University of Southern Queensland
Schools of Engineering, Built Environment, and Agricultural & Environmental Science

Reactive Power Contribution of Grid-Connected Solar

A dissertation submitted by

Mr Dylan Edward PERRETT

In fulfilment of the two-part requirements of

ENG4111 (Part one) and ENG4112 (Part two) Research Project

Towards the degree of

Bachelor of Engineering (Honours)
Majoring in Electrical and Electronic Engineering

Submitted 18th October 2023

Abstract

A government commitment to net zero emission in Australia by 2050. An increase in renewable energy generation throughout the electricity grid and a reduction of synchronous thermal generation. The exponential growth in rooftop grid-connected solar photovoltaic penetrations. The characteristic of renewable energies can often be intermittent, and diurnal output fastened by direction to go green can lead to temporal mismatches between the balance of power generation and consumption. One of the fastest-growing forms of renewables is end-user grid-connected rooftop solar. The use of solar photovoltaics has many advantages but, correspondingly, comes with disadvantages that require significant attention. Additionally, there is a significant increasing trend in capacitive reactive power flow in the electricity grid, especially within the distribution network. The research in this dissertation focused on examining and analysing the reactive power contribution of grid-connected solar.

The project aimed to examine industry monitoring data from an aggregate end-user load and isolated solar system perspective to establish reactive power contributions from the solar system. In addition, the objectives were to detect and model operational reactive power within compliance with Australian standard 4777.2, Grid connection of energy systems via inverters, Part 2: Inverter requirements. In particular, the reactive contributions during low irradiance conditions. The method applied was to link industry monitoring and irradiance data together and compare three scenarios to highlight any significant trends to benefit regression modelling.

The initial literature review discovered that while there is considerable knowledge regarding the optimal operation of grid-connected solar photovoltaic systems, there is less research during the less-than-optimal operating conditions. While the impact of one solar system may not be significant, the cumulative effect of increasing penetrations demands more attention. Dissertation results established during these low irradiance events; the cumulative impact is generating substantial trends in the capacitive reactive power flow. While most contributions are within compliance, there are moments of un-constrained reactive contribution. Finally, regression models were generated to characterise the impact of reactive power contributions during these low irradiance conditions. The simulations will support understanding the contributions of increasing penetrations of photovoltaic systems in the distribution network.

Keywords: Reactive Power, Capacitive Power Flow, Low Irradiance, Solar photovoltaic.

Limitations of Use Disclaimer

University of Southern Queensland

Faculty of Health, Engineering and Sciences

ENG4111 & ENG4112 Research Project

Limitations of Use

The Council of the University of Southern Queensland, its Faculty of Health, Engineering and Sciences, and the staff of the University of Southern Queensland, do not accept any responsibility for the truth, accuracy, or completeness of material contained within or associated with this dissertation.

Persons using all or any part of this material do so at their own risk, and not at the risk of the Council of the University of Southern Queensland, its Faculty of Health, Engineering and Sciences or the staff of the University of Southern Queensland.

This dissertation reports an educational exercise and has no purpose or validity beyond this exercise. The sole purpose of the course pair entitles “Research Project” is to contribute to the overall education within the student’s chosen degree program. This document, the associated hardware, software, drawings, and any other material set out in the associated appendices should not be used for any other purpose: if they are so used, it is entirely at the risk of the user.

Certification of Project Work

I certify that the ideas, designs and experimental work, results, analyses, and conclusions set out in this dissertation are entirely my own effort, except where otherwise indicated and acknowledged.

I further certify that the work is original and has not been previously submitted for assessment in any other course or institution, except where specifically stated.

Dylan Perrett

Student Number: [redacted]

Acknowledgements

I want to take this opportunity to express my heartfelt appreciation to all those who have been instrumental; without their support, this research project would not have been achievable.

I thank Joel Hockey and Energy Queensland Limited for supporting access to an industry project and any others throughout EQL who have indirectly contributed.

I am deeply grateful to Dr Joel Kennedy from EQL for sharing your project, invaluable expertise, guidance, and mentorship throughout the process.

I am also immensely thankful to Dr Jason Brown from The University of Southern Queensland for your tremendous patience and unwavering support.

It has been an honour to complete this research under the supervision of these two great supports.

I give special thanks to the two extraordinary ladies in my life, my wife Amanda and daughter Bethannie. Words cannot describe the remarkable dedication you have both endured during some of these most challenging times in our journey together.

Lastly, I want to thank all my friends and family for their untiring kindness, support and understanding during all those missed moments and weekends spent on this dissertation.

D.Perrett

Table of Contents

Table of Figures

Table of Tables

Table of Equations

List of Appendices

Acronyms and Nomenclature

- Alternating Current (AC)
- Amendment (AMD)
- Australian and New Zealand standards (AS/NZS)
- Apparent Power (S)
- Behind the Meter (BTM)
- Comma-Separated Values (CSV)
- Coordinated Universal Time (UTC)
- Direct Current (DC)
- Distribution Network Operator (DNSP)
- Distributed Generation (DG)
- Distribution Network Operator (DNO)
- Distributed Energy Resource (DER)
- Direct Normal Irradiance (Dni)
- Diffuse Horizontal Irradiance (Dhi)
- Data Frame (df)
- Energy Storage (ES)
- Energy Queensland Limited (EQL)
- Front of Meter (FOM)
- Global Horizontal Irradiance (Ghi)
- Global Tilted Irradiance (Gti)
- High Voltage (HV)
- Local Network Service Providers (LNSPs)
- Lack of Reserve (LOR)
- National Metering Identifier (NMI)
- National Electricity Rules (NER)
- Overhead (OH)
- Powerlink Queensland’s (PLQ)
- Photovoltaic (PV)
- Pulse Width Modulation (PWM)
- Reactive Power (Q)
- Real Power (P)
- Small-scale Technology Certificate (STC)
- Transmission Network Operator (TNO)
- Underground (UG)
- Visual Studio (VS)

Chapter One – Introduction

Background

In recent decades, there has been a considerable shift in societal awareness toward environmental concerns; a significant section is using non-renewable resources for electricity generation. This has sparked massive growth in the renewable energy sector, especially towards wind and solar generation. The traditional electrical uniform energy system from generation to load is transitioning to include diverse and complex arrangements including but not limited to distributed energy resource (DER) like household rooftop solar and energy storage (ES) like community battery storage. Figure one demonstrates a generalised present view of what the future electrical grid might comprise. These variations to the traditional energy system come with unfavourable challenges to maintaining a stable and efficient electricity network. This electrification transition is driven by numerous factors happening throughout different levels of society, focusing on reducing environmental impacts.

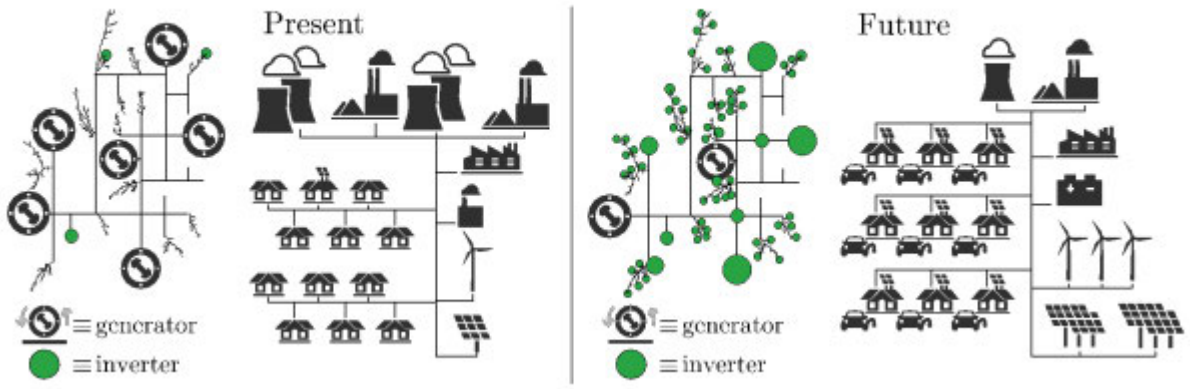


Figure 1: Present and future power system [1]

Australia's government has committed to target net zero greenhouse gas emissions by 2050. The current road-mapped goal of reducing pre-2005 levels by 2030 has been predicted to be overachieved by at least five percentage points. The electricity sector is a strength to meet this emission decline with an anticipated reduction of 60% pre-2005 levels; this is imparted due to the strong uptake in the transition to renewables [2]. One vital renewable energy source that is versatile and essentially does not cause environmental harm during operation is solar photovoltaic (PV) energy. Over three million rooftop solar PV systems were installed across Australia on 31st January 2022 [3]. This growth in solar PV has been propelled by several influential factors, including rebates from the Australian government as financial

incentives in the form of small-scale technology certificates (STC), allowing savings on eligible installed solar PV and solar hybrid systems—the reduction in the cost of solar panels and manufacturing of system technologies. Lastly the energy is sourced from the sun, has no direct sourcing costs and generally abundant throughout Australian seasons.

Until the turn of the century, a radiality approach was adopted in distribution network design. From a distribution network operator (DNO) perspective, widespread high penetrations of DERs influence power flow throughout the distribution network. One of those strategic energy resources is solar PV, which is undergoing unprecedented growth worldwide, especially here in Australia [2]. The optimal operation of PV is highly dependent on the flux in radiant energy per unit area, more commonly known as irradiance, rendering these systems highly reliant on weather and seasonality conditions. Increasing solar PV integration throughout the distribution network directs importance toward adequate power quality management and that end-user voltages are not adversely influenced [4].

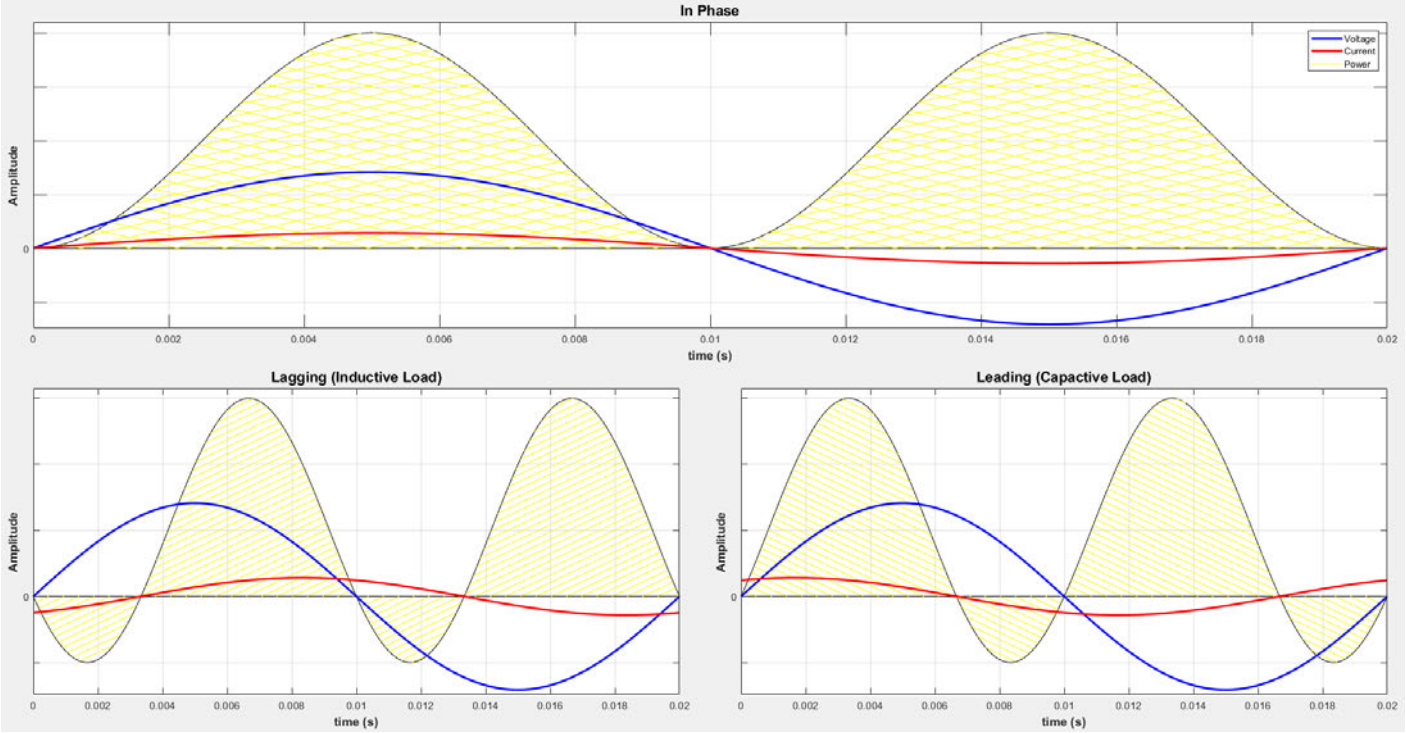


Figure 2: One cycle (@50Hz) of Voltage, Current and Power diagrams [5].

