plan for power system and community resilience to extreme weather events in developing countries.

Chapter 2

Literature Review and Research Questions

As illustrated in Chapter 1, the power system paradigm is evolving and must take into consideration the inclusion of growing number of DER at consumer point, mainly Solar PV. Until the adoption and penetration of households DER are contained to low digits values, the effects on the network are considered negligible.

In Australia, the actual penetration nationwide was already more than 20% in September 2019 and fast-growing. The new uncertainty and variability introduced by a forecasted large number of DER in the power system appear at two levels. Firstly there is uncertainty regarding the total number, the individual nominal capacity and location of the DERs that will be installed over time in the next few years. Secondly, there is an important variability of the power generated at all time due to intrinsic intermittency of the energy source.

Whilst the effect of a single Solar PV remain negligible for the surrounding network, the aggregated effect of multiple DER concentrated in parts of the network can impact the security and reliability of the power supply. Acknowledging that the typical lack of inertia in Solar PV also affects the frequency stability of the entire system, this research focuses principally on voltage nodes and reverse power flows in steady-state, assuming a stable frequency.

The traditional planning tools and methods in use by power network operators are ill-equipped to forecast the combined effects of DER penetration uncertainty and individual generation variability on their existing power network assets and configurations.

2.1 Current risk management practises for power network planning in the Australian context

In general, there are two different approaches to classical network planning: the deterministic planning standards and the probabilistic approach [10].

The underlying philosophy behind the deterministic approach is creating worse case failure or outage scenario of individual or double network elements (known as the N-1 and N-2 conditions) and designing the network with sufficient resilience so that zero interruptions of energy supply is caused by that condition. An evolution of the deterministic approach is a ‘value-based’ approach that has been developed by the California Independent System Operator (CAISO) for the transmission network planning. The method assesses the economic benefits for the transmission upgrade, including the reduction in the cost of constructing and operating per plants along with changes in market price [11].

In the probabilistic approach, the N-1 and N-2 criterion are relaxed and simulations studies are undertaken to assess the amount of energy that would not be supplied if an element of the network was out of service. A probabilistic analysis is performed evaluating the system under a vast amount of future initial conditions and then the impact of the failure of individual components (not multiple) is evaluated. The failures are based on the past observed failure rate of specific types of equipment. Reliability analysis methods for distribution network planning have been explored in [12–14]. Equipment loadings and voltages that violate acceptable ranges are resolved by re-dispatching generation and, if not adequate to relieve the violation, by disconnecting the customer load.

For planning purpose, in Victoria, the DNSP are applying probabilistic and risk-based approaches to make appropriate decisions related to network augmentation [15–18] and the main risk events considered are the failure or outage of key assets. To determine the 'economically optimal' level and configuration of network capacity (and hence the supply reliability that will be delivered to customers), it is necessary to place a value on supply reliability from the customer's perspective and AEMO has conducted a review and published a final report that quantifies in dollar terms the estimated aggregated value that customers place on reliable supply of electricity [19].

The planning frameworks, even though leveraging risk and probabilistic methods and techniques are still based on the premise that the power generation is centralised and that the main cause of disruption within the system that causes reliability and security issue is the sudden failure of an asset. To assess EUE, vast amount of power system simulations are performed which are already very computer-intensive.

In the new paradigm, as discussed in Chapter 1 there is a totally new phenomenon that can disrupt the security and reliability of power supply which is not an asset failure. There is a growing penetration of DER within the system, which aggregated impacts must be forecasted and considered in power system planning. DNSPs are responsible for connecting consumers DER in the distribution network and have predisposed automated pre-approval processes based on predefined thresholds for maximum power exported [20]. Connections requests are managed individually according to those criteria and no aggregated effects are taken into consideration for every single approval, which is mostly automatically approved if the DER nominal capacity is under the threshold.

The current probabilistic approach and simulation techniques that are considering the failure of one asset at the time are insufficient to manage the new complexity. New tools and frameworks must be created to forecast the combined effect of future DER-rich scenarios on the power network security and reliability. Those tools and framework would have the objective to identify the intensity, probability and location of potential negative effects on the power system due to future scenarios of DER penetration. Key effects and their impacts of are summarised in 1.4 and this research focuses on two main ones: the effect of DER on node voltage and the presence of reverse power flows.

2.2 Voltage stability and voltage reliability

Voltage stability aspects, and especially voltage collapse, have been studied since the early 1980s. Voltage collapse is a condition, generally caused by a disturbance such as sudden loss of an asset that implies a non-viable voltage value in some part of the system which magnitude is decreasing fast in time and cannot be recovered to acceptable values. Multiple books and IEEE Technical report have been written to describe and model the phenomenon [3, 21]. Methods such as continuation power flow, modal analysis of the reduced Jacobian matrix, sensitivity analysis, P–V curves, Q–V curves and use of voltage stability indices are widely used to analyse the system and identify or attempt to predict voltage instability events. Most of those methods are model-based and therefore depend to some extent on the model accuracy. Recently, in [22] a novel voltage collapse prediction framework based on stochastic differential equations to describe the system is proposed as an alternative. In [23], Glavic et al proposed a survey for methods in voltage instability detection based on indices or decision trees and real-time state variables measurements. Interestingly the decision tree methods initially explored by van Custem and all [24] may be viewed as a precursor to using Machine Learning techniques in assessing the voltage stability of a power system. Recently, new indices have being explored [25] based on graph and network response theory and network structural characteristics to identify potential weak nodes of power networks based on their topology and some specificities.

Most of the illustrated methods are attempting to design indicators that would measure the vicinity of a current system state to the maximum loadability of its nodes and therefore attempting to intercept conditions that could cause a voltage collapse within the system to trigger preventative actions. In the "load following" centralised generation paradigm where load profile is mostly "predictable", such indicators are useful to contribute to a safe operation of the network. In that paradigm, the main origin of disruption within the system that could cause a voltage stability risk event was an unexpected asset outage or contingency (eg. tripping on lines, generators,..). In a DER-rich paradigm, additionally, there are also fast and steep ramp-up and downs of net-demand due to the impact of intermittent renewable generation (eg. duck curve - [26]) that are increasing the risk of destabilising the power grid.

In a DER-rich power network paradigm, the intermittency of renewable generation can near-instantly influence the voltage reliability, or containment of values within statutory ranges [27, 28].

The integration of DERs and their impact on the node voltage has been subject to research in recent years, but for the limited scope of the amount of DER embedded in the grid. The papers [29, 30] provide an overview of the consequences of decentralised generation for typical voltage stability issues encountered at transmission level, and in [31] sensitivity analysis (SA) methods are studied for the identification of critical parameters affecting voltage stability in network with intermittent generation; similarly, in [32] an approach to produce voltage sensitivities in case of DER is proposed. In [33–35] new or modified voltage stability indexes considering uncertainties associated with intermittent generation source. [36, 37] introduce the study of new voltage control capabilities, including security-constrained operational power flow (SCOPF) control strategies. That research is still principally focused on the impact of intermittency on voltage stability at generation and transmission level, while the local stability of voltage level, in particular, the over-voltage, is often assumed to be managed with individual voltage control at DER.

The manifestation of the renewable generation variability has a stochastic nature and consequently so is its impact on the voltage. Resulting voltage stability and reliability issues can therefore be analysed using quantitative risk management techniques. In [38], a probabilistic voltage risk analysis is proposed to assess the value at risk associated with decision making of installing SVC in the power system and [39] proposes a probabilistic method to evaluate the over-voltage risk in a distribution network with different, but still limited PV capacity sizes under different load levels, and [40] proposes a voltage collapse risk approach. In [41–44], planning frameworks for distribution grids are broadly discussed, highlighting the need for new quantitative analysis tools.

With the rise in adoption of Solar PV installed at the edge of the network, in sunny days, voltage rises are becoming the key issue to mitigate. In [45–47], various mitigation and regulation options are discussed including Volt-Watt mode and Reactive power control.

2.3 Reverse power flows impact on power system

The reverse power flows (RPF) are flows of electricity that are in the reverse direction from the normal flow that may occur when electrical power is injected in the power system from a point which is usually a consumption point. RPF. are well know phenomenon and their potential consequences extensively described in the literature and technical reports [48, 28, 49]. In their summary of impacts, Walling et al [50] conclude that: 'Without careful engineering, DER penetration can potentially have adverse system effects, including exposing system and customer equipment to potential damage, decrease in power quality, decrease in reliability, extended time to restoration after outage, and potential risks to public and worker safety'. Clearly, DNSPs must appropriately prepare and plan mitigation actions on their network prior to accommodate the rapidly increasing adoption of DERs amongst consumers. In [51, 52], the authors illustrate optimised planning models for distribution network with DERs under the assumption that DNSPs can decide the location and capacity of the DERs. In the case of consumer driven expansion such as residential solar PV systems, DNSPs must be able to assess multiple scenarios and identify the resulting potential adverse effects on their network using traditional simulation tools or specifically designed tool and models, such as [53, 54]. Simulation tools are powerful but their utilisation of real scale power networks is computationally intense and therefore might restrict the capability to assess a vast number of scenarios. In [55–57], Artificial Intelligence techniques have been experimented to perform power flow analysis of power networks with good performances. The advantage of those techniques is the modest computational cost required to use the models to solve the load flow once they are trained. In [58], the authors propose a data-driven probabilistic power flow analysis in response to stochastic solar irradiance.

Hornik [59] established that "multilayer feed-forward neural networks are, under very general conditions on the hidden unit activation function, universal approximators provided that sufficiently many hidden units are available.", best known as the "Universal Approximation Theorem". The theorem doesn't provide strict criteria regarding the structure of the network or number of hidden layers and neurons that would perform optimally but states that structure can be found such that it can provide an arbitrarily accurate approximation of a function. In [60, 61, 56], the authors are experimenting the use of Deep Neural Networks (DNN) to solve the Power flow equations and showing the effectiveness of the method to predict the voltages and power angles mainly on the transmission network. In [62], the authors generalise the use of physic-informed DNN in power system application, both in steady-state and dynamic. Very recently, in [63] the authors propose a Deep Learning approach to the Optimal Power Flow (OPF) solution. The proposed model is evaluated on a large collection of realistic medium-sized power systems. The experimental results show that its predictions are highly accurate with average errors as low as 0.2%. We are proposing to extend the method and train a regression DNN to predict the currents at any node of a distribution network in response to given net-loads using well established supervised learning techniques as described in [64, 65].

2.4 Community and power system resilience to extreme weather event

A conceptual framework to assess and improve the resilience of power system versus extreme weather events is presented by Pantelli et al [7] that discuss three major dimensions of improvement: Stronger versus Bigger versus Smarter Power Grid. In [66] the framework is applied to the Northern Chilean electric power system in the context of exposure to seismic events and in [67, 68] fragility model of individual component and their failure probabilities are introduced and correlated with resilience metrics such as Expected Energy Not Serve (EENS) and Loss of Load Frequency (LOLF). The model is sophisticated, but requires the individual fragility assessment of each asset and the resulting failure probabilities curve in response to extreme weather events intensity (eg. wind speed,..). This is applicable for a restricted number of key assets such as main transmission lines but unrealistic for assessing the vast number of distribution network assets fragility and resilience. In [69] Mukhi et al investigated scenario-based methodologies to build climate resilience into power system planning in Bangladesh and compared two approaches to manage the long term climate change uncertainty into planning decision, the Stochastic Linear Programming (SLP) and Robust Decision Making (RDM). The uncertainty managed by those models are the long term uncertainty connected to unknown changes in climate conditions impacting the severity and frequency of multiple extreme weather events that must be considered in long term energy mix and infrastructure investments. In [70], Zhou et al investigate the role that Microgrids can play in providing reserve service to the main electricity grid in response to resilience oriented contingencies, demonstrating a clear synergetic effect.

From a social and community resilience point of view, which acquire specific importance in developing countries where the level of electricity supply is poorly guaranteed even is normal operation, [71] examine the concept of resilience for understanding the gendered experience of women to extreme climate events through a qualitative case study in Bangladesh. This highlight the need for specific regards to community and social aspect when assessing and planning for resilient power supply.

2.5 Key identified gaps

Key gaps encountered in the literature and current practise are listed below:

1. **A scalable method to forecast the voltage excursion and reverse power flow risks at any node associated with any future extreme DER penetration scenarios within power system for system planning purpose**

   The current practises and standards in terms of power system planning and contingency management are principally focused on mitigating the effect unexpected failure of assets to maintain security and reliability of the service, and the probabilistic approach consider the failure of one or two individual assets at the time.

   The emergence of DER within the power system is bringing new challenges that were traditionally not present, such as node voltage surge and reverse power flows. The effect of reverse power flows on the power system are well known and described in the literature. So far the penetration of household DER has been limited and consequently, their impact neglected. The literature mostly focuses on the operational stability issues encountered due to intermittent large or medium scale generation such as Solar Farms or Wind Farms. The effects of household solar PV embedded in the distribution network is neglected or assumed to be controlled or controllable at the individual inverter level.

The traditional deterministic methods to predict and analyse the effect of perturbation in the system are founded upon full simulation of the system that requires solving of the load flow equations. In order to assess the impact of future DER penetration such as individual rooftop Solar PV, a vast amount of scenarios must be considered (How many units, where are they located, what is their individual capacity) and a probabilistic approach must be adopted to their generation profile. The number of scenarios to consider might be in the orders of 10^6, if not more. Applying the brute force of solving the load flow equations for each scenario is impractical.

No identified literature addresses alternative methods to estimate the aggregated probabilistic effect of future extreme DER penetrations (theoretically up to 100%) scenarios at every node of the system under any given stochastic generation and load values.

2. **A method to assess community and power system resilience that can be linked to long-term power system planning**

   The literature proposes rich and complex frameworks and models to assess the resilience of power system assets to extreme weather events. Linkage with planning is offered where the resilience metrics used to weight re-enforcement cost are the Loss of Load and Energy Not Served, which are purely economic. This applies to developed country where customers are accustomed to very reliable electricity, where the power system is privatised, where the extreme weather events are rare and all the communities equally dependent to electricity and paying for it.

In developing countries, ordinarily, the reliability of electricity supply might only be guaranteed for some hours during the day for some communities and the power systems are not fully privatised or heavily subsidised. Additionally, some countries like Bangladesh are particularly vulnerable to multiple sorts of extreme weather events (eg. cyclones, flooding, sea-level rise, etc.) that happen quite frequently. In those countries, the intrinsic community resilience must also be taken into consideration when planning for resilient power supply. No identified literature addresses an aggregated method to assess the resilience of communities and power system in combined metrics that can be used for planning purpose.

2.6 Key research questions

In an attempt to bridge the identified gaps, this research is addressing the following main questions:

1. "How can we built risk frameworks that enable power network operators to assess any future DER penetration scenarios and support decision making for their long-time network planning?"

2. "How can we propose a linkage between energy community resilience to extreme weather events and power system planning in developing countries?"

The first question is addressed in Part. I and will develop two frameworks to assess the impact on power system of :

- Voltage excursions at any node of the system in any scenario of future DER penetration (up to 100%)

- Reverse power flows at any node of the system in any scenario of future DER penetration (up to 100%)

The second question is addressed in Part.II and will propose a methodology to establish a linkage between energy resilience in case of extreme weather events and long term planning of power networks in Bangladesh.

2.7 Key research contributions

As a contribution to the first question and novelty, this research has produced two data-driven risk frameworks that enable the assessment of a large amount of DER penetration scenarios at a fraction of the computational cost that is required for traditional load flow equation solving. In particular, the first framework assesses the probability distribution of three different voltage risk events (voltage swell, voltage sag and voltage collapse) in response to any scenario of DER (such as residential rooftop solar PV) penetration in the system. The scenarios are defined by the number, the location, the individual capacity of each DER and by a large stochastic set of individual net-loads profiles. Also, a financial model has been developed to assess the cost of those risks to support future network planning decisions.

The second framework assesses the probability distribution of reverse power flow intensity in the system in response to any scenario of DER. Similarly, The scenarios are defined by the number, the location, the individual capacity of each DER and by a large stochastic set of individual net-loads profiles. The knowledge of the probability distribution of RFP intensity enables the DNSPs to investigate the adequacy of their network assets ratings and the adequacy of their protection configuration and rating.

For both frameworks, our simulations have shown that they can lead to multiple orders of magnitude savings in computational time while retaining an acceptable accuracy compared to classical brute force simulations.

As a contribution to the second question, this research has produced a risk-based methodology to extreme weather event resilience. The novel methodology integrates traditional measures of power system technical resilience with quantifiable sociological criteria that combine into a community energy resilience measure. The methodology is also linked to power system planning through an optimisation formulation.

2.8 Scope and Limitations

In its first part, the research produced two computationally-efficient risk frameworks using machine learning technique that enable the assessment of a large amount of DER penetration scenarios. The approach that is used in this research relies on training machine learning models using simulated data. Under this approach, there is an underlying assumption that the power system parameters and configuration (eg. line impedances, transformer configuration, reactive loads, etc) are well known and static. We know that it is not fully realistic in the case of real and large scale distribution power networks for which an improved sampling method would need to be investigated to create the training dataset. A possibility would be to mix the simulated data with observed past data recorded from the system's operation, as recently investigated in [72]. This research will focus on developing end-to-end frameworks tested on simulated data from IEEE test systems. The sensitivity analysis of the models to real and large scale network configuration will be left for future work.

Other variabilities to consider are the intermittency of the DER generation conditions and the penetration levels of DER within the network (how many, where in the network and what capacity for each DER ?). Future penetration levels, being hypothetical yet, can only be simulated to generate training data points for the models. In order to create the data points that cover most of the possible scenarios of intermittency and DER penetration, a probabilistic approach is adopted in this research. For each simulated training data point, in a Monte Carlo simulation, the input loads and DER generation values are individually and randomly extracted from a range between zero and the max nominal capacity of the DER. With an elevated number of simulations, this approach will generate a training data set that is vastly representative of all the possible combination to be encountered in real situations. Sampling reduction and optimisation, for instance taking into consideration the spatial correlation of DER generation, are left for future work.

In its second part, the research produced a risk-based quantitative methodology that enables to assess the combined community and power system resilience to extreme weather events. At this stage, the resilience assessment is still at ‘methodology’ level. The next phase is intended to be a ‘pilot’ during which data will be collected on the terrain and models will be created based on the methodology. The calibration of the model based on collected data could well use ML techniques if appropriate depending on the availability of data.

Part I

Risk frameworks for DER-rich power network planning using AI techniques

This part is addressing the first research question: "How can we built risk frameworks that enables power network operators to assess any DER penetration scenario and support decision making for their long-term network planning?"

Traditionally long-term planning uses techno-economic models to assess the future infrastructure and configuration of the power network that would support scenarios of a forecasted increase in peak demand, at the lowest cost and maintaining security and reliability of supply even under contingencies. The main uncertainties and risk factors were traditionally the future peak demand value to serve and the worst-case scenario contingencies to mitigate.

In a hypothesis where the future penetration of household DER such as Solar PV is extremely high (above 50%, hypothetically close to 100%), their combined effect on the power network cannot be ignored and should be considered while doing long-term planning.

The variability introduced is twofold at two different levels:

- the rate of adoption by household is unknown and not directly controlled by the network operator who needs to consider a vast amount of scenarios regarding the number, the location and the individual capacity of each DER.

- the individual generation of each DER is intermittent and varies daily as well as seasonally

The elevated level of uncertainty necessitates the use of a stochastic method to assess the impact of DER on the grid, and therefore a vast amount of simulations that justify the use of data-driven approaches.

Chapter 3

Approach and methodologies

3.1 Data-Driven Approach

Modern Power systems are hugely complex engineering systems with thousands of interconnected devices (centralised power generators, power transmission lines, switches, protection devices, voltage transformation stations and substations, distribution lines, distributions transformer, consumer loads, decentralised power generators,...) that needs to be balanced at any point of time. The instant power that is demanded at any time by consumers needs to be exactly generated and transported by the infrastructure. The system is in balance when every consumer is connected, the exact amount of power consumed is generated, the frequency and amplitude of the AC voltage are within statutory limits at any point of the system. Unexpected events can happen in the system that would tend to throw it off balance, those events can be an accidental fault in assets, a fast and broad change in load/generation or any combination of those. The configuration of the system will play an important role in the way it will react to the unexpected change and will remain in balance or will drift off balance.

Broadly, power system infrastructure planning is a process that assesses the current and expected future power requirements of communities and identify options for a power grid configuration and investments that would support the long-term interests of consumers for safe, secure, reliable electricity, at the least cost, across a range of plausible future scenarios. The outcome of this process is a cost-based complex engineering optimisation plan that forecasts the overall system requirements over an agreed future period. This requires extensive modelling of the system in a vast number of scenarios.

Traditionally, this modelling is operated at bulk generation and transmission network level and not at the distribution network level and mainly assesses the adequacy of the overall generation and transmission capacity to address the aggregated demand at the substation level, ignoring the details of what is happening within the distribution network. In this old paradigm, the distribution network is considered as a passive element (set of poles and wires) that unidirectionally transport the energy from main substations to the consumer and configured in a "fit-and-forget" way.

In the new paradigm, the distribution network is not a passive element any more as DER's are generating power within, and the most common DER that is nowadays inserted at consumer point are renewable energy sources such as Solar PV systems, with or without storage unit. By nature the renewable energy sources are intermittent, and the power generated is not strictly predictable and therefore stochastic. Additionally, the future number of individual Solar PV generation within the system itself is an important factor to consider because not directly planned by the power system operators but the decision to install the system is made by the consumer. The expected penetration value and its rate of increase are both unknown to the system planner and are stochastic to the extent that they cannot be predicted with certainty.

The penetration of those DER can have dynamic impacts that were traditionally not present: sudden voltage rises or drops due to sudden excess/drop of decentralised intermittent generation as well as reverse power flows (RPF) or excess currents generated at consumption point and that are circulating upstream within the power network. Those phenomena are stochastic by nature and the number of scenarios to consider to predict and assess their impact is significant.

It appears therefore that power system planners are facing new challenges, they must be able to assess the aggregated impact of intermittent generation deep within the distribution network in multiple scenarios of DER penetrations (how many DER and where).

The stochasticity of both the point-on-time intermittent generation values for each DER and the future penetration of those DERs within the network dictate the use of statistical methods to assess the impacts on the grid. Vast Monte Carlo simulations are required to derive the probabilities of those impact varying the number of DERs, their location and the power generation values time series.

The complexity of modelling the distribution network behaviour is dictated by the elevated number of the elements to consider in the solving process (eg. substations, lines, transformers, etc to serve thousands/millions of point of connection). Assuming that the system parameters are known, solving the entire system (active/reactive power flows, node voltages, node currents, line losses, voltage angles,..) is computationally intense. The traditional simulation techniques that entice solving the load flow equations for the entire system. As there exists no known analytical solution to this problem that must be solved using well established numerical iterative methods (e.g. Gauss-Seidel, Newton-Raphson, Fast-decoupled Method, etc.) [73]. Typically, the rate of convergence of those methods is quadratic, but in [74] it is shown that the convergence region and number of iteration can significantly vary depending on the initialisation and the actual loading condition of the system. For a system of n-buses, each iteration requires to calculate and inverse the $2n \times 2n$ Jacobian matrix of the system which time complexity is $\sim O((n)^3)$.

Therefore, for a large-scale system, and also depending on the loading condition of the system, solving the system has high computational complexity. It is consequently not practical to analyse the effect of vast numbers of stochastic net-load scenarios (10^5 - 10^6) via Monte Carlo simulations using the brute force of solving the load flow equations for each combination of net-loads input.

For that purpose, we are proposing to use data-driven methods that are based on Machine Learning and Artificial Intelligence techniques that will offer a prediction of the voltage and currents within the system in response to any net-loads with an appropriate precision at a fraction of the computational cost required to solve the load-flow equations.

The data-driven prediction methods will be used in Monte Carlo simulations to infer probability distributions of DER rich impacts on the grid. The impacts under examination in the scope of this research or the voltage fluctuations and the reverse power flows. Risk Management techniques will be adopted to construct two major risks frameworks that are derived from an elevated penetration of DER within power networks:

1. The voltage risk framework

2. The reverse power flow risk framework

Both frameworks can support power system planners to assess the impact of DER penetration scenarios in their network and consequently adopt appropriate measures to maintain security and quality of supply in response to those scenarios. Specific Machine Learning and Artificial Intelligence methods have been employed to construct those frameworks and the following sections are describing them.

3.2 Method: Risk management model

3.2.1 Risk framework

According to ISO 31000 [75], risk is the “effect of uncertainty on objectives” and an effect is a positive or negative deviation from what is expected

Uncertainty is a potential, unpredictable, and uncontrollable outcome; risk is an aspect of action taken in spite of uncertainty. *Risk management* is the identification, assessment, evaluation, and prioritization of risks (defined in ISO 31000 [75] as the effect of uncertainty on objectives) followed by coordinated and economical application of resources to minimize, monitor, and control the probability or impact of unfortunate events or to maximize the realization of opportunities.

The generic method to address risk is the following:

1. Identify the threats/hazards (**Risk Identification**)
2. Assess the vulnerability and impact on critical assets from each threat/hazard (**Risk Assessment**)
3. Identify and plan ways to reduce those risks (**Risk Evaluation**)
4. Prioritize and monitor risk reduction measures (**Risk Treatment**)

Risk identification requires expert domain knowledge to answer the basic question: "What could go wrong ? ". For engineering systems, the set of contingencies to consider is usually the failure of assets (physical or algorithmic), external attacks or conditions, instability of internal state physical quantities, etc.

Once identified and listed all the possible threats to the system, a precise analysis of each is undertaken to assess the likelihood and impact of such threats. The *Quantitative Risk Assessment* [76] introduces probabilistic frameworks for assessing the risks (uncertainty of threat) and their impact (damage induced by threat). When the damage can be quantified in monetary terms, the *Risk Value* for a thread is simply:

$$\begin{gathered} Risk = probability \cdot impact \\ \text{or} \\ Risk = \int_S probability(s) \cdot impact(s)\, ds \end{gathered} \tag{3.1}$$

If there exists a risk probability distribution of the thread across a continuum of scenarios (S).

A typical example of risk probability is the failure rate of an asset. Examples of risk distribution that are relevant to this research are:

- the intensity of reverse power flows at one node of the system caused by aggregated Solar PV power generation in response to solar irradiation
- the destruction probability function of assets in response to environmental conditions such as wind speed

Risk assessment act as a foundation of the next stage, i.e. the risk treatment stage. Risk evaluation includes weighting specific risk values against the costs of mitigation measures to:

- **Treat** the risk implementing some controls which could be technical, administrative, physical or environmental
- **Transfer** the risk, e.g. buy an insurance e cover
- **Accept** the risk as it may be difficult or uneconomical to mitigate the risk
- **Avoid** the risk by stop conducting the activity

Continuous monitoring must be undertaken at all time to monitor and measure the risk occurrence against the predictions or intercept new conditions that might require adjustment to the overall risk assessment and evaluation results.

3.2.2 Risk probabilities via Monte Carlo simulation

Monte Carlo simulation [77] is a technique that performs risk analysis by building models of possible results by substituting a range of values—a probability distribution—for any factor that has inherent uncertainty. During a Monte Carlo simulation, values are sampled at random from the input probability distributions. Each set of samples is called an iteration, and the resulting outcome from that sample is recorded. Monte Carlo simulation does this hundreds or thousands of times, and the result is a probability distribution of possible outcomes. In this way, Monte Carlo simulation provides a much more comprehensive view of what may happen. It tells you not only what could happen, but how likely it is to happen.

3.3 Method: Machine Learning classifiers

3.3.1 The classification problem

In machine learning, the classification is the problem of identifying to which set of categories (class labels) a new observation belongs, based on a training set of data containing observations whose category membership is known. The classification is said binary when the possible categories are two.

Definition The problem of binary classification is defined as:

- **Input:** a set of m examples $(x^j, y^j), j = 1, 2..m$ (the *training set*) sampled from some distribution D, where $x^j \in R^n$ and $y^j \in \{-1, 1\}$. The i-th component of x^j, x_i^j, is named *feature i* and the i-th component of y^j, y_i^j, is named *class label i*.

- **Output:** a function $f : R^n \to \{-1, 1\}$ which classifies additional samples $\{x^k\}$ in their respective labels from the same distribution D.

- **Supervised Learning:** the task consisting in inferring the function f from a given *training set* of data.

In the context of this research, we will identify multiple classification problems where the *features* will be the vector of net-loads at every node of the system and the *class labels* will be 'voltage within statutory range' and 'voltage out-of statutory range'.

3.3.2 Support Vector Machine

The goal of binary classification is to take an input *feature* vector x and assign it to one of the two *class labels*. If the classes are taken to be disjoint, each input is assigned to one and only one class. The input space is therefore divided into two *decision regions* whose boundary is called *decision boundary* or *decision surface*.

Support Vector Machine (SVM) is a Machine Learning technique adapted to solve classification problems where the model parameters are defined to maximise the *margin*, which is the smallest distance between the decision boundary and any of the training samples.

The classification function f is described by the following linear model:

$$f(x) = w^T \times \phi(x) + b \tag{3.2}$$

where $\phi(x)$ denotes a fixed feature-space transformation and b is the bias.

The determination of the model parameters corresponds to a quadratic convex optimization problem as described by Bishop in [2].

Fig.3.1 illustrates a 2-class linear SVM (where $\phi(x) = x$). The support vectors for the blue class are the vector x for which $w^T \times x + b = 1$ and the support vectors for the green class are the vector x for which $w^T \times x + b = -1$. The decision boundary is given by: $w^T \times x + b = 0$.

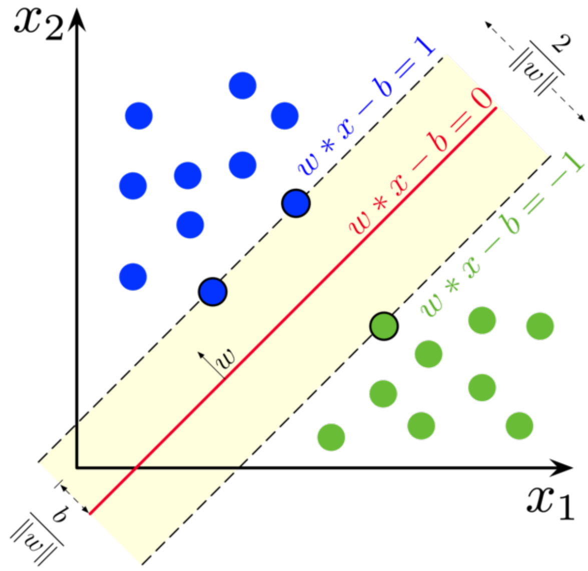


Fig. 3.1 Example of linear SVM for binary classification (blue and green) [2]

As discussed in Bishop in [2], the way to create non-linear classifiers involves applying the *kernel* trick, a concept introduced in pattern recognition by Aizerman et al [78]. For models that are on a fixed nonlinear feature space mapping $\phi(x)$, the kernel function is given by: $k(x,x') = \phi(x)^T \phi(x')$. The kernel function is used directly in the dual representation of the maximum margin optimisation problem that is solved to identify the decision boundary that maximises the *margin*.

The most common kernel functions are:

- Polynomial kernel: $k(x,x') = (x \cdot x' + 1)^d$ where d is the degree of the polynomial
- Gaussian Radial Basis (RBF): $k(x,x') = e^{(\frac{\|x-x'\|^2}{\gamma})}$ for $\gamma > 0$
- Hyperbolic tangent: $k(x,x') = \tanh(\kappa x \cdot x' + c)$ for some $\kappa > 0$ and $c < 0$

Fig.3.2 illustrates an example of a non-linear binary classifier that uses RBF kernel function. Even though the data are non linearly separable, the use of the kernel function enables the identification of non-linear boundary that separates the data.

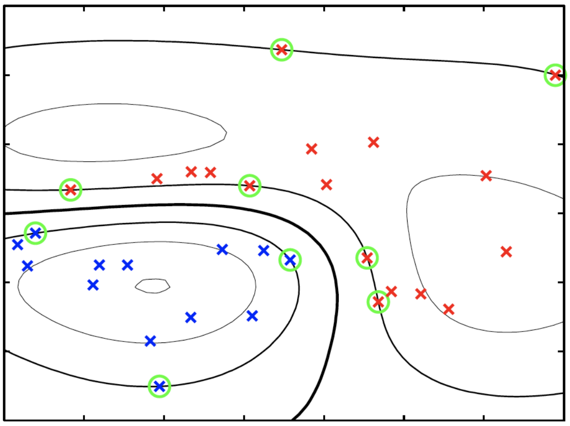


Fig. 3.2 Example of synthetic data from two classes (red and blues) in two dimensions showing contours of constant y(x) obtained from a support vector machine having a Gaussian kernel function. Also shown are the decision boundary (in bold), the margin boundaries, and the support vectors (circled) [2]

SVM is a classification methodology that is particularly well suited when there exists a decision boundary between the classes. The effectiveness of non-linear SVM depends on the selection of the kernel, the kernel parameters $\gamma$and the bias b. A common choice is the RBF kernel and the best combination of γ and b is often selected by grid search with exponentially growing sequences of both parameters. The selection of the best set of parameters is checked with cross-validation of the model where the parameters leading to the best cross-validation accuracy are picked. Cross-validation is a model validation technique that assesses the performance of the model on independent data set.

3.4 Method: Neural Network regression model

As described by Russel and Norwig in [64], the neural network is a series of functional transformation aimed at approximating a non-linear multivariate function $f : R^D \to R^K$. It is inspired by biological neural networks and based on a collection of *units* (called artificial neurons) that are connected by directed *links* which loosely model the functioning of animal brains. Each connection (like the synapses in the brain) can transmit a signal to the connected neurons and each artificial neuron that receive a signal processes it via an *activation function* and transmit the processed value to the next connected neurons. A *feed-forward network* has connections only in one direction, that is forms directed acyclic graph. [64]

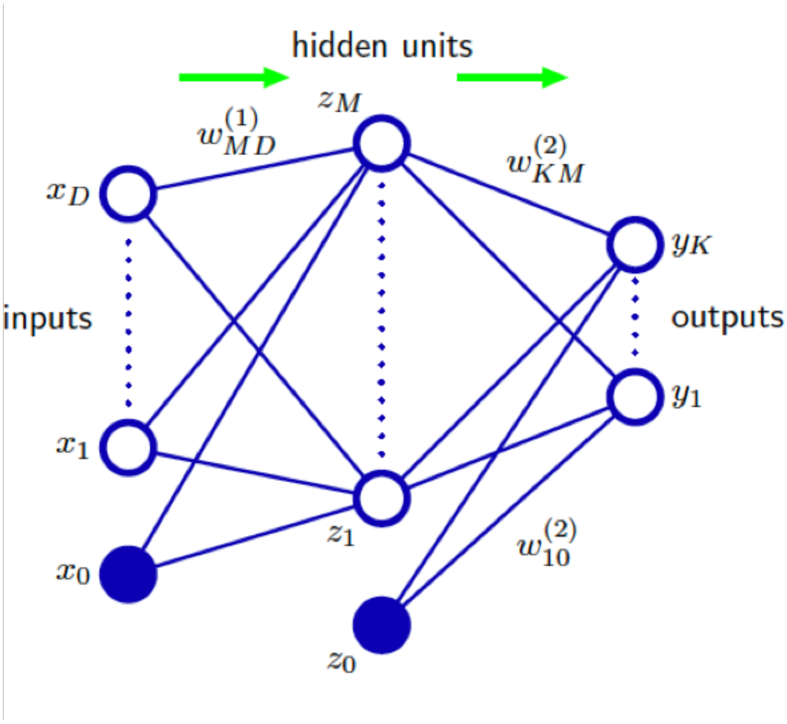
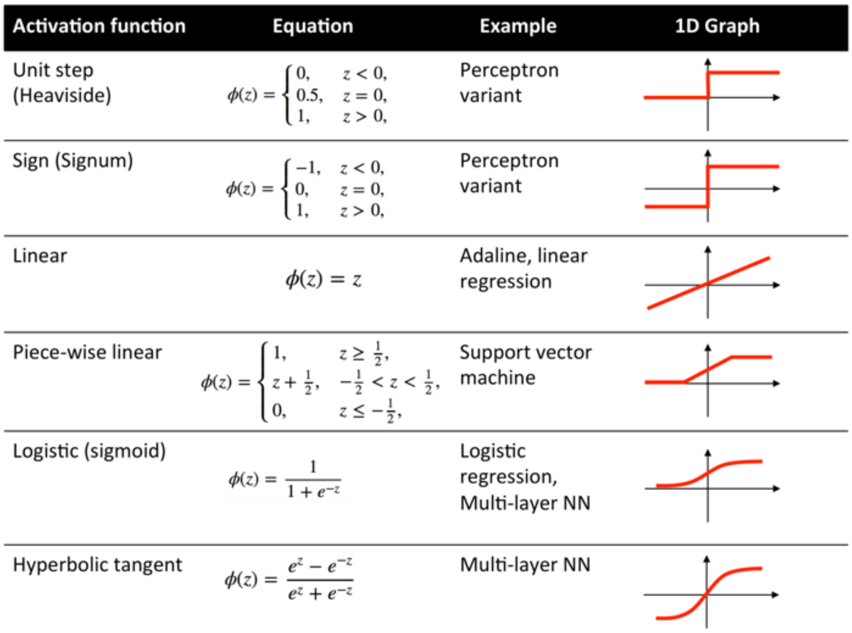


Fig. 3.3 Example of an ANN with one hidden layer [2]

Fig.3.3 illustrates the structure of a feed-forward ANN with D inputs, K outputs and one hidden layer with M neurons and (3.4) describes the relationship between the outputs and inputs.

$$y_k(x,w) = \sigma\left(\sum_{j=1}^{M} w^{(2)}_{kj} h\left(\sum_{i=0}^{D} w^{(1)}_{ji} + w^{(1)}_{jo}\right) + w^{(2)}_{k0}\right) \tag{3.3}$$

where w are the weights and σ, h are the respective neurons *activation* functions. The most common activation function in use when building an ANN model as described in Fig.3.4.



Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$	Multi-layer NN	

Fig. 3.4 Commonly used activation functions (Source: RIP Tutorials)

The function represented by a network can be highly nonlinear, as composed of nested nonlinear soft threshold activation functions, the ANN is ,therefore, a strong tool for doing *nonlinear regression*.

ANN is called Deep Neural Network (DNN) when the number of hidden layers is multiple.

Hornik in [59] proofs the *Universal Approximation Theorem* stating that multilayer feed-forward neural networks are, under very general conditions on the hidden unit activation function, universal approximators provided that sufficiently many hidden units are available.

3.5 Mapping of methodologies to research frameworks

This address the research questions, this research is developing 3 main outputs (illustrated in details in the next chapters):

- A Voltage Risk Framework
- A Reverse Power Flow Risk Framework
- A Community Energy Resilience Methodology

The approach and methods illustrated in this chapter will be leveraged to develop the 3 main outputs. Fig.3.5 maps the methods used in each output and indicate the main rationale that justifies the use of each method.

Research Output	*Method Used*	*Rationale*
Voltage Risk Framework		
- Scenario Analysis	Monte-Carlo simulations	Consolidated technique to yield probability distributions
- Voltage Risk identification	Support Vector Machine classifier	Technique particularly adapted and well-accepted to identify boundaries and the voltage risks are defined by hyper-surface boundaries
- SVM training data sets creation	Load Flow solving	An alternative would be using past data recorded from the network operation in addition to simulated data
- Voltage Risk financial impact	Risk Management model	Combines the voltage risk with their financial impact
RPF Risk Framework		
- Scenario Analysis	Monte-Carlo simulations	Consolidated technique to yield probability distributions
- RPF Risk identification	DNN Regression models	The load flow equation are non-linear and DNN are adapted and well-accepted to approximate any non-linear function (universal approximation theorem)
- DNN training data sets creation	Load Flow solving	An alternative would be using past data recorded from the network operation in addition to simulated data
Community Energy Resilience Methodology		
- Resilience Assessment	Risk Management model	A risk-based approach to resilience is considered (highest the risk, lowest the resilience and vice-versa)

Fig. 3.5 Mapping of the Methods to research outputs (Source: Author)

For both Risk Frameworks (Voltage Risk and RPF Risk), a data-driven approach is adopted to predict and identify the risk respective risk events. The data-driven approach (versus full simulation and load-flow equation solving for the entire power system) is justified by the elevated number of scenarios to consider in order to obtain the risk probability distributions (estimated in the orders of 10^5-10^6). In fact, the variability is twofold: firstly the number, location and capacity of DER that will be connected to the grid in future are unknown and secondly, the actual net-load profile of each DER is intermittent and widely variable in function of the weather conditions. Therefore, in order to assess the impact of multiple scenarios of future DER penetration in the grid, a large number of Monte-Carlo simulations must be accomplished. Solving the load-flow equations for each simulations would be impractical, while to use of Machine Learning models decreases the computational burden by orders of magnitude while retaining an acceptable accuracy.

SVM are respectively well-accepted machine learning methods for classification [2] and regression [64]. Therefore, they have been the respective first choice for the voltage risk (within or without statutory range classification) and RPF risk (current intensity). In both cases, the performances demonstrated were high enough to make further analysis a redundant endeavour.

This part is addressing the first research question: "How can we built risk frameworks that enables power network operators to assess any DER penetration scenario and support decision making for their long-term network planning?"

Traditionally long-term planning uses techno-economic models to assess the future infrastructure and configuration of the power network that would support scenarios of a forecasted increase in peak demand, at the lowest cost and maintaining security and reliability of supply even under contingencies. The main uncertainties and risk factors were traditionally the future peak demand value to serve and the worst-case scenario contingencies to mitigate. In a hypothesis where the future penetration of household DER such as Solar PV is extremely high (above 50The variability introduced is twofold at two different levels:

• the rate of adoption by household is unknown and not directly controlled by the network operator who needs to consider a vast amount of scenarios regarding the number, the location and the individual capacity of each DER. • the individual generation of each DER is intermittent and varies daily as well as seasonally The elevated level of uncertainty necessitates the use of a stochastic method to assess the impact of DER on the grid, and therefore a vast amount of simulations that justify the use of data-driven approaches.

on the initialisation and the actual loading condition of the system. For a system of n-buses, each iteration requires to calculate and inverse the $2n \times 2n$ Jacobian matrix of the system which time complexity is O((n)3).

Chapter 4

The voltage risks framework

4.1 Introduction

Established power networks were designed and built to transport power energy from high-capacity power generators to the consumers where the generation value at any time is controlled to exactly match the aggregated consumer load and the power is flowing unidirectionally from upstream generators to downstream consumers across transmission, substations and distribution networks with decreasing voltage levels.

Increasing the share of intermittent Distributed Energy Resources (DERs) to significant levels would create some effects on the power networks that need to be appropriately planned for and managed under unknown scenarios. The intermittency of some renewable sources (solar PV and wind) characterised by large fluctuation and uncontrollability of the power generated can cause rapid unbalance between power generated and consumed locally. Consequently, this unbalance would cause voltage excursions and reverse power flows within the power network, especially the part of the network closely connected to the intermittent generation [79].

Increasing the share of Distributed Energy Resources (DERs) to significant levels would create some effects on the power networks that need to be appropriately planned for and managed under unknown scenarios. The intermittency of some renewable sources (solar PV and wind) characterised by large fluctuation and uncontrollability of the power generated can cause rapid unbalance between power generated and consumed locally. Consequently, this unbalance could cause voltage excursions and reverse power flows within the power network, especially the part of the network closely connected to the intermittent generation.

In those circumstances, in addition to traditional voltage collapse [3, 80, 73, 21], the containment of voltage within statutory limits would become a relevant issue [81, 82, 50, 30, 83]. Because unbalances between power generated and consumed locally can be either positive or negative, both threshold limit value is susceptible to be breached (under and over-voltage excursion). Distribution Network Service Providers (DNSPs) must plan to adapt their network assets and configuration to control and manage the voltage where it is susceptible to exceed the threshold limits.

The fluctuation of generation and local imbalance of net-demand are both of stochastic nature, and therefore, so are the resulting voltage excursion risks. It is consequently appropriate to analyse and assess those phenomenons using quantitative risk management techniques (3.2) that combine the likelihood and the severity of event to quantify the risk [84].

The likelihood of a voltage excursion event at any node of the distribution network will depend on the particular network configuration, the DER penetration (number, nominal capacity and location) within the network and the combined stochastic values of net-demand values over time. The likelihood at any node can, therefore, be modelled through a marginal probability that will depend on the net-demand joint distributions at every node given a network configuration and DER penetration scenario. The severity of voltage excursion event can be modelled through financial impact for the customers. To that respect, AEMO has conducted a review and published a final report that quantifies in dollar terms the estimated aggregated value that customers place on the reliable supply of electricity [19] for the Australian market. In a scenario where the penetration of intermittent generation within the low voltage network (eg rooftop PV) increases significantly, the relevance of voltage excursion risk will strengthen. As an example, in Australia, rooftop PV is expecting to increase from 5GW in 2016 to 20 GW of installed capacity by 2036-2037 according to AEMO (the "Australian Market Energy Operator") Electricity Forecasting – June 2017. In this context, network operators must plan for adaptation on their network.

4.2 Voltage risk definitions and models

4.2.1 Voltage risks definition

The voltage risk framework that we are proposing is intended to determine marginal voltage risk event probability at any selected bus of a power network in response to any scenario of intermittent DER penetration and net-loads profile in the system where :

- The *penetration scenario* is defined by the number of DERs, their individual nominal capacity and location within the network
- The *net-load* is defined as the power generated minus the power load at any node of the system

The voltage risk events and criteria that are considered for this research are:

1. *"Over-voltage"* - or breach of the upper statutory voltage limit at one node, (eg. 1.1 $p.u.$ in Australia),
2. *"Under-voltage"* - or breach of the lower statutory voltage limit at one node at one bus (eg. 0.94 $p.u.$ in Australia).

3. *"Critical voltage"* - or potential voltage collapse due to maximum loadability reached at one node [3, 21, 85]

The two first risks are referred to net-load conditions within the system that would cause the voltage to breach at one node respectively the upper or lower voltage statutory limit that is set according to local regulation. The last one refers to extreme load condition within the system that would cause the voltage to reach its bifurcation point at one node and initiate a collapse event within the power system.

To accomplish this, we firstly need a model that assesses the occurrence of voltage risk events in response to any particular net-loads values within the system. Secondly, we need the ability to run the model on a vast amount of net-loads values and DER penetration scenarios (through Monte-Carlo simulations) to determine marginal voltage risk probabilities.

The proposed method is decoupling the assessment of voltage risk events and the actual stochastic distribution of net-loads that is dictated by both the DER penetration scenario and intermittent generation profile (eg. Solar irradiation, wind speed,...). This offers a framework that can be applied to any intermittent DER penetration level and net-load profile, knowing that distributions of the net-loads vary according to the geography, the time of the year/day and are highly dependent on external factors such as weather forecast.

4.2.2 Voltage risk events models

Critical Voltage

The proposed voltage risk assessment model describes the steady-state power system by a set of $2B$ load-flow equations in $2B$ algebraic variables V_i, θ_i :

$$0 = -P_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik} \cos \vartheta_{ik} + b_{ik} \sin \vartheta_{ik}), \forall i \in B \tag{4.1}$$

$$0 = -Q_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik} \sin \vartheta_{ik} + b_{ik} \cos \vartheta_{ik}), \forall i \in B \quad (4.2)$$

where

- P_i is the net injected real power (power generated minus power consumed) at bus i
- Q_i is the net injected reactive power (power generated minus power consumed) at bus i
- v_i is the voltage at bus i
- ϑ_{ik} is the difference in voltage angle between bus i and k
- y_{ik} is the admittance of the line between bus i and k
- g_{ik} is the conductance , or the real part of the admittance of the line between bus i and k
- b_{ik} is the susceptance, or the imaginary part of the admittance of the line between bus i and k

For the sake of abbreviation, we denote the set of $2B$ load flow equations (4.1),(4.2) for the entire system as:

$$\phi(V, \theta, L_n) = 0 \quad (4.3)$$

where

- V is the B-dimensional vectors of voltages v_i at each node of the system
- θ is the B-dimensional vectors of voltage angle differences ϑ_{ik} at each node of the system
- L_n is a $2B$-dimensional vector of net-load values P (active power) and Q (reactive power) at each node of the system

$$L_n := [P_1, .., P_B, Q_1, ..., Q_B] \quad (4.4)$$

In that model, the loadability limit or critical points are the points where the demand values reaches an extremum value after which there is no solution of the load flow equations. As demonstrated in [3], in the B-dimensional parameter space, the points satisfying those conditions belong to a manifold of dimension $B-1$ which is called the bifurcation surface or *critical surface*, or:

$$L^*_{n,crit} = \{L_n \mid g(L_n) = 0\} \tag{4.5}$$

where $g(L_n)$ is the critical surface.

For illustration, Fig.4.2.2 shows a representation in the P-Q-V space of the load flow equations solutions for a simplified single load system [3]. Each point of the surface "equator" of this surface corresponds to the maximum power points or critical points for that system.

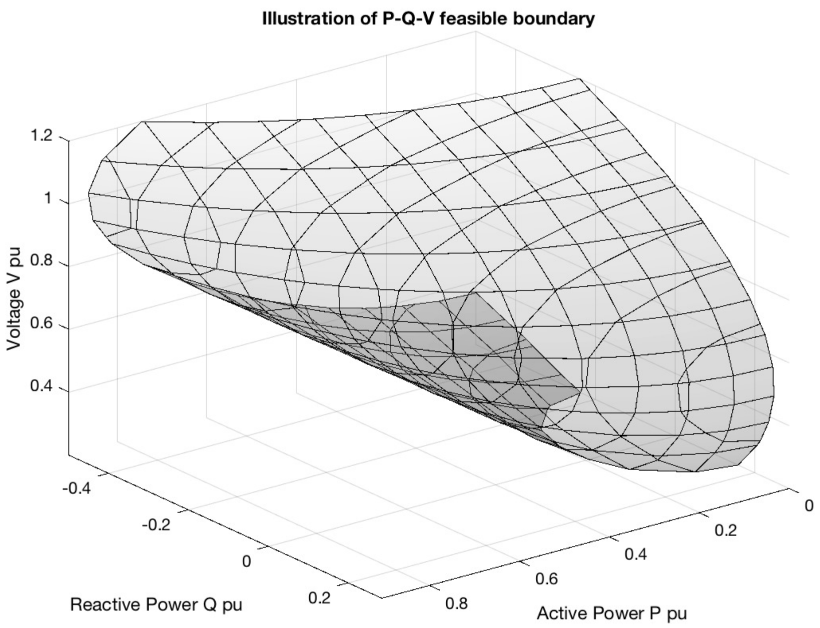


Fig. 4.1 Illustration of load flow equations solutions for a simplified single load system [3]

As demonstrated in [24], a necessary condition for loadability limits derived from the Kuhn-Tucker conditions for the optimisation problem is: $det J(\phi)_V = 0$ where $J(\phi)_V$ is the

Jacobian of ϕ with respect to V. In addition, in the B - dimensional parameter space, the points satisfying that conditions belong to a manifold of dimension $B-1$ which is called the bifurcation surface or *Critical Surface*.

Let's define the *Stable Region* (hyper-volume) inside the critical surface in the $2B$ - dimensional parameter space:

$$S_{stable} = \{(V,\theta,L_n)|L_n < L^*_{n,crit}, \phi(V,\theta,L_n) = 0\} \tag{4.6}$$

The model of the *Critical voltage* set of risk events is defined by the set of net-loads values L_n^{crit} that not in the stable region, or:

$$\{L_n^{crit} \mid L_n \notin S_{stable}\} \tag{4.7}$$

Under normal operating conditions, the net-loads in the power system are well within the stable region and the multi-dimensional critical surface can be reached from any interior stable point by 'pushing' the net-load value of one bus at the time until it reaches the critical surface as illustrated in Fig.4.2.2 and the approach can be adopted independently of the number of buses B in the network. The new net-load conditions will be on the critical surface, meaning that any further 'push' in the net-load would lead to a voltage collapse.

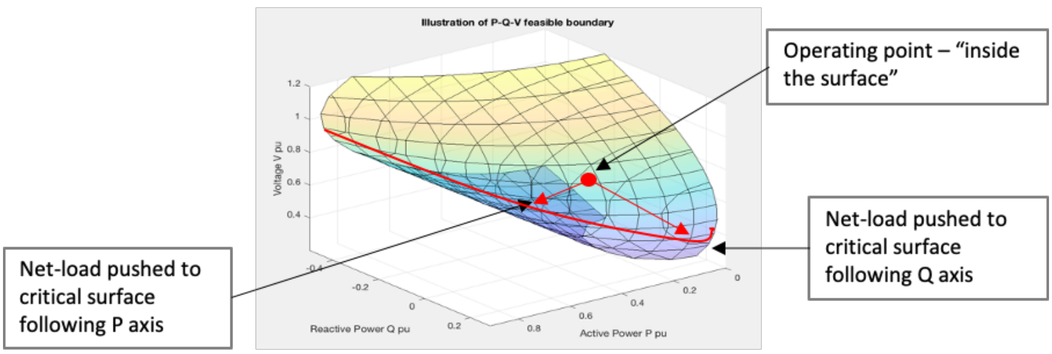


Fig. 4.2 Approach to identify 'Critical' points from any operating stable point (Source = Author)

Voltage excursion risks models

Similarly, inside the *Stable Region* we can define the under-voltage boundaries ($L^*_{n,i,under}$) and over-voltage boundaries ($L^*_{n,i,over}$) as the points where the net-load values reach the extremums causing the voltage v_i at node i to reach the respective extremum statutory limit respectively due to excessive local generation ($L_n < 0$) or load ($L_n > 0$), or:

$$L^*_{n,i,under} = \{L_n \mid \phi(V,\theta,L_n) = 0, v_i = U_v, \forall i\} \tag{4.8}$$

$$L^*_{n,i,over} = \{L_n \mid \phi(V,\theta,L_n) = 0, v_i = O_v, \forall i\} \tag{4.9}$$

where

- O_v is the upper regulatory threshold for the voltage at any node
- U_v is the lower regulatory threshold for the voltage at any node

Similarly to the *Critical Surface* the points satisfying those conditions belong both to a manifold of dimension $B-1$ that we call respectively *Under-Voltage Surface* and *Over-Voltage Surface*.

Therefore, inside the stable region (hyper-volume), we can define a smaller *Acceptable Region* (hyper-volume) which is upper bounded by the over-voltage multi-dimensional surface and lower-bounded by the under-voltage multi-dimensional surface and that can be represented by the set:

$$
\begin{aligned}
S_{acceptable} = \quad & \{(V,\theta,L_n)| \quad & | L_n < L^*_{n,i,under}, \\
& & \phi(V,\theta,L_n) = 0, \\
& & v_i \geq U_v\} \\
\cup \quad & \{(V,\theta,L_n)| \quad & | L_n > L^*_{n,i,over}, \\
& & \phi(V,\theta,L_n) = 0, \\
& & v_i \leq O_v\}
\end{aligned}
\tag{4.10}
$$

The model of the *Over-voltage* risk event for the node i is defined by the set of net-loads values (L_n^{over}) that are not in the stable region and for which the voltage at node i over the upper regulatory threshold or:

$$\{L_n^{over} \mid L_n \notin S_{acceptable}, v_i \geq U_v\} \tag{4.11}$$

The model of the *Under-voltage* risk event for the node i is defined by the set of net-loads values (L_n^{under}) that are not in the stable region and for which the voltage at node i under the lower regulatory threshold or:

$$\{L_n^{under} \mid L_n \notin S_{acceptable}, v_i \leq O_v\} \tag{4.12}$$

4.3 Voltage risk assessment methodologies

The previous models declined the criteria to use in order to establish if a net-load scenario leads a risk event at any node of the system. The proposed voltage risk framework will require the assessment of a vast amount of net-load conditions given in input in Monte Carlo simulations to derive the voltage risk probability for multiple DER penetration scenarios. For each net-load condition, it will be necessary to identify if any risk event is verified, meaning: critical-voltage condition or over-voltage/under-voltage at any node of the system.

In order to accomplish this, we can:

- Simulate the entire system and solve the load flow equations for each net-loads input.
- Train and use a Machine Learning predictor

4.3.1 Load flow equation solving

A traditional way to predict the voltage values at every node of a power system is to solve the load flow equation (4.1)(4.2) for each input net-load condition, assuming that all the parameters are known. By comparing the simulated voltage at each node with the regulatory thresholds U_v and O_v one could identify if a risk event would occur in response to that net-loads condition.

There exist no known analytical solution to this problem that must be solved using well established numerical iterative methods (e.g. Gauss-Seidel, Newton-Raphson, Fast-decoupled Method, etc.) [73]. Typically, the rate of convergence of those methods is quadratic, but in[74] it is shown that the convergence region and number of iteration can significantly vary depending on the initialisation and the actual loading condition of the system. Each iteration requires to calculate and inverse the $2n \times 2n$ Jacobian matrix of the system which time complexity is $\sim O((n)^3)$. As soon as the system's size increase, and also depending on the loading condition of the system, solving the system becomes computationally intense. It is therefore poorly practical to analyse the effect of vast numbers of scenarios using the brute force of solving the load flow equations for each combination of net-loads input.

4.3.2 Machine Learning classifiers

An alternative method to the load flow equation solving could be to use a data-driven approach to treat the problem as classification problem (3.3.1) for each risk, where the inputs are the net-loads values at every node of the system, and the labels are binary values for each node i : 'risk event verified' or 'risk event not verified'.

Support Vector Machine (SVM) is a Machine Learning technique adapted to solve classification problems where the model parameters are defined to maximise the *margin*, which is the smallest distance between the decision boundary and any of the training samples (3.3.2).

In our case, the decision boundaries should coincide with the *Critical Surface*, *Under-voltage surface* and *Over-voltage surface* surfaces (4.2.2) and we train multiple two-class (*Stable* versus *Unstable*, *Acceptable* versus *Unacceptable*) linear models:

$$Y(L_n) = w^T \phi(L_n) + b \tag{4.13}$$

where $\phi(L_n)$ denotes a fixed feature-space transformation and b is the bias.

The preliminary step of this method requires to create a meaningful set of training data for the classifiers. In order to achieve that, a network configuration and related network parameters are identified and the set of data generated through Monte-Carlo simulations as described in the next sections.

4.3.3 Training the classifier for over-voltage and under-voltage risks

As mentioned in the previous paragraph, the under-voltage risk and over-voltage risk, the decision boundaries coincide with the *Under-voltage surface* and *Over-voltage surface* surfaces and we train 2 distinct two-class linear models per nodes:

1. Over-voltage model that discriminate between acceptable region and over-voltage region for that node (4.11)

2. Under-voltage model that discriminate between acceptable region and under-voltage region for that node (4.12)

For each model type (over-voltage and under-voltage) and bus i, we will require a set of training data that will comprise M input net-loads vectors $L_{n1},..,L_{nM}$ with corresponding target value $Y_1,..,Y_M$ where $Y \in \{-1,1\}$ corresponding to the voltage of node i being respectively within threshold ($Y = -1$) or not ($Y = 1$), meaning being under-voltage or over-voltage respectively depending on the model type .

In order to generate those training sets, a network configuration and related network parameters are identified, and a Monte-Carlo simulation on the system is executed extracting net-load values for every node from a stochastic distribution, solving the load-flow and comparing the voltage at each bus with the thresholds to capture the 'true' target values for each model type as described in Fig.4.3.

The inputs to the training dataset generation framework are::

- The network configuration and parameters. It is to be noted that the selection of the network configuration and parameters will influence the construction of the training data sets and therefore the classifiers. The classifiers will be used for long term probabilistic scenario analysis of DER penetration. Therefore, in first approximation, the most likely network configuration should be used. To account for multiple network's configurations, the presented data set generation framework should be repeated and specific classifiers trained for each configuration. A detailed sensitivity analysis of the classifier versus different network configuration changes (eg. switch of feeders, compensating devices,..) is left for future work.

- The buses that include DERs. In a conservative approach, as the models will be used to assess the impact of any future penetration of DERs, the training dataset should include a good representation of every possible cases (from 0 to 100 % of bus with DERs). Practically, the penetration of DER to consider while creating the training dataset must be compatible with the penetration level for which the models will be used.

- Expected individual loads time series. This can be derived from past data on the network, characterised by an observed average daily profile and standard deviation

- Expected individual generation time series. Those must include the worst case scenarios in terms of intermittency an that can be synthesised with large stochastic fluctuations from average daily generation profiles observed in the past.

- In case of combined solar PV and storage capacity, the individual loads and generation time series would be adapted to account for the given battery charge-discharge profiles

The outputs are:

- A set of trained SVM classifiers for each bus

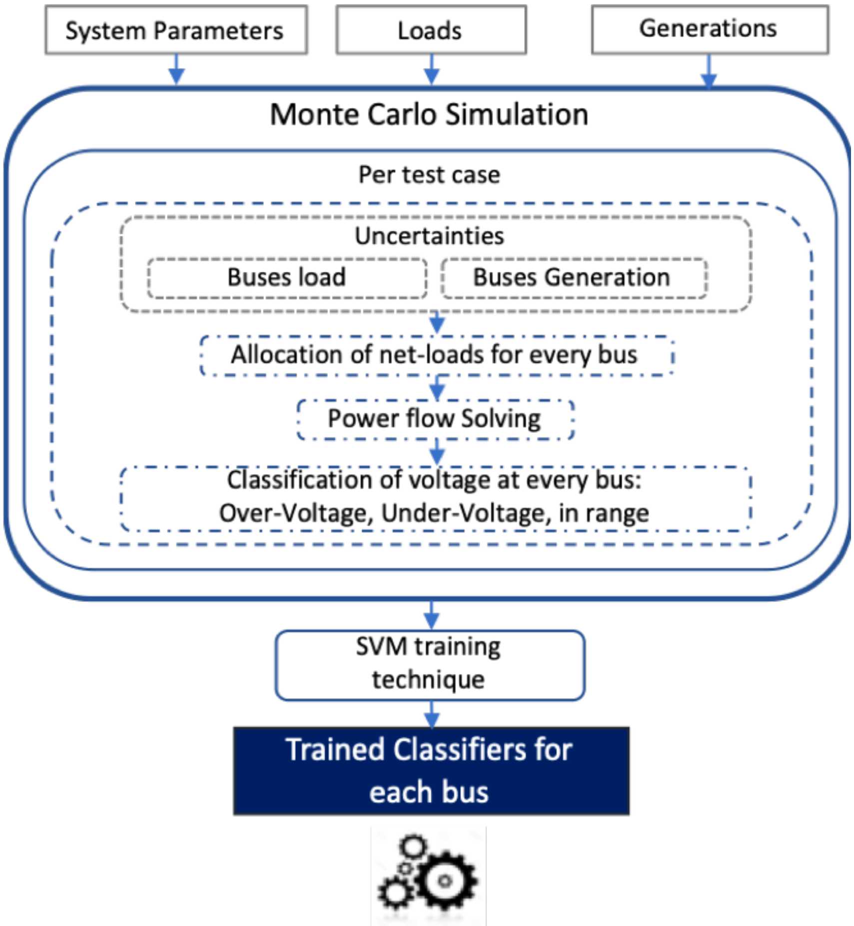


Fig. 4.3 Training data set generation framework (Source: Author)

Each training set is used to train the respective classifier, meaning identifying the model parameters: $w, \phi(L_n)$ and b in ((4.13)) using the standard training techniques described in [2] and [64].