An understanding of instantaneous and average power must be explained to understand the difference between true power factor and displacement power factor. Consider the scenario of a purely resistive circuit; the instantaneous value of power (P) in watts (W) is equal to the product of its voltage (V) in volts and current (I) in amperes (A) at a moment in time. In an in-phase alternating current (AC) circuit, as depicted by the waveform in Figure two, power continually oscillates between zero and peak values; therefore, the average power value is

taken over the wave cycle [5]. In an inductive or capacitive circuit, the energy is not dissipated the same as in purely resistive components; the energy transfer process reflects to the generator [5]. In Figure two, the reflective characteristic of each type of circuit can be demonstrated by two lower graphs: positive is useful power and negative is reflective or dissipated by the circuit components.

Let instantaneous voltage defined be:

$$v = V_i \sin \omega t \quad \text{Equation 1}$$

And instantaneous current is defined as:

$$i = I_i \sin(\omega t - \varphi) \quad \text{Equation 2}$$

The current wave is a phase shift relative to the voltage waveform, i.e., zero or in-phase, leading or lagging. By multiplying together Equations one and two to determine instantaneous power, the corresponding equations can be defined as:

$$power_i = V_i \sin \omega t \cdot I_i \sin(\omega t - \varphi)$$

$$= \frac{1}{2} V_i I_i [\cos \varphi - \cos(2\omega t - \varphi)]$$

$$= \frac{1}{2} V_i I_i \cos \varphi - \frac{1}{2} V_i I_i \cos(2\omega t - \varphi) \quad \text{Equation 3}$$

This can be further refined by removing the second component of Equation three that can be established as non-contribute to the average active power used from generation, where V and I become r.m.s. values:

$$P = \frac{V}{\sqrt{2}} \cdot \frac{I}{\sqrt{2}} \cdot \cos \varphi$$

$$P = VI \cos \varphi \quad \text{Equation 4}$$

Therefore, further defining the reactive (Q) components of power termed volt-amperes (vars):

$$Q = VI sin \varphi \quad \text{Equation 5}$$

The apparent power (S) in voltampere (VA) can be further defined by Equation six:

$$S = \sqrt{P^2 + Q^2} \quad \text{Equation 6}$$

Furthermore, the ratio between active and apparent power is characterised as the power factor defined by Equation seven:

$$power\ factor\ (\lambda) = \frac{P}{S} = \frac{P}{VI} = \cos\varphi \qquad \text{Equation 7}$$

Therefore, the displacement power factor and the current are either lagging or leading in reference to the supply voltage at a fundamental frequency with the increase of non-linear loads, such as invertors onto the electrical network, introducing harmonics impacting the effective power factor and the calculation of Equation seven [5]. Therefore, the distortion factor Equation eight should be combined to determine the total or true power factor Equation nine [6].

$$Distrotion\ factor\ (\lambda_D) = \frac{1}{\sqrt{(1 + THD^2)}} \qquad \text{Equation 8}$$

$$\therefore\ True\ Power\ Factor\ (\lambda_T) = \lambda_D\ \cdot \lambda \qquad \text{Equation 9}$$

The intermittent and diurnal output of renewable generation can lead to temporal mismatches between the balance of power generation and consumption [7]. Though this occurrence is characteristic of renewable energies, it is more noticeably observed during the daytime, when solar PV is more susceptible to environmental conditions, for example, localised irradiance [7]. Increasing solar-PV penetration throughout the distribution network contributes to the falling demand during daytime periods. The increase in leading power factor trends of the end-user and the reduction of synchronous thermal generation, especially during the early morning hours, limits adequate management of voltage excursions, especially overvoltage events [8].

Power quality is essential in all electrical networks; in an AC system, this represents energy efficiency. The power factor is a ratio between real power and the apparent power delivered to a load, numerically expressed from zero (0) to one (1.0). Any deviation from a unity power factor 1.0 increases generation, transmission, and distribution costs. This becomes increasingly obvious since real power is proportional to the square of apparent power and the current being supplied [5]. This increase would be represented in the current carry capabilities of all layers throughout the electrical system, like generation, conductors, and transformers. Some commercial and industrial customers with additional load considerations can have a typical load power factor between 0.9 and 0.98; however, this is represented in the Tariff charges. Therefore, it is advantageous for those industries to implement power factor correction for site considerations.

Problem Statement

The increase in solar-PV penetration throughout the low-voltage distribution network imposes challenges for distribution network operators (DNO) as it enables bidirectional power flow [9]. With this growth of renewable generation, like solar-PV, utility operators worldwide are observing an increasing trend in the capacitive reactive power flow throughout distribution networks [10; 11]. Depending on the level of PV density, this trend can impact the network locally or system-wide and can be either dynamic or steady state in nature [12]. This trend is particularly evident and challenging in ensuring adequate management of distribution network power quality and end-user voltages are not adversely influenced [4].

Most existing inverter control mechanisms like maximum power point tracking (MPPT) and pulse width modulation (PWM) controllers with phase lock loops (PLL) techniques already manage voltage and current variances during optimal solar PV operating conditions [13]. Legacy grid-connected inverters were typically operated at unity power factor for all outputs between 20% to 100% of unit rating and shall operate between ranges of 0.8 leading or lagging [14; 15; 16]. However, there is minimal control to counteract reactive power contributions during low irradiance conditions. It is acceptable for the reactive power contribution of a system to be controlled that the vars are less than the defined lowest operating percentage of unit rating [15; 16]. These functional limits depend on the relevant version of grid connection of energy systems via inverters standards part two as outlined in Australian and New Zealand standard (AS/NZS) 4777.2 [16].

While the reactive power contribution of a singular small-scale rooftop solar-PV system during low irradiance conditions may be negligible in an isolated instance, the cumulative effect in an increasing capacitive load environment is of greater significance. Particularly when coupled with other trends impacting reactive power flows, like changes in end-user household appliances from purely resistive load characteristics and increases in underground electric development (URD), especially within the distribution network. The effects, particularly towards voltage management, are more apparent and problematic where solar-PV densities are more significant [12].

Research Project Objectives and Scope

The project aims to identify reactive power contributions of grid-connected end-user solar PV. Due to solar PV's proliferation, analysis of lower limits of AS/NZS4777.2 during low irradiance conditions is to be achieved. An investigation will examine the cumulative contributions towards the disturbance of voltage variation from a steady-state nature.

Project achievements to be captured by the conclusion of ENG4112:

- Literature review investigating reactive power impact challenges focusing on an increasing capacitive load environment.
- Verification of modelling techniques to apply a purposeful methodology to the project.
- Identify local end-user solar-PV reactive power contributions leveraging Energy Queensland Limited (EQL) supplied data focussed on low irradiance conditions.
- Build or apply verified methodological techniques for regression modelling of identified contributions.
- Determine and report on potential corrective measures.

Advancing achievements if time and resources permit:

- Extrapolate method to define potential future reactive contributions with the intent of a forecasting algorithm.
- Apply algorithm/s to future forecasts for increased solar penetrations.
- Application with the use of specialised software to a known feeder.

The project scope will focus on the cumulative reactive power contributions from end-user grid-connected solar PV in the distribution network. The review will cover the impacts of an increasing capacitive load environment from a transmission network operators (TNO) view but focus on the DNO perspective. The key will be to represent the implications of the cumulative contribution towards voltage management for the DNO. The review will also seek to include any implications that may apply to grid-connected equipment, whether physical or financial. There are minimal inverter control mechanisms that counteract reactive contributions during low irradiance; therefore, the aim will be to model a forecast with various levels of local feeder solar PV penetrations.

Table 1 below covers the AS/NZS related to the project's scope.

Document ID	Description	Date Published
AS4777.2-2005	Grid connection of energy systems via inverters, Part 2: Inverter requirements	20/05/2005
AS4777.2:2015		10/09/2015
AS4777.2:2020 AMD 1:2021		12/18/2020 AMD: 10/01/2021
AS/NZS IEC 60904.1:2023	Photovoltaic devices - Part 1: Measurement of Photovoltaic current-voltage characteristics	06/2023

Table 1: AS/NZS Standards scoped throughout the project [14; 15; 16; 17]

List of critical assumptions or standard definitions applied through the project:

1. Linking between NMI and Irradiance was through postcode/substation connection using a secure network (Desensitised correlation)
2. Correlated data points were within a 10km radius of Irradiance locations.
3. Inverter output ratings are linked through a secure network. (Desensitised correlation)
4. The low irradiance condition is $< 200\ {}^{W}/_{m^2}$ defined by AS/NZS IEC 60904.1:2023
5. Nominal (steady state) Voltage: 230V AC
6. Mitigation of measurement data error where the power factor is greater or equal to $\pm$ unity assumption was adjusted the power factor as unity.
7. Negative solar current events were found within the measurement data due to reverse magnitude; the assumption is the CT polarity was reversed.
8. The power factor for data collection was implemented using the IEEE convention.
9. The power factor applied to the aggregate load will be used for solar power calculations.
10. Singular Inverter connections were assumed.
11. The latest AS/NZS 4777.2:2020 was used for compliance comparisons.

Ethical and Confidential Considerations

The confidentiality agreement entered with Energy Queensland Limited certifies that any results that may identify individuals are not to be shared with any third party. Restricted access to the data set is to be processed via secured network admittance and distributed under the supervision and support of Energy Queensland Limited. Effective oversight is in place to protect confidential customer information and reduce the capability of inferencing customer locations or subsets of customers.

Significance of Project

TSO and DSO of electrical networks are responsible for adequately managing the electrical infrastructure for a safe and reliable electricity network. The increase in renewable distributed energy generation throughout different levels of the network rather than a traditional centralised generation is altering the temporal delivery approach and operations throughout the grid network. One significant difference is the changes in reactive power flow, particularly the increase of capacitive flow. This variance of reactive power flow during low-demand periods can significantly alter the network operation's reliability, especially the susceptibility to overvoltage events. The research project will examine the contribution of solar PV systems on reactive power flow and research in modelling solutions to apply to the forecasted increases in PV penetrations on the distribution network.

Report Overview

The report is divided into five chapters.

CH 1. – Project Background introduces the basics of the project elements, Problem Statement outlines the project, Objectives, Scope, and miscellaneous aims to clarify the bounds in which the

CH 2.– Literature Review comprehensively evaluates applicable literature to establish critical analysis.

CH 3. – Methodology derives the method undertaken using existing fundamental methods to frame the reactive power contribution of grid-connected solar PV.

CH 4. – Concept Development and Results analyse the data related to various project applications and present the corresponding models.

CH 5. – Discussion and Conclusions critically reviewed the results in the context of the literature evaluated, including future work that could advance the project outcomes.

Chapter Two – Literature Review

Introduction to the LR Chapter

This chapter studies the trends of reactive power contributions impacting voltage rise and fluctuations. The literature review will report on subsequent contributing factors towards reactive power and its additional cumulative effect on the transmission and distribution networks. The aim will pivot to focus on the contributions of grid-connected solar-PV and review methodologies towards real and reactive power demands that the utility experiences from a distribution network perspective.

Contributors to Reactive Power

Power Lines

The transfer of electrical energy from the generation to the end-user has been accomplished using OH power lines. Reactive power in high voltage (HV, >110kV) transmission lines contribute to total reactive power demands transmitted. A power line's reactive contribution is dependent on characteristics of length, composition, and power level transmitted [18]. The voltage levels in the line are influenced by operating outside the acceptable loading conditions and will consequently generate a more significant reactive power contribution when in an underloaded state [18]. An article from Slovakia found the apparent contribution to capacitive reactive power from OH lines increased when operated at higher voltage levels; this is visualised through Figure three below comparing reactive power of low voltage (LV, 400V) and medium voltage (MV, >400V <110kV) lines.