It should be noted that the values of the net-loads used in the Monte Carlo simulation at each node are extracted from a uniform distribution within their maximum operating range. Using uniformly distributed net-loads within their maximum range of operation enables us to uniformly sample the surface boundaries when building the training set for the SVM classifiers.

The number of samples is empirically chosen so that the SVM classifiers are sufficiently accurate in their prediction. The accuracy will be defined as the proportion of true results (true positives and true negatives) among the total number of cases examined and arbitrarily set to be close to 95%.

Once the classifiers are trained they can be used to assess any given set of new intermittent net-load vector for the same configuration of the power system and predict potential voltage risk events. The creation of the training set data for the classifiers is computationally intense, but is a one-off exercise. Once trained and tested, the classifiers which are simple linear functions can be used to assess the voltage risk for any given set of net-loads.

4.3.4 Train the classifier for voltage collapse risks

Creating the training set of data for the *Critical voltage* risk is different from the *Under-voltage* and *Over-voltage* risks because by definition, the *Critical voltage* risk corresponds the maximum loadability condition of any node as described by (4.7), after which the load flow equations (4.3) have no solution and it is therefore impossible to identify solutions for those equations out of the stable region (4.6).

Using appropriate methods (eg. continuation load flow method (Appendix B), it is possible to calculate for each node i the net-load value that exactly sits on the *Critical Surface* (4.5), or the maximum loadability value or *Critical points* of node i given constant net-loads value for all the other nodes of the system.

With this method, we can therefore identify input vectors $L-n$ that sit exactly on the *Critical Surface* for the node i and that surface corresponds exactly to the decision boundary surface of the classification problem (where $Y=0$ in (4.13). But technically, the points belonging to that surface are still 'feasible' and we must identify conditions that are unstable. In order to artificially generate the points that sit 'beyond' that boundary (points for which there is no solution to the load-flow equations), we will add a small ε to the critical net-load value of node i to create a point that leads to an unstable load condition. This point will be labelled as $Y = 1$ or *Critical voltage* risk reached. Similarly, ε will be subtracted to the critical net-load value of node i to create a point that sits in the stable region and this point will be labelled as $Y = -1$ or *Critical voltage* risk is NOT reached.

We remember that a binary SVM classifier identifies a decision boundary between the two decision regions that maximises the margin (smallest distance between the decision boundary and any of the training samples). In our case, this margin will correspond to the arbitrary ε.

Employing this method and in order to create a set of training data, a network configuration and related network parameters are identified, and a Monte-Carlo simulation on the system is executed extracting net-load values for every node from a stochastic distribution. The load flow equations are solved for the entire system at each simulations and at each simulation, for each node i, the *Critical points* for that node are systematically calculated keeping anchored the net-loads values L_n of the other nodes. This is equivalent to having a sparse random set of multi-dimensional state values for the system and systematically "pushing" one value at the time (each node i) to its stability boundary following its directive axis. The method is illustrated in Fig.4.4.

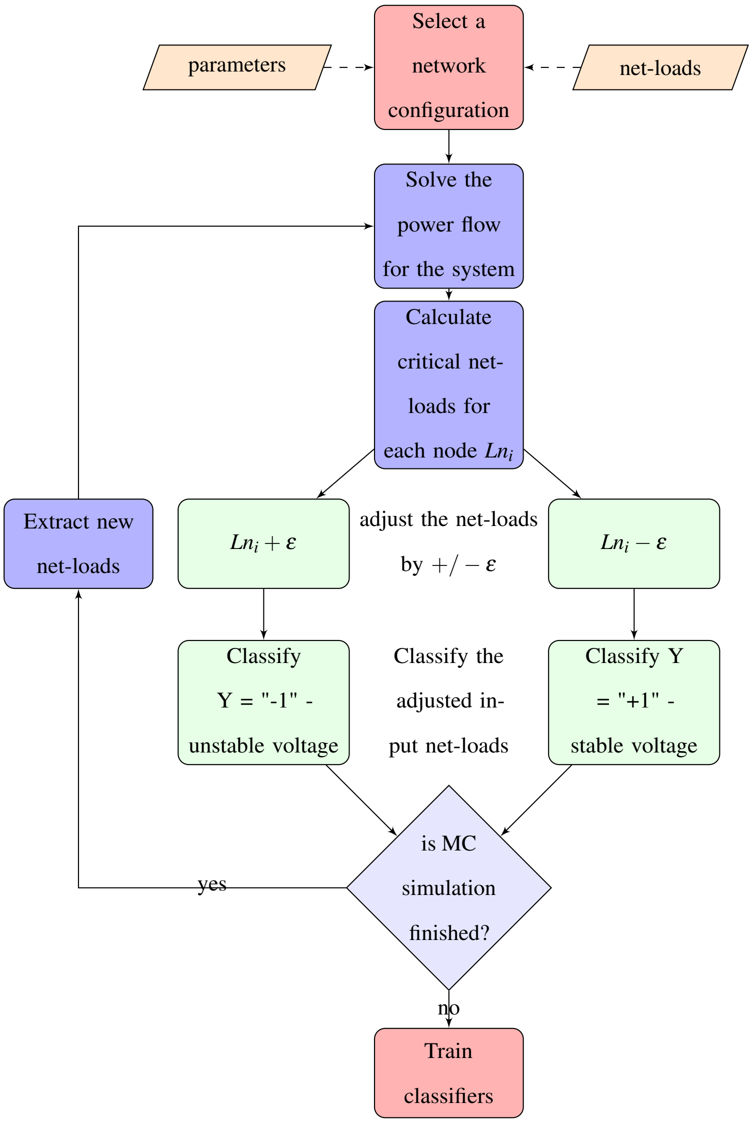


Fig. 4.4 Flow chart to generate SVM training data set (Source: Author)

At each simulation and for each node i the boundary points are systematically calculated for that node keeping invariant the net-loads of the other nodes. It is assumed that the power factor pf_i at every node i is constant and known. Therefore, when extracting the net-loads in the Monte-Carlo simulation, the active power P_i of node i is randomly extracted and the reactive power Q_i set to $Q_i = P_i \cdot pf_i$. Similarly, when searching for the critical load at node i, the power factor pf_i is considered as constant.

Therefore, for each of the M Monte-Carlo simulation and for node i of the B buses system, we calculate 3 sets of vectors with the respective labels:

The critical net-loads ($L_{i,crit}$) :

$$\begin{aligned} \text{Input:} \quad & L_{i,crit} = [P_1, ., P_{i,crit}, .., P_B, Q_1, .., pf_i \cdot P_{i,crit}, .., Q_B] \\ \text{Label:} \quad & Y = 0 \end{aligned} \tag{4.14}$$

where:

- $P_{i,crit}$ is the active power that sits exactly on the critical surface (calculated using Continuation Load Flow method - Appendix B))
- pf_i the power factor of node i

The 'stable' net-loads input ($L_{i,stable}$) and label:

$$\begin{aligned} \text{Input:} \quad & L_{i,stable} = [P_1, ., P_{i,stable}, .., P_B, Q_1, .., pf_i \cdot P_{i,stable}, .., Q_B] \\ \text{Label:} \quad & Y_{i,stable} = -1 \end{aligned} \tag{4.15}$$

where :

- $P_{i,stable} = P_{i,crit} - \varepsilon$
- ε the arbitrary margin
- pf_i the power factor of node i

The 'unstable' net-loads input ($L_{i,unstable}$) and label:

$$\begin{aligned} \text{Input:} \quad & L_{i,unstable} = [P_1, ., P_{i,unstable}, .., P_B, Q_1, .., pf_i \cdot P_{i,unstable}, .., Q_B] \\ \text{Label:} \quad & Y_{i,unstable} = +1 \end{aligned} \tag{4.16}$$

where:

- $P_{i,unstable} = P_{i,crit} + \varepsilon$
- ε the arbitrary margin
- pf_i the power factor of node i

Only the 'stable' and 'unstable' data sets are used to train the SVM classifier, while the critical net-load and margin ε are used to construct them.

The number of samples (M) and the margin ε is empirically chosen so that the SVM classifiers trained on the data set are sufficiently accurate in their prediction. The accuracy will be defined as the proportion of true results (true positives and true negatives) among the total number of cases examined and arbitrarily set to be close to 90%. It has been observed that very small ε that creates a very thin margin between the positive and negative classes lead to less accurate models that are trained on the same data but with ε in the range $5-10\%$.

Once the classifiers are trained they can be used to assess any given set of new intermittent net-loads vector for the same configuration of the power system and predict potential voltage risk events. The creation of the training set data for the classifiers is computationally intense, but is a one-off exercise. Once trained and tested, the classifiers which are simple linear functions can be used to assess the voltage risk for any given set of net-loads.

4.3.5 Voltage risk probabilities - scenario assessment

The previous subsections describe the method identified to build and train SVM classifiers that can assess the occurrence of a voltage risk event at any node in response to any given point in time net-loads within the power system.

The strength of the method is that the SVM classifiers are decoupled from the stochastic distributions of the net-loads and can repetitively be utilised to quickly assess the occurrence of a risk event in response to load and generation profile within the system.

Leveraging the law of big numbers, the proposed risk framework will use large Monte Carlo simulations to determine the marginal voltage risk event probability for every node in response to any scenario of intermittent DER within the system as illustrated in Fig.4.3.5.

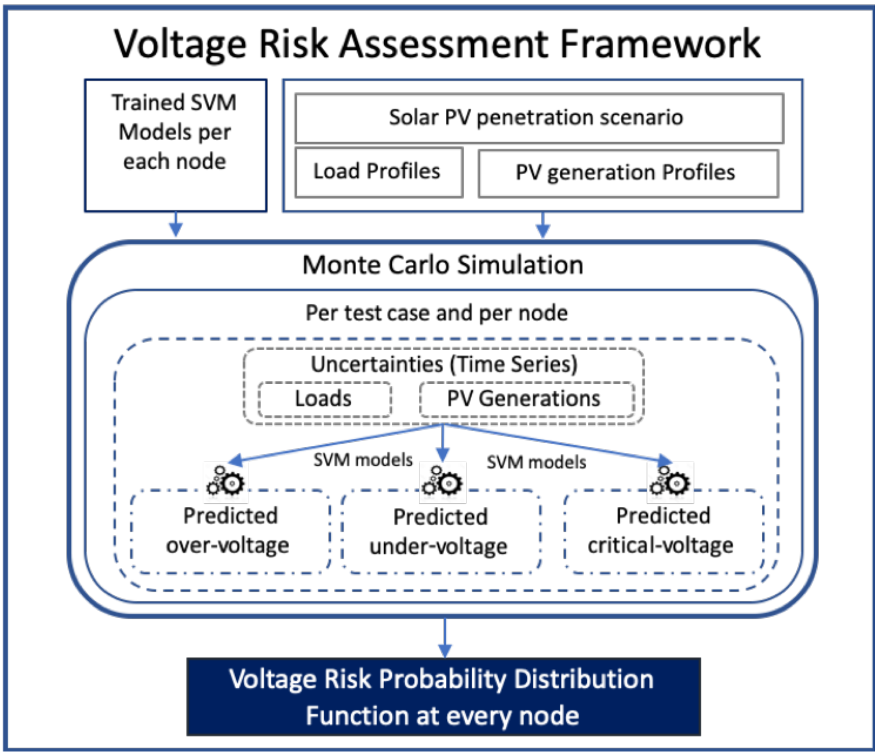


Fig. 4.5 DER scenario voltage risk assessment framework (Source: Author)

The inputs to the voltage risk assessment framework are the SVM models and the scenarios:

- Scenarios of DERs penetration, where each scenario is defined by: number of DER connections, location of connections in the grid and capacity of connected DER
- Expected individual loads time series. This can be derived from past data on the network, characterised by an observed average daily profile and standard deviation

- Expected individual generation time series. This can be derived from solar irradiation statistics and differentiated by different geographical area of the network with any chosen level of granularity

- In case of combined solar PV and storage capacity, the individual loads and generation time series would be adapted to account for the given battery charge-discharge profiles

The outputs of the voltage risk assessment framework are:

- Time series of predicted voltage risk events (over-voltage under-voltage, critical-voltage) at every node of the power system for each scenario

- Aggregated probability distribution functions (pdf and cdf) of the risk occurrence at every node of the system for every risk type

The accuracy of the outputs is determined by the accuracy of each model to assess the effect of the scenarios given in input. For the models to be accurate, they must have been trained on data that are in similar range as the the scenarios. Practically, as an example, if we want to assess the impact of 80 % DER penetration in the network, we must have simulated similar level of penetration when creating the training dataset for the models. Similarly, the generation intermittency must be considered when generating the training dataset for the models. This can be done including worst-case scenarios of intermittency across the individual DERs when simulating the training dataset.

The voltage risk framework enables to quickly assess the impact of different net-load distributions or scenarios in terms of voltage risk event probability of occurrence, our simulations have shown that the assessment is 800 times faster with our method compared to traditional load flow equation solving. The comparison is made using the same laptop (MacBook Pro) and assessing the voltage risk for 10.000 scenarios using two methods. In the first case, the load-flow equations are solved using the Newton-Raphson method in Matlab. In the second case, the SVM classifiers are used to predict the voltage risk events also in

Matlab. Assessing the voltage risk through the SVM is 800 times faster than solving the load-flow equations for the same scenarios.

4.4 Voltage risk severity - Financial Impact

The previous section has described a framework usable to assess the voltage risk events probabilities in response to given stochastic net-loads were the risk events are *Critical voltage*, *Over voltage* and *Under voltage*.

As described in [84], [86], in probabilistic risk assessment the risk value associated to a scenario S_i over a period of time T is the product of its *probability* Pr_i and *consequence* X_i integrated over the period of time or

$$Risk_{S_i} = \int_T Pr_i \cdot X_i \, dt$$

In our case, we have identified three different types of events or scenarios and developed a methodology to estimate their probabilities of occurrence but need to identify metrics to assess their *consequence* in monetary value.

For the first scenario listed (critical voltage), the *consequence* is a total or partial collapse of the system, the extend of which would depend on the action taken by the operator while it occurs as described by van Cutsem [24].

In Australia, the *"Value of Customer Reliability (VCR)* [19] provides a monetised value to the reliability of energy supply differentiated by type of consumers. With $k \in \{residential, commercial, industrial\}$, VCR_k estimates the dollar value of the energy not supplied to each type of customer (in $\$/Kwh$) .

Using this metric, we can assess the risk value associated to a *Critical Voltage* scenario over a period of time T for each node i by:

$$Risk_{crit} = \int_T Pr_{crit} \cdot \sum_k (VCR_k \cdot C_k \cdot ENS)\, dt \tag{4.17}$$

where:

- k is the type of consumers: residential, commercial, industrial
- Pr_{crit} is the probability of a voltage collapse risk event happening during the period
- VCR_k is the value of customer reliability as defined by the market operator
- C_k is the number of consumers impacted for each type.
- ENS the energy not supplied during the voltage collapse risk event at node i

In addition to estimating Pr_{crit}, the proposed voltage risk frameworks identify the *weak* bus that would collapse first and therefore helps to localise the portion of network impacted and consequently the type and number of consumers connected. Some assumptions need to be made in order to estimate the amount of energy not supplied and it is proposed to use recorded past service interruption data and expected loads profiles at similar time of the day and period of the year.

For the second and third scenario identified (under-voltage and over-voltage), the *consequence* is mainly felt by the consumers and the excursion of voltage can disrupt or damage the connected appliances or machinery. In Australia, regulations are in place to enforce the DNSP's (Distribution Network Service Provider) to maintain the voltage within statutory limits and a non-respect, even temporary, entitles the consumers to lodge claims that must be managed and cleared by the DNSP's. The cost of those complaints can be used as a metric to assess the *consequence* of over-voltage and under-voltage scenarios at every node $i/$ as follows :

$$Risk_{over} = \int_T Pr_{over} \cdot \sum_k (Pr_{compl} \cdot C_{compl} \cdot C_k)\, dt \tag{4.18}$$

where:

- k is the type of consumers: residential, commercial, industrial
- Pr_{over} is the probability of an over- voltage risk event happening during the period
- Pr_{compl} is the probability of receiving a complain from the consumer type k due to an over-voltage event
- C_{compl} is the average cost of complaint for consumer type k
- C_k is the number of customer of type k connected at node i

$$Risk_{under} = \int_T Pr_{under} v \sum_k (Pr_{compl} \cdot C_{compl} \cdot C_k)\, dt \tag{4.19}$$

- k is the type of consumers: residential, commercial, industrial
- Pr_{under} is the probability of an under-voltage risk event happening during the period
- Pr_{compl} is the probability of receiving a complain from the consumer type k due to an under-voltage event
- C_{compl} is the average cost of complaint for consumer type k
- C_k is the number of customer of type k connected at node i

The proposed voltage risk framework estimates the point in time Pr_{over} and Pr_{under} and identifies the bus where it occurs, and therefore helps to assess the number of customers impacted. C_{compl} can be extrapolated by DNSP's from past events and similarly, the Pr_{compl} can be extrapolated by DNSP's from historical data depending on the severity, duration and location of the event.

4.5 The end-to-end risk framework

By using machine learning classifiers, the presented method decouples the problem of assessing the occurrence of a risk given one net-load profile and the capability of assessing vast amounts of scenarios. The initial problem, which is connected to training the machine learning classifiers, is computationally intensive and uses traditional methods to solve the load flow equations for the system, but this is a one-off exercise per network. Once the classifiers are trained and sufficiently accurate, they can be repetitively and swiftly use to assess any net-load and penetration scenario for the system. Our simulations have shown that using the classifiers is 800 times faster than solving the load flow equations.

This proposed methodology opens the way to assess wide range of 'what-if' analyses for the system and derive voltage risk probabilities that can then be linked to financial impact and constitute a quantitative voltage risk assessment framework as illustrated in Fig.4.5.

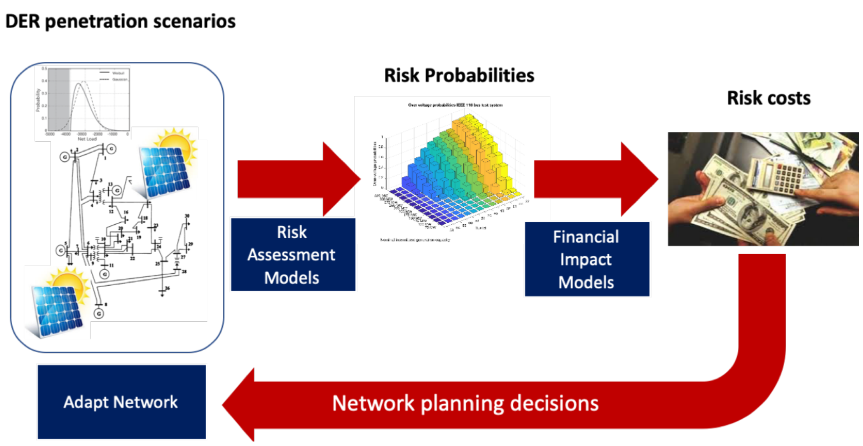


Fig. 4.6 End-to-end voltage risks framework (Source:Author)

This end-to-end framework is useful for DNSPs to assess the probabilistic financial impact of any scenario of DER penetration within their network. The financial impact can be compared to mitigation actions (eg. change of network configuration, assets upgrade, network reinforcement, etc) in order to select the most appropriate ones.

4.6 Case Study - the IEEE118 buses test network

In this section we illustrate on a test case (the IEEE 118 buses test network [4]) end-to-end and step-by-step the methodology required to build the voltage risk framework. This test case represents a transmission network and therefore we develop the analysis by aggregating all the loads and solar PV generations from the downstream distribution networks at the respective connection bus to the 118 buses test network. Without loss of generality, a set of nodes has been identified to accommodate solar PV generation and arbitrarily selected in pockets to simulate high concentration of intermittent loads in some geographic area (circled in fig. 4.7). Those nodes will be referred to *weak* nodes. Those nodes have been selected arbitrarily to focus the discussion and analysis on a limited set of cases. The proposed methodology would be applicable to accommodate and analyse the impact of solar PV generation at each node of the system.

The risk events considered in this simulation are the critical voltage, over-voltage and under-voltage.

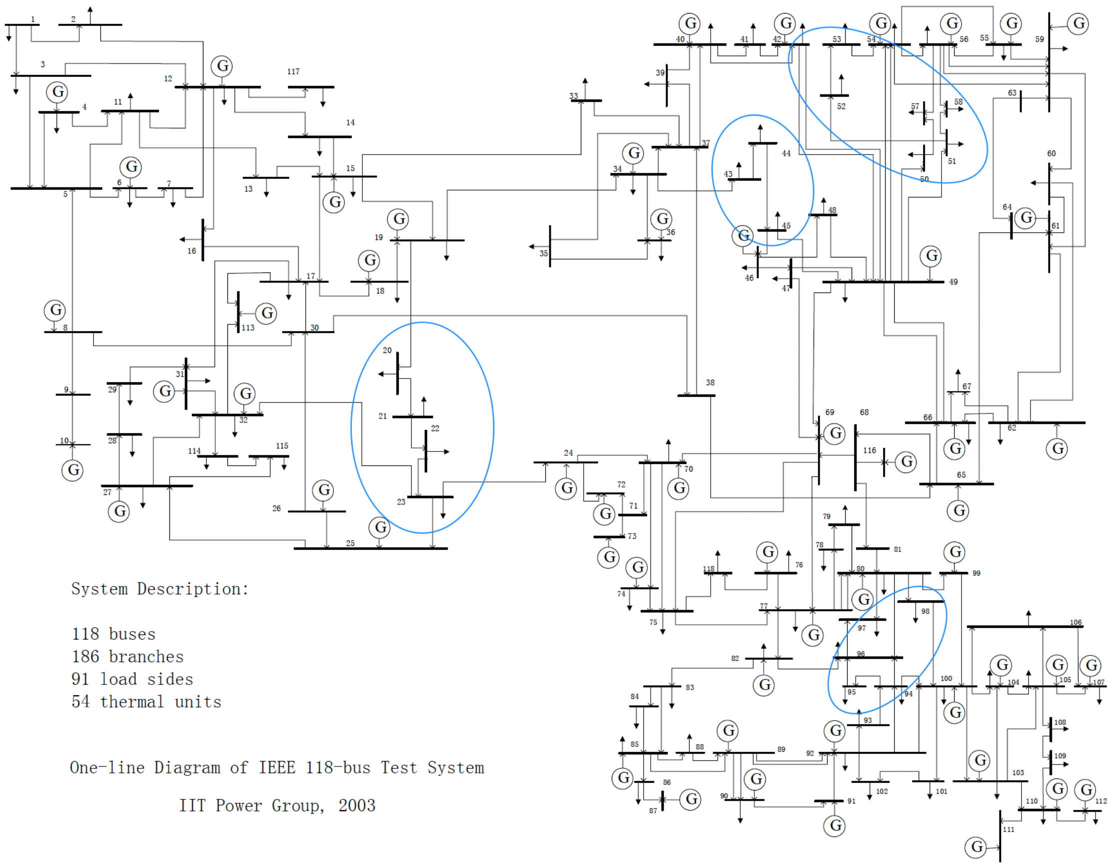


Fig. 4.7 IEEE 118 test system and weak buses [4]

4.6.1 Train the SVM classifiers

The initial step of the method illustrated in 4.3.3 and 4.3.4 requires the creation of training data sets in order to train SVM classifiers. To achieve that a 100k Monte Carlo iteration has been performed varying each load and embedded generation independently, randomly extracted from uniform distributions.

For each iteration, the load flow equations (4.1) and (4.2) have been solved and the state of the system (V, θ) recorded.

In order to create training sets of data for the *Critical Voltage* risk event (Vectors described in (4.15) and (4.16), for each simulation and *weak* node (Fig.4.7), the *Critical Loads* has been calculated implementing the continuation load flow algorithm [87, 88] that is described in AppendixB.

To create training sets for *Over-Voltage* or *Under-Voltage* risk events, for each simulation and weak node i the recorded voltage v_i has been compared to the statutory limits ($1.1 p.u.$ or $0.94 p.u.$) and classified accordingly. A data set is classified as *positive* when the risk event occurs (breach of threshold or voltage collapse) and *negative* otherwise.

For each *weak* node i, 3 SVM classifiers have therefore been fitted, one for each risk events that can predict an event occurring at that node using a Gaussian Radial Basis Function (RBF) as kernel as described in (3.3.2)

Following standard practice in Machine Learning, the models have been trained and validated using 10-fold cross-validation technique [89]. This technique is commonly used to protect against over-fitting of the model.

It is to be noted that the training procedure proposed requires solving the load flows for the entire system at each Monte-Carlo simulation. That is computationally intense, but this preliminary operation is required only once given a power network configuration.

4.6.2 Measure the SVM classifiers performance

The performance of the classifiers is assessed using a *confusion matrix* Tab.4.1 which records correctly and incorrectly classified risk events and where true positive (tp) refers to the count of classifier correctly predicting of a risk event that occurs, true negative (tn) refers to the count of classifier correctly predicting the absence of a risk event that doesn't occur, false positive (fp) refers to the count of classifier wrongly predicting the absence of a risk event while it occurs and false negative (fn) refers to the count of risk that are wrongly predicted by the classifier while it doesn't occur[90].

Class \ Recognised	as Positive	as Negative
Positive	tp	fn
Negative	fp	tn

Table 4.1 Confusion matrix (Source: Author)

The classification performance will be assessed via commonly-accepted measures [91]: *accuracy* and *F-1 score*:

$$accuracy = \frac{tp+tn}{tp+tn+fp+fn} \tag{4.20}$$

$$precision = \frac{tp}{tp+fp} \tag{4.21}$$

$$sensitivity\ or\ recall\ = \frac{tp}{tp+fn} \tag{4.22}$$

$$F1\ score = 2 \times \frac{precision \times recall}{precision + recall} \tag{4.23}$$

where the precision is the number of true risk events predicted divided by the total number of predictions and therefore captures the model "exactness". The sensitivity is the number of true risk events predicted divided by the total amount of risk events in the test data and therefore capture the model "completeness". The F-1 score conveys a balance between precision and sensitivity.

Over-voltage and under-voltage risks classifiers

To test the performance of the over-voltage and under-voltage risks classifiers, an independent validation data set has been generated from a new 100k Monte Carlo simulation on the same system configuration as the one used to create the training data set but with newly extracted input net-loads. For each iteration Monte Carlo simulation and respective set of input net-loads, the load flow equations have been solved and the state of the system (V, θ) recorded.

On the same set of input net-loads, the previously trained classifiers have been applied and respective prediction of voltage excursion risk made. The Tab.4.2 and Tab.4.3 show the resulting confusion matrix and accuracy measure for the voltage excursion risks (over-voltage and under-voltage respectively)

Bus	tp	tn	fp	fn	acc.%	prec. %	sens. %	F1 %
20	2251	97532	80	137	99,7	96,6	94,3	95,4
21	24286	75381	167	166	99,7	99,3	99,3	99,3
22	27792	71902	154	152	99,7	99,5	99,5	99,5
43	11724	87959	131	186	99,7	98,9	98,4	98,7
44	39652	60139	119	90	99,8	99,7	99,7	99,7
45	29084	70732	94	90	99,8	99,7	99,7	99,7
51	19027	80688	144	141	99,7	99,3	99,3	99,3
52	33870	65882	129	119	99,8	99,6	99,6	99,6
53	3486	96316	89	109	99,8	97,5	96,9	97,2
94	466	99434	44	56	99,9	91,4	89,3	90,3
95	19068	80624	181	127	99,7	99,1	99,3	99,2
98	14265	85453	145	137	99,7	98,9	99,0	99,0

Table 4.2 Accuracy of SVM classifiers per node and Over Voltage excursion risk event (Source: Author)

Bus	tp	tn	fp	fn	acc. %	prec. %	sens. %	F1%
20	31625	67922	218	235	99,6	99,3	99,3	99,3
21	34789	64763	236	212	99,6	99,3	99,3	99,3
22	9508	90037	211	244	99,5	97,8	97,4	97,7
43	31220	68403	186	191	99,7	99,4	99,4	99,4
44	37694	62000	155	151	99,7	99,6	99,6	99,6
45	19886	79758	178	178	99,6	99,1	99,1	99,1
51	18848	80984	84	84	99,8	99,6	99,6	99,6
52	36368	63474	72	86	99,8	99,8	99,8	99,8
53	46590	53295	61	54	99,9	99.9	99.9	99.9
94	10723	88939	143	195	99,7	98,7	98,2	98,4
95	35584	64098	142	176	99,7	99,6	99,5	99,5
98	16435	83169	171	225	99,6	98,9	98,7	98,8

Table 4.3 Accuracy of SVM classifiers per node and Under voltage excursion risk event (Source: Author)

Voltage collapse risk classifiers

In order to create an independent validation set of data for the voltage collapse risk classifier, some extreme net-loads has been generated as input for each sequentially for each weak nodes and an attempt to solve the load flow equations made. When the load flow equations for such input conditions presented no solutions, it has been assumed that the maximum loadability for the weak node has been exceeded under such conditions. Therefore those net-load conditions for node i has been labelled as 'critical voltage reached' ($Y = +1$), whilst the 100k Monte-Carlo simulations used for the voltage excursion risks (which all could find a solution) has been labelled as 'critical voltage not reached' ($Y = -1$) and those cases correspond to input net-loads values that are exceeding the maximum loadability.

The independent validation data set contains 14633 records of true voltage collapse events and 100k records of true non-voltage collapse. On those 114633 net-loads, the SVM Critical voltage classifiers have been applied to predict the voltage collapse event. Tab.4.4 shows the resulting confusion matrix and accuracy measure for the voltage collapse risk.

Bus Id	tp #	tn #	fp #	fn #	acc. %	prec. %	sens. %	F1 %
20	14623	81746	18254	10	84,07	99,99	81,75	89,95
21	14595	87703	12297	38	89,24	99,96	87,70	93,43
22	14591	74301	25699	72	77,52	99,90	74,30	85,22
43	1732	99695	305	12901	88,48	88,54	99,70	93,79
44	14542	99762	238	91	99,71	99,91	99,76	99,84
45	14391	81216	18784	242	83,40	99,70	81,22	89,51
51	14629	87452	12548	4	89,05	100.00	87,45	93,30
52	12502	98745	1255	3131	97,05	97,89	98,75	98,31
53	5255	99993	7	9378	91,81	91,43	99,99	95,52
94	5270	100000	0	9363	91,83	91,44	100,00	95,53
95	0	99990	10	14633	87,23	87,23	99,99	93,18
98	16435	83169	171	225	99,12	99,98	99,01	99,49

Table 4.4 Accuracy of SVM classifiers per node and voltage collapse risk event (Source: Author)

Performances discussions

We observe that the *F1 scores* of the trained classifiers for over-voltage and under voltage tested on independent are distinctively high (on average $99,2\%$ and 98.8%), even if tested on completely independent validation data. The reason of this good performance is attributed to the elevated size of the training data sets used ($100k$) to train the model. An attempt to train the models with only $10k$ data points has been done and in that case the average *F1 score* of those "limited" classifiers drops to $89,5\%$.

The elevated accuracy makes the classifiers trustworthy and useful to predict voltage excursion (over-, under-) risk events.

We denote that the performances of the critical voltage risks ar noticeably lower than the performances of the excursion risks. The attributed reasons for that are two:

1. The training data sets are unbalanced (class distribution not uniform). In fact, we have 100k observations for "non-risk" event and only 14633 observations for "risk" event. Therefore the size of the useful training data set is dictated by the lower range, e.g. $15k$

2. The construction of the training data set relies by construction on an arbitrary margin ε (Ref. 4.3.4) that was set to 10% in those simulations.

Nevertheless, the accuracy is still mostly above > 90 which makes the classifiers trustworthy and useful to predict voltage collapse risk events.

4.6.3 Scenario Risk Assessment using the SVM classifiers

Appropriately trained classifiers are now used to predict risk events on respective nodes in response to given net-load input. Illustratively, we are constructing multiple scenarios for our IEEE118 test buses system where we progressively introduce solar PV generation capacity on the selected pocket of weak buses (Fig.4.7) and similarly vary the peak loads values to analyse the effect on the voltage risks.

For each scenario:

- the maximum solar PV capacity is set to values increasing from 75MW up to 200MW (extreme situation corresponding to the peak nominal generation capacity equal to 3 times the average peak load) and in each case we analyse the effect of those potential new solar PV capacities on the over-voltage risk within the system over a full day.

- the maximum peak load for each bus is set to values increasing from 20% up to 250% of the reference peak value from the IEEE118 test case baseline data [4] and in each case we analyse the effect of those loads on the under-voltage risk within the system over a full day.

In order to generate the hourly net-load of the system, for every hour of the day we:

- use the relative generation profile mean and sigma values indicated in the profile illustrated in Fig.4.8 for that hour to randomly (according to a Gaussian distribution) construct a standardised generation factor to simulate intermittency

- multiply the previously constructed standardised generation factor by the maximum solar PV capacity of the scenario

- similarly use the relative load profile mean and sigma values indicated in the profile illustrated in Fig.4.8 for that hour to randomly (mean, sigma) construct a standardised load factor

- multiply the previously constructed standardised load factor by the maximum peak load of that bus and scenario

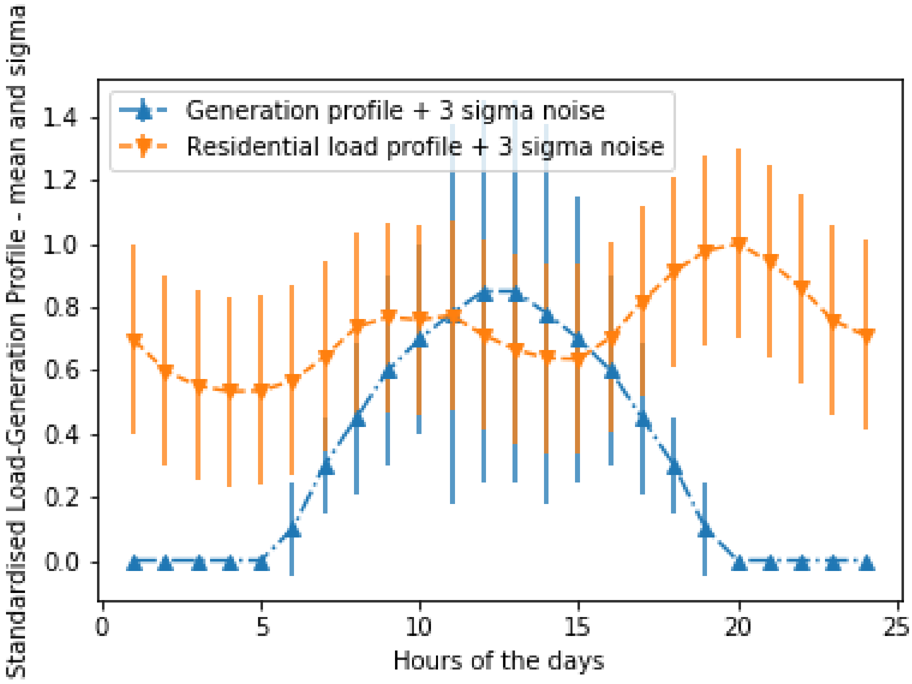


Fig. 4.8 Load and DER generation profiles - mean and sigma values (Source: Author)

For each scenario, we executed $10k$ daily Monte-Carlo simulations to generate hourly net-load series and the over-voltage risk events assessed using the respective trained classifiers. The hourly probability of risk event in each scenario has been inferred from the calculated frequency $(\frac{\#PositiveClass}{10{,}000})$.

Over-voltage risk

Over-voltage at buses within the system is caused by elevated values of power generated downstream the network that is not absorbed by the local loads and therefore cause reverse power flows. Even if the single value of power generated at one house is modest, when the penetration of solar PV within a system is high within a portion of the grid, all the single power generated and not consumed aggregate and can have more significant impact upstream the system. In this simulation, every bus represents a distribution network sub-station and the net-load considered are the aggregated values of elevated number of solar PV concentrate din the downstream distribution grid.

Fig.4.9 illustrates the over-voltage probabilities per bus and scenario at 12PM.

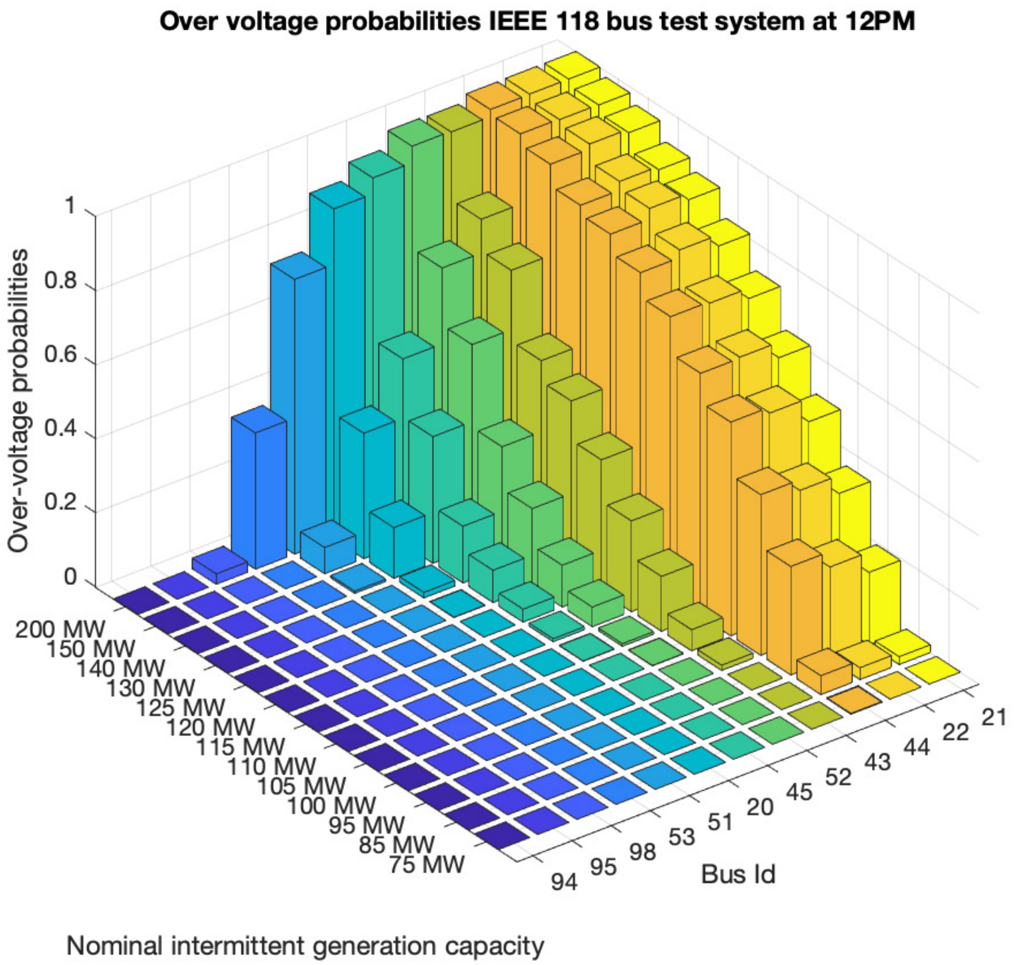


Fig. 4.9 Over-voltage risk probability for each weak buses at 12PM (Source: Author)

The graph shows that some buses are more sensitive to intermittent generation than others and that is dictated by the network topology and internal parameters (line admittances, transformer tap ratios). Interestingly, the proposed method allows to identify and map the voltage risk's sensitivity of each bus under any set of given scenarios.

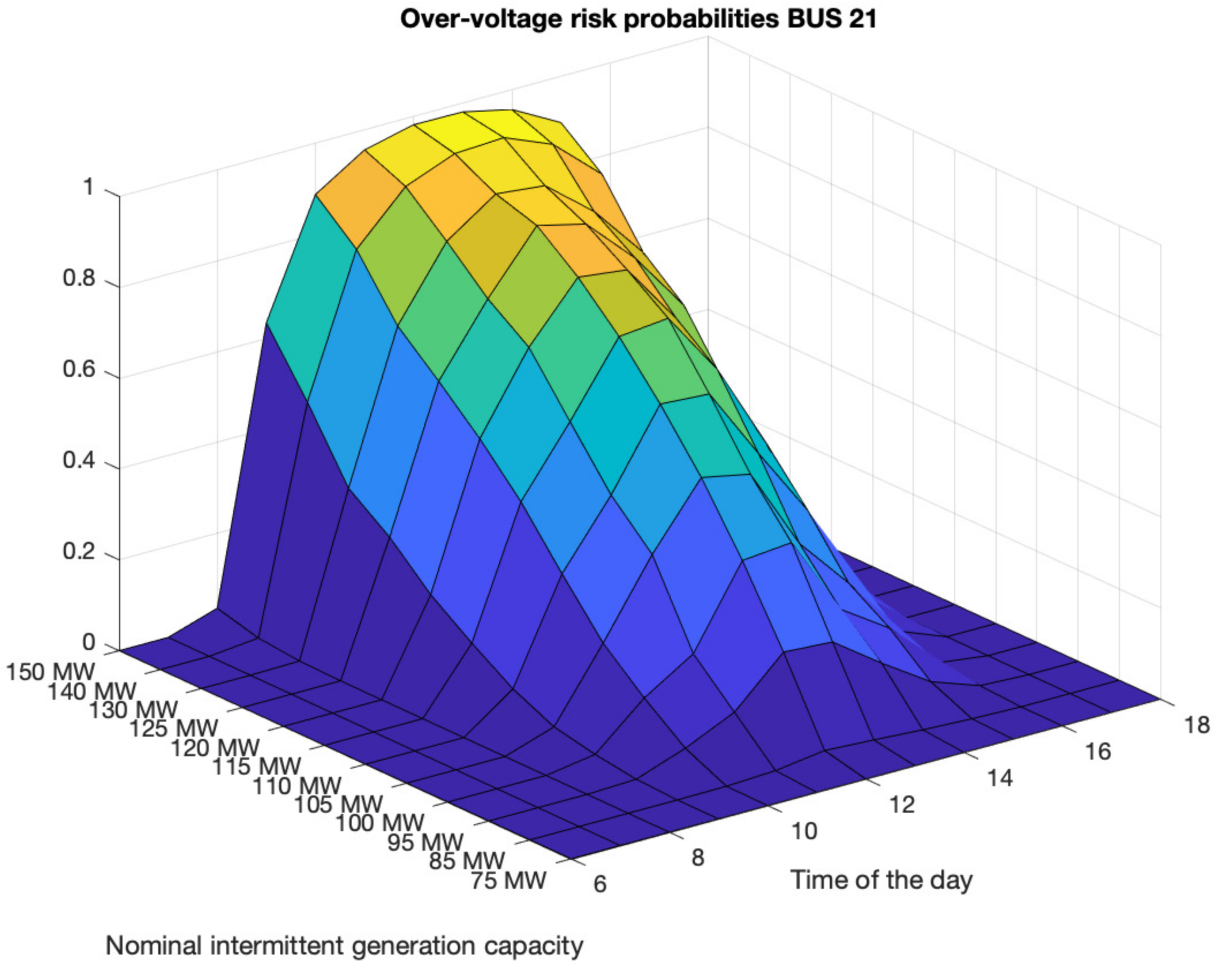


Fig. 4.10 Hourly over-voltage risk probabilities for BUS 21 (Source: Author)

Fig.4.10 illustrates for one bus the hourly over-voltage risk probability per scenario which remarkably illustrates the good sensitivity of the classifiers that are used to generate the probabilities which unsurprisingly follows the shape of the generation profile as over-voltage is caused by an excess embedded generation within the system.

Under-voltage risk

Conversely, under-voltage at buses within the system is caused by elevated loads values downstream the network that is not compensated by the local generation.

Fig.4.11 illustrate the under-voltage probabilities per bus and scenario at 8PM for the same Monte-Carlo simulations.

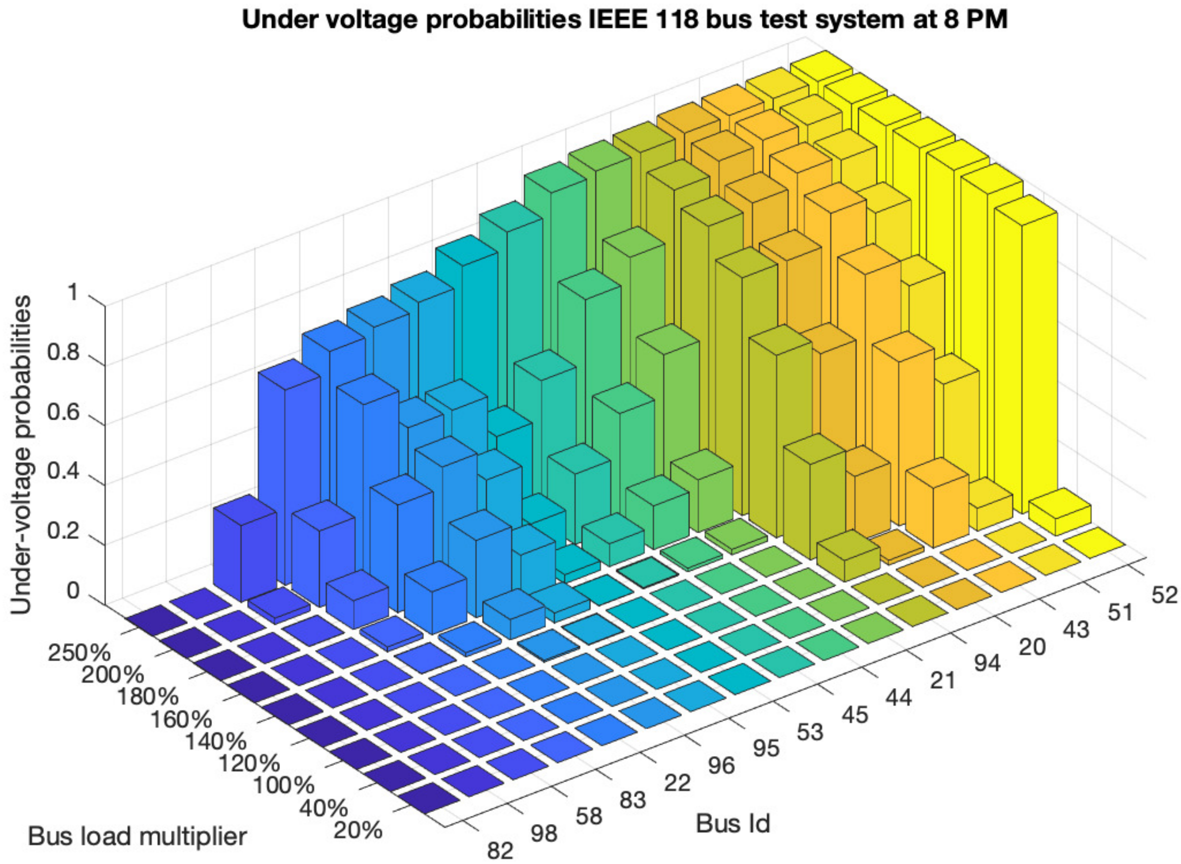


Fig. 4.11 Under-voltage risk probability for each weak buses at 8 PM (Source: Author)

The graph shows that some buses are more sensitive to elevated loads than others and that is dictated by the network topology and internal parameters (line admittances, transformer tap ratios). Interestingly, the proposed method allows to identify and map the voltage risk's sensitivity of each bus under any set of given scenarios.

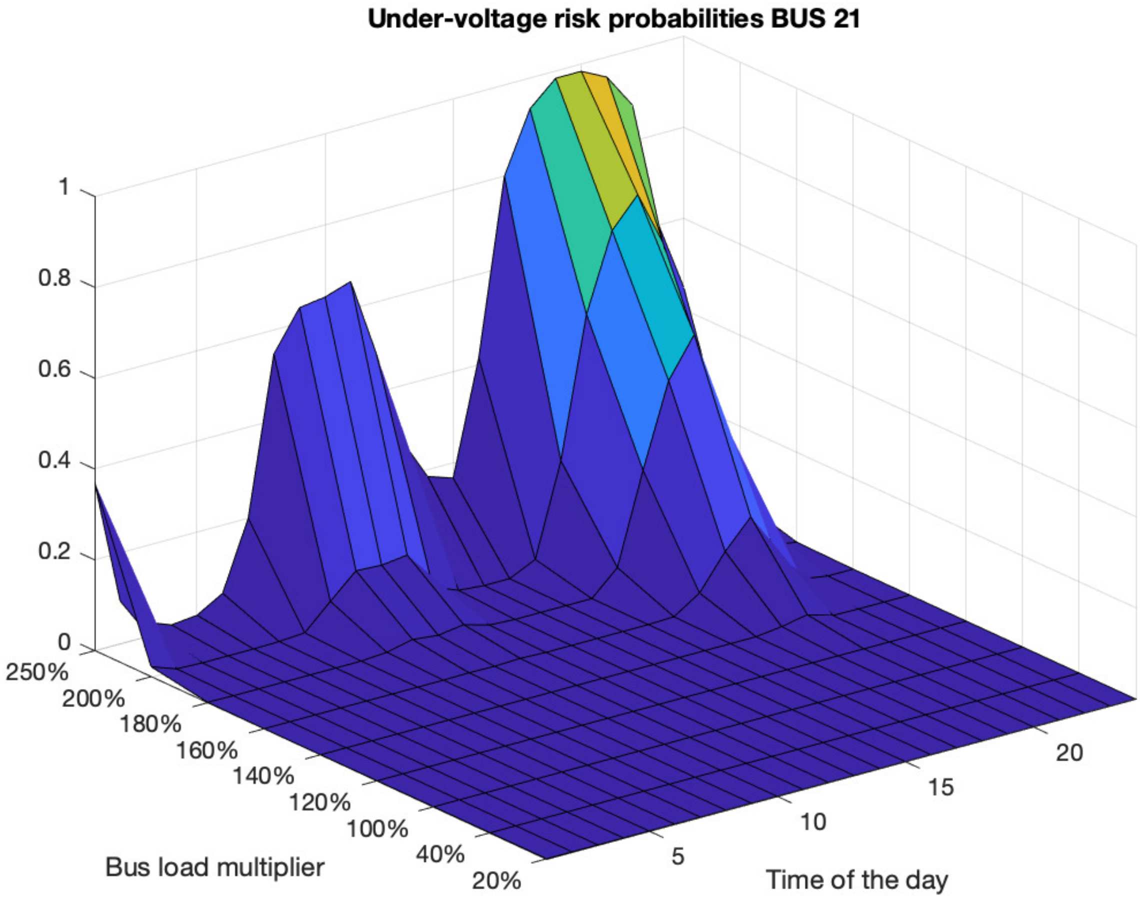


Fig. 4.12 Hourly under-voltage risk probabilities for BUS 21 (Source: Author)

Fig.4.12 illustrate for one bus the hourly over-voltage risk probability per scenario which remarkably illustrate the good sensitivity of the classifiers that are used to generate the probabilities which unsurprisingly follows the shape of the net-loads profile as under-voltage is caused by elevated loads within the system.

Voltage collapse risk

Under the conditions used for this Monte-Carlo simulation, there are no voltage collapse events predicted. This is not surprising as the conditions simulated are mostly in the normal operational range of the network test case. In order to generate the conditions that would trigger a voltage collapse, extreme loads must be applied to the system, in the order of 10 times the test case peak load.

4.6.4 Risk Assessment - estimate the financial impact

To complete our risk assessment for this simulation, we illustratively calculate the daily impact of one risk type for each scenario if no mitigation action is undertaken. The over-voltage risk type is selected because more likely to trigger complaints. In E(4.18), Pr_{over} is the probability of over-voltage illustrated in Fig. 4.10, $Pr_{compl.}$ is arbitrary set to 5%. To calculate the cost of complaint term we need to multiply an average cost of complaint with the number of customer impacted. The DNSP (Distribution Network Service Provider) usually precisely know that information. For the sake of this simulation, we will use an average cost of \$1.500. For each scenario, we will estimate the number of customers dividing the nominal intermittent generation of the scenario by $5kW$, this equivalent to saying that every customer connected to the bus is contributing to the intermittent embedded generation with $5kW$ Solar PV and the scenarios are simulating an increase of penetration within the system. Tab. 4.5 summarises the daily risk value for 4 buses: bus 21, bus 22, bus 43 and bus 44.

Nominal Intermittent Capacity scenario	Bus Id			
	21	22	43	44
150 MW	$15,58 M	$16,19 M	$10.68 M	$16,17 M
140 MW	$13,19 M	$13,81 M	$7,86 M	$14,02 M
130 MW	$10,56 M	$11,36 M	$5,34 M	$11,57 M
125 MW	$9,09 M	$9,99 M	$3.97 M	$10,00 M
120 MW	$7,62 M	$8,55 M	$2,77 M	$8,70 M
115 MW	$5,96 M	$6,90 M	$1,67 M	$7,51 M
110 MW	$4,39 M	$5,30 M	$0,88 M	$5,93 M
105 MW	$3,09 M	$4,04 M	$0,29 M	$4,52 M
100 MW	$1,71 M	$2,45 M	$48 k	$3,18 M
95 MW	$0,74 M	$1,29 M	-	$1,92 M
85 MW	$31 k	$89 k	-	$0,33 M

Table 4.5 Simulated over-voltage risk values for 4 buses (Source: Author)

The risk values are to be used by the DNSP to assess the proper mitigation actions and investment on their network in order to appropriately manage those risks (accept, reduce, avoid or transfer).

4.7 Cost of SVM model training discussion

The key advantage to use SVM models to assess the voltage risk is the significant decrease in computational complexity and time versus the traditional solving of the load-flow equations. But the use of SVM comes at the cost of training the models and generating the respective training data sets. This cost is a ‘one-off’ cost for each system configuration, and once the models are trained they can be used on vast (10^6) of DER penetration scenarios. The size of the training data sets influences the accuracy of the trained model. Intuitively, the bigger the size, the higher is the accuracy, up to a certain point after which it becomes asymptotic. In our simulation (4.6.1), we have tested two training data set dimension as indicated in Tab.4.6.

Training data set size	Average model accuracy
10^5	$99,2\%$
10^4	89.5%

Table 4.6 Model accuracy versus training set size (Source: Author)

This illustrates that even by reducing the training data set dimension by an order of magnitude, the accuracy remains close to the limit of 90% that was arbitrarily set. This illustrates that even by reducing the training data set dimension by an order of magnitude, the accuracy remains close to the limit of 90% that was arbitrarily set. In that case, the cost of creating the training data sets is 2 orders of magnitude lower than the equivalent cost of using traditional simulation techniques to execute the vast Monte-Carlo scenario analysis.

Further optimisation of the sampling of the training data set (eg [92]) is left for future work.

4.8 Conclusion

This chapter presents a novel end-to-end methodology to assess 3 types of voltage risks and financial impact within a power system. The benefits and contribution of the presented risk framework are discussed below:

- **Versatility** to assess any different scenario's type of nominal intermittent generation capacities and net-loads profiles as developed in Section 4.3.5. In effect, the presented method decouples the problem of assessing the occurrence of a risk given one net-load profile and the capability of assessing vast amounts of scenarios.

- **End-to-end** relevant for network planning decision supported by financially quantified risk values as developed in Section 4.6.4

- **Discrimination** to precisely identify weak bus within the system in response to scenarios, as the presented method in Section 4.3.3 trains individual classifiers by bus and risk event;

- **Scalability** under multiple dimensions:
  - Scalability to *large scale distribution system* due to intrinsic characteristics of SVM classifiers that are well suited and performing in high-dimension space [2];
  - Scalability to manage *multiple risk classes*. The presented methodology can be extended to manage a discretised number of classes for each risk. Each class would capture a distance range from the statutory limit and each class would have proper impact parameters in E(4.18) or E(4.19);

Chapter 5

The reverse power flow risk framework

5.1 Introduction

Increasing the share of renewable Distributed Energy Resources (DERs) to significant levels impacts the power networks in multiple ways, one-off which is the Reverse Power Flows (RPFs) [79]. The RPFs are currents generated from DERs that are travelling upstream in power networks that were originally designed to accommodate only downstream currents. The intensity of such currents depends on the penetration, concentration and individual capacity of the DERs that are generating them. The phenomenon is well known and the potential consequences are described in the literature [48, 28, 49].

Until the penetration and concentration of the DERs is contained to small levels, the aggregated intensity of the RPFs remains negligible compared to the peak load current and ratings of the assets in place to distribute it. Conversely, if the penetration and concentration of the DERs were to raise significantly in certain portion of the network, the aggregated RPF intensity would become significant and cause adverse system effects such as equipment damage and security issues [50] resulting in the need to alter the original design of the network to maintain its reliability and security levels. In [93], the authors propose an Optimal Distributed Generation Placement (ODGP) problem solving algorithm to minimise the RPFs in the grid, and [94] propose a load blinding stabilisation for protection relays subject to RPF.

The penetration rate and pace are mostly driven by the end-consumer and fuelled by incentives, ideology and affordability of the DERs themselves. This can't be predicted with certainty and therefore can be viewed as a stochastic process. Additionally, the factors influencing the adoption of DER are mostly out of control for the Distribution Network Service Providers (DNSP) who have a duty to maintain security and reliability in their network in any future scenario of DERs adoption and penetration.

Similar to the previous chapter on Voltage Risk, this research propose a scenario analysis framework that can be integrated in network planning decisions in order to control, prevent or mitigate the effects of increased DER penetration on the network.

5.2 Reverse power flow model and impacts

The power flows for a power system with B nodes are described by the set of $2B$ load-flow equations in $2B$ algebraic variables V_i, θ_i :

$$0 = -P_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik}\cos\vartheta_{ik} + b_{ik}\sin\vartheta_{ik}), \forall i \in B \tag{5.1}$$

$$0 = -Q_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik}\sin\vartheta_{ik} + b_{ik}\cos\vartheta_{ik}), \forall i \in B \tag{5.2}$$

where

- P_i is the net injected real power (power generated minus power consumed) at bus i
- Q_i is the net injected reactive power (power generated minus power consumed) at bus i

- v_i is the voltage at bus i
- ϑ_{ik} is the difference in voltage angle between bus i and k
- y_{ik} is the admittance of the line between bus i and k
- g_{ik} is the conductance , or the real part of the admittance of the line between bus i and k
- b_{ik} is the susceptance, or the imaginary part of the admittance of the line between bus i and k

The currents $\begin{bmatrix} I \end{bmatrix}$ at every nodes are given by:

$$\begin{bmatrix} I \end{bmatrix} = \begin{bmatrix} Y_{bus} \end{bmatrix} \times \begin{bmatrix} V \end{bmatrix} \tag{5.3}$$

When P_i is positive (more power is generated than consumed at node i), the current I_i at node i is reversed, meaning the instead of coming into the node and being absorbed by the load, it will flow out off the node and distribute upstream in the system according to the loads or form a *Reverse Power Flow - RPF* condition. The node i that was typically a load node acts like a generation node.

For our analysis, the real (P_i) and reactive (Q_i) net-power at every load node will be known and given, while the currents $\begin{bmatrix} I \end{bmatrix}$ are the unknowns that need to be solved for.

As highlighted in the High-Penetration PV Integration Handbook [28], the reverse power flows (RPFs) that are caused by non-negligible penetration of DERs such as residential solar PV embedded within the distribution network may have significant impacts on existing network's asset and their configuration. Amongst the RPF impacts, this research will consider the following:

1. Overload of assets (ampacity ratings of asset that are exceeded) due to aggregated peak generation values exceeding the aggregated peak load values for which the assets where

originally dimensioned. Fig.5.1 illustrate a situation where an elevated concentration of high capacity residential solar PV can contribute to significant reverse currents (bigger than the peak load current) at the upstream transformer. This situation is also true at the substation level where all the RPF aggregate towards the transmission network.

2. System protection impacts, that are: change in current levels and placement and coordination of protection devices to account for new sources of current that are embedded deep within the distribution network. This include new potential islanding situations within the network that must be prevented. The new current levels and direction (reverse power flows) might also affect the coordination rules and type of protection assets as discussed by Che et al in [95] or more recently in [96, 97]. Fig.5.2 illustrate two types of malfunctions in protective devices (PD) that is caused by RPFs. In Fig.5.2a PD1 might not trip if I_{der} is large enough as it contributes to the fault current but doesn't transit via PD1. In Fig.5.2b, PD 2 might trip if I_{der} that contributes to the fault current is big enough and therefore unnecessarily disconnect feeder 2 from the grid (sympathetic tripping).