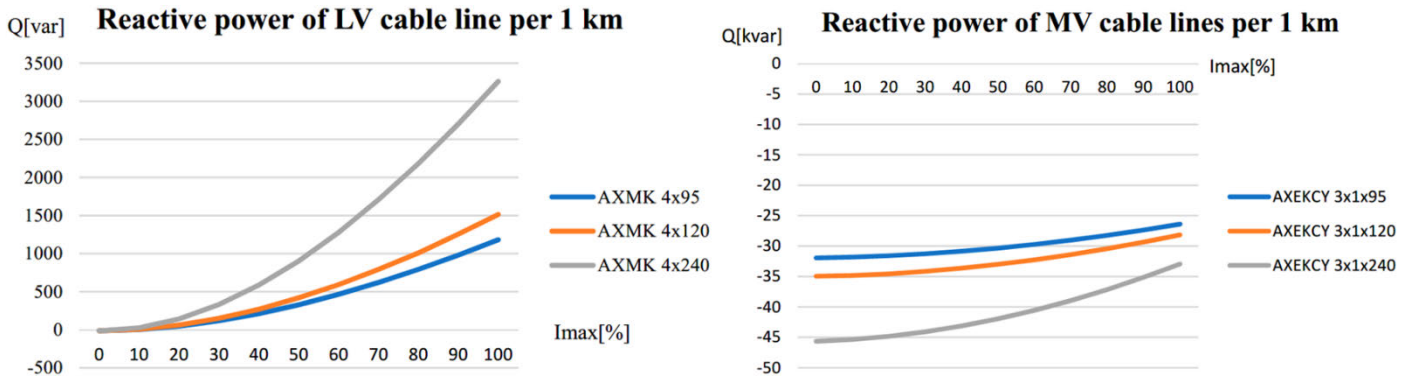


Figure 3: Comparison of LV and MV reactive power per 1 km [11]

Compared with traditional OH lines, there are advantages in the integration of UG cabling. One main advantage is the decrease in external influences the line must endure. Minimising the above-ground environmental impact allows considerable grid security and reliability

increases. The disparity between the two-line types is the sizeable capacitive characteristic of UG cabling. Karki et al. [19] found that UG cabling inherent capacitive reactive power contribution can accomplish functionality similar to shunt capacitors when used with OH lines. Therefore compensating the reactive power contribution of a loaded OH line [19]. Research by Christos et al. [10] investigating reactive power demands on UK distribution networks similarly found that a higher percentage of UG cabling penetrations resulted in extra injections of reactive power flow. Correspondingly, as Karki et al. [19] stated, the voltage increases from MV (33 and 66kV) to HV (132kV), and so too the capacitive nature of reactive power flow experienced from the cable [10]. Reactive power contributions from the lines in the network have increased due to the construction of new cabling being extensively UG cables. This is becoming more evident throughout the distribution network as municipal zones expand and some urban zones draw the same approach. This trend aligns with the increased capacitive power flow characteristics experienced by both the TSO and DSO.

The electrical energy generated flows throughout the grid through transmission and distribution network lines. These lines produce energy loss due to the electrical resistivity of the conductor material. In Australia, these network losses account for approximately ten percent between generation sources and end-users [20].

Trends in End-User Energy

The trend of usage characteristics of the typical end-user has changed, mainly throughout the distribution network layer during the last couple of decades. This occurrence is developing due to numerous factors; one of the primary considerations, yet small, is the change from purely resistive devices. The design methodology of household devices has undergone some minor and considerable development changes. This change can be linked to the increased number of devices in the typical household having more features integrated into the overall operation [21]. Domestic devices increasingly seem more interconnected, and function extensions come with increased integration of semiconductor componentry and power electronics to many modern-day household appliances [11]. Tazky et al. [11] monitored the secondary side of a modern housing estate 22/0.4 kV transformer with a compact power quality meter. Figure four exhibits the active and reactive power flow over a chosen 24-hour period. The significance recognised is that the power factor shifted capacitively during

minimum load times. It must be noted that these transformer measurements were carried out in parallel with the MV cables of Figure three [11].

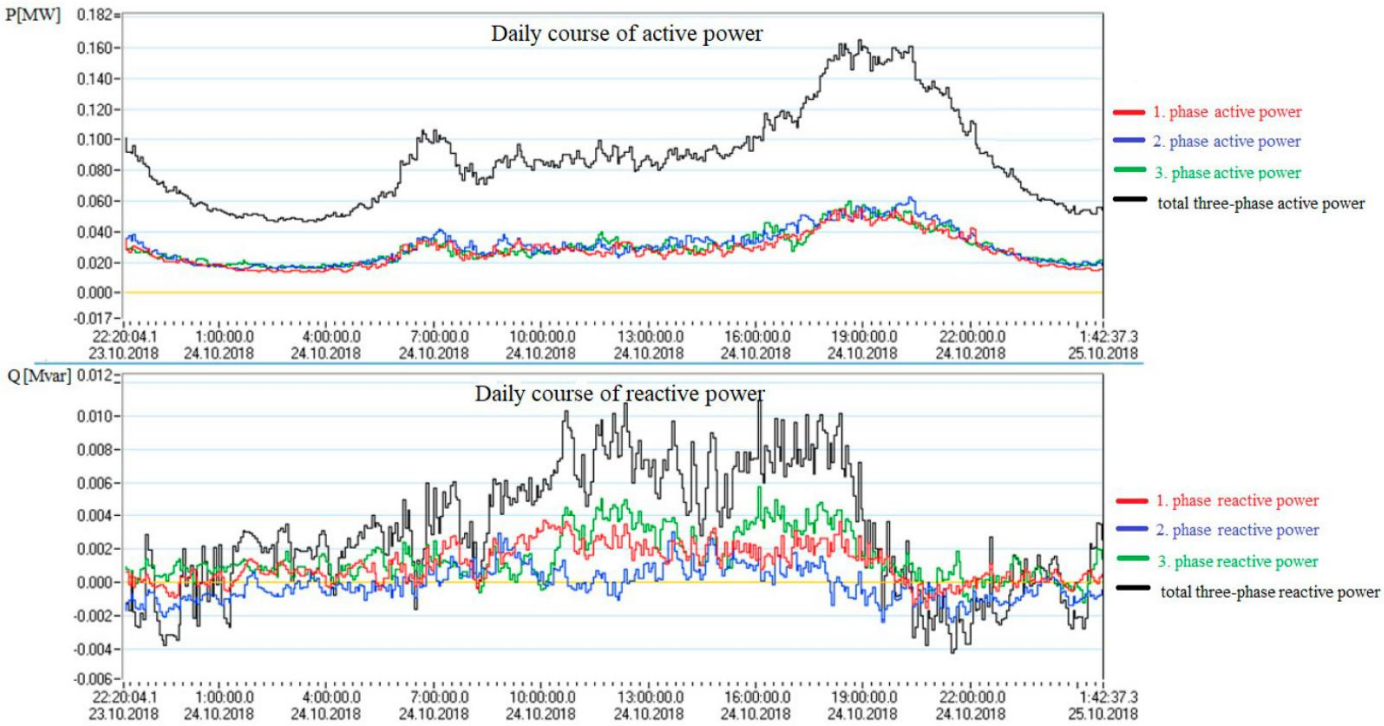


Figure 4: Transformer station active and reactive power flow, observing capacitive reactive power trend onto the MV network [11].

There is an undesirable and parasitic nature to appliance energy consumed behind the meter (BTM). The effect of this was experienced regardless of whether the appliance was actively displaying something like time or on standby awaiting input from a remote or storing memory for future uses [21]. Many studies have previously been focussed on active standby power consumed as it is the kWh usage measurement that is Tariffed and is of importance to the household end-user perspective. Another influence affecting relevant studies into reactive power contributions is the input from a single device has historically been minimal and insignificant in impacting power quality concerns front-of-meter (FOM). Niitsoo and Palu [21] found that seven out of the ten devices measured on standby mode resulted in capacitive power flow. The other significant result noticed from the study is the net reactive power of all the tested appliances while under primary operating conditions. They observed that the appliances' net reactive power contributes to the overall household as a capacitive load.

End-users are shifting away from traditional lighting designs towards alternatives to become more energy efficient. These alternatives include gas halogen lamps, compact fluorescent lamps (CFLs) and light-emitting diodes (LEDs). Wlas and Galla observed that the vast implementation of LEDs throughout new facilities influences quantifiable and cumulative

reactive power contributions beyond designed characteristics [22]. The reductions in real power demand were found to be around 30%. Still, the four LED light sources subjected to analysis exhibited considerable reactive power contributions measuring between 27 and 21 vars. However, one bulb may not be significant; the cumulative effect of several leads to be considered. For example, if an arbitrary value of 15 bulbs per residence in a new complex of 100 units is used, the accumulative effect if 20% of those LEDs are operating accounts for 7.2kvars of capacitive load. That would influence the overall power quality management of those local connections.

Distributed Energy Resources

There are three main design and planning areas in the approach to a traditional distribution network philosophy: construction design (near-immediate), network planning (near future) and strategic or long-term planning [23]. A traditional radial approach to grid design between generation and load has been the norm until the turn of the century. The transformation to renewable DERs integrated throughout transmission and distribution networks is causing temporal mismatches between demand and supply. The challenge is amplified by intermittency, diurnal nature, and decreased synchronous thermal generation like coal or gas. There is a saturation limit on renewable energy sources able to supply adequate grid stability [24]. Observations by Xu et al. [24] of grids studied throughout Japan could only reach solar-PV and wind turbine penetrations of 25 and 60 per cent, respectively [24]. This is evidence that transitioning to sole renewable energy sources using traditional implementation processes requires diverse rotational generation to ensure stable and reliable energy distribution.

The energy transition on the network commands essential technical forecasting techniques applied to historical and trending data to anticipate upcoming periods of concern. A fundamental area of concern of any energy supply system is operational reserve. Reserve energy can come in many forms and be used for different response concerns but has traditionally been related to the availability of synchronous generation. Renewable energy sources like solar and wind cannot typically be activated in response to shortfalls outside operating conditions. Queensland underwent a lack of reserve (LOR) conditions over nine days due to various forecasted and unanticipated timing of energy demands [25]. This change in uniformity between generation and load imposes reliable energy supply concerns but can have significant financial implications.

Temporal mismatches related to supply and demand are a constant challenge with distributed renewables. Transmission and distribution networks are experiencing a growing surplus of reactive power conditions. This can be attributed to the decline of daytime minimum demand periods, end-user trends to capacitive natured load, and reduced reactive absorption [8]. These are leading to increased over-voltage occurrences requiring appropriate and timely management. Figure five illustrates the increasing net average reactive power injection between 2015 and 2020 [8]. The resultant, if unchecked, can lead to equipment damage, reliability of supply and safety consequences. Powerlink Queensland's (PLQ) transmission network operator (TNO) assessment found to manage the over-voltage events, the switching of 275kV feeders and dispatch of reserve generation has become an emerging concern.

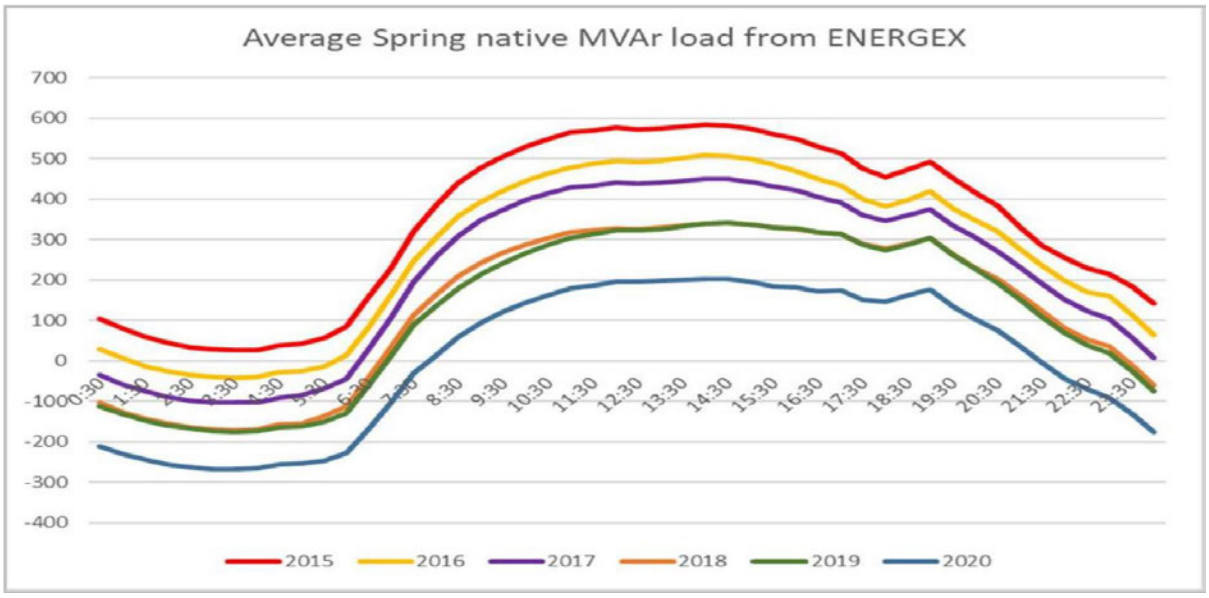


Figure 5: Queensland distribution network MVAr mean load over Spring [8].

Grid Connected Inverter

The purpose of the inverter is integral to a grid-connected solar PV system. Using designed internal hardware and software specifications, a solar inverter has two operation zones to control and interact. The first zone of operation on the DC side is to convert the current source generally supplied from various arrangements of solar panel arrays. A popular approach to DC/AC conversion is the technique of Pulse Width Modulation, typically implemented through the operation of a full bridge rectifier configuration with four switches (PWM) [26]. However, using the PWM method through high-frequency switching components produces higher-order current harmonics, influencing poor power quality events at the output [27].

The second zone of operation on the AC side is the output filtering control. The functional purpose of the designed filter is to attenuate those current harmonics produced through switching [28]. Illustrated in Figure six is a typical third-order LCL-type filter configuration. The inverter side (reference Li, Ri) of the L filter reduces the inevitable output in current ripples, whereas the second stage LC filter (reference Rf, Cf & Lg, Rg) diminishes the inevitable voltage ripples[28]. However, resonances are generated on the output AC voltage through the LC filter stage with the insertion of a time delay over the capacitance. They are controlled by switching the PWM [27].

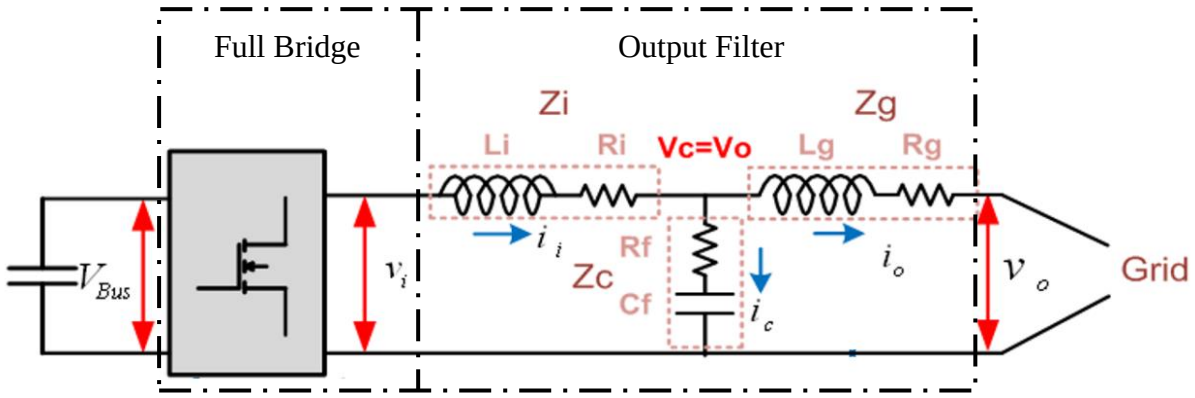


Figure 6: Typical Single Inverter Configuration [26]

Standards

Standards are norms established with a technical perspective to apply definitions around rules and guidelines related to a process, procedure, or production while improving safety, reliability, and performance. It is common practice that technical standards are generally published with inclusion and recognition from the applicable industry or institution locally or globally [29]. Two leading associations that contrast electrical and electronic implementation methods, including the power and energy industry, are:

- Institute of Electrical and Electronics Engineers Standards Association (IEEE SA)
- International Electrotechnical Commission (IEC)

The divergence between the two standard approaches can be visualised when evaluating the variance in the bi-directional metering approach between watts, vars and power factor. As introduced in Chapter One, when the current lags the reference voltage, the power factor is lagging, therefore an inductive load. When the current leads to the reference voltage, the power factor leads to a capacitive load. Figure seven illustrates the differences between the IEC and IEEE. The sign of the power factor as per IEC is indicative of the relationship between real power flow and independent of the load characteristics [30]. However, the IEEE

approach is meaningful in the relationship of the load characteristics and independent of the direction of real power flow [30].

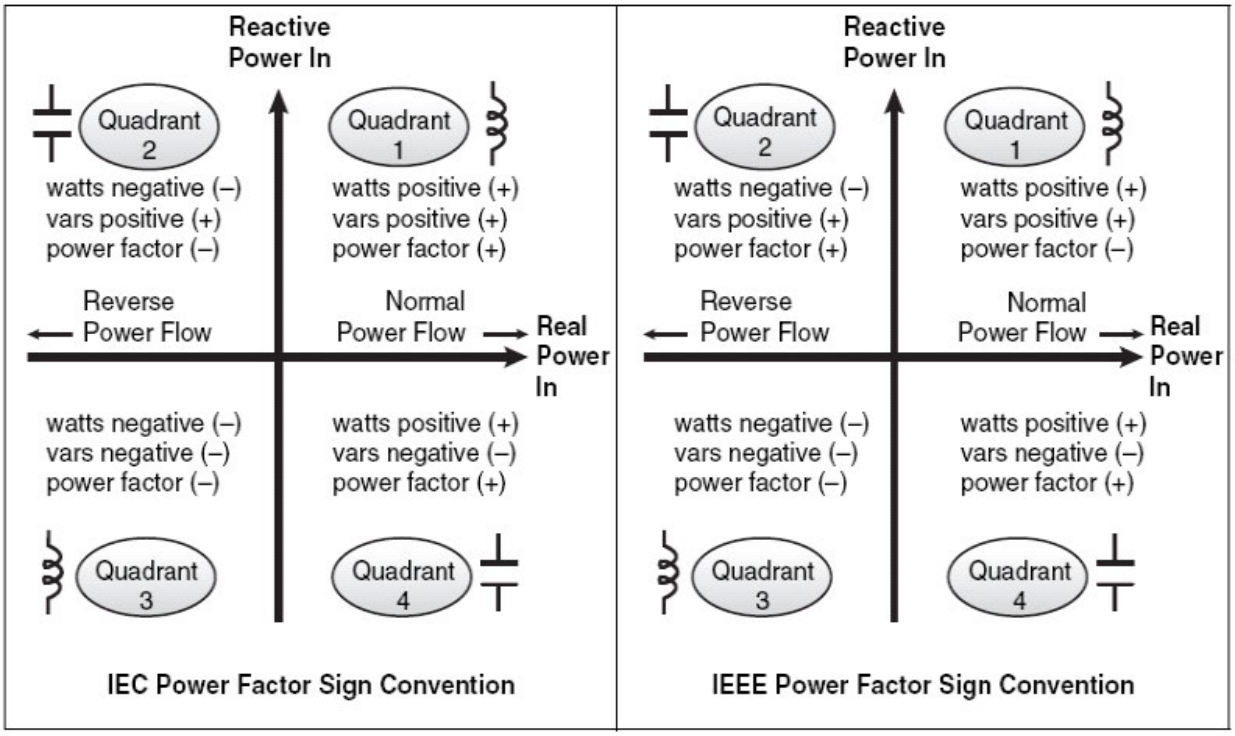


Figure 7: IEC or IEEE standard sign conventions [30]

4777.2 Grid Connection of Energy Systems via Inverters, Part 2: Inverter Requirements

Traditionally, end-user behind-the-meter (BTM) has been considered a uniform load with minimal operating complexities. Electrical technology trends are changing households' power flow interactions with the distribution network. Couple this with the uptake in rooftop embedded generation being established and operating in parallel throughout the network; BTM influences are more significant. This is indicative considering the version evolution of the AS/NZS 4777.2 Grid Connection of Energy Systems via Inverters, Part 2: Inverter Requirements. Since the inception of a three-part series of 4777 Grid connections of energy systems via inverters in 2002, part two has had four publication replacements over the past two decades. Two of these were released within five years of each other: AS/NZS 4777.2-2015 and AS/NZS 4777.2-2020, correlating with the increases in solar PV. There are many differences between these editions, but what is crucial to this project is the inclusions and revisions toward demand response and power quality modes [15; 16].

All three versions consider inverter operation from the perspective of the grid when evaluating either real or reactive power flow. Therefore, the lagging power factor sinks

reactive power such that the inverter acts as an inductive load and leading sources reactive power such that the inverter acts as a capacitive load [14; 15; 16].

- AS 4777.2-2005 outlined that for the range of operation between 20 to 100 per cent of rated output, inverters shall be in the range from 0.8 leading to 0.95 lagging [14].
- AS/NZS 4777.2-2015 revision outlined that for all current outputs of operation between 25 to 100 per cent of rated current outputs, inverters shall function at unity within the range from 0.95 leading to 0.95 lagging [15].
- The latest version AS/NZS 4777.2-2020 revision specified two levels rated apparent power output reactive power capability constraints as illustrated in Figure eight [16].
    - The reactive power absorption and supply shall be at least 60 per cent of the rated apparent power down to a 0.8 power factor [16].
    - The reactive power absorption and supply shall be at least 44 per cent of the rated apparent power between 0.2 to 0.6 power factor [16].

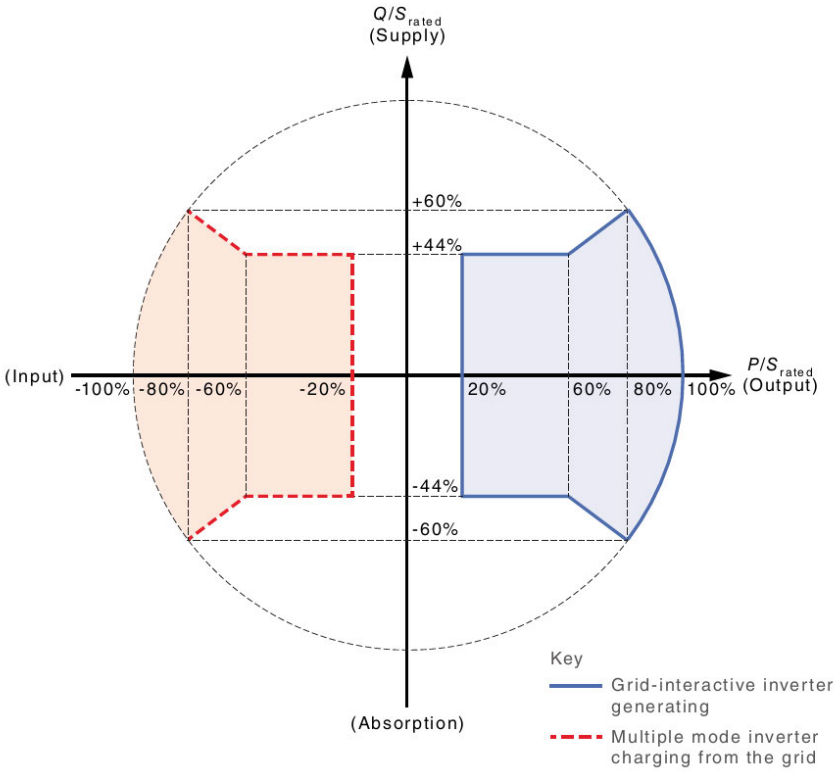


Figure 8: AS/NZS 4777.2-2020 Reactive Power Capability Constraints [16]

Though the earlier versions did not define the amount of acceptable vars that were drawn or supplied, in the AS/NZS 4777.2-2015 and AS/NZS 4777.2-2020, compliance was to be less than the vars provided at 25 and 20 per cent, respectively [15; 16]. There is a concise progression relating to the reactive power response capability between each AS/NZS 4777.2 publication. The AS 4777.2-2005 release states that with relevant DNO approval, inverter

operation can deviate from defined limits to deliver voltage support [14]. The later releases represent various inverter power quality modes intended to facilitate consistency towards managing power quality to localised grid connections. The different power quality modes available between AS/NZS 4777.2-2015 and AS/NZS 4777.2- AS/NZS 4777.2-2020 are illustrated in Table two.

AS/NZS 4777.2-2015	AS/NZS 4777.2-2020
i. Volt response modes	a) Volt-var response mode
ii. Fixed power factor or Reactive mode	b) Volt-watt response mode
iii. Power response mode	c) Fixed power factor
iv. Power rate limit	d) Reactive power mode
	e) Power rate limit

Table 2: Invertor Power Quality modes between 2015 & 2020 [15; 16].

Volt-Watt Response

The maximum real power output variation will be adjusted in response to a voltage deviation measured at the grid-interactive port [16]. These upper and lower set-point values can be seen in Table three, configured by four regions of operational compliance. Figure nine illustrates an example of a theoretical response. As of the AS/NZS 4777.2-2015 revision, if available, this mode shall be enabled by default [15; 16].

Region	Default value	V_{W1}	V_{W2}
Australia A	Voltage	253 V	260 V
	Inverter maximum active power output level (P) % of S_{rated}	100 %	20 %
Australia B	Voltage	250 V	260 V
	Inverter maximum active power output level (P) % of S_{rated}	100 %	20 %
Australia C	Voltage	253 V	260 V
	Inverter maximum active power output level (P) % of S_{rated}	100 %	20 %
New Zealand	Voltage	242 V	250 V
	Inverter maximum active power output level (P) % of S_{rated}	100 %	20 %
Allowed range	Voltage	235 to 255 V	240 to 265 V
	Inverter maximum active power output level (P) % of S_{rated}	100 %	0 % to 20 %
NOTE Australia C parameter set is intended for application in isolated or remote power systems.			