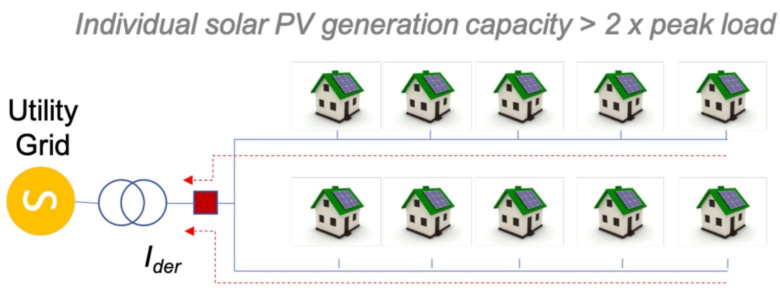
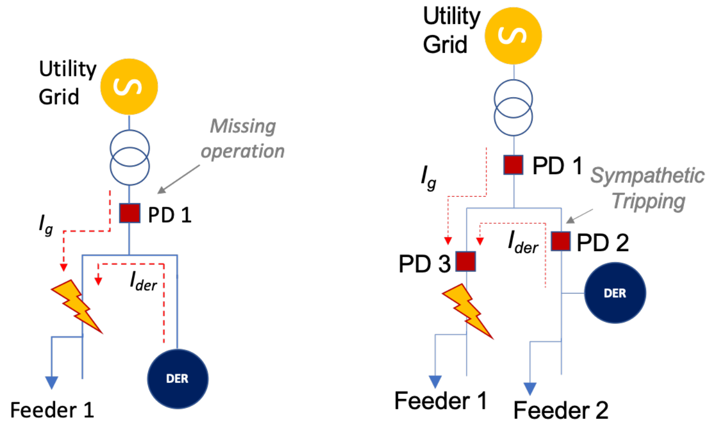


Fig. 5.1 Reverse Power Flows (Source: Author)



(a) Malfunction of protective device (PD)

(b) Sympathetic tripping

Fig. 5.2 Example of DERs impact on fault detection (Source: Author)

The extent of those impacts would depend on the penetration and geographical concentration of installed DERs within the system. Those factors dictate the existence, intensity and probabilities of reverse power flows at any node of the system. Individual solar PV installation decision is originated by the end customer and not strictly controlled by the DNSPs. The extent to which individual solar PV are adopted depends on multiple factors such as the incentive policies in place, the overall convenience or the desire to substitute part of their energy source to renewable and sustainable energy. The adoption of solar PV over time (e.g. how many in total and where will they be located) is not strictly predictable and therefore has elements of long-term stochasticity. Additionally, the individual solar PV generation are intermittent throughout the year/day and therefore also stochastic by nature.

We, therefore, propose to adopt a probabilistic approach to assessing the impacts on the network derived from reverse power flows that are originated by any scenario of increased adoption of DERs such as residential solar PV embedded within the distribution system.

5.3 RPF risk assessment methodology

The intent of the proposed framework is to enable a fast assessment the current's probability (direction and intensity) within distribution network's in response to any scenario of embedded intermittent DERs penetration such as residential solar PV. Similarly to the previous voltage risk framework, to establish a stochastic prediction of the RPFs in response to any scenario of DER penetration within the system, we can:

- Simulate the entire system and solve the load flow equations for each net-loads input.
- Train and use a regression model

5.3.1 Load flow equation solving

Exactly like 4.3.1, a traditional way to predict the current values at every node of a power system is to solve the load flow equation (5.1)(5.2) (5.3) for each input net-load condition, assuming that all the parameters are known. As we already know, there exists no know analytical solution to this problem that must be solved using well established numerical iterative methods (e.g. Gauss-Seidel, Newton-Raphson, Fast-decoupled Method, etc.) [73]. Typically, the rate of convergence of those methods is quadratic, but in[74] it is shown that the convergence region and number of iteration can significantly vary depending on the initialisation and the actual loading condition of the system. Each iteration requires to calculate and inverse the $2n \times 2n$ Jacobian matrix of the system which time complexity is $\sim O((n)^3)$. As soon as the system's size increase, and also depending on the loading condition of the system, solving the system becomes computational intense. It is therefore poorly practical to analyse the effect of vast numbers of scenarios using the brute force of solving the load flow equations for each combination of net-loads input.

5.3.2 Neural Network regression model

For that purpose, we are proposing to use a data-driven method that is based on Artificial Intelligence techniques that will offer a prediction of the currents within the system in response to any net-loads with an appropriate precision at a fraction of the computational cost, and therefore enabling to assess vast amounts of scenarios and derive probability distribution of currents value at any node of the system. In our simulations, we experienced that solving the systems using a trained Deep Neural Network (DNN) model is ~ 1250 times faster than using the established simulation tool OpenDSS [98].

In [60, 61, 56], the authors are experimenting the use of Deep Neural Networks (DNN) to solve the Power flow equations and showing the effectiveness of the method to predict the voltages and power angles mainly on the transmission network. We are proposing to extend the method and train a regression DNN to predict the currents at any node of a distribution network in response to given net-loads using well established supervised learning techniques as described in [64, 65]. The key objective of training DNN models resides in the fact that once the model is trained and performs with adequate accuracy, it is much faster to use compared to solving the entire system.

Training the DNN model using supervised learning techniques requires an adequate training data set where the true value of currents in response to net-loads are known. We are generating such training dataset employing Monte-Carlo simulation as described in Fig.5.3.2 where the inputs are:

- Random allocation of DER within the network with increasing penetration from 0% to 100%
- A standard individual load profile from which the individual loads will be randomly extracted at each Monte-Carlo simulation.

- A standard individual DER generation profile from which the individual generation values will be randomly extracted at each Monte-Carlo simulation.

The Monte-Carlo simulation will generate a vast training dataset of net loads and simulated currents at each node.

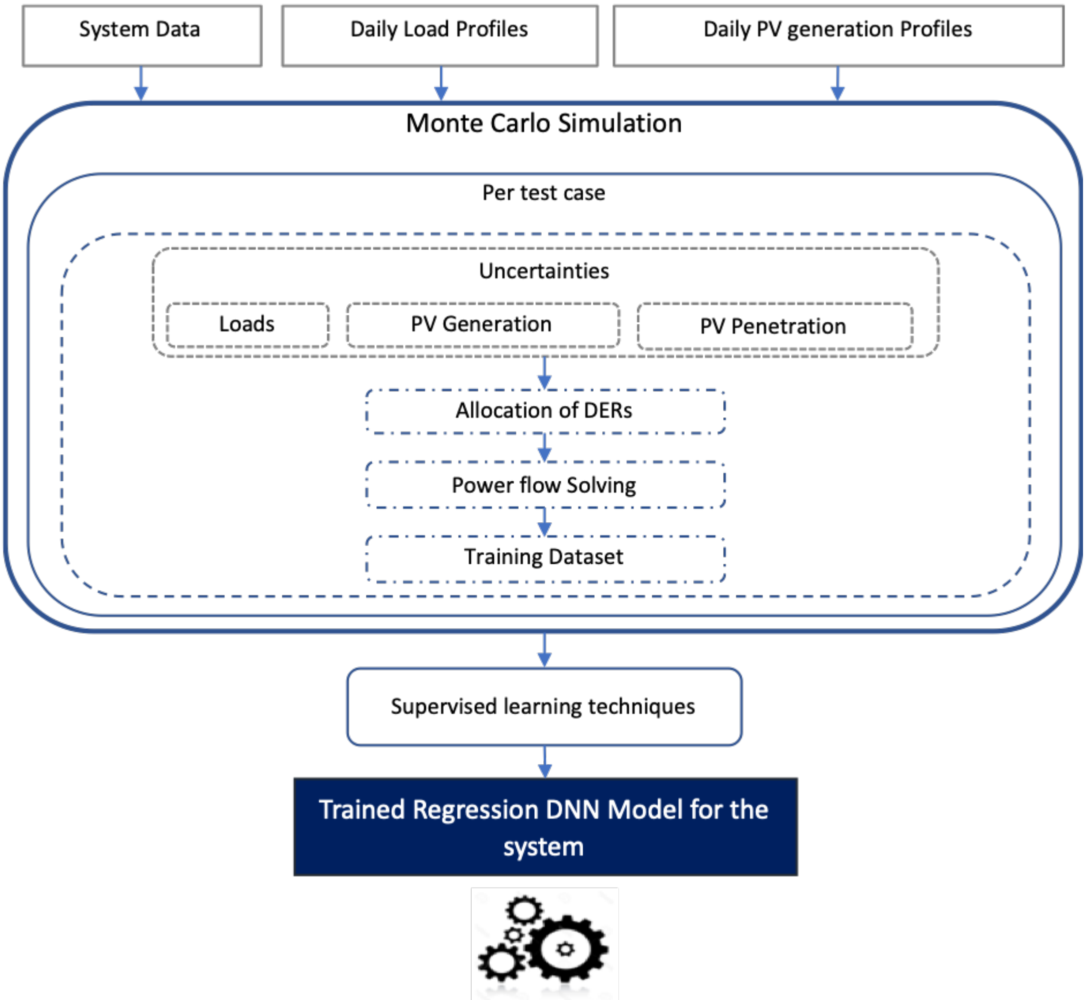


Fig. 5.3 DNN model training framework (Source: Author)

Once the training dataset is created, using the net-loads at each node as input and the respective currents at each nodes as output, we employ well established supervised learning techniques as described in [64, 65] to train a DNN regression model that will predict the current at each node of the system in response to the net-loads at each node.

It is to be noted that the selection of the network configuration and parameters will influence the construction of the training data sets and therefore the DNN regression model. The DNN model will be used for long term probabilistic scenario analysis of DER penetration. Therefore, in first approximation, the most likely network configuration should be used. To account for multiple network's configurations, the presented data set generation framework should be repeated and specific model trained for each configuration. A detailed sensitivity analysis of the models versus different network configuration changes (eg. switch of feeders, compensating devices,..) is left for future work.

5.3.3 RPF scenario analysis framework

The intent of the proposed framework is to enable a fast assessment the current's probability (direction and intensity) within distribution network's in response to any scenario of embedded intermittent DERs penetration such as residential solar PV. The framework identifies the 'hotspot' within the network in terms of location and probabilistic intensity of RPFs. The DNSP will leverage this information to assess the adequacy of their assets to support those reverse power flows on a case by case basis. Those assessments will lead to identifying the most adequate reinforcement and/or protection measures that would maintain security and reliability in the network in response to the increased DER penetration scenarios.

The framework uses vast Monte Carlo simulations and derives probability distribution functions of the current at any node of the system for each penetration scenario, where the currents are predicted by a trained DNN 5.3.2.

The use of a DNN model to predict the currents time series in response to net-loads profiles is instrumental due to the vast amount of simulations required and the limited computational cost of prediction using the DNN model compared to a full solving of the system.Fig.5.3.3 illustrates the end-to-end analysis framework.

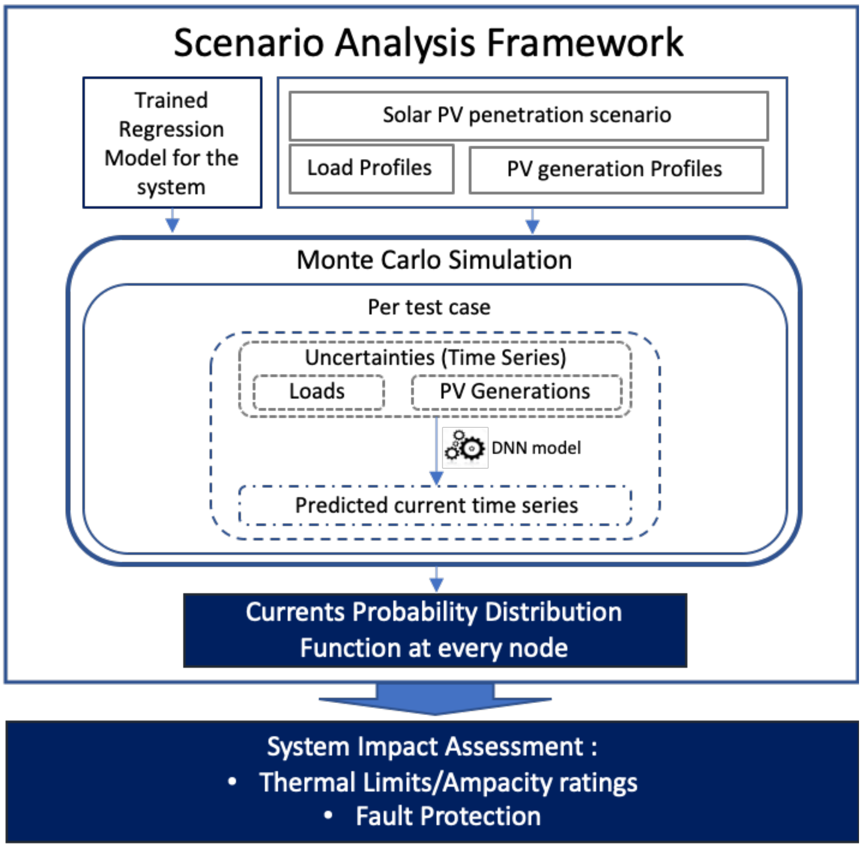


Fig. 5.4 DER's penetration scenario analysis framework (Source: Author)

The inputs to the analysis framework are the scenarios:

- Scenarios of DERs penetration, where each scenario is defined by: number of DER connections, location of connections in the grid and capacity of connected DER
- Expected individual loads time series. This can derived from past data on the network, characterised by an observed average daily profile and standard deviation
- Expected individual generation time series. This can derived from solar irradiation statistics and differentiated by different geographical area of the network
- In case of combined solar PV and storage capacity, the individual loads and generation time series would be adapted to account for the given battery charge-discharge profiles

The outputs of the scenario analysis framework are:

- Time series of predicted currents (intensity and direction) at every node of the power system for each scenario
- Aggregated probability distribution functions (pdf and cdf) of the currents at every node of the system for each scenario

5.4 Case Study - the IEEE123 buses test network

As a baseline, the IEEE123 bus test system (Fig.5.5) and its parameters [5] has been used to illustrate the proposed framework.

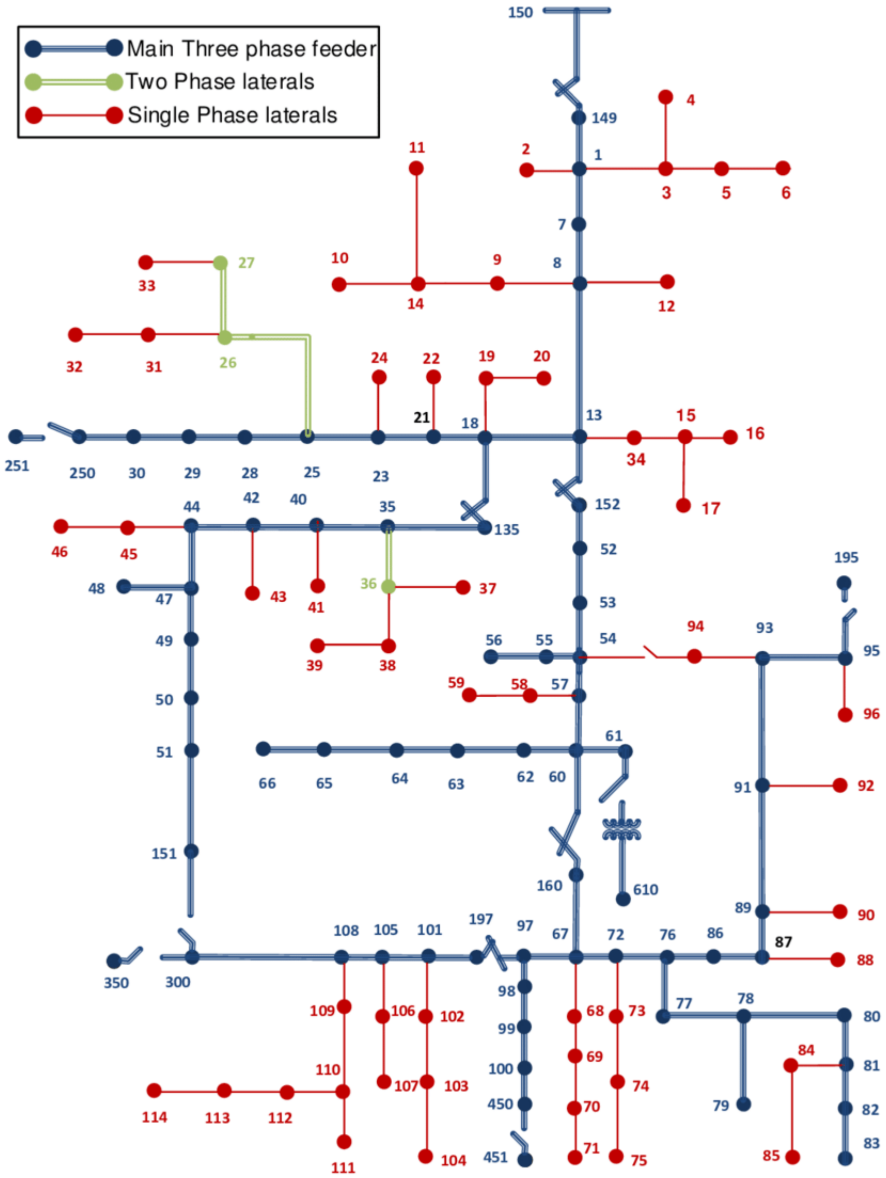
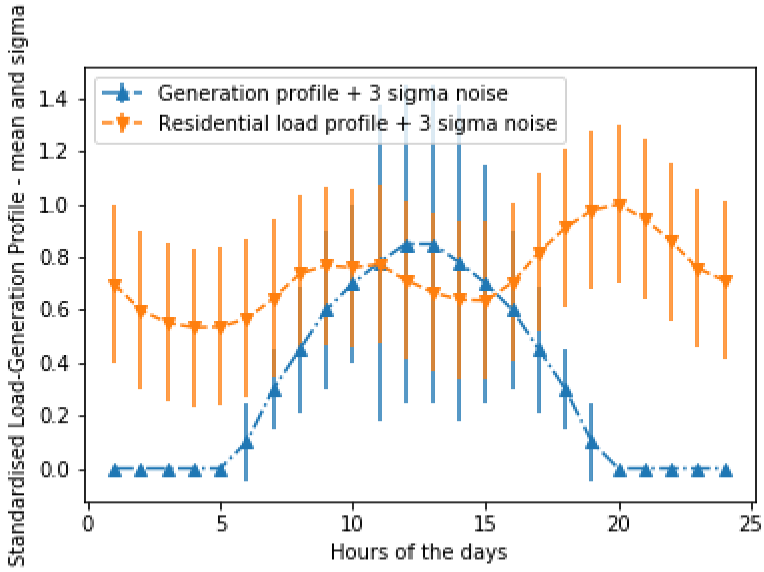


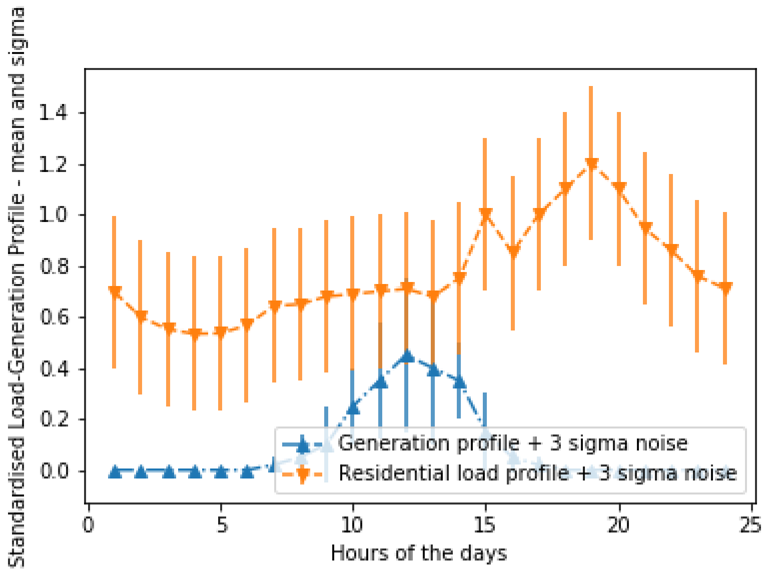
Fig. 5.5 IEEE 123 Bus test distribution network [5]

In addition to the baseline network and load parameters, individual solar PV generation units have been embedded in the system randomly in order to simulate various level of penetration (20%, 40%, 60% and 80%) in the system where $X\%$ represents the number of nodes equipped with Solar PV systems divided to the total number of nodes in the system.

In accordance to recent values observed in Victoria-Australia, the average capacity of the solar PV system installed is set to $5KW$ and the average daily energy residential consumption considered of $24KWh$. Typical hourly load and generation profiles have been considered for our simulations and in order to generate randomness in the loads and solar PV generation and an artificial noise is added at every single point as shown in Fig.5.4.



(a) Typical summer day



(b) Typical winter day

Fig. 5.6 Standardised load and generation daily profiles + noise (Source: Author)

5.4.1 Simulation setup and DNN training

Following the training framework illustrated in Fig.5.3.2, the set of data required to train the DNN model has been generated through Monte Carlo simulation for the entire system in 5 different configurations where :

- Each configuration corresponded to a scenario of solar PV penetration: 0%, 20%, 40%, 60% and 80%. The percentage of penetration corresponding to the percentage of load buses where Solar PV is installed, placed randomly
- For each configuration, 10.000 days have been simulated using the load/generation profiles illustrated in (Fig.5.4) . Per each day, the system has been fully solved hourly between 9am and 8pm, corresponding to 12 full simulations per day
- For every hourly simulation, the loads and generation values at every node has been randomly extracted using a normal distribution with the mean and variance illustrated in (Fig.5.4) capped by zero and the max nominal capacity ($5KW$ per unit) where applicable
- OpenDSS [98] integrated with Python has been used to solve the multi-phase power flow and the voltage, angles and currents of every nodes and hourly simulation have been recorded in a data base.

The vast amount of data set generated ($600k = 5 \times 12 \times 10.000$) has been used to train an DNN regression model that predicts the node's currents (Output) in function of the node's net (generation - load) real and reactive power (Input). ssion ma The trained model structure includes 3 fully connected hidden layers, the input layer has a number of neurons equal 2 times the number of nodes (as both the net active power P and net reactive power Q are input for each node) , the first hidden layer has 4 times the number of nodes, the second has twice the number of nodes and the last hidden layer has as many neurons as the nodes in the system. The hidden layers activation functions are *tanh* and the output is *linear* .

Fig.5.4.1 illustrates the structure of the DNN model for one node, where In_1 and In_2 are the net P and net Q respectively and Out is the predicted current I at the node.

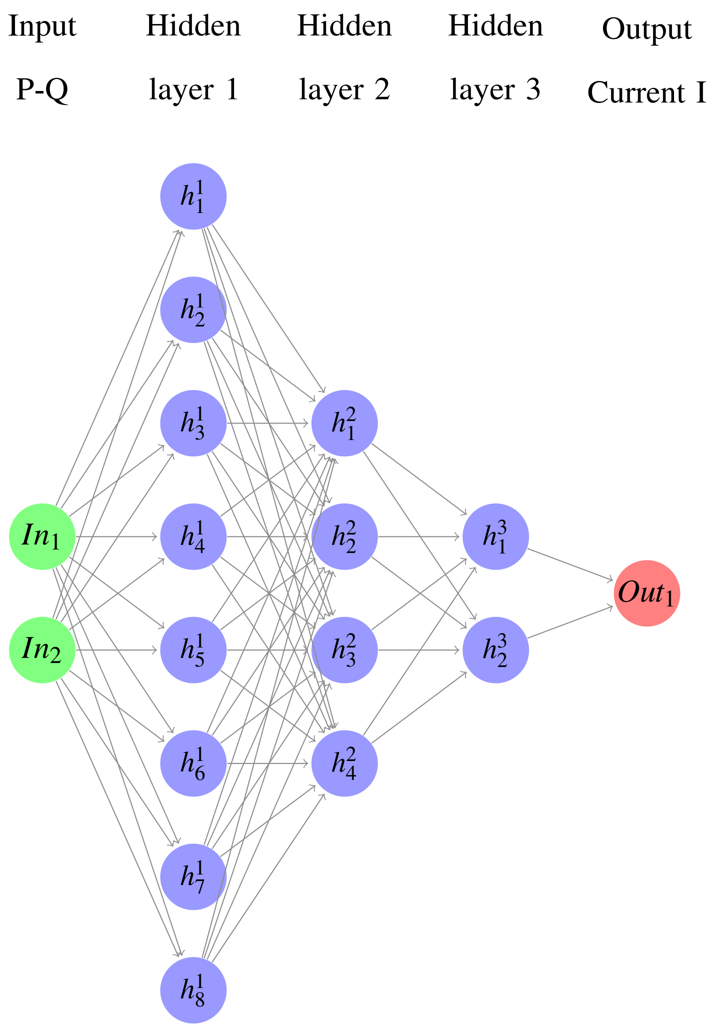


Fig. 5.7 DNN regression model structure per node

Tested on independent validation dataset, and comparing the predicted value with the true value, the Mean Absolute Percentage Error (MAPE) (5.4) of the trained model for all outputs is equal to 5.97%, that is within the 10% tolerance range that is commonly used when evaluating the rating current of assets.

$$MAPE = \frac{100}{n}\sum_{0}^{n} \left| \frac{predicted - true}{true} \right| \times 100\% \tag{5.4}$$

5.4.2 Reverse Power Flows Analysis

To appreciate and illustrate the effect of simulated increased solar PV penetration in the system, we will firstly use the 'true' data that was generated solving the system with OpendDSS. Fig.5.4.2 illustrates the average current for each penetration scenario at the main feeder: phase 1 of node 149 in Fig5.5 on a typical summer day (Fig. 5.6a) and where the individual solar PV capacity is set to $5kW$ with no storage or curtailing involved. Conventionally, the reverse current is positive and load current negative.

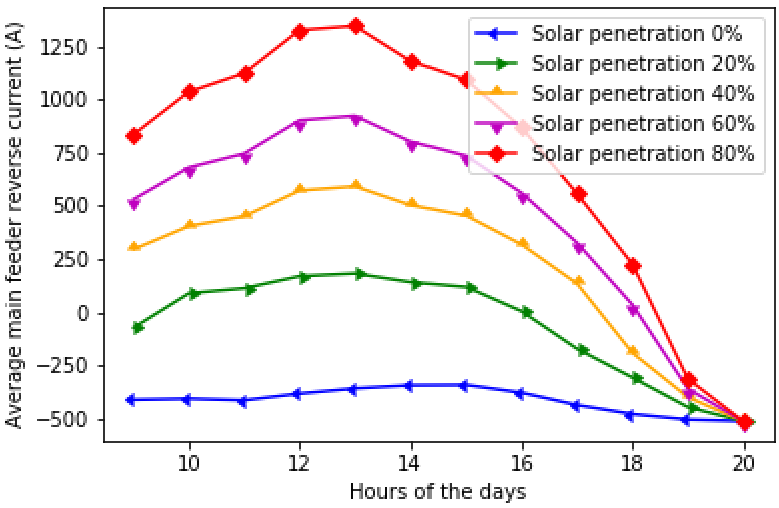


Fig. 5.8 Main feeder phase 1 current - "true" data (Source: Author)

Without solar PV embedded generation, the average load current stands in the range ~ -300 to $500A$, while in an extreme case of 80% solar PV penetration (without storage or curtailing), the reverse current exceeds $+1250A$ at 1pm. Naturally, while the day progresses, the effect of any Solar PV is fading and the currents in each scenario are converging to the same values. We denote that not only the direction of the current reverses but its intensity increases by a factor of $\sim 250\%$ during the central hours of the day.

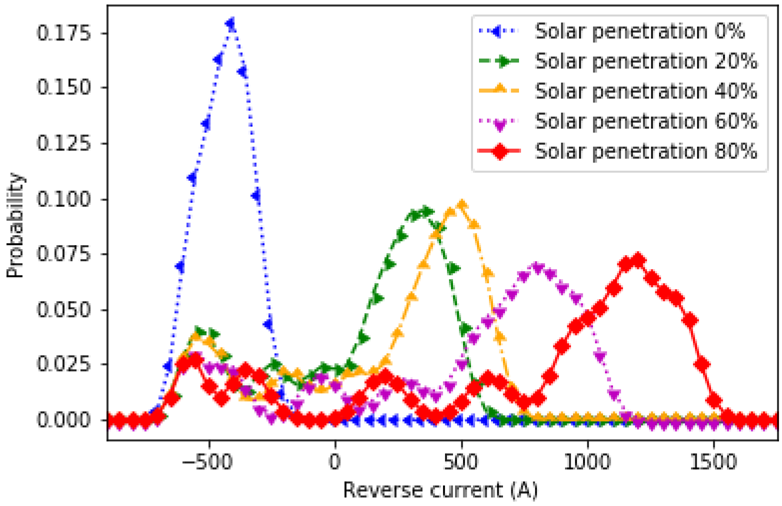
This illustrates the potential intensity of reverse power flows within the power system due to high penetration of solar PV and justify the need for an analysis framework that can help the DNSPs to assess impacts and make mitigation decisions to protect the network regarding fast increasing residential solar PV system penetration.

The average system size in Australia peaked to 7.95kW in 2019 (Solar Report Sept 2019 - Energy Council Australia [99]) , and therefore with penetration of rooftop Solar PV above 60 % the peak reverse power flow can exceed the peak load (which is estimated close to 5kW per household) for which the grid was originally designed. Similarly, in that case voltage over-limits will occur. Both the RPF and Voltage Risk framework must be used to analyse their respective impact (RPF, Voltage) on each node of the network in response to future Solar PV penetration.

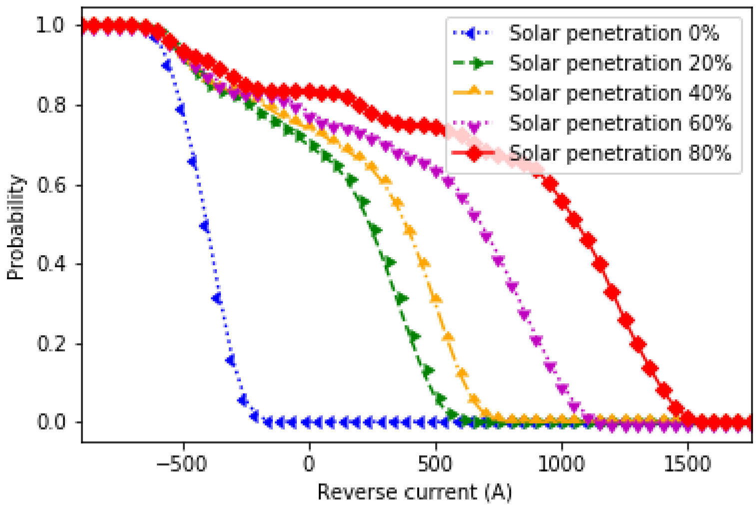
5.4.3 Scenario Analysis

While OpenDSS is certainly a reliable way to simulate and solve power flows, it also comes with a computing cost that restricts the number of scenarios that can be assessed. For example, on a MacBook Pro it took more than 14 hours to simulate the 600k scenarios described in 5.4.1. We are therefore proposing a framework where the previously trained DNN model described in 5.4.1 is used to predict the currents at each node in response to vast amounts of net-loads scenarios. Using the DNN model is a good compromise between precision and speed of simulation, while the average model accuracy is close to 6%, the speed to predict 600k on the same MacBook Pro is 40 seconds, or 1250 times faster than system solving. Applying the framework illustrated in 5.3.3, a vast Monte Carlo simulation that randomly extracts the loads from a typical summer day (hourly from 9am to 8pm) for 100k days in

4 distinct solar PV penetration scenario has been performed and the currents at every node predicted using the DNN model. We have defined current buckets, ranging from $-850A$ to $+1750A$ by step of $50A$ and have assigned each predicted current in their respective bucket. Leveraging the law of big numbers, we derived probabilities from the frequency of current being in buckets for each scenario. Fig.5.9a shows the probability distribution of the current in the main feeder for the 4 scenarios, and Fig.5.9b shows the cumulative probability distribution of the current in phase 1 of the main feeder in the 4 scenarios.



(a) Probability distribution function (pdf)



(b) Cumulative probability function (cpf)

Fig. 5.9 Probabilities of current in phase 1 of the main feeder - DNN predicted data between 9am and 8pm (Source: Author)

Remarkably, the current prediction made with the DNN show good sensitivity to the solar PV scenarios and consistency with the "true" data simulated (Fig.5.4.2). It also demonstrates the progressive increase of reverse current that is most likely to happen with increased penetration of solar PV within the system. From a DNSP point of view, this node by node probability analysis is a piece of important information when associated with the thermal limits of the network asset installed at that node (e.g. Line, protection device, transformer, etc.) and the threshold current that is can safely withstand. The combination of those two information can enable to identify and map the assets within the network that are most likely to be utilised at full thermal capacity (or above) in any scenario of increased solar PV penetration.

5.4.4 Impact Analysis - Ampacity Rating

In order to support planning decision making, the current rating of each asset must be mapped with the cumulative current probability distribution of the node where the asset is located. A systematic construction and analysis of such curves for each nodes would help the system planner to assess the adequacy of each asset rating to sustain a future increase of Solar PV. It would help to identify the weak spots in the grid and plan for remedial before the solar PV penetration reach critical levels.

As an example, if the transformer located at node 149 is rated at $500A$, we know that with 80%, 60% or 40% solar penetration, there is respectively approx 75%, 60% or 30% of chance that the rating will be breached at any time between 9am and 20pm, but no risk if penetration is kept below 20% as Fig.5.4.4 shows.

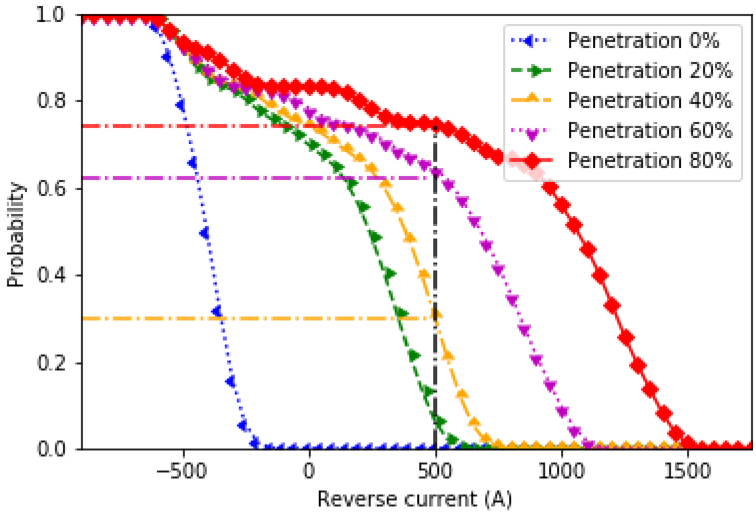
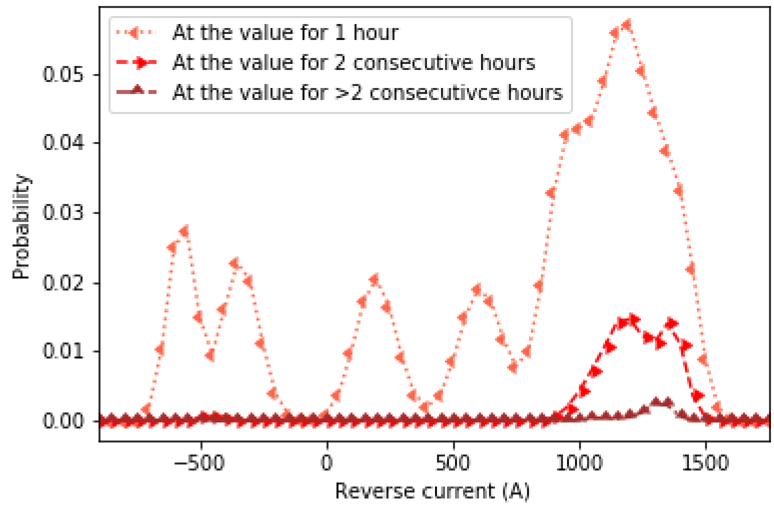


Fig. 5.10 Control asset rating with current probabilities (Source: Author)

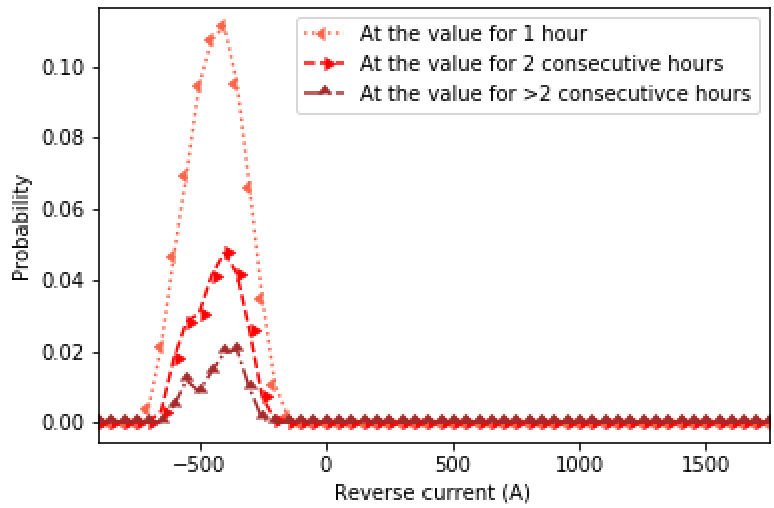
We see how the analysis framework enables to assess every asset located in the network against scenarios of increased penetration of intermittent DER within the system. By establishing some risk acceptance thresholds (max risk probability), the network operators can swiftly assess the rating adequacy of their asset pool and/or prioritise the assets that must be upgraded in response to an increased penetration of intermittent DER (eg. solar PV).

5.4.5 Impact Analysis - Protection device

In addition, we are introducing the time dimension in the analysis and deriving for each current bucket the probability of being in that bucket for one time period (1 hour), 2 consecutive time periods or more than 2 consecutive time periods. (Fig.5.4.5).



(a) 80% Solar PV penetration



(b) 00% Solar PV penetration

Fig. 5.11 Main Feeder - Probability distribution functions of being at current value for consecutive time period between 9am and 8pm (Source: Author)

Even though for most of the times the predicted peak current would last only for 1 time period (1 hour in our simulations), we observe that the probability of having a sustained high current for more than one time period is not insignificant under our test conditions. This information is valuable when mapped to the characteristic time-current curve of protection devices embedded in the network such as circuit-breakers or transformers thermal protection relays. Fig.5.4.5 illustrates a typical tripping curve of protection devices such as circuit breakers that connects the time to trip to the amplitude of current relative to the current rating of the device.

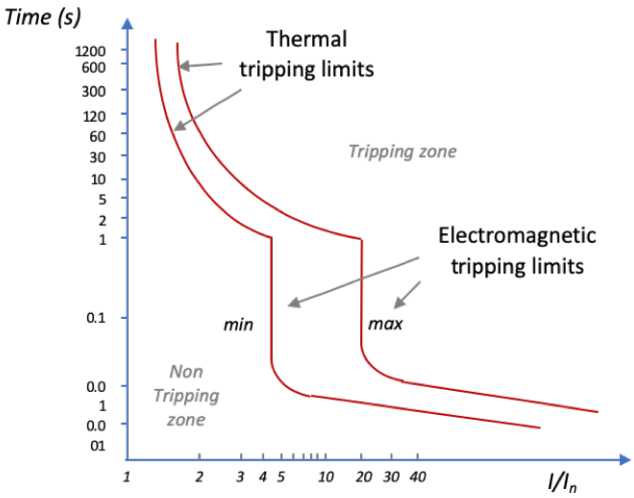


Fig. 5.12 Typical protection device tripping curve (Source: Author)

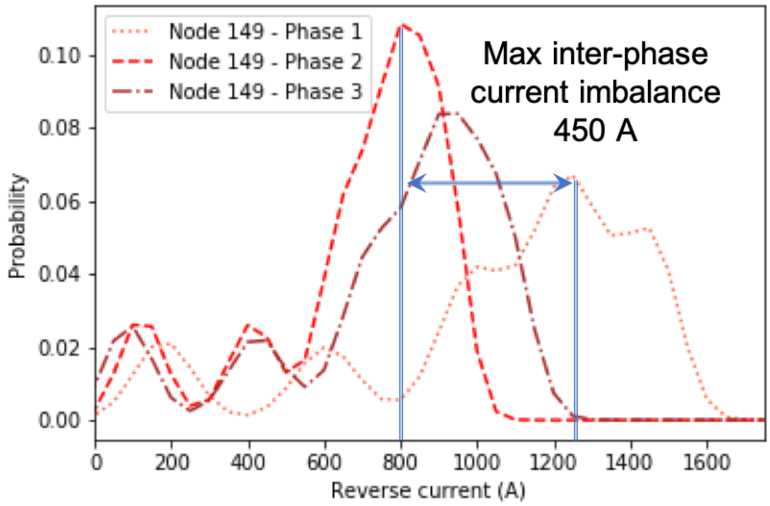
Therefore, for each protection device that is placed in the network or embedded in assets such as transformers, the probabilities of sustained current illustrated in Fig.5.4.5 and obtained as an output of the RPF scenario analysis framework are valuable to assess the risk of thermal tripping due to sustained current that is multiple of the device current rating value.

Following our previous example, if the transformer located at node 149 is rated at 500A, using the probabilities in Fig.5.11a, we see that the probability of a current that is 250% of the transformer rating for at least 2 consecutive hours is not negligible and might trip its own protection in a scenario of 80% solar PV penetration.

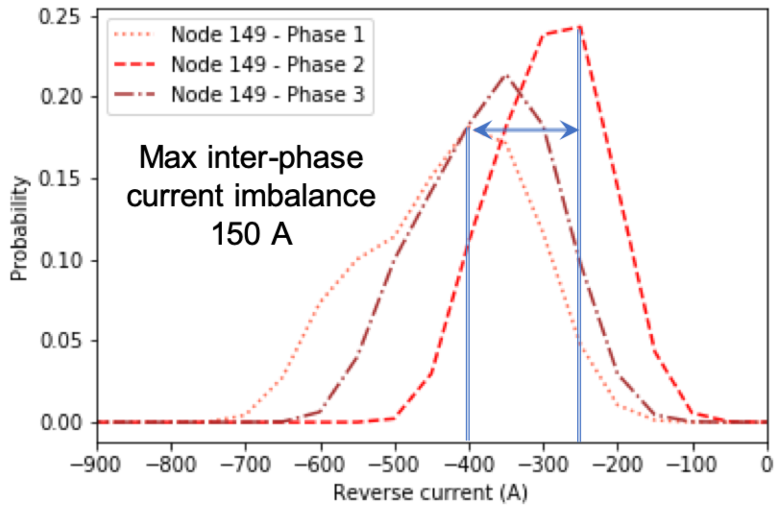
We see how the analysis framework enables to assess every asset located in the network against scenarios of increased penetration of intermittent DER within the system. By establishing some risk acceptance thresholds (max risk probability), the network operators can swiftly assess the rating adequacy of their asset pool and/or prioritise the assets that must be upgraded in response to an increased penetration of intermittent DER (eg. solar PV).

5.4.6 Impact Analysis - Phase imbalance

Additionally, the analysis framework enables to analyse the effect of increased penetration of solar PV on the overall balance of the 3-phase system. As shown in Fig.5.4.6, for the configuration used in our simulations, we denote that the unbalance of currents between the 3 phases of the main feeder is exacerbated by high penetration of solar PV.



(a) 80% Solar PV penetration



(b) 0% Solar PV penetration

Fig. 5.13 Main Feeder - Probability distribution functions of currents in each phases of the main feeder (Source: Author)

Fig.5.13b shows that without any solar PV within the system, the unbalance gap between the 3 phases is contained and mainly due to the configuration of the IEEE 123 Test case which is slightly unbalanced. Fig.5.13a shows that for an elevated penetration of solar PV in the system (80% in that case), the gap between the phases widens significantly. This demonstrates the versatility of the analysis framework that enables to assess granular DERs penetration scenarios and identify precisely the impacts of the resulting RPF on installed assets on the network.

5.4.7 Storage and curtailment considerations

The DNN model is generalised and enables to predict the current within the system in response to any load and generation profiles. In our simulations, for illustration purpose, we have derived the RPF probabilities from a vast Monte Carlo simulation using a typical summer net-load profile with Solar PV and without storage or curtailment capability. To simulate the effect of storage or curtailment rules, an adapted net-load profile that considers the charge and discharge cycle of the batteries would need to be used in the Monte Carlo simulation, but the core DNN model would remain identical.

5.5 Cost of DNN model training discussion

The key advantage to use DNN regression model to assess the RPF risk is the significant decrease in computational complexity and time versus the traditional solving of the load-flow equations. But the use of DNN comes at the cost of training the models and generating the respective training data sets. This cost is a 'one-off' cost for each system configuration, and once the models are trained they can be used on vast (10^6) of DER penetration scenarios. The size of the training data sets influences the accuracy of the trained model. Intuitively, the bigger the size, the higher is the accuracy, up to a certain point after which it becomes asymptotic. In our simulation (4.6.1), we have tested four training data set dimension as indicated in Tab.5.1

Training data set size	Mean Absolute Percentage Error
6×10^5	5.97%
1×10^5	7.82%
6×10^4	8.57%
2×10^4	10.71%

Table 5.1 Model accuracy versus training set size (Source: Author)

This illustrates that even by reducing the training data set dimension by an order of magnitude, the tolerance remains close to the limit of 10% that was arbitrarily set. In that case, the cost of creating the training data sets is 2 orders of magnitude lower than the equivalent cost of using traditional simulation techniques to execute the vast Monte-Carlo scenario analysis.

5.6 Conclusion

Recent data from the Clean Energy Regulator (CER) shows that over 8 GW of rooftop PV capacity is installed in Australia as of 2019 corresponding to slightly more than 20% of the dwellings, and fast-growing as 1.55 GW has been added only in 2018. The AEMO (Australian Energy Market Operator) in its "Integrated System Plan - July 2018" [100] forecasts that the National Energy Market (NEM) will host between 21 GW and 56 GW of rooftop PV capacity by 2040. Clearly, the solar PV penetration is increasing fast and reaching non-negligible amounts, and therefore DNSPs must assess the impact of future penetration scenarios on their network.

The presented framework provides a versatile and powerful methodology to predict the intensity and probabilities of reverse power flows in any given scenario of intermittent DER penetration embedded within the distribution network. The key strengths of the proposed framework are the speed and versatility to analyse the impact of any penetration and net-load profile in the system. The use of a trained DNN model substitute the requirement to solve the load flows equations for the system and enables to assess a vast number of scenarios.

The resulting cumulative probability of current intensity at every node is instrumental to assess the impact on installed assets or configurations within the network for given solar PV penetration scenarios and plan for adequate interventions that would maintain the security and reliability of the power network in those scenarios.

The framework relies on the training of DNN models and future work would investigate the use of Transfer Learning methods to adapt the current models to predict different configurations of the power network and its internal parameters (such as impedances).

Part II

Extreme weather events risk and energy community resilience method in Bangladesh

Chapter 6

Community and power system resilience assessment method

6.1 Background and context

Bangladesh is ranked sixth for extreme weather events in the last 20 years (Global climate index, 2016), fifth-most vulnerable among 170 countries [101] and sixth most flood and erosion-prone country in the world [102]. It is considered as a disaster-prone country due to flood, riverbank erosion, heavy rainfall, cyclone, storm and tidal surge, etc. The benefits from continued investments in basic infrastructure like improving access to electricity could outweigh the climate-related loss in living standards [103]. Similarly, technological advances, coupled with expanded irrigation systems, work to make agriculture less sensitive to climate change in the long-term [104]. Climate change poses a great challenge to society with urgent responses needed that contribute to building and strengthening energy and community resilience.

Rising average temperatures can affect living standards through negative impacts on agricultural and labour productivity, health, migration, and other factors that affect economic growth and poverty reduction. It can decrease agricultural productivity, leading to a decline in living standards for agriculture-dependent households. A warmer climate can also increase the propagation of vector-borne and other infectious diseases. About 9.9% of female-headed households will be directly affected by climate change [103] . Supplying clean and modern energy services to the rural community and disaster-prone areas can mitigate vulnerability. The economic growth of Bangladesh demands more energy. However, the economic enhancement of Bangladesh has been held back due to climate change impacts and disasters and extreme weather events.

In Bangladesh, more than 88 million people live in low-lying flood-prone areas. However, 112 million people have access to early warning and disaster-related information. UNDP Bangladesh has been working with the government and development partners to reduce risk and promote resilience. The United Nations Office for the Coordination of Humanitarian Affairs played an important role increasing preparedness and it saved numerous lives during the cyclone 'Mahasen' in 2013 [102].

Local initiatives for disaster resilience plan and approach were very substantial, and community-led risk reduction initiatives assisted to reduce vulnerabilities to climate change impact and disaster. UNDP targeted the poorest communities; vulnerable women, including female-headed households, widows, and other excluded groups such as transgender people, minorities and indigenous communities.

Broadly, the resilience is defined as: “The ability of a system, community or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions” [105].

In adaptation to extreme weather events, that are most likely to worsen in frequency and intensity of the next period due to Climate Change, we need to consider two main dimensions of resilience that are connected to power supply.

1. The first one is associated with the people and community who have an intrinsic capacity to adapt and react to disasters broadly and more specifically in relation to power supply and usage.

2. The second one is the technical resilience of the power grid (main grid and mini-grid or individual power systems) when faced with destructive power of extreme weather events.

We are introducing a method that can encompass and aggregate both dimensions in quantifiable terms to measure an aggregated community and system resilience that is linked to the power system planning.

6.2 Risk-based approach to resilience: the greater the resilience, the less the risk impact

6.2.1 The approach

In an attempt to capture quantifiable measures to characterize the energy community and power grid resilience, we adopt a risk-based approach and inherit from this approach some methods that enable to aggregate and incorporate various measures, or risk factors, into broader risk classes. Additionally, a linkage method between those classes and the system planning framework is explored.

First, we link the resilience to extreme weather events to the impact that such event is causing to communities and power system assets in an inverse relationship. The greater the resilience, the less the risk impact and vice-versa. Second, we propose a method to quantify the energy risk impact of extreme weather events based on selected metrics assessments and define a precise mathematical formulation for the energy risk impact.

6.2.2 The energy risk impact

Definition A *Extreme Weather Risk* is defined as shock or stress due to extreme weather event.

In our specific case, the risk event is the loss of electricity supply due to extreme weather event. For each extreme weather event considered, there exists a spatial probability distribution of that event happening. Commonly this information is collected through spatial risk maps or "heat maps" and connected to a pre-defined time horizon that we suggest being 1 year. For example: What are the chances of a Category-5 Cyclone to hit this territory in the next year? What are the chances of a Summer Storm to hit this territory in the next year?

Definition *Vulnerability* is the propensity to be adversely affected by a risk event. It is represented by a number $\in 0, 1$

In our specific case, we will consider two interrelated components: the first is the propensity of the community to be affected, the second is the propensity of the power system assets to be affected. In each case the vulnerability is specific to the community and or the assets implied to provide the energy. A more adaptable and resourceful community will be less vulnerable and similarly bigger, or smarter or stronger power grid will be less vulnerable. This measure, like the probability of the risk event is distributed spatially.

Definition *Exposure* is the amount of power Energy (kWh) that is Not Supplied (ENS) to the communities, and that can't be supplied from other sources of energy. This depends on the size of the community, their average daily usage of energy from the power system and the elapsed time of supply interruption.

Definition *Energy Risk Impact (ERI)* is a measure of the power energy at risk from extreme weather event across a territory and for a pre-defined time-horizon.

The mathematical formulation of the power ERI is:

$$ERI = \sum_W \int_S (P_w(s) \cdot Vulnerability_w(s) \cdot Exposure_w(s))\, ds \tag{6.1}$$

where

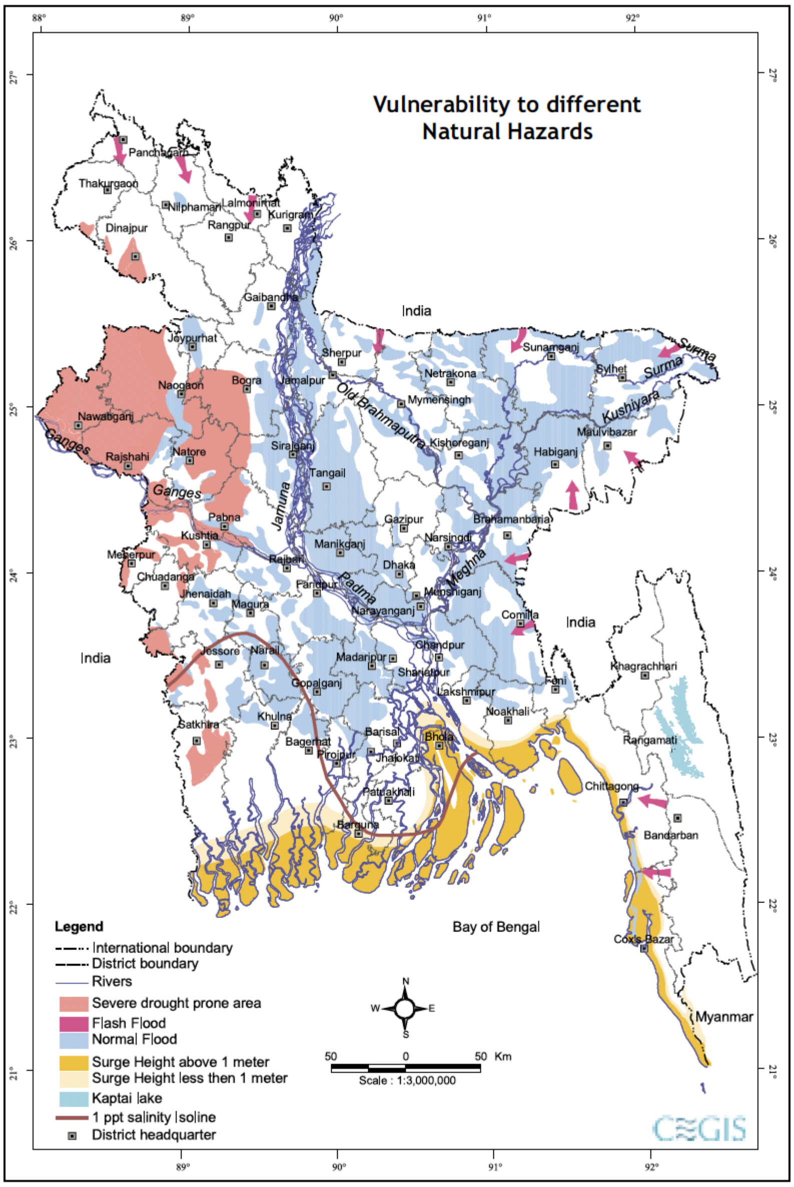
- S is territory or geographical area considered
- w is the weather event type, and $W = \{Cyclone, Flood, SummerStorms, etc\}$ the set off all events considered
- $P_w(s)$ is probability distribution of the extreme weather event w happening over the territory s and over the predefined time-horizon (eg. next year, next 10 years,..)
- $Vulnerability_w(s)$ is an indicator $\in 0, 1$ that quantifies the combined community and power system vulnerability to weather event w across the territory s. Section 6.3 illustrates in details the proposed method to assess this indicator.
- $Exposure_w(s)$ is the total power energy in kWh that would be not supplied in the territory s if a weather event w happens.

Under that definition, the risk impact is a quantifiable measure that takes into consideration and aggregates the local specificities of communities, power system assets and landscape. The granularity to consider estimating the risk impact will depend on the availability of data. The more granular, the more refined will be the spatial risk impact, but even with low granularity the proposed methodologies enables important regional differentiation.

6.2.3 Extreme weather risk probabilities

Dastagir et al in [106] explore the modelling of climate change induced extreme events in Bangladesh and the Ministry of Environment and Forests of the Government of the People's Republic of Bangladesh in [6] has built an action plan to address the potential impacts of climate change and consequent extreme weather events.

Illustratively, Fig.6.2.3 shows the geographic distribution of zones that are potentially affected by climate-related disasters.More specifically, Fig.6.2.3 and Fig,6.2.3 illustrate respectively the risks of Cyclones and Floods.



Source: CEGIS, Dhaka.

Fig. 6.1 Areas affected by different types of climate-related disaster of Bangladesh [6]

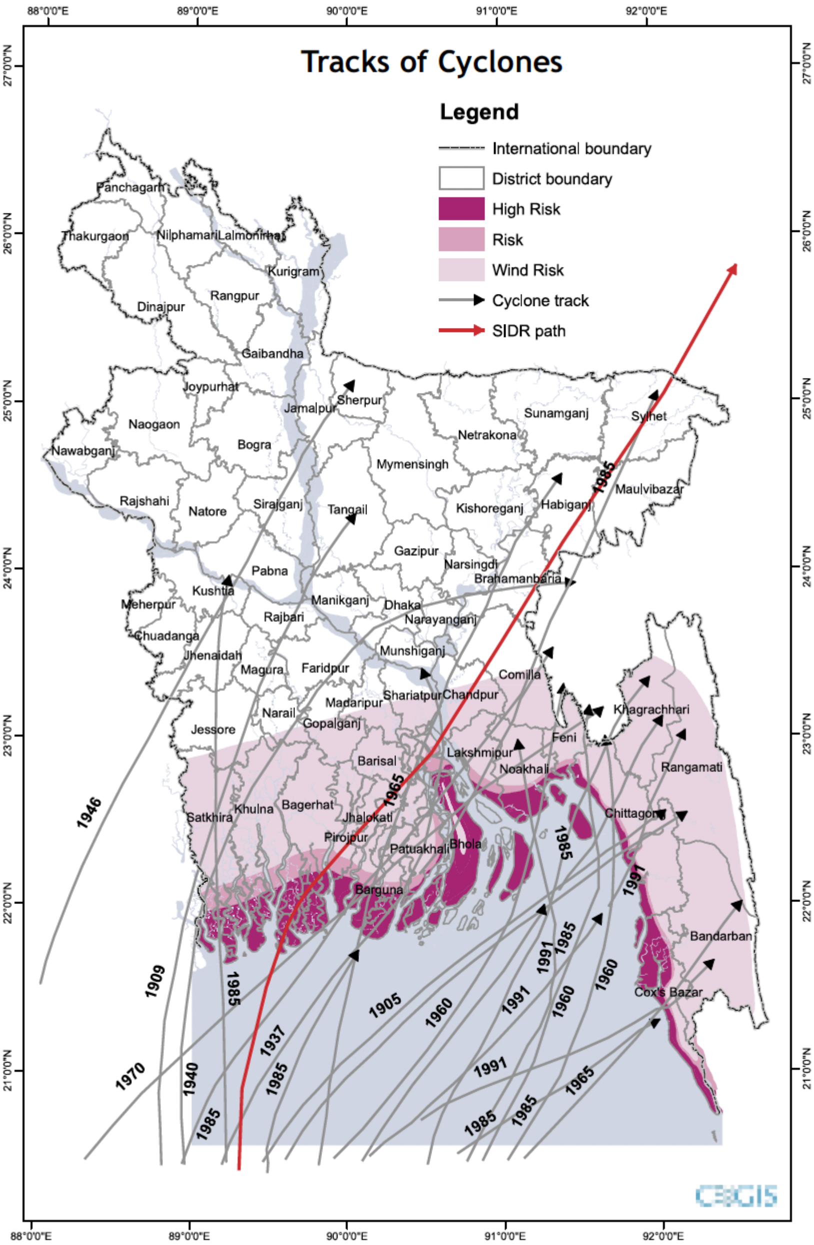


Fig. 6.2 Track of cyclones over the last 50 years [6]

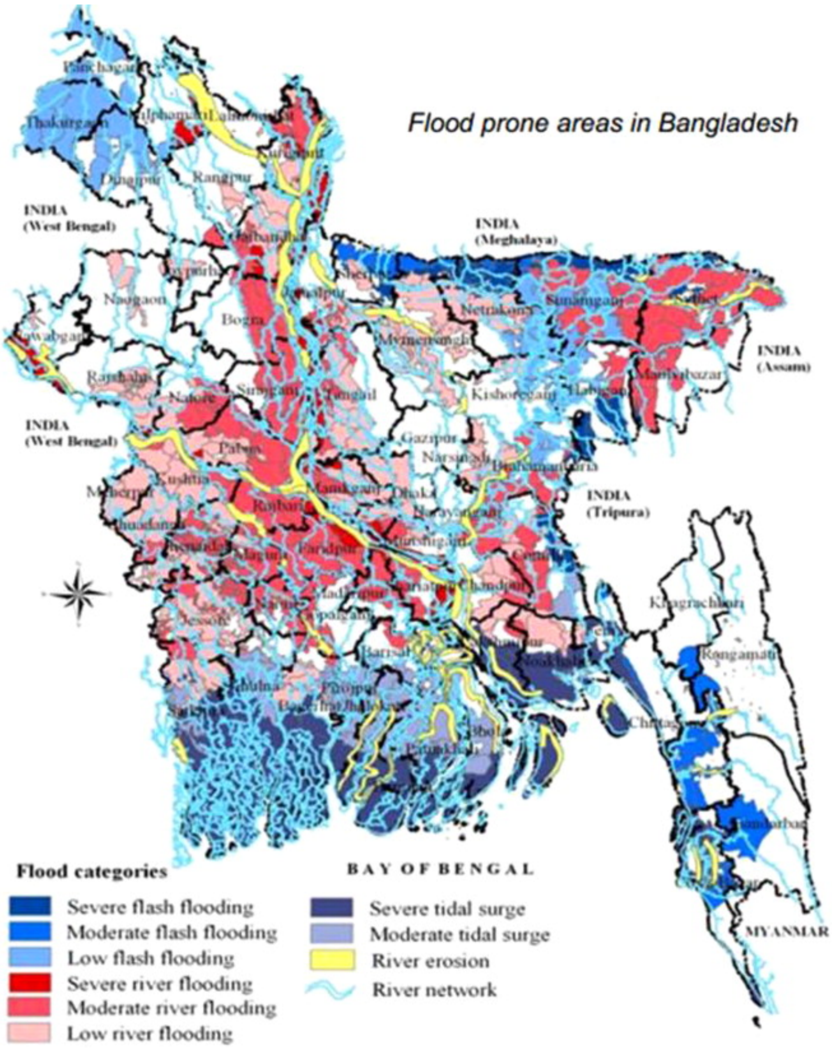


Fig. 6.3 Flood-prone areas in Bangladesh [6]

The use of past data, combined with climate modelling can extrapolate probabilities of any extreme weather events happening in a given time window with some level of confidence. The result of such modelling can then be used as the $P_w(s)$ or probability distribution of the extreme weather event in (6.1).