Table 3: Default set-points for Volt-Watt response mode [16]

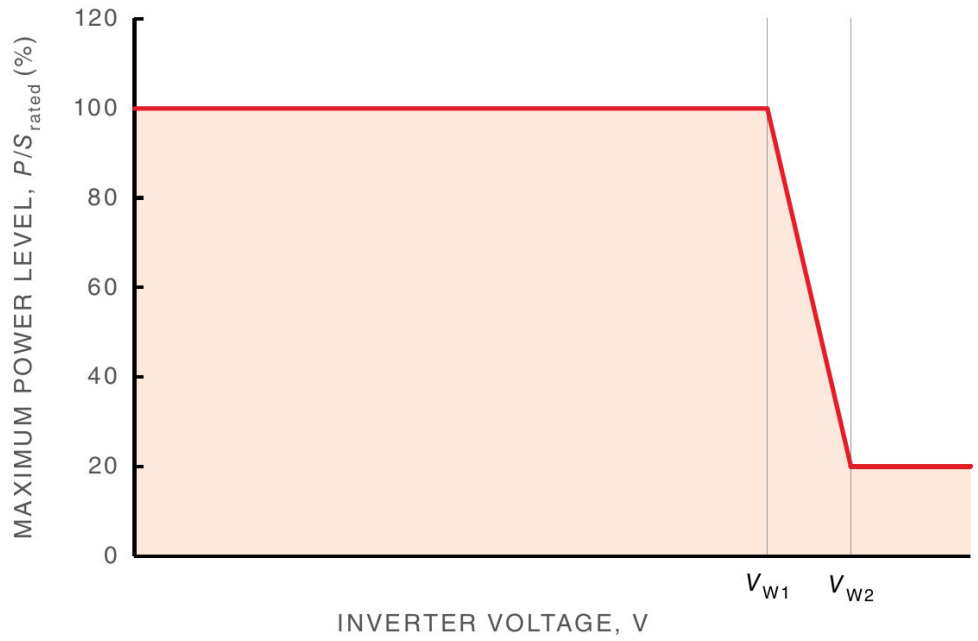


Figure 9: Volt-Watt response curve [16]

Volt-Var Response

A variation in the reactive power absorbed or supplied will be adjusted in response to a voltage deviation measured at the grid-interactive port [16]. These four set-point values corresponding to a reactive power response can be seen in Table four, configured by four regions of operational compliance. Figure ten illustrates an example of theoretical response as a percentage of VA-rated output. In the AS/NZS 4777.2-2015 revision, if this mode was available, it was to be by default disabled, but as of the latest release, it shall be a default enabled mode [15; 16].

Region	Default value	V_{V1}	V_{V2}	V_{V3}	V_{V4}
Australia A	Voltage	207 V	220 V	240 V	258 V
	Inverter reactive power level (Q) % of S_{rated}	44 % supplying	0 %	0 %	60 % absorbing
Australia B	Voltage	205 V	220 V	235 V	255 V
	Inverter reactive power level (Q) % of S_{rated}	30 % supplying	0 %	0 %	40 % absorbing
Australia C	Voltage	215 V	230 V	240 V	255 V
	Inverter reactive power level (Q) % of S_{rated}	44 % supplying	0 %	0 %	60 % absorbing
New Zealand	Voltage	207 V	220 V	235 V	244 V
	Inverter reactive power level (Q) % of S_{rated}	60 % supplying	0 %	0 %	60 % absorbing
Allowed Range	Voltage	180 to 230 V	180 to 230 V	230 to 265 V	230 to 265 V
	Inverter reactive power level (Q) % of S_{rated}	30 to 60 % supplying	0 %	0 %	30 to 60 % absorbing
NOTE 1 Inverters may operate at a reactive power level with a range up to 100 % supplying or absorbing.					
NOTE 2 Australia C parameter set is intended for application in isolated or remote power systems.					

Table 4: Default set-points for Volt-Var response mode [16]

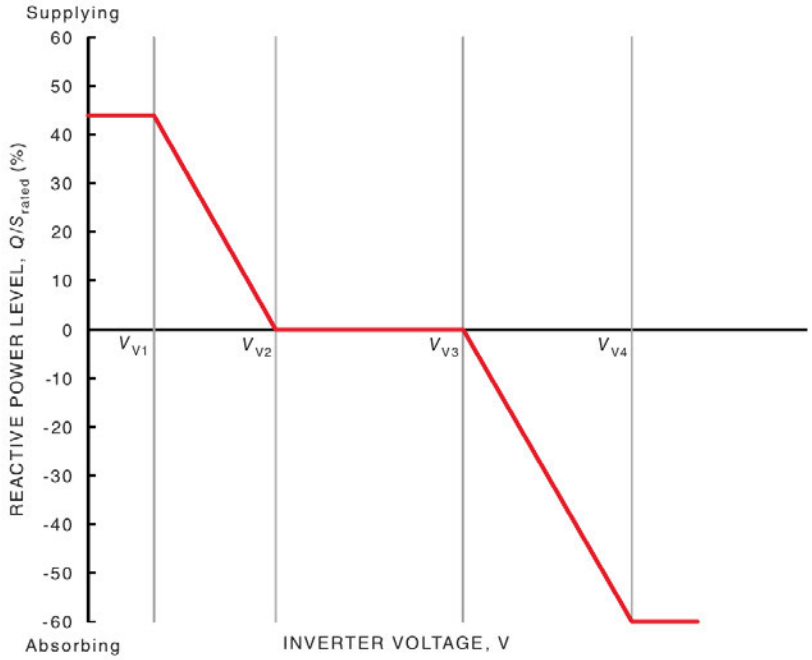


Figure 10: Volt-Var response curve [16]

Fixed Power Factor and Reactive Power

In some circumstances, the DNO may use the option to enable the fixed power factor mode or reactive power mode to stabilise localised grid demands. Both modes shall be able to operate concurrently with the volt-watt mode enabled but will be disabled by default [16]. AS 4777.2-2005 power factor requirements initially could be considered as a fixed mode for all ranges between 20 to 100 per cent. Whereas, from AS/NZS 4777.2-2015 onwards, it is defined by the optimal operation of the inverter. Figure eight illustrates an example of both modes included in AS/NZS 4777.2-2020. The default factor mode setting shall be unity for all other ranges constrained within 0.8 to 1 supplying or 1 to 0.8 absorbing [16]. Subsequently, the default reactive power (vars) mode setting shall be 0 per cent, while contributions shall be at least 60 per cent supplying or absorbing inverter-rated apparent power [16].

Methods and Models

Real and Reactive Power

Power forecasting is an integral part of predicting the changes in the electrical network to achieve a secure and reliable operation. Different models and methods are used for various purposes of electrical forecasting produced for either short-, medium- or long-term understanding [31]. Xiong et al. [32] found there are different electricity demand forecasting techniques available, for instance, exponential smoothing, regression, and adaptive models; fuzzy logic or interference systems; kernel-based or semi-parametric methods; grey-based

approaches, wavelet transforms, support vector machines, artificial neural networks, and generally any combination creating a hybrid model approach [32]. Throughout the past two decades, there has been an increase in the deployment of reactive power resources in all design stages, from network planning to construction. Power forecasting enables fundamental knowledge of network planning and anticipates appropriate demand responses, which include voltage control measures and reactive power ancillary services [33].

While modelling for generation and load has been significantly advanced in informing both real and reactive power considerations, research in power quality conditions at the distribution level has risen over the past decade. However, power quality has always been an essential consideration within electrical power systems but has become popular due to growth in trends that affect a secure and reliable supply. In the early 2010s, Great Britain's transmission network experienced unprecedented voltage excursions compared to previous years [34]. Kaloudas et al. [10] investigated and modelled the effect of different cable penetrations in 132 and 33kV circuits using monitoring data from 2012-13 [10]. The results highlighted an increasing trend in reactive power injection indicated at the grid supply points back to the transmission network [10]. In the years following, Kaloudas et al. [35] emphasised this growing trend was not an isolated scenario and that a combination of end-user interactions, DER, and cable penetrations was pressuring the challenges forecasted [35].

Power Quality

The relationship between supply and demand has influenced an evolution in identifying increased power quality concerns of end-user rooftop PV penetration on the distribution network. The characteristics of a PV energy system connected via an inverter and the intermittent, diurnal output can be challenging in achieving rigorous models. This unpredictable nature can be attributed to variations in irradiance, as Patsalides et al. [36] established. The study developed two cases of average winter and low irradiation cases in Cyprus. The report found during low irradiance events, the power factor reacts linearly with reactive contributions between one to two kvars [36]. The findings cover the daytime between 0630 and 1700 [36]. Another report by Cangi and Hasan [37] found a similar power factor linear relationship during the ramp-up in the morning when irradiance is lower than 220 wats per metre squared [37].

Irradiance

Solar-PV output depends on the irradiance levels sourced from the sun at an instantaneous moment. The modelling of the spectral irradiance incorporates many atmospheric factors that

alter the final value that may be expected at the earth's surfaces [38]. The term irradiance can be generalised by the sum of energy in watts per unit area as meters squared (units: W/m^2) at any instantaneous moment in time. A common approach in recording the mean irradiance values over a specified period. All magnitudes of irradiance involve a relationship between the sun and the relative surface; this is indicative of the defined terms below:

- **Normal** is a reference to the radiation beam being experienced on the surface [39]
- **Horizontal** as a reference to the ground surface [40].
- **Titled** the radiant energy received on the surface at a set tilt angle and azimuth [38].

In application, the following solar energy parameters used are:

- Direct Normal Irradiance (Dni) – Includes the extra-terrestrial spectral irradiance received at the atmospheric level minus loss factors corresponding to Rayleigh scattering, wave attenuation, water, gas and ozone absorption [38].
- Diffuse Horizontal Irradiance (Dhi) – is determined using the Dni with the cos(Z) of the solar zenith (Z) angle, including all values of radiation from the atmospheric level [38].
- Global Horizontal Irradiance (Ghi) – is the sum of both Dni and Dhi, where the solar zenith (Z) angle is accounted [40].
- Global Tilted Irradiance (Gti)
    - Fixed Tilt –a set tilt angle and azimuth or
    - Tracking – sun-tracking for optimisation of plant output.

Conclusion

The race to net zero emissions and environmental awareness revolution grows with every passing year. One of the critical sectors going through significant changes is the energy sector, particularly electrical energy generation [2]. These environmental aims for the coming decades are impacting all layers of the traditional grid system. The literature reviewed suggests considerable changes in reactive power flow throughout grid networks; however, this variation is not isolated and is being recorded in electrical grids worldwide. The phase shifts are not isolated to one area of the network but are an aggregate of recent upgrades occurring with increased speed. One of the most apparent areas of expansive integration is in end-user rooftop solar PV. Given this growth, minimal research is still linking irradiance variability and solar PV operation.

This dissertation will endeavour to bridge the gap between solar PV and low irradiance events with the analysis to focus on reactive power contributions during these conditions. Optimal operations are well-defined within the AS/NZS standards. Still, the lower operation ranges are less defined below 20 or 25% of rated demand from 2015 or 2020, respectively. This research aims to produce a concise understanding of cumulative reactive power contributions for solar PV penetration levels. Confirm PV compliance during less-than-optimal operating conditions and develop models to evaluate the increase in PV penetration in the distribution network.

Chapter Three – Methodology

Introduction

The data accumulated through EQL must have an initial examination to meet the defined assumptions and scope of the project. A list of assumptions from Chapter One has evolved and used to determine the methodology's implementation.

Data Procurement & Initial Examinations

With the support of EQL, access to a range of datasets has been authorised for use in this dissertation. The primary dataset includes end-user monitoring measurements from a previous strategic venture with industry partner recording changes aggregate to the end-user and solar system installed. The secondary dataset includes irradiance data at locations throughout Queensland; secure (desensitised) access was required to refine further information to link the two data locations.

The monitoring data has been procured over a recent three-year period and consists of over 10,000 separate files in Parquet format, encompassing around 4,500 unique end-user sets. Table five outlines the data points available for analysis:

Data	***Description***	***Unit***
Time Stamp	Displayed as: [year-month-day hour:minute:seconds]	μs
Grid Voltage	Bus Voltage	V
Grid Current	Aggregate Current for Load & Solar	A
Power Factor	Overall, Angle for Household [cos φ]	Ratio, (-1) to 1
Solar Current	Current for solar only	A
Impedance	Not used for project	Ω
Voltage Variation	Not used for project	V

Table 5: Available variables for analysis.

Initial examination established the datasets supplied by EQL as raw state and inclusive of additional variables that are out of scope for this project. Some preliminary data sorting and amendments were used for adjusting unnecessary erroneous data. The data process applied was unique to the project scope and, therefore, was designed in isolation and specific for data amendments in assisting data analysis within the dissertation. Figure 11 is an example of raw data from a randomly selected NMI.