6.3 A method to assess and quantify the community and system vulnerability

The main lever at disposition to minimize the risk impact (or optimize the resilience) stands in the vulnerability component $Vulnerability_w(s)$ of the equation (6.1). External forces dictate the probability of extreme weather events happening, but planners can build smart or strong or big networks and the communities can develop adaptative skills and strategies that allow them to prevent, adapt or recover quickly and wisely from power supply interruption.

In this section we propose a methodology to assess and quantify the overall vulnerability for a given geographic area s, in three steps:

1. use a multi-criteria approach to assess a community vulnerability and assign a respective 'community vulnerability' class. The communities assessed are the communities residing in the given geographic area s.
2. use an attribute-based approach to assess a power system vulnerability and assign a respective 'system vulnerability' class. The power grid assessed is the grid supplying the given geographic area s in electricity
3. use a commonly used technique in risk management to combine both vulnerability classes in one indicator using a scoring matrix.

6.3.1 Community vulnerability assessment

Background

There are a few key socio-economic demographics that can map the correlation between access to electricity and increased well-being/empowerment of people in the community. The most obvious one that shows a direct relation is increased per capita income. Increased access to electricity translates into more opportunities for income-generating activities, using the available electricity and alternative energy resources to generate higher incomes. This can also mean more employment opportunities as more businesses and organizations can now run efficiently and offer people jobs. Increase in hours of employment/income generation activity can also map how access to electricity creates better employment opportunities and increase women's labour force participation. It is also important to track how access to electricity and energy resources changes the employment/income generation opportunities of vulnerable groups, to identify whether access results in increased formal job market participation, increased ownership of small and medium businesses, increased farming versus non-farming activities in the community, etc. The analysis thus needs to also map the increase in SME ownership for women and vulnerable groups. This can also be tracked by looking at increased access to mobile financial services and banking. One key indicator for development is an increase in the use of mobile phones, radios and televisions that directly translates into increased access to information. This is especially important in coping with natural disasters where early warning systems can disseminate messages effectively and efficiently to all in a community including the vulnerable and hard to reach people, and also in coping after the disaster where information about relief and post-disaster rehabilitation is crucial.

There is a strong link between education and reducing vulnerability at the individual and community level. In the past, disaster management was mainly focused on post-disaster management such as relief, rehabilitation and reconstruction, but now the concept has been changed, and the government and other development agencies are focusing on pre-disaster management including planning, preparedness activities, organizational planning, training, information management, public relations and awareness raising. Disaster education is basically mutual interaction among people and institutions, and it improves the societies regarding disaster resilience and creates awareness and increase knowledge and skills among individuals and at the community level [107]. It includes more than formal education at educational institutes of different levels, and recognition of traditional knowledge or indigenous knowledge. Furthermore, education also encompasses information technology, staff training, established learning arrangements, experience sharing, electronic and print media and facilitating the sharing of information, skills and knowledge with other stakeholders (e.g. professionals, non-professionals, organizations and policymakers and community people).

The poor and socially excluded are particularly vulnerable to climate variability and stresses [108]. Factors such as ethnicity, religion, caste status, and profession are common root causes for social marginalization [109], leading to the exclusion of poor women, female-headed households (including those headed by widows and divorced women), indigenous communities, sex workers, transgender people, religious minorities, low caste people (Dalit and Methor), the disabled, etc. In particular, these socially excluded groups are characterized by a lower access to energy deriving from social norms and traditions including hierarchical values; inadequate attention to the needs of women and other socially excluded groups; and lack of understanding of social inequalities. Indigenous communities also have very limited access to electricity due to the remoteness of most settlements and lack of information on existing technologies and financial support available from different agencies for means to acquire them. The cost of energy services and inability of poor and marginalized groups to pay is another factor affecting energy access.

Successive policies have outlined the government's commitment to disaster management, including support for vulnerable groups. The government designed the Bangladesh Climate Change Strategy and Action Plan 2009, defined a National Plan for Disaster Management for 2010-2015, and set up a dedicated Department of Disaster Management (DDM) under the Ministry of Disaster Management in 2012. Bangladesh has disaster management committees from union to national levels. The Disaster Management Act of 2012 provides preferential support to ultra-poor and underprivileged communities such as tribal groups, small ethnic groups and anthropological communities that are deprived of socio-economic and other facilities, especially older persons, women, children, and people with disabilities.

The Government of Bangladesh and UN agencies including UNDP, WFP, and other national and international NGOs are implementing the Cyclone Preparedness Programs (CPP) for early warning system development and the National Resilience Program to:

- improve national-level capacity for risk-informed, gender-responsive and disability inclusive development planning;
- strengthen national capacity to address disasters, including those of exceptional magnitude due to climate change, in a gender-responsive and disability-inclusive manner;
- improve the capacity of selected public institutions to achieve resilient outcomes through risk-informed, gender-responsive infrastructure systems;
- enhance women's leadership capacity for gender-responsive national and local disaster management decision-making, investment and policy;
- strengthen community-level preparedness, response and recovery capacity for disasters, including those of exceptional magnitude due to climate change.

Bangladesh has disaster management committees from union to national levels. Each committee has standing orders during disasters to coordinate and support government and other humanitarian agencies.

Assessment method

Leveraging on the framework developed by Nerini et al [110] and adapted by Mohideen [111], we propose to develop a multi-criteria approach to assess the vulnerability of communities. The vulnerability of communities regarding access to electricity supply will depend on multiple criteria's :

- Socio-economic resourcefulness of households and communities
- Availability of early warning systems and shelters equipped with standalone power generation capability
- Access to information and communication -such as radio, TV, mobile phones, and the internet
- Availability of local skilled pool of labour to protect based on early warning, and repair and restore the power system autonomously
- Availability of alternative source of energy that can meet the basic life requirements and basic business/farming requirements while the original power supply is being restored
- Access to services—government, private, NGO, and other
- Access to markets
- Governance, that is the capacity of the local government institutions to deliver necessary information and services;

Tab.6.1 indicates in summary the key measurable indicators and metrics interlinked to the above criteria.

Table 6.1 Community vulnerability criteria (Source: Author)

Main criteria	Sub-criteria
Income (I)	Income, Land ownership, Productive Assets ownerships, etc.
Knowledge (K)	Education, Technical skills, etc.
ICT (C)	Ownership or access to TV, Radio, Internet, etc.
Access (A)	Protection Infrastructure, Early Warning systems, power generation assets, etc.
Governance (G)	Capacity, Information management, service delivery, etc.

A multi-criteria analysis can be applied to these criteria. The criteria and sub-criteria will be weighted, through a participatory consultation process based on qualitative methodologies, such as focus group discussions and the gathering of stories, with the community and the target groups identified as especially vulnerable. Each criterion (I,K,C,A and G) value will be standardised to numbers $\in [0,1]$. Through this methodology a final community resilience 'composite index' can also be calculated:

$$CVI = w_i \cdot I + w_k \cdot K + w_c \cdot C + w_A \cdot A + w_g \cdot G \tag{6.2}$$

where:

- CVI is the Community Vulnerability Index, and index arbitrarily set to be $\in [0,100]$
- w_x are the weights for each criterion that must be calibrated through collection of community data and regression analysis. The sum of the weight must be equal to 100.

For the purpose of this method, the CVI range is arbitrarily divided into 5 classes and each community ranked against those as illustrated in Fig.6.3.1.

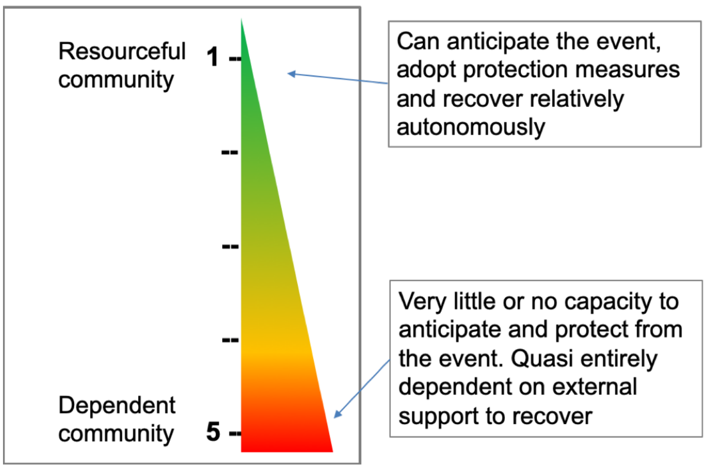


Fig. 6.4 Community vulnerability classes ranking (Source: Author)

The proposed method illustrates the steps and rationale to enable a quantitative classification of community energy resilience to extreme weather events based on sound criteria.

In order to influence the community vulnerability for the most dependent communities, a set of incremental actions to improve the criteras indicated in Tab.6.1can be undertaken. Some actions, such as infrastructure protection, activation of early warning systems, access to information and communication can be designed and implemented in a relative short time frame (provided they are sufficiently funded), but some others (such as education, technical skill) require long term investment in human capital.

6.3.2 Power system vulnerability assessment

Pantelli et al [7, 67, 68] have presented a conceptual framework of power system resilience where the enhancement measures can conceptually be summarised in: Bigger, Stronger and Smarter as illustrated in Fig.6.3.2

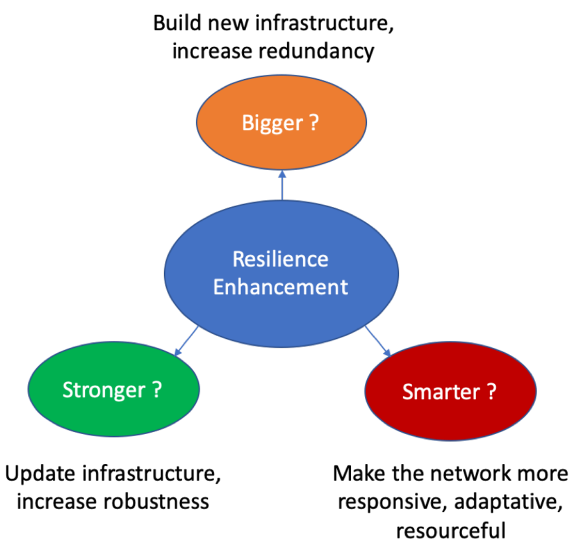


Fig. 6.5 Power system resilience enhancement [7]

’Stronger’ measures the robustness of the system facing extreme weather events. The following attributes can qualify the robustness of the system:

- Presence of underground distribution and transmission lines
- Use of strong and robust material used in equipment construction (e.g. Stronger poles, lines)
- Presence of elevated/protected substations and lines
- Use of distributed energy resources (e.g. Distributed Generation such as Solar PV Systems, Energy Storage Systems)

’Bigger’ measures the level of structural redundancy within the system. A grid with redundant transmission routes to specific location will be measured as “bigger” than if the transmission route is unique. Similarly the presence of distributed back-up generation will be a positive factor of redundancy.

’Smarter’ measures the intrinsic adaptability and controllability of the system. For example, decentralized energy systems that can function autonomously islanded play are intrinsically less vulnerable and more adaptable than the main grid. Similarly, some adaptative wide-area protection and control schemes can minimize the vulnerability of the overall system.

Assessment method

For the purpose of this research, we adopt an ’attribute-based’ approach to measure resilience [112]. Those attributes will be targetted at each extreme weather event type and strength considered (eg. cyclone-category, wind-speed,flood-height, etc.). The first step of the method requires an assessment of the robustness attributes of the existing power system assets in each given geographic area and for each considered weather event type.

The robustness of the power system is usually enhanced by “hardening” the assets so that their fragility to the extreme weather event is decreased. It is assumed that the technical specifications and cost of each “hardened” assets are known to the network operators and adequate to significantly decrease the fragility of the assets to the extreme weather event.

In order to characterize the robustness of the power system assets in response to different extreme weather events, we propose to quantify the proportion of hardened assets that are installed in a given region and expressed in percentage of the total assets in that region. (e.g. km of hardened poles and lines divided by the total length of installed lines in that region, etc.). An illustrative set of attributes diversified by weather event type is proposed in Tab.6.2.

Table 6.2 Power grid assets robustness attribute. (Source: Author)

Grid Asset Type	Illustrative Robustness attribute by weather event type		
	Floods	Summer Storms	Cyclones
Poles and wires	$\%elevated$	$\%elevated$ $\%hardened^1$ $\%clearance$	$\%elevated$ $\%hardened^2$ $\%clearance$ $\%underground$
Substations	$\%elevated$ $\%hardened^3$	$\%elevated$ $\%hardened^3$	$\%elevated$ $\%hardened^3$
Transmission lines		$\%hardened^1$ $\%clearance$	$\%hardened^2$ $\%underground$
Bulk Generators	$\%elevated$ $\%hardened^3$ $\%fuel\ reserve$	$\%fuel\ reserve$	$\%hardened^3$ $\%fuel\ reserve$
DER	$\%elevated$ $\%hardened^3$ $\%fuel\ reserve$	$\%elevated$ $\%hardened^3$ $\%fuel\ reserve$ $\%clearance$	$\%elevated$ $\%hardened^3$ $\%fuel\ reserve$ $\%clearance$

- $hardened^1$ = poles and wires hardened to sustain mild/strong wind
- $hardened^2$ = poles and wires hardened to sustain extreme wind
- $hardened^3$ = presence of flood barriers or waterproof equipment
- $clearance$ = trimmed vegetation near the assets

The method requires identifying a set of vulnerability classes that would characterize the broad power system strength for a selected geographical area. An illustrative scale of vulnerability ranging from 1 to 7 is proposed Tab.6.3 where the vulnerability class of a system in a given area and for a given weather event type is determined by its overall percentage of "hardened" assets for that event type as per Tab.6.2.

Importantly, for each weather event type, the scale also connects the vulnerability of the power system to a respective average time range of system restoration after such weather event happens. Data from past extreme weather event must be collected, such as : Strength of the weather event, robustness of the power system at the time of the event and actual restoration time after the event. It is to be noted that the restoration time might be influenced by the topography or remoteness of the geographic area, and therefore each geographic area will have its own table.

The calibration and fine-tuning of this scale to enable its operationalisation would require the use of historical data that correlates the time of system restoration with its overall robustness.

Table 6.3 Power grid vulnerability classes given a weather event type. (Source: Author)

Vulnerability Scale [1]	*%Robustness* (up to) [2]	Energy Supply Downtime (up to #days) [3]
1	100%	2
2	80%	7
3	70%	14
4	50%	30
5	30%	90
6	20%	180
7	10%	360

The integrated system vulnerability considers two additional dimensions, its level of redundancy and "smartness". The redundancy of a power system in one given region depends on the redundancy of the upstream network (mainly the feeding transmission lines). The way this would be taken into consideration is by notching up or down the stand-alone vulnerability of the power system in each region to take into consideration the configuration of the upstream feeding network. Similarly, if the "smartness" of a local power system gave it capability to operate in islanding mode, that would notch-down the stand-alone vulnerability of the system. The comprehensive assessment of the system vulnerability is illustrated in fig.6.3.2

[1] The number of classes in the scale is arbitrary, only after collection and analysis of data from local network operators could the most appropriate number of classes in the scale be determined

[2] The percentage ranges shown in this table are indicative, only after collection and analysis of data from local network operators could the most appropriate ranges be determined.

[3] The values shown in this table are indicative, only after collection of data from the local network operators could the most appropriate downtime values be determined and correlated with the identified classes and robustness ranges.

flood	Stronger	Bigger	Smarter	
	Stand-alone vulnerability class	Upstream Transmission Network vulnerability / redundancy	System Island-able	Overall Power System Vulnerability class
Geographical Area 1	1-7	notch-up or down	notch down	1-7
Geographical Area 2	1-7	notch-up or down	notch down	1-7
Geographical Area 3	1-7	notch-up or down	notch down	1-7
...	1-7	notch-up or down	notch down	1-7
Geographical Area n	1-7	notch-up or down	notch down	1-7

cyclone	Stronger	Bigger	Smarter	
	Stand-alone vulnerability class	Upstream Transmission Network vulnerability / redundancy	System Island-able	Overall Power System Vulnerability class
Geographical Area 1	1-7	notch-up or down	notch down	1-7
Geographical Area 2	1-7	notch-up or down	notch down	1-7
Geographical Area 3	1-7	notch-up or down	notch down	1-7
...	1-7	notch-up or down	notch down	1-7
Geographical Area n	1-7	notch-up or down	notch down	1-7

Storms (wind+rain)	Stronger	Bigger	Smarter	
	Stand-alone vulnerability class	Upstream Transmission Network vulnerability / redundancy	System Island-able	Overall Power System Vulnerability class
Geographical Area 1	1-7	notch-up or down	notch down	1-7
Geographical Area 2	1-7	notch-up or down	notch down	1-7
Geographical Area 3	1-7	notch-up or down	notch down	1-7
...	1-7	notch-up or down	notch down	1-7
Geographical Area n	1-7	notch-up or down	notch down	1-7

Fig. 6.6 Power grid vulnerability classes ranking. (Source: Author)

6.3.3 Combined vulnerability assessment

The overall vulnerably to extreme weather event in a selected geographical area is the combination of the power grid assets vulnerability and the local community vulnerability explained respectively in (6.3.2) and (6.3.1).

The grid vulnerability Classes 1-7 are described in Fig.6.3.2 and the community vulnerability classes in Fig.6.3.1.

An aggregation approach must be identified in order to provide a single vulnerability measure ($Vulnerability_w(s)$) for each extreme weather events that combine the community and grid vulnerability in a defined geographical area for the Energy Risk Impact as described in (6.1).

An illustrative aggregated scoring matrix is proposed in tab.6.4

Table 6.4 Aggregated vulnerability scoring matrix for one weather event and given geographical area (Source: Autor).

Community / Grid	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Class1	0.01	0.05	0.2	0.5	0.75
Class2	0.05	0.1	0.3	0.6	0.8
Class3	0.2	0.3	0.6	0.75	0.90
Class4	0.5	0.6	0.75	0.90	0.95
Class5	0.6	0.75	0.80	0.95	1
Class6	0.75	0.80	0.90	0.95	1
Class7	0.8	0.90	0.95	1	1

In this approach, the aggregated vulnerability is a scalar number between 0 and 1 that is weighting the exposure in the calculation of the Energy Risk Impact (6.1). Similarly, to community and grid vulnerability, the aggregated vulnerability itself can be subdivided in discrete classes. For example, 3 classes: Low, Medium and High vulnerability (Green, Orange and Red in fig.6.3.2). The meaning that could be given to a "Low Vulnerability" class could be the desired state for each community. Those could be the intersections between community vulnerability and grid vulnerability that are deemed desired and satisfactory. Conversely, the "High Vulnerability" class could be an indicator of prioritization for action for those communities.

To operationalize the approach, a dedicated calibration activity must be undertaken in order to assess aggregated vulnerability scalars within the matrix. In order to do so, data point must be collected in the field to qualify and quantify local community vulnerability and the intersection with the grid vulnerability in that location.

The importance of the vulnerability is dictated by the fact that both components (community and grid) can be separately or jointly addressed with specific measures in order to reduce its aggregated value and this is also the key lever that government and planners can use to mitigate the impact of extreme weather conditions when it comes to power energy supply and usage.

6.4 A method to estimate the exposure or Energy Not Supplied

The exposure in the Energy Risk Impact $Exposure_w(s)$ is the total power energy in kWh that would be not supplied in the territory s if a weather event w happens.

As per the vulnerability, this measure varies for each geographical area and can be estimated by multiplying the local population average daily usage of energy by the expected number of days before restoration. That last number can be extrapolated from the Power Grid Vulnarability Class Tab.6.3 following the assessment of the local power system vulnerability. To operationalize the approach, a dedicated calibration activity must be undertaken in order to average time of power supply interruption for each weather events in function of the power system robustness.

6.5 Assignment of monetary value to the energy risk impact

In order to enable a homogeneous and coherent linkage between the energy risk impact and investments needed to minimize it, it is proposed to assign a monetary value to that risk. By proposed definition, the energy risk impact is a risk weighted measure of an amount of energy not served, expressed in kWh. In his study , Mahdi Masuduzzan [113] from the Bangladesh Ministry of Finance, shows the existence of unidirectional causality running from electricity consumption to economic growth in Bangladesh. It therefore assumed that a proportional relationship between one kWh consumed by the population and one dollar of resulting economic activity (GDP) can be identified and utilized to assign a pseudo monetary economic value to each kWh consumed.

By association, we can argue that the energy risk impact value that is a measure of an energy not supplied in response to extreme weather events carries a similar impact on the GDP, but with a minus sign, that is therefore identifying the monetary cost of that energy risk impact.

6.6 Linkage to power system planning

The proposed methodology can inform a framework that assesses the Energy Risk Impact of extreme weather under multiple "vulnerability" scenarios and inform the planning and design of the power network in such a way that would minimize the energy risk impact (and consequently increase the overall community and power grid network resilience) in a given geographical area under budget and social benefit constraints. The framework can be applied iteratively for network extension and network reinforcement for all regions $s \in S$

The proposed mathematical formulation of the problem is:

$$
\begin{aligned}
min_{x_i(s)} \quad & \left(ERI_{S,W}\left[Vulnerability_w\left(x_i(s)\right), Exposure_w\left(x_i(s)\right)\right]\right) \\
s.t. \quad & \left[\sum_{i\in A, s\in S}\left(c_i \cdot x_i(s)\right)\right] \leq budget \\
& x_i(s) \in \{0,1\}, \forall i \in A, \forall s \in S
\end{aligned}
\tag{6.3}
$$

where

- S is the geographical area considered
- s is a region $\in S$
- $W = \{Cyclone, Flood, SummerStorms, etc.\}$ the set off all weather events considered
- w is the weather event type $\in W$
- A is the set of possible grid assets or robustness enhancement for grid assets (such as flood barrier)
- $x_i(s)$ is a grid assets or robustness enhancement for grid assets $\in A$ that is considered to be installed in region s
- $Vulnerability_w(x_i(s))$ is the combined vulnerability (Tab.6.4) in the configuration of the power grid where assets x_i are in place in region s
- $Exposure_w(x_i(s))$ is the total power energy in kWh that would be not supplied in the configuration of the grid where assets x_i are in place if a weather event w happens in region s
- $ERI_{S,W}$ is the power Energy Risk impact in the configuration of the grid where assets x_i are in place in the area S for all the weather events considered

The optimisation result will select the most effective investment in terms of assets but also community to serve via the combined vulnerability (Tab.6.4). With this formulation, it is expected that resilience enhancement assets will be located to serve the most vulnerable communities for each extreme weather event type.

A sensitivity analysis of the budget value would need to be performed in order to support the policymakers in assessing an optimal budget to allocate in the resilience enhancement. The sensitivity analysis would progressively increment the budget and minimize the Energy Risk Impact for each new budget value until when the pseudo monetary value of the ERI reduction resulting from the increased budget is lower than the monetary value of the increase in budget.

At a country level, the development of the proposed methodology and baseline calculation of energy risk impact across the entire territory, even with moderate granularity, is useful to identify the stronger and weaker areas and communities of the country. This can inform high-level prioritization of resources to certain area to decrease their specific vulnerability (community and grid). The distinction between community vulnerability and grid vulnerability is also helpful to identify the type of measures that need to be undertaken for improvement and therefore the type of resources to be provided.

At local government level, the proposed methodology can offer a fine and granular analysis of the vulnerability, at community and grid level. From an original baseline assessment, multiple improvement scenarios can be identified and their impact on the aggregated vulnerability assessed to inform investment decision, community support and power system design.

Part III

Conclusion and future work

Chapter 7

Conclusion and Future Work

This research presents two novel data-driven frameworks that leverage Machine Learning techniques developed to solve the load flows and integrates it with stochastic DER penetration scenario analysis.

The frameworks can be integrated into long-term network planning to manage the forecasted effects any DER penetration scenario, limited to voltage excursions and reverse power flows assuming a stable frequency.

Unlike previous approaches, this research combine ML techniques with probabilistic analysis to infer Voltage excursion or RPF probabilities in response to a vast number of DER penetration and individual generation scenarios. The use of ML models substitutes the need for heavy computations required to solve the load flow equations for the entire system. The following sections discuss in more details the strengths, limitations and future work connected to each framework.

7.1 The voltage risk framework

In summary, the voltage risk framework uses Support Vector Machine (SVM) classifiers to infer voltage excursions (out-off the statutory limits) probabilities in response to any future scenario of DER penetration. A financial model is also proposed to estimate the monetary impacts of such events.

7.1.1 Strengths

The key strengths of the proposed data-driven end-to-end voltage risk framework can be summarised as:

- The SVM models demonstrate good accuracy in predicting the over- and under-voltage risks
- The SVM models demonstrate excellent performance in assessing large amounts of scenarios
- No assumptions are made regarding the DER daily/seasonal generation profile and load profile. The framework enables to assess any type of complex scenarios for each individual DER.
- The financial impact models are sufficiently realistic and the proposed parameters easy to estimate based on historical claims data

Therefore the statistic approach, based on ML instead of simulations to derive the risk probabilities and the simple but realistic financial impact model makes it a reliable and scalable tool for DNSP to assess and forecast the probabilistic impacts of any future scenario of DER penetration in large networks. The impact assessment is valuable to support the respective network long term planning decisions necessary to maintain reliability and quality of service under those scenarios.

7.1.2 Limitations

Within the scope of this research that assumes a stable frequency and investigates the effect of DER on node voltages, three key limitations of the developed framework are the following:

1. The discussed construction of the training data-set required to train the SVM models that is simulation-based assume that all the network parameters are known with precision. We know that this is a fully realistic assumption and that DNSPs have partial knowledge of such parameters. An alternative method would be to use real data sampled from the network operation (net-loads and voltage at every node) under multiple conditions to train the models. Additionally, the discussed models have been constructed on a limited test case,

2. Additionally, a network configuration is not entirely static as, for instance, transformers can have their load tap changed to adapt to load conditions which affect the network parameters, therefore multiple power system configurations must be considered while training and using the SVM models.

3. The financial model proposed assumes that a complaint is made as soon as a voltage statutory limit breach happens. This binary approach could be over-simplistic as the severity of the breach (extend of the breach and duration of the breach) would most likely influence the probability of those complaints happening and also the associated cost. A more sophisticated approach where the cost of a complaint and the probability of complaints are a function of the breach severity could be developed.

7.1.3 Future work

In order to address the three above-listed limitations, the following respective future works would be necessary:

1. Test the proposed framework on a large-scale distribution network and train the SVM leveraging a mix of observed and simulated data (net-loads and node voltages).

2. Explore reduced sampling methods ([92]) to contain the effort required to generate the training data sets and training the models

3. Perform a theoretical sensitivity analysis of the SVM model accuracy in response to local changes in parameters and test new supervised training methods to train the existing model to recognise the new configurations.

4. As per RPF, train regression models (DNN instead of SVM) that would not only recognise the breach in binary terms but could also predict the severity of the breach. Such models would enable the design of a more complex financial impact modelling.

7.2 The RPF scenario analysis framework

In summary, the RPF scenario analysis framework uses a Deep Neural Network (DNN) regression models to forecast RPF intensity probabilities in response to any future scenario of DER penetration. The forecasted RPF intensity at any point of the network is used to assess potential security and reliability impacts on network assets (eg. asset rating, protection configuration, etc.).

7.2.1 Strengths

Similar to the voltage risk framework, the key strengths of the proposed data-driven RPF scenario analysis framework can be summarised as:

- The DNN models demonstrate good accuracy in predicting the currents at every node of the system.
- The DNN models demonstrate excellent performance in assessing large amounts of scenarios.
- No assumptions are made regarding the DER daily/seasonal generation profile and load profile. The framework enables to assess any type of complex scenarios for each individual DER.

Therefore the statistic approach, based on DNN instead of simulations to derive the RPF intensities and probabilities at any node makes it a reliable and scalable tool for DNSP to assess any future scenario of DER penetration in large networks. The RPFs intensities are valuable to assess the long-term impact on network assets ratings and circuit protection configuration that are necessary to maintain the reliability and security of power supply.

7.2.2 Limitations

Within the scope of this research that assumes a stable frequency and investigates the effect of DER on RPFs, two key limitations of the developed framework are the following:

1. The discussed construction of training data-set required to train the DNN models that is simulation-based assume that all the network parameters are known with precision. We know that this is an unrealistic assumption and that DNSPs might have only partial knowledge of such parameters. An alternative method would be to use real data

sampled from the network operation (net-loads and currents at every node) under multiple conditions to train the models.

2. Additionally, a network configuration is not entirely static as, for instance, transformers can have their load tap changed to adapt to load conditions which affect the network parameters, therefore multiple system configurations must be considered while training and using the DNN models.

7.2.3 Future work

To address the two above-listed limitations, the following respective future works would be necessary:

1. Test the proposed framework on a large-scale distribution network and train DNN leveraging a mix of observed and simulated data

2. Perform a theoretical sensitivity analysis of the DNN model accuracy in response to local changes in parameters and test new supervised training methods to train the existing model to recognise the new configurations

3. Explore the use of reinforcement learning technique to adapt the model to new conditions, such as change in system parameters and configuration.

7.3 The extreme weather event risk and community resilience method

In summary, the resilience method adopts a probabilistic risk-based approach that capture and aggregate quantifiable measures to characterize the community and power grid resilience to extreme weather events. Additionally, a linkage method between those classes and the system planning framework has been explored.

7.3.1 Strengths

Overall the presented risk-based approach to extreme weather event resilience offers quantifiable measures that can be used and integrated into optimised power grid planning decisions. Remarkably, the approach integrates sociological aspects of intrinsic community resilience with the technical resilience of power system that serves the same community in quantifiable terms. The method can be integrated with Geographic Information System (GIS) for any available geographical granularity to merge localised weather events risks, local community vulnerability indicators and power system assets robustness. Even if developed for the Bangladesh case, the methodology is transferable to other parts of the world.

7.3.2 Limitations

Within the scope and context of this research that is focused on the Bangladesh case, two key limitations of the resilience method are the following:

1. Some vulnerability criteria data might not be available or monitored by the local government. The application and roll-out of the method would require a vast monitoring framework to be activated and rolled-out. In the meantime, some proxy data could be used as the best approximation

2. The availability of extreme weather event probability distribution over the geographical area might not be available with sufficient accuracy and granularity. Some further modelling effort might be required

7.3.3 Future work

To calibrate and operationalise the proposed method, the following future work must be done:

- Evolve the proposed approach from method to model. Pilot test some selected communities, gather and monitor over time vulnerability data point to calibrate the weights and vulnerability classes
- Test and calibrate the models on a larger scale (at wide region or country level)
- Build a geographic information system (GIS) framework that connects and aggregates multiple data sources. That would enable to determine and visualise the vulnerability/resilience of communities and power system on a map
- Extend and integrate the extreme weather event risk modelling with climate change modelling

References

[1] MEI. Smarter, inclusive energy solutions, 2019.

[2] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, August 2006.

[3] Thierry Van Cutsem and Costas Vournas. *Voltage stability analysis of electric power systems*. Vluwer Academic Publications, 1998.

[4] R. Christie. 118 bus power flow test case - power systems test case archive - august, 1993.

[5] K. P. Schneider, B. A. Mather, B. C. Pal, C.-W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W. Kersting. Analytic considerations and design basis for the IEEE distribution test feeders. *IEEE Transactions on Power Systems*, 33(3):3181–3188, may 2018.