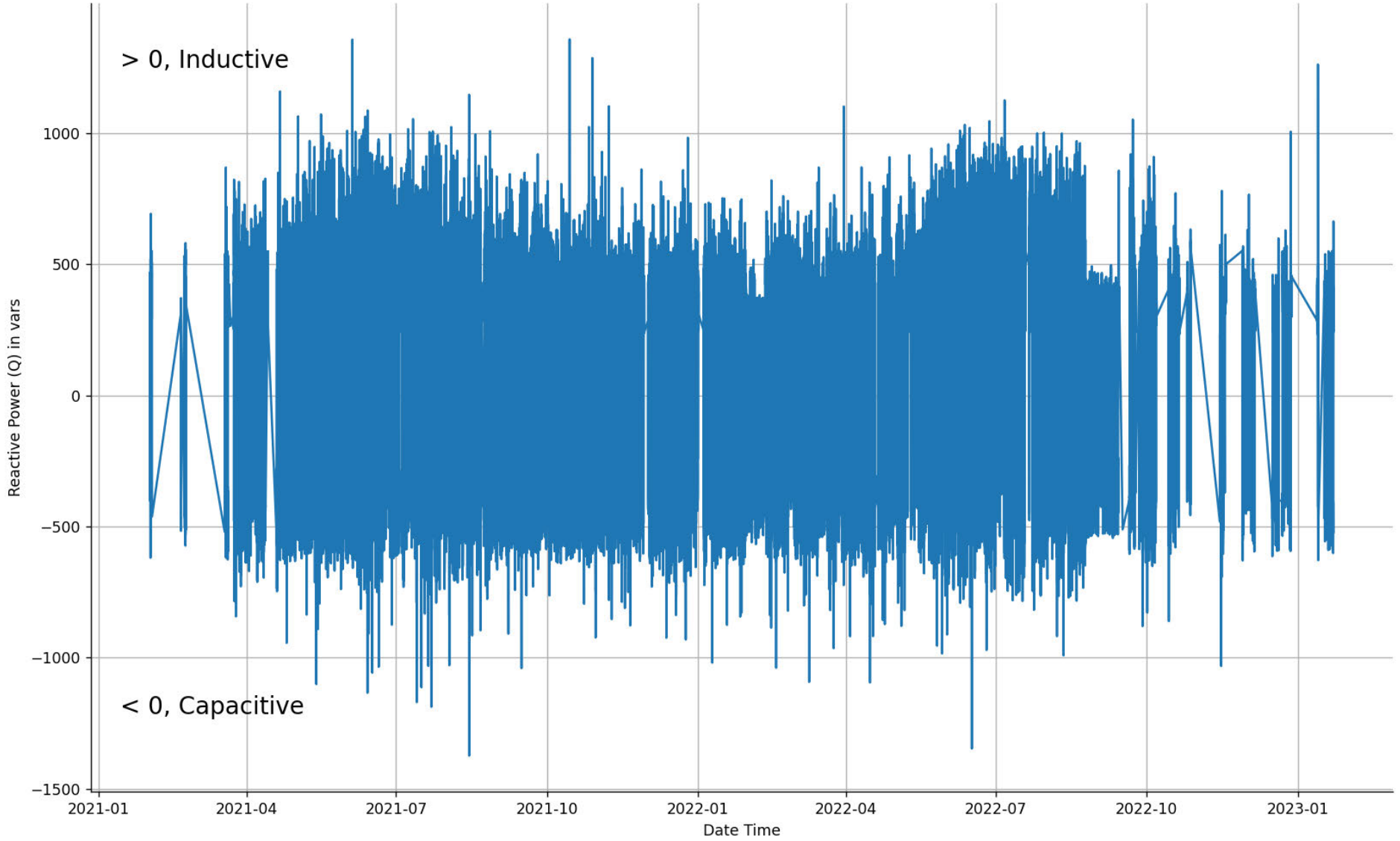
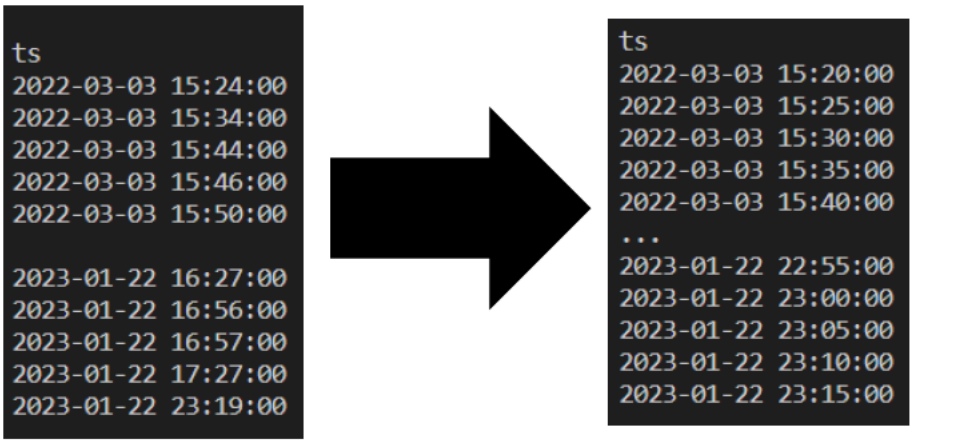


Figure 11: Example: Household Aggregate Reactive Power (Q) Contributions

A five-minute interval period was decided to connect the temporal resolution between the two datasets. The five-minute intervals were reflected in the irradiance dataset, but as illustrated in Figure 12, there are missing data points. Also, on initial examination, the temporal resolutions on the measurement data were inconsistent, as is visualised in Figure 11.



Unorganised Dataset Organised into 5-min Intervals

Figure 12: Initial Temporal Nature of Measurement Data

Irradiance

The second or irradiance data has been procured between 2007 – and 2022, encompassing 20 different locations throughout Queensland, and stored in a comma-separated values (CSV) file format. Table six outlines the data points available for analysis:

Data		***Description***	***Unit***
Location		The longitude and Latitude of each monitoring location	Angular °
Time Stamp	Period Start	Displayed as: [year-month-day hour: minute: seconds]	
	Period End		
Period		Descriptor (PT5M) period time, five minutes	
Cloud Opacity		How opaque the clouds covering divided into eighths of the sky.	oktas
Dhi		Diffuse Horizontal Irradiance	W/m^2
Dni		Direct Normal Irradiance	W/m^2
Ghi		Global Horizontal Irradiance	W/m^2
Gti Fixed Tilt		Global Tilted Irradiance	W/m^2
Gti Tracking		Global Tilted Irradiance	W/m^2
Zenith		Solar Angle	Radians

Table 6: Available variables for Irradiance analysis

Using an open-source mapping program with the Longitude and Latitude values, Figure 13 outlines available locations for the project. At the same time, there is data for all sites, but not all are in scope for this project. The geolocation technique linking the two datasets will be

managed between the Zone -substation, end-user National Metering Identifier (NMI), and closest irradiance monitor within an acceptable radius.

Figure 13: Irradiance Locations available

Radiant solar energy, or irradiance, refers to the amount of power per unit area at a given location. The quantification of this radiant energy is vital to the project, where variations in energy magnitude are impacted by environmental factors that transpire during seasonal changes. Introduced in the literature review, it was determined that Ghi encompasses tolerable loss factors, both extra-terrestrial and atmospheric; therefore, it will be used for modelling purposes. Figure 14 represents seasonal characteristics from a chosen site.

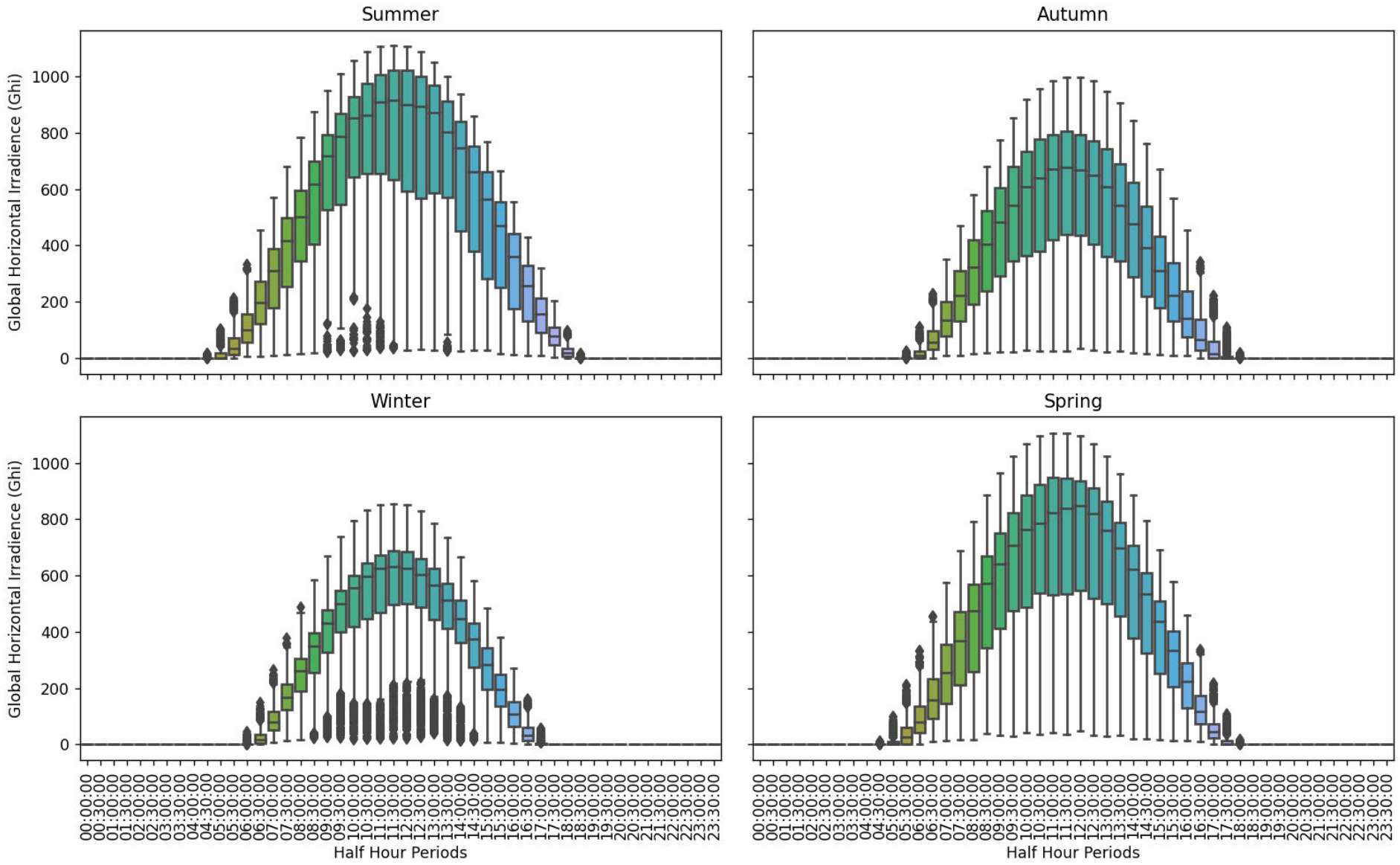


Figure 14: Ghi Irradiance Seasonal Example

Parquet files

Apache Parquet is an open-source data storage system designed to reframe traditional data organisation and uses columnar data representations [41]. This file format provides benefits where standard record-orientated or row-based data sets are lacking when handling larger data sets and leverages compression and encoding techniques implemented on the data [41,; 42]. The strategic function leveraged in the Parquet file format is the storage and use of metadata in the file footer. Figure 15 visually represents the metadata, column (chunk), and schema information included [41].

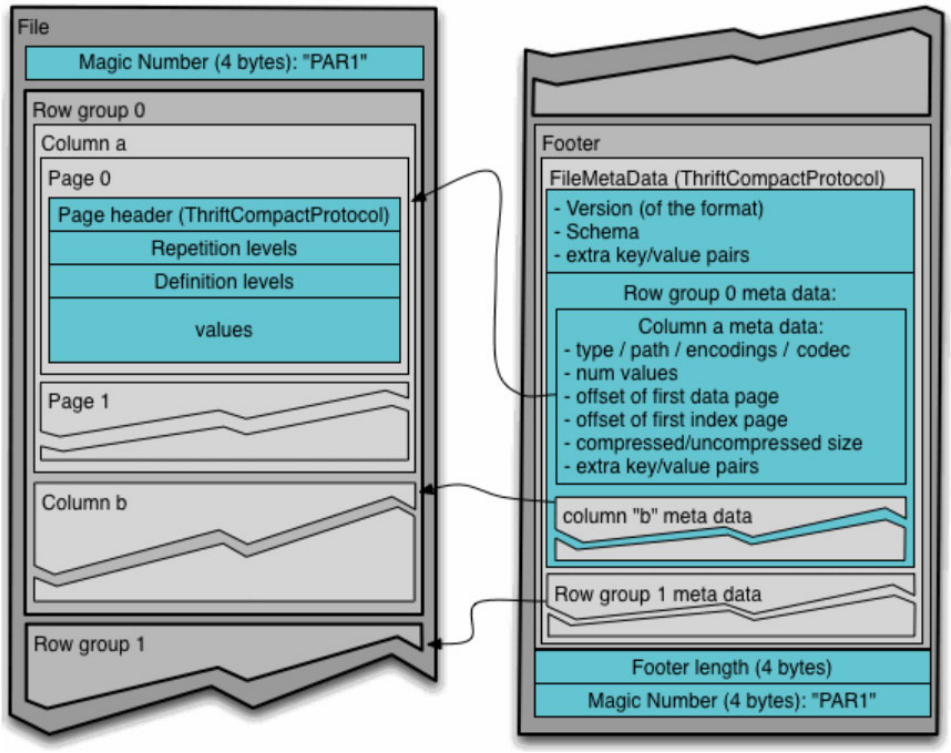


Figure 15: Parquet file representation, including metadata [41].

Another motivation and critical advantage of using the Parquet file format is language agnostic, giving flexibility in programming language choice. The project will use the latest version of Python as the primary language, using the Pandas library to manipulate data read/write and analysis processing—irradiance locations.

Energy Flow Guideline

The convention of energy flow is a determinant of viewpoint; Figure 16 illustration on the left details the market settlement and transfer solution database (MSATS) procedure and is specified from the DNO perspective and on the right from an end-user perspective which will be applied methodology throughout the project [43].

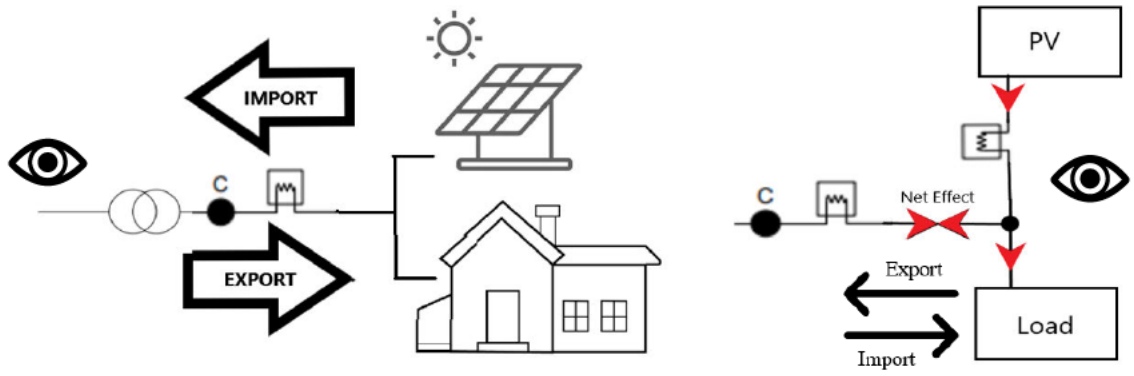


Figure 16: Simplified example of energy flow convention from DNO and the inverse project perspective (eye position == point of perspective) [43].

$$Net\ Energy = Export - Import$$ Equation 10

Explanation for Figure 16 example:

- All energy consumed by the end-user is considered an export from the grid network; therefore, the same convention export is used for the vars supplied.
- All energy the end-user generates is considered an import to the grid network; therefore, the same convention import is used for vars received.

What is a National Metering Identifier?

An NMI is a ten-digit unique character identifier for metered and unmetered electricity network connection points throughout Australia [43; 44]. It is the responsibility of the Local Network Service Providers (LNSPs) to assign an identifier under the NMI procedure, authorised by the National Electricity Rules (NER) at clause 7.3.1(d), (da) and (db) [43]. The primary purpose of NMI standing data is to positively identify a connection point geographical location, applicable network tariff and associated loss factors [45]. Some secondary functions include:

- Collation of metering data
- Work requests via the LNSP
- Transfers between retail electricity providers

There are numerous variations of connection points and registered NMI configurations. Still, Figure 17 illustrates one customer, one connection point, and NMI on the LV side of the transformer as an example [43]. It must also be noted that the NMI is not the same as the assigned metering number.

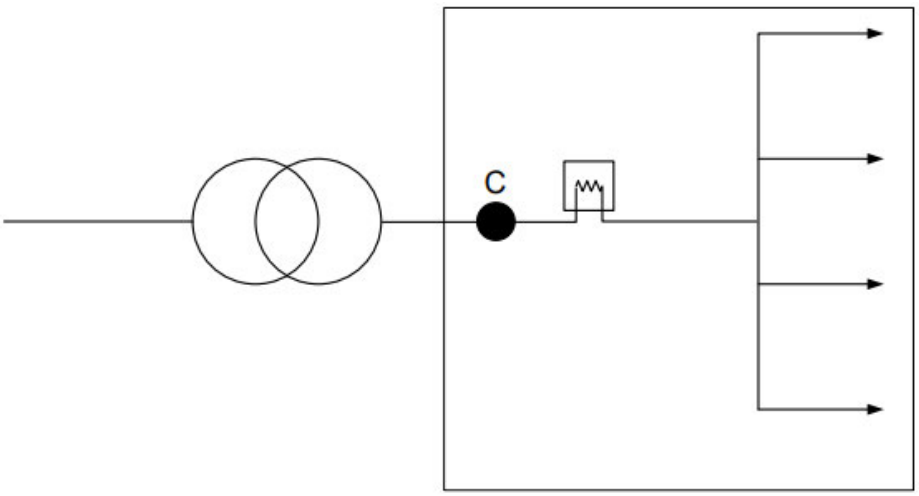


Figure 17: One customer connection point on the LV side transformer [35]

Calculations for S, P, Q & per-unit

To represent clarity, the sign applied on calculations represents the household's perspective and, therefore, the solar-PV system as portrayed on the right of Figure 16. Figure 18 illustrates an adopted representation of the IEEE power factor convention from Figure seven with slight adjustments to the sign of both the real and reactive power flow.

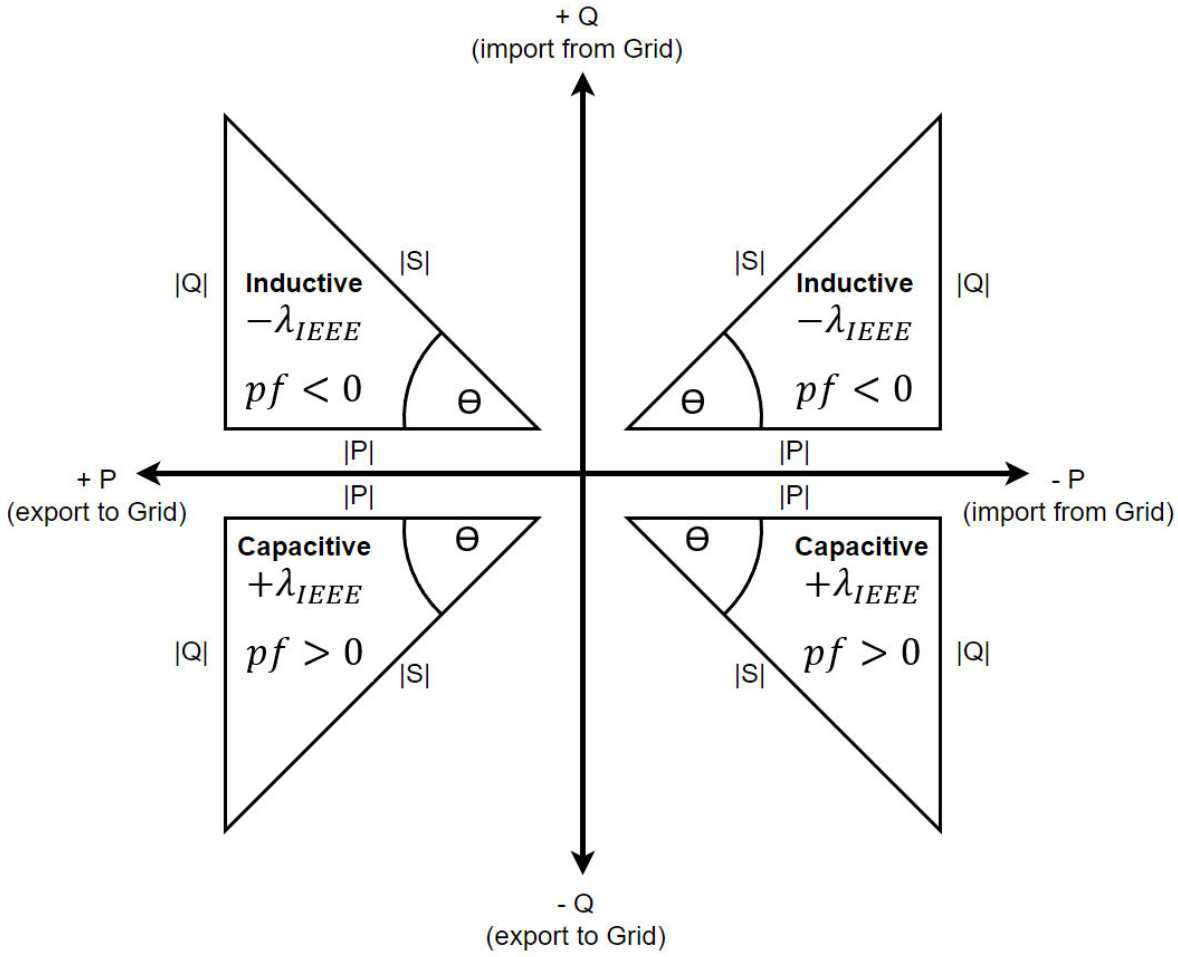


Figure 18: Project Power Factor and Power Triangle Convention [30]

Equation 11 represents the applied calculation for apparent power.

$$|S| = |V| \cdot |I_{NET}| \quad \text{Equation 11}$$

Using Equation 4 with substitution where $S = VI$, Equation 12 represents calculation for real power, including directionality.

$$P = -(|S| \cdot |\lambda| \cdot sign(I_{NET})) \quad \text{Equation 12}$$

Using a rearrangement of Equation 6, Equation 13 represents the calculation for reactive power, including flow directionality. However, the negative sign was removed when calculating PV power contributions due to the assumption that PV output current flow is positive within some negligible tolerance.

$$Q = -\sqrt{|S|^2 - P^2 sign(\lambda)} \quad \text{Equation 13}$$

Per Unit calculations will be applied represented as:

$$Per\ Unit = {Present\ Value}/{Base\ Value}$$

Where base value is reflective of site inverter-rated output as illustrated in Equation 14:

$$S_{pu}, P_{pu}\ or\ Q_{pu} = {Data\ Point}/{PV_{Rated}} \quad \text{Equation 14}$$

De-identification

De-identification and anonymisation are paramount in protecting clients' identifiers [46]. De-identified data has personally identifiable information removed or encrypted, so there is no reasonable likelihood of re-identification to any individual or a smaller subset of a group [46]. This is an important aspect when managing any controlled personal information. Throughout this project, the access to personal data will be stored and modified through a secure access-only source. Data sharing and results published throughout this dissertation will be generalised to a limited overview. However, a clear inference can be made between the data and Queensland at this project stage. Therefore, oversight will be with myself and the supervisor regarding the accuracy of the geolocation to be shared. If there are to be data components to be shared, a censoring and exclusion methodology will be implemented [46].

Methodology Flow Diagrams

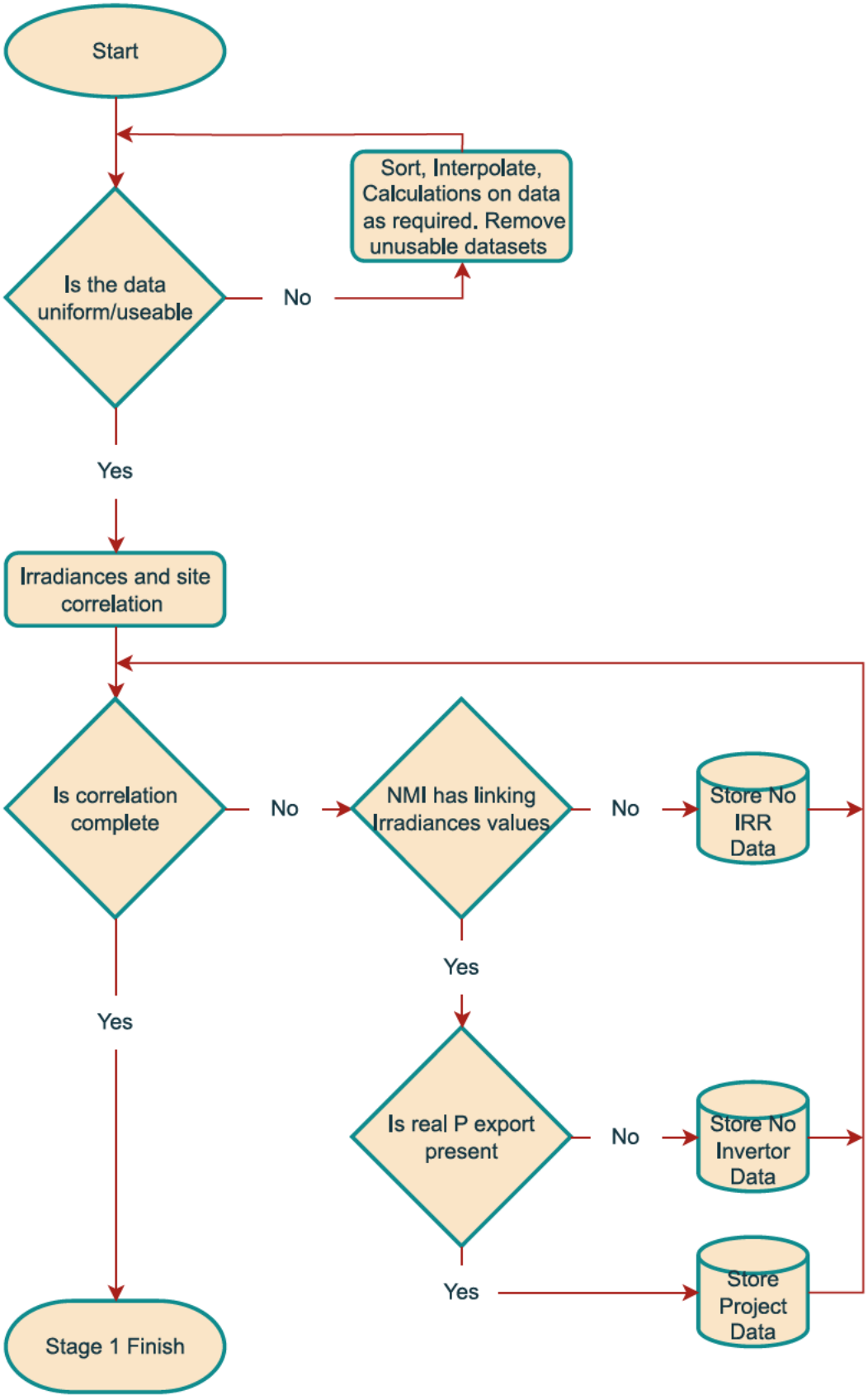


Figure 19:Stage One Methodology Flow Chart

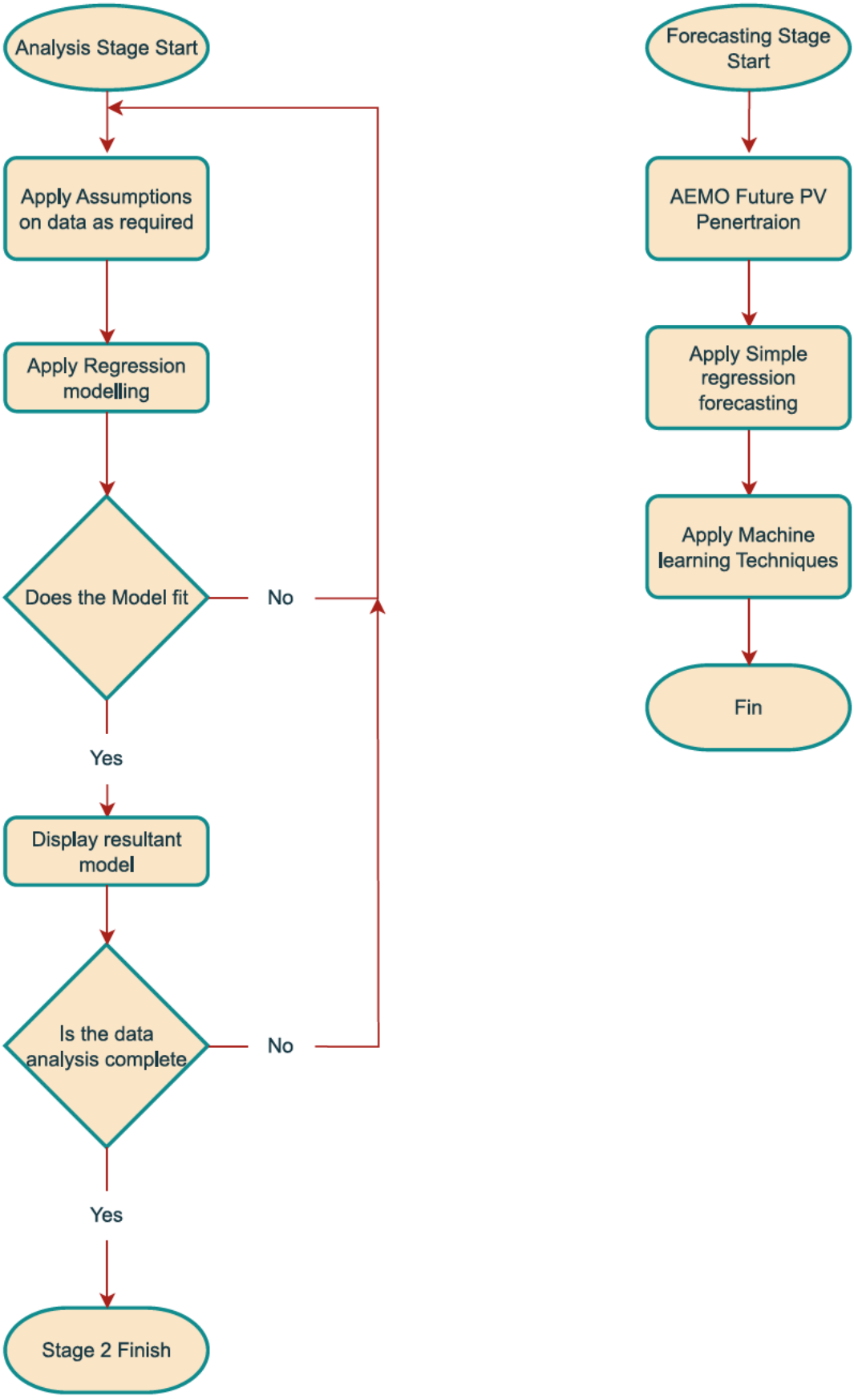


Figure 20: Stage two and three Methodology Flow Chart

Regression Modelling

Regression modelling is a powerful statistical technique used to understand relationships between variables. It is a fundamental tool in data analysis and performs a crucial role in many other professional areas, such as economics, finance, and health [47]. The primary objective of regression modelling is to provide a quantitative framework for making predictions and reveal associations between one or more independent variables and a dependent variable. The project looks to explore and leverage the fundamental types of regression analysis to make practical inferences towards reactive contributions.

Linear (simple Linear)

Linear regression is a widely used modelling technique; it is a fundamental method in directing variable relationships and serves as an initial building block towards more complex regression techniques.

A simple linear regression can be recognised in the general form of Equation 15 [48]:

$$Y_d = \beta_0 + \beta_1 X_i + \epsilon \qquad \text{Equation 15}$$

Where:

- Y_d is the dependent variable.
- X_i is the independent variable.
- β_1 represents the slope coefficient between the change in Y_d for one-unit of X_i.
- β_0 represents the intercept value of Y_d when X_i equates to zero.
- ϵ represents the error where Y_d cannot be explained by the linear relationship of X_i.

Non-linear (complex Linear)

When the performance of a model cannot be recognised by a straight line in the form of Equation 15, non-linear models allow more complex and curved relationships to be depicted.

Polynomial

The general form of a polynomial regression equation is recognised in Equation 16 [48]:

$$Y_d = \beta_0 + \beta_1 X_i + \beta_2 {X_i}^2 + \beta_3 {X_i}^3 + \cdots + \beta_n {X_i}^n + \epsilon \qquad \text{Equation 16}$$

Where the new terminology is:

- $\beta_1, \beta_2, \beta_3, \ldots, \beta_n$ represents the estimated coefficient between Y_d for a unit of X_i.

Exponential

Exponential regression models express situations where a variable's growth or decay is proportional to the current value. The general form of an exponential regression equation is recognised in Equation 17 [48]:

$$Y_d = A \cdot e^{BX_i} \quad \text{Equation 17}$$

Where the new terminology is:

- A represents the initial value Y_d when X_i equates to zero.
- B represents the rate of growth (if $B > 0$) or decay (if $B < 0$).
- e is the base of the natural logarithm ($e \cong 2.71828$)

Logarithmic

Logarithmic regression models express situations where the dependent variable changes are proportional to the logarithm of the dependent variable. The general form of a logarithmic regression equation is recognised in Equation 18 [48]:

$$Y_d = A \cdot \ln(BX_i) \quad \text{Equation 18}$$

Where the new terminology is:

- A represents the scaling factor via stretch or compression of the curve.
- B represents the slope coefficient of the logarithmic curve.
- ln is the natural logarithm.

Evaluating Models

Evaluating the model is a crucial step in assessing the performance of the applied model/s and the suitability for making predictions on the data. The techniques and metrics that are leveraged in the project to evaluate model performance are:

- The coefficient of Determination (R-squared) represents the variance between Y_d and X_i of the model, values are between 0 (no-fit) and 1 (perfect-fit) [49].
- Calculating the average differences squared between predicted and actual values allows evaluation usage of the Mean Squared Error (MSE) [49].

Chapter Four –Concept Development and Results

Introduction

The concept development and results analysis chapter endeavours to culminate the previous chapters' learnings concerning grid-connected solar reactive power contributions. The uptake in rooftop solar is growing exponentially, and with it, the constructive and adverse characteristics of PV systems influence. The contribution of solar systems operating under optimal conditions is evident and well-recognised. While the contributions of a singular, isolated instance during low irradiance may be negligible from a singular small-scale rooftop solar-PV, the cumulative effect in an increasing capacitive load environment is of greater significance.

Chapter Four will utilise the stages outlined in the flow diagrams illustrated in Figures 19 and 20 as a structural outline in addressing the project's aims. A considerable project component was applied using the open-sourced code editor Visual Studio (VS) Code program to conform and sort the supportive datasets supplied by EQL. Access to the two datasets has granted correlations between the household connection data, including isolating PV current and localised irradiance levels. The links between household and irradiance substation connections were undertaken on the secured network to mitigate confidential concerns as a desensitisation technique. Reactive power contributions from PV systems at low irradiance events will be examined and modelled in three forms: Q vs S, Q vs Irradiance, and Q vs Time (in half-hour intervals throughout 24 hours).

Stage One – Pseudocode

Introduced in Chapter Three, the dataset supplied was in parquet format, which requires specific code editors to read and write, which was mainly done by setting up a virtual environment in Microsoft VS code using Python 3.11.2 as the programming language. The program was created at this initial stage of the project for isolated purposes to handle data points as required while leveraging open-sourced libraries for visual and statistical modelling purposes. The program source code can be examined in Appendix F.

Initial Setup

The initialisation of any libraries needed for data processing:

Data Handling			
Pandas	Numpy	Glob	Pathlib
Datetime	Os	Pytz	Math
Visual and Statistical Modelling			
Seaborn	Matplotlib	Sklearn (various packages)	

Due to the program size, the designation of any constant literals used throughout the program can be referenced in Appendix F, the final source code for visualisation. The initialisation of any necessary file read paths included setting up empty data frames. These empty data frames stored the linking NMIs that allowed confirmations of source and removal of any duplications within merging processes.

Functions

convert_datetime_timezone(dt, tz1, tz2):

This function accepts a datetime descriptor as **dt** accompanied by time zone one **tz1** and time zone 2 **tz2**. Where:

- **dt** is the datetime to be converted, which can be in either datetime64[ns] or string (str) format.
- **tz1** & **tz2** are time zone library descriptors as a str; some examples include: 'UTC' (Coordinated Universal Time) or 'Australia/Queensland'.

The imported datetime library modifies the time stamp (**ts**) to be the same format as the measurement data, allowing **merge()** on the five-minute intervals.

get_half_interval (minutes):

Accepts the minutes in datetime64[ns] format, then is divided by a half-hour period, using a conditional on the remainder, allowing a determination of:00-:29 and:30-:59.

get_season (month):

Accepts the month in datetime64[ns] format, then using a conditional on the remainder of a mathematical calculation allows the determination if the remainder is one, two, three or four, returning the season as a str.

get_time_ref (hour):

Like get_season(), the function accepts the hour in datetime64[ns] format, using a conditional on the remainder of a simple mathematical calculation that allows the determination of the remainder [1-6]. This return is (Late Night, Early Morning, Morning, Noon, Evening or Night) a predetermined generalisation period of the day as a str.

pf_assumption_fix (pf):

The function addresses an inconsistency in the measurement data where the power factor (**pf**) was greater than unity and where unity was recorded as -1.00. The conditional function accepts the pf as a float and returns the corrected assumption.

solar_current_assumption_fix (df):

Like pf_assumption_fix(), the function is to modify an error found in the measurement data solar current where the polarity of the solar current is returned positive with some degree of tolerance.

negate (number):

Is as the name suggests and returns the input as a negative.

sort_dfs (df, NMI, columns_to_drop, column_arrangement):

The function uses four arguments to correctly arrange the data before any calculations or grouping to confirm that all data frames (**df**) are correlated. Using the lists, columns_to_drop removes unused values like 'Voltage Variation' and column_arrangement for uniformity of column naming convention and agreement, returning the sorted df.

cleanse_df (df):

The cleanse function is designated to fix any assumptions through functions while returning a **groupby()** df by five-minute intervals prepared for **merge()**.

power_calc_df (df):

The function is designated to apply Equations 11 – 13 to calculate the power values for the aggregate at the connection point and the assumptions placed on the solar PV system output. While the function accepts the df for calculations, it also joins other critical data columns used throughout the analysis and returns the completed **join()** df.

Main

Irradiance

The irradiance data was supplied in the format of CSV Excel files and, therefore, needed differing library functions to read the data with VS code. Once the file path is read, the same process as per the measurement