[6] Ministry of Environment and Government of the People's Republic of Bangladesh Forests. Bangladesh climate change strategy and action plan 2008, September 2008.

[7] Mathaios Panteli and Pierluigi Mancarella. The grid: Stronger, bigger, smarter?: Presenting a conceptual framework of power system resilience. *IEEE Power and Energy Magazine*, 13(3):58–66, may 2015.

[8] IRENA. Global energy transformation: A roadmap to 2050. Technical report, 2018.

[9] APVI Australian PV Institute. Australian pv market since april 2001. online, September 2019.

[10] David Korinek P. Jeffrey Palermo, Mike Wilks. International review of transmission reliability standards. Technical report, Australian Energy Market Commission Reliability Panel, 2008.

[11] M. Awad, S. Broad, K.E. Casey, Jing Chen, A.S. Geevarghese, J.C. Miller, A.J. Perez, A.Y. Sheffrin, Mingxia Zhang, E. Toolson, G. Drayton, A.F. Rahimi, B.F. Hobbs, and F.A. Wolak. The california ISO transmission economic assessment methodology (TEAM): principles and application to path 26. In *2006 IEEE Power Engineering Society General Meeting*. IEEE, 2006.

[12] J. Paulinder. Reliability analysis of distribution network investments. In *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*. Institution of Engineering and Technology, 2013.

[13] Xianing Jin, Yi Song, Weihong Yang, Yangyang Lv, and Zhanjun Li. Research on planning strategy to achieve high reliability for urban distribution systems. In *2019 IEEE 3rd International Electrical and Energy Conference (CIEEC)*. IEEE, sep 2019.

[14] Wei Liu, Yan Wang, Jingran Xue, and Zhicheng Sha. Study on distribution network planning method based on reliability of cost-benefit analysis. In *2018 China International Conference on Electricity Distribution (CICED)*. IEEE, sep 2018.

[15] AUSNET Services. Distribution Annual Planning Report 2017-2011. Technical Report December 2016, 2016.

[16] Powercor. Distribution Annual Planning Report. Technical Report December, 2016.

[17] Citipower. Distribution Annual Planning Report. Technical Report December, 2016.

[18] Victorian Distribution Businesses. Transmission Connection Planning Report 2016. Technical report, 2016.

[19] AEMO. Value of customer reliability – application guide. Technical Report December, November 2014.

[20] Standard Operating Procedure. Export limits for Embedded Generators up to 200 kVA connected at Low Voltage. Technical Report July, 2017.

[21] IEEE Power and Energy Society. Voltage stability assessment: Concepts , practices and tools. Technical report, IEEE PES, August 2002.

[22] Le Zheng, Wei Hu, Yong Min, and Jin Ma. A Novel Method to Monitor and Predict Voltage Collapse: the Critical Transitions Approach. *IEEE Transactions on Power Systems*, 8950(c):1–1, mar 2017.

[23] Mevludin Glavic and Thierry Van Cutsem. A short survey of methods for voltage instability detection. *IEEE PES General meeting Conference*, pages 1–8, jul 2011.

[24] T. van Cutsem, L. Wehenkel, M Pavella, B Heilbronn, and M Goubin. Decision tree approaches to voltage security assessment. *IEE Proc. C - Generation, Transmission and Distribution*, 140(3):189, 1993.

[25] Isaiah Adebayo, Adisa A. Jimoh, and Adedayo Yusuff. Voltage stability assessment and identification of important nodes in power transmission network through network response structural characteristics. *IET Generation, Transmission & Distribution*, 11(6):1398–1408, 2017.

[26] Paul Denholm, Matthew O'Connell, Gregory Brinkman, and Jennie Jorgenson. Overgeneration from Solar Energy in California: A Field Guide to the Duck Chart (NREL/TP-6A20-65023). *Technical Report*, (November):46, 2015.

[27] S. Papathanassiou et all. *Capacity of Distribution Feeders for Hosting Distributed Energy Resources*. CIGRE (International Council on Large Electric Systems) - WG C6.24, June 2014.

[28] Rich Seguin, Jeremy Woyak, David Costyk, Josh Hambrick, and Barry Mather. High-penetration PV integration handbook for distribution engineers. Technical report, jan 2016.

[29] J. Freek Baalbergen, Madeleine Gibescu, and Lou Van Der Sluis. Voltage stability consequences of decentralized generation and possibilities for intelligent control. *IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT Europe*, pages 1–8, 2010.

[30] P.P. Barker and R.W. De Mello. Determining the impact of distributed generation on power systems. i. radial distribution systems. In *2000 Power Engineering Society Summer Meeting (Cat. No.00CH37134)*, volume 3, pages 1645–1656 vol. 3. IEEE, July 2000.

[31] Buyang Qi, Student Member, and Jovica V Milanovic. Identification of Critical Parameters Affecting Voltage Stability in Networks with Renewable Generations using Sensitivity Analysis Methods. In *PowerTech*, 2017.

[32] Z. Zhang, L. F. Ochoa, and G. Valverde. A novel voltage sensitivity approach for the decentralized control of dg plants. *IEEE Transactions on Power Systems*, 33(2):1566–1576, March 2018.

[33] Mariana Magdy Mounir Kamel, Abdelrahman A Karrar, and Ahmed H Eltom. Development and Application of a New Voltage Stability Index for On-Line Monitoring and Shedding. *IEEE Transactions on Power Systems*, 33(2):1–1, mar 2017.

[34] Juan Muñoz, Claudio Cañizares, Kankar Bhattacharya, and Alfredo Vaccaro. An affine arithmetic-based method for voltage stability assessment of power systems with intermittent generation sources. *IEEE Transactions on Power Systems*, 28(4):4475–4487, nov 2013.

[35] M. R. Monteiro, A. C. Zambroni de Souza, and B. I. L. Lopes. The influence of renewable generation in voltage collapse indexes. In *2017 6th International Conference on Clean Electrical Power (ICCEP)*, pages 513–518. IEEE, jun 2017.

[36] S. Veerapen and R. T. F. Ah King. Integrating distributed energy resources in the electrical grid considering resource variability for reliable power planning. In *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, pages 1356–1361, Oct 2012.

[37] P. D. F. Ferreira, P. M. S. Carvalho, L. A. F. M. Ferreira, and M. D. Ilic. Distributed energy resources integration challenges in low-voltage networks: Voltage control limitation and risks of cascading. *IEEE Trans. Sust. Energy*, 4(1):82–88, 2013.

[38] T. Lakkaraju, a. Feliachi, and J.E. Saymansky. Voltage Stability Risk Analysis. *2007 IEEE Power Engineering Society General Meeting*, jun 2007.

[39] Saeed Alyami, Yang Wang, and Caisheng Wang. Overvoltage Risk Analysis in Distribution Networks with High Penetration of PVs. In *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, oct 2016.

[40] T. Alpcan, S. Saha, and M. Aldeen. Assessment of voltage stability risks under intermittent renewable generation. In *2014 IEEE PES General Meeting | Conference Exposition*, pages 1–5. IEEE, July 2014.

[41] Eric Martinot, Lorenzo Kristov, and J. David Erickson. Distribution system planning and innovation for distributed energy futures. *Current Sustainable/Renewable Energy Reports*, 2(2):47–54, Jun 2015.

[42] Paul De Martini. More than smart: A framework to make the distribution grid more open, efficient and resilient. *Resnick Sustainability Institute, California Institute of Technology*, page 28pp., 2014.

[43] A. Lawson, M. Goldstein, and C.J. Dent. Bayesian framework for power network planning under uncertainty. *Sustainable Energy, Grids and Networks*, 7:47–57, sep 2016.

[44] E.A. Martínez Ceseña and P. Mancarella. Practical recursive algorithms and flexible open-source applications for planning of smart distribution networks with demand response. *Sustainable Energy, Grids and Networks*, 7:104–116, sep 2016.

[45] Ashwini Patil, Rucha Girgaonkar, and Shravana Kumar Musunuri. Impacts of increasing photovoltaic penetration on distribution grid — voltage rise case study. In *2014 International Conference on Advances in Green Energy (ICAGE)*. IEEE, dec 2014.

[46] Jinah Noh, SeokJu Kang, Jaewoo Kim, and Jung-Wook Park. A study on volt-watt mode of smart inverter to prevent voltage rise with high penetration of PV system. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, aug 2019.

[47] W. Moondee and W. Srirattanawichaikul. Study of coordinated reactive power control for distribution grid voltage regulation with photovoltaic systems. In *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*. IEEE, mar 2019.

[48] Lou van der Sluis Pieter Schavemaker. *Electrical Power System Essentials*. John Wiley & Sons Inc, July 2017.

[49] Johanna Myrzik Edward Coster and Wil Kling. *Distributed Generation*. IntechOpen, February 2010.

[50] R. A. Walling, R. Saint, R. C. Dugan, J. Burke, and L. A. Kojovic. Summary of distributed resources impact on power delivery systems. *IEEE Transactions on Power Delivery*, 23(3):1636–1644, July 2008.

[51] Nima Amjady, Ahmad Attarha, Shahab Dehghan, and Antonio J. Conejo. Adaptive robust expansion planning for a distribution network with DERs. *IEEE Transactions on Power Systems*, 33(2):1698–1715, mar 2018.

[52] Xian Guo, He Guo, and Haozhong Cheng. Coordinated planning of distributed energy resources and microgrid network. In *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. IEEE, may 2016.

[53] M. J. E. Alam and K. M. Muttaqi and D. Sutanto. A comprehensive assessment tool for solar pv impacts on low voltage three phase distribution networks. In *2nd International Conference on the Developments in Renewable Energy Technology (ICDRET 2012),*, 2012.

[54] Kalpesh Joshi and Naran Pindoriya. Advances in distribution system analysis with distributed resources: Survey with a case study. *Sustainable Energy, Grids and Networks*, 15:86–100, sep 2018.

[55] A. Augugliaro, V. Cataliotti, L. Dusonchet, S. Favuzza, and G. Scaccianoce. Load flow solution in electrical power systems with variable configurations by progressive learning networks. In *PowerTech Budapest 99. Abstract Records. (Cat. No.99EX376)*. IEEE.

[56] Meriem Fikri, Bouchra Cheddadi, Omar Sabri, Tonria Haidi, Belfqih Abdelaziz, and Meriem Majdoub. Power flow analysis by numerical techniques and artificial neural networks. In *2018 Renewable Energies, Power Systems & Green Inclusive Economy (REPS-GIE)*. IEEE, apr 2018.

[57] Anna Larsson, Alain Germond, and Boming Zhang. Application of neural networks to the identification of steady state equivalents of external power systems. In *2006 International Conference on Power System Technology*. IEEE, oct 2006.

[58] Gonzalo Esteban Constante-Flores and Mahesh S. Illindala. Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using monte carlo simulation. *IEEE Transactions on Industry Applications*, 55(1):174–181, jan 2019.

[59] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, 1991.

[60] Wael Abdullah Alsulami and R Sreerama Kumar. Artificial neural network based load flow solution of saudi national grid. In *2017 Saudi Arabia Smart Grid (SASG)*. IEEE, dec 2017.

[61] A. Rathinam, S. Padmini, and V. Ravikumar. Application of supervised learning artificial neural networks [CPNN, BPNN] for solving power flow problem. In *IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007)*. IEE, 2007.

[62] George S. Misyris, Andreas Venzke, and Spyros Chatzivasileiadis. Physics-informed neural networks for power systems.

[63] Ferdinando Fioretto, Terrence W.K. Mak, and Pascal Van Hentenryck. Predicting AC optimal power flows: Combining deep learning and lagrangian dual methods. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01):630–637, apr 2020.

[64] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (3rd Edition)*. Pearson, 2009.

[65] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[66] Sebastian Espinoza, Rafael Sacaan, Hugh Rudnick, Mathaios Panteli, Pierluigi Mancarella, Alan Poulos, Juan Carlos De La Llera, Alejandro Navarro, and Rodrigo Moreno. Seismic resilience assessment and adaptation of the Northern Chilean power system. *IEEE Power and Energy Society General Meeting*, 2018-Janua:1–5, jul 2018.

[67] Mathaios Panteli and Pierluigi Mancarella. Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events. *IEEE Systems Journal*, 11(3):1733–1742, sep 2017.

[68] Mathaios Panteli, Cassandra Pickering, Sean Wilkinson, Richard Dawson, and Pierluigi Mancarella. Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures. *IEEE Transactions on Power Systems*, 32(5):3747–3757, sep 2017.

[69] Neha Mukhi, Deb Chattopadhyay, Benjamin Hobbs, and Evangelia Spyrou. Building climate resilience into power system planning : the case of Bangladesh. *World Bank Report*, ACS23320(July), 2017.

[70] Yutian Zhou, Mathaios Panteli, Rodrigo Moreno, and Pierluigi Mancarella. System-level assessment of reliability and resilience provision from microgrids. *Applied Energy*, 230(August):374–392, nov 2018.

[71] J. C. Jordan. Deconstructing resilience: why gender and power matter in responding to climate stress in Bangladesh. *Climate and Development*, 11(2):167–179, feb 2019.

[72] Xiujie Zheng, Ning Wu, and Jianbo Yi. Development of a hybrid simulation technology to optimize parameters of multiple power system stabilizers. In *2020 5th Asia Conference on Power and Electrical Engineering (ACPEE)*. IEEE, jun 2020.

[73] Xi-Fan Wang, Yonghua Song, and Malcolm Irving. *Modern Power Systems Analysis*. Springer, 2008.

[74] Jiao-Jiao Deng and Hsiao-Dong Chiang. Convergence region of newton iterative power flow method: Numerical studies. *Journal of Applied Mathematics*, 2013:1–12, 2013.

[75] the International Organization for Standardization. Iso 31000 risk management.

[76] Norway) Aven Terje (Universitet i Stavanger. *Quantitative Risk Assessment*. Cambridge University Press, 2011.

[77] Dirk P. Kroese Reuven Y. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley & Sons, 2016.

[78] L. I. Rozonoèr M. A. Aizerman, È. M. Braverman. Theoretical foundations of the potential function method in pattern recognition learning. *Avtomat. i Telemekh.*, 25:917–936, 1964.

[79] The Electricity, Supply Regulations, The Electricity, Supply Regulations, Electricity Safety, Continuity Regulations, The Electricity Safety, Continuity Regulations, Distribution Network Operator, and Electricity Supply Regula. Voltage rise the big issue when c nnecting embedded generat on to long 11 kV overhead lines. *POWER ENGINEERING JOURNAL*, pages 5–12, 2002.

[80] C.W. Taylor. *Power system voltage stability*. McGraw-Hill, New York, Inc., NY, USA, 1994.

[81] Sara Eftekharnejad, Vijay Vittal, Brian Keel, and Jeffrey Loehr. Impact of increased penetration of photovoltaic generation on power systems. *IEEE Transactions on Power Systems*, 28(2):893–901, may 2013.

[82] Joel Kennedy, Phil Ciufo, and Ashish Agalgaonkar. Over-voltage mitigation within distribution networks with a high renewable distributed generation penetration. *ENERGYCON 2014 - IEEE International Energy Conference*, pages 1107–1114, 2014.

[83] M. A. Akbari, J. Aghaei, M. Barani, M. Savaghebi, M. Shafie-Khah, J. M. Guerrero, and J. P. S. Catalão. New metrics for evaluating technical benefits and risks of dgs increasing penetration. *IEEE Transactions on Smart Grid*, 8(6):2890–2902, Nov 2017.

[84] Tim Bedford and Roger Cooke. *Probabilistic Risk Analysis: Foundation and Methods*. Cambridge University Press, 2001.

[85] V. Ajjarapu and B. Lee. Bifurcation theory and its application to nonlinear dynamical phenomena in an electrical power system. *IEEE Transactions on Power Systems*, 7(1):424–431, Feb 1992.

[86] Stanley Kaplan, B. John Garrick, S Kaplin, and G John Garrick. On the quantitative definition of risk. *Risk Analysis*, 1(1):11–27, 1981.

[87] V. Ajjarapu and C. Christy. The continuation power flow: a tool for steady state voltage stability analysis. *IEEE Transactions on Power Systems*, 7(1):416–423, Feb 1992.

[88] MZ Laton, I Musirin, and TK Abdul-Rahman. Voltage stability assessment via continuation power flow method. *Int. journal of electrical and electronic system research (IEESR)*, 1:71–78, 2008.

[89] John Langford. The cross validation problem. In Peter Auer and Ron Meir, editors, *Learning Theory*, pages 687–688, Berlin, Heidelberg, June 2005. Springer Berlin Heidelberg.

[90] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427 – 437, jul 2009.

[91] Marina Sokolova, Nathalie Japkowicz, and Stan Szpakowicz. Beyond accuracy, f-score and roc: A family of discriminant measures for performance evaluation. In Abdul Sattar and Byeong-ho Kang, editors, *AI 2006: Advances in Artificial Intelligence*, pages 1015–1021, Berlin, Heidelberg, November 2006. Springer Berlin Heidelberg.

[92] Greg Schohn and David Cohn. Less is more: Active learning with support vector machines. In *ICML*, 2000.

[93] D.I. Doukas, P.A. Gkaidatzis, A.S. Bouhouras, K.I. Sgouras, and D.P. Labridis. On reverse power flow modelling in distribution grids. In *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016)*. Institution of Engineering and Technology, 2016.

[94] A. Wixon, K. Venkataraman, P. Newman, and K. Maleka. Load blinding stabilisation for DOC relays subjected to reverse load flow. In *12th IET International Conference on Developments in Power System Protection (DPSP 2014)*. Institution of Engineering and Technology, 2014.

[95] Liang Che, Mohammad E. Khodayar, and Mohammad Shahidehpour. Adaptive protection system for microgrids: Protection practices of a functional microgrid system. *IEEE Electrification Magazine*, 2(1):66–80, mar 2014.

[96] K. N. Bangash, M. E. A. Farrag, and A. H. Osman. Manage reverse power flow and fault current level in LV network with high penetration of small scale solar and wind power generation. In *2018 53rd International Universities Power Engineering Conference (UPEC)*. IEEE, sep 2018.

[97] Juan Pablo Holguin, David Celeita Rodriguez, and Gustavo Ramos. Reverse power flow (RPF) detection and impact on protection coordination of distribution systems. *IEEE Transactions on Industry Applications*, 56(3):2393–2401, may 2020.

[98] Electric Power Research Institute. Opendss, 2019.

[99] AEC. Solar reportquarter 3, 2019, September 2019.

[100] AEMO. Integrated system plan for the national electricity market. Technical report, 2018.

[101] Sonke Kreft and David Eckstein. Global climate risk index 2014. Technical report, Germanwatch e.V., November 2013.

[102] United Nations Development Programme. *Human Development Report 2014*. United Nations Publications, 2014.

[103] Muthukumara Mani, Sushenjit Bandyopadhyay, Shun Chonabayashi, Anil Markandya, and Thomas Mosier. *South Asia's Hotspots: The Impact of Temperature and Precipitation Changes on Living Standards*. The World Bank, jul 2018.

[104] Farzad Taheripour, Thomas W. Hertel, Badri Narayanan, Sebnem Sahin, Anil Markandya, and Bijon Kumer Mitra. Economic and land use impacts of improving water use efficiency in irrigation in south asia. *Journal of Environmental Protection*, 07(11):1571–1591, 2016.

[105] UNISDR. *2009 UNISDR Terminology on Disaster Risk Reduction*. United Nations International Disaster Disaster risk Disaster risk management Strategy for Disaster Reduction (UNISDR), 2009.

[106] M. Rehan Dastagir. Modeling recent climate change induced extreme events in bangladesh: A review. *Weather and Climate Extremes*, 7:49–60, mar 2015.

[107] Umma Habiba, Md. Anwarul Abedin, and Rajib Shaw. Disaster education in bangladesh: Opportunities and challenges. In *Disaster Risk Reduction Approaches in Bangladesh*, pages 307–330. Springer Japan, 2013.

[108] Zayeda Sharmin and Mohammad Samiul Islam. Consequences of climate change and gender vulnerability: Bangladesh perspective. *SSRN Electronic Journal*, 2013.

[109] Naila Kabeer. Poverty, social exclusion and the MDGs: The challenge of 'durable inequalities' in the asian context. *IDS Bulletin*, 37(3):64–78, jul 2006.

[110] Francesco Fuso Nerini, Mark Howells, Morgan Bazilian, and Maria F. Gomez. Rural electrification options in the brazilian amazon. *Energy for Sustainable Development*, 20:36–48, jun 2014.

[111] Reihana Mohideen. energy technology innovation in south asia: implications for gender equality and social inclusion. ADB SOUTH ASIA WORKING PAPER SERIES, oct 2018.

[112] Cesar Silva-Monroy Eric Vugrin, Anya Castillo. Resilience metrics for the electric powersystem: A performance-based approach. Technical report, Sandia National Laboratories, 2017.

[113] Mohammad Masuduzzaman. Electricity consumption and economic growth in bangladesh: Co integration and causality analysis. *Global J. Manag. Bus. Res.*, 12:31–48, January 2012.

[114] P. W. Sauer and M. A. Pai. Power system steady-state stability and the load-flow jacobian. *IEEE Transactions on Power Systems*, 5(4):1374–1383, Nov 1990.

[115] S. Greene, I. Dobson, and F. L. Alvarado. Contingency ranking for voltage collapse via sensitivities from a single nose curve. *IEEE Transactions on Power Systems*, 14(1):232–240, Feb 1999.

[116] S. Greene, I. Dobson, and F. L. Alvarado. Sensitivity of the loading margin to voltage collapse with respect to arbitrary parameters. *IEEE Transactions on Power Systems*, 12(1):262–272, Feb 1997.

Appendix A

Traditional load flow Solution method

An alternating current load-flow model is a model used in electrical engineering to analyse power grids. It provides a nonlinear system which describes the energy flow through each power line. For the scope of this research, steady-state operation is assumed, with no transient changes in power flow or voltage due to load or generation changes. The system frequency is also assumed to be constant. For e steady state power system that comprises B buses, the load-flows model is a a set of $2B$ equations in $2B$ algebraic variables V_i, θ_i :

$$0 = -P_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik}\cos\vartheta_{ik} + b_{ik}\sin\vartheta_{ik}), \forall i \in B \tag{A.1}$$

$$0 = -Q_i + \sum_{k=1}^{B} |v_i||v_k|(g_{ik}\sin\vartheta_{ik} + b_{ik}\cos\vartheta_{ik}), \forall i \in B \tag{A.2}$$

where

- P_i is the net injected real power (power generated minus power consumed) at bus i
- Q_i is the net injected reactive power (power generated minus power consumed) at bus i
- v_i is the voltage at bus i

- ϑ_{ik} is the difference in voltage angle between bus i and k
- y_{ik} is the admittance of the line between bus i and k
- g_{ik} is the conductance , or the real part of the admittance of the line between bus i and k
- b_{ik} is the susceptance, or the imaginary part of the admittance of the line between bus i and k

The goal of a power-flow study is to obtain complete voltages angle and magnitude information for each bus in a power system for specified load and generator real power and voltage conditions.[3] Once this information is known, real and reactive power flow on each branch as well as generator reactive power output can be analytically determined. Due to the nonlinear nature of this problem, numerical methods are employed to obtain a solution that is within an acceptable tolerance.

The solution to the power-flow problem begins with identifying the known and unknown variables in the system. The known and unknown variables are dependent on the type of bus. A bus without any generators connected to it is called a Load Bus. With one exception, a bus with at least one generator connected to it is called a Generator Bus. The exception is one arbitrarily-selected bus that has a generator. This bus is referred to as the slack bus.

In the power-flow problem, it is assumed that the real power P_i and reactive power Q_i at each Load Bus are known. For this reason, Load Buses are also known as PQ Buses. For Generator Buses, it is assumed that the real power generated P_i and the voltage magnitude $|v_i|$ is known. For the Slack Bus, it is assumed that the voltage magnitude $|v_i|$ and voltage phase ϑ_{ik} are known. Therefore, for each Load Bus, both the voltage magnitude and angle are unknown and must be solved for; for each Generator Bus, the voltage angle must be solved for; there are no variables that must be solved for the Slack Bus.

There are several different methods of solving the resulting nonlinear system of equations. The most popular is known as the Newton–Raphson method. This method begins with initial guesses of all unknown variables (voltage magnitude and angles at Load Buses and voltage angles at Generator Buses). Next, a Taylor Series is written, with the higher order terms ignored, for each of the power balance equations included in the system of equations . The result is a linear system of equations that can be expressed as:

$$\begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix} = - \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial V} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial V} \end{bmatrix}^{-1} \times \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \tag{A.3}$$

The linearized system of equations is solved to determine the next guess $(t+1)$ of voltage magnitude and angles based on

$$\begin{bmatrix} \theta \\ V \end{bmatrix}^{t+1} = \begin{bmatrix} \theta \\ V \end{bmatrix}^{t} + \begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix} \tag{A.4}$$

The process continues until a stopping condition is met. A common stopping condition is to terminate if the norm of the mismatch equations is below a specified tolerance.

Appendix B

Continuation load flow method

Calculating the *Critical Point* at a node using the load flow equations is not a straightforward exercise as the Jacobian of the system becomes singular as it approaches its maximum value. As described by [114],[115],[116] the Jacobian can reach a critical point in multiple ways, such as the increase of impedance in a key tie line; the increase of generation level at a generator with weak transmission while decreasing generation at all other generators; increase the load at a single bus; increase the load at all buses. In all cases, we are looking the "nose" point of the $V-\lambda$ curve, where λ is the load parameter (P or Q) that is being increased. To model this, we increase the power flow equations so that they are explicitly dependent a selected the load parameter λ [87, 88]. The resulting Jacobian of the Load Flow equations take the form:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial V} & \frac{\partial P}{\partial \lambda} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial V} & \frac{\partial Q}{\partial \lambda} \end{bmatrix} \times \begin{bmatrix} \Delta \theta \\ \Delta V \\ \Delta \lambda \end{bmatrix} \tag{B.1}$$

We note that the vector $[d\theta dV d\lambda]$ denotes the tangent direction of small changes in the Power Flow equations at any operating condition θ, V, λ and we have:

$$\begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial V} & \frac{\partial P}{\partial \lambda} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial V} & \frac{\partial Q}{\partial \lambda} \end{bmatrix}_{\theta,V,\lambda} \times \begin{bmatrix} d\theta \\ dV \\ d\lambda \end{bmatrix} = 0 \tag{B.2}$$

In order to determine the critical point of the $V - \lambda$ curve at a node i, starting from an operating point, we will use an iterative "predictor-corrector" algorithm as follows:

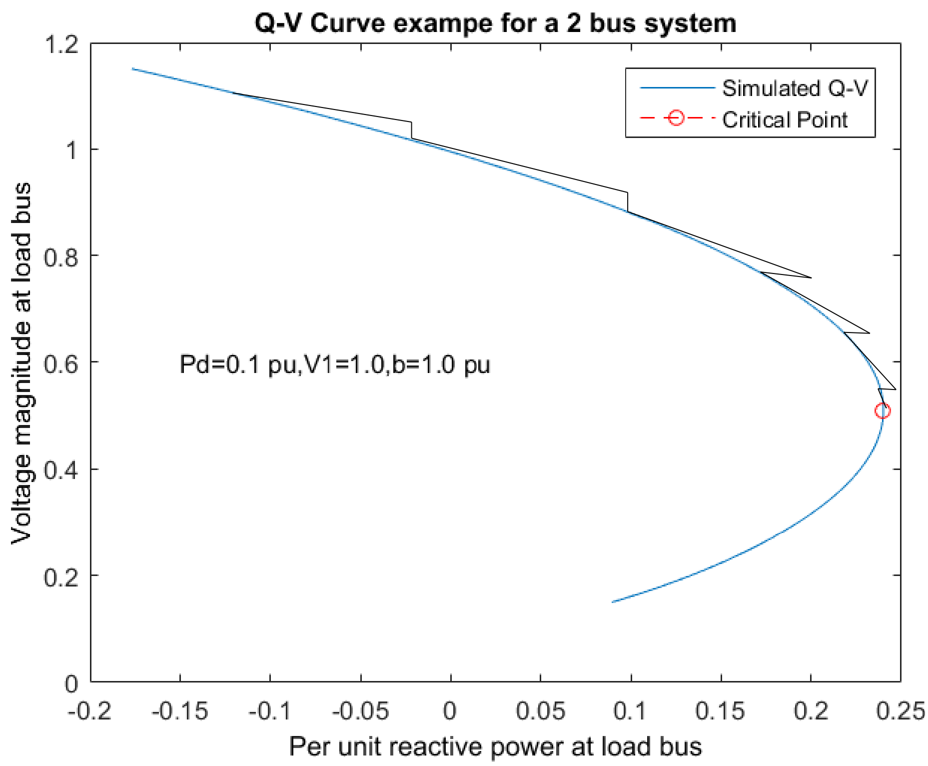


Fig. B.1 Predictor-Corrector iterative process (Source: Author)

- Step 1: We select an initial load parameter (λ) for node i, assume Q_{load} in this case, and calculate the tangent vector $[d\theta dV d\lambda]$ using (B.2)

- Step 2: We define a "step size" σ and calculate a "predicted state" $(\theta_p, V_p, \lambda_p)$ using:

$$\begin{bmatrix} \theta_p \\ V_p \\ \lambda_p \end{bmatrix} = \begin{bmatrix} \theta \\ V \\ \lambda \end{bmatrix} + \sigma \begin{bmatrix} d\theta \\ dV \\ d\lambda \end{bmatrix} \tag{B.3}$$

- Step 3: We correct the error found on the predictor process using the modified Newton-Raphson load flow algorithm as follows:

$$\begin{bmatrix} \Delta P \\ \Delta Q \\ \lambda_p \end{bmatrix}_{k+1} = \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial V} & \frac{\partial P}{\partial \lambda} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial V} & \frac{\partial Q}{\partial \lambda} \\ & e_k & \end{bmatrix}_{V_k, \theta_k, \lambda_k} \times \begin{bmatrix} \Delta \theta_p \\ \Delta V_p \\ \Delta \lambda \end{bmatrix}_k \tag{B.4}$$

  Where λ_p is the fixed value of the load parameter and e_k a row vector of the form $e_k = [0..1..0]$ with all elements equal to zero except the element corresponding to the selected load parameter which is valued at 1. The solution of the Newton-Raphson algorithm will identify the corrected values $(\theta_c, V_c, \lambda_c)$ that solve the Load flow equations for the given load parameter λ that is closer to the SNB point. If the Newton-Raphson doesn't converge, that means that the step size σ chosen in step 2 was to big and overshoot the SNB point. Therefore, we step-back to step 2 halving the step size until convergence is found.

- Step 4: Before re-iterating Step 2, we define the new continuation parameter that will be selected among the state variables according to the one that is changing the most with λ. Typically, as we approach the SNB point, the voltage itself V_i will become the new continuation parameter and we set λ to be the selected continuation parameter for the next iteration

- Step 5: We iterate the steps 2 to 4 until $\Delta\lambda$ becomes negative, meaning that we have overtaken the "nose" point of the curve $V-\lambda$.

The presented iterative method enables us to calculate the critical points at any selected node for any chosen load parameter λ (P or Q) and provide us the state values of the entire system under that condition given the directive values of P, Q, V for the other PV and PQ nodes of the system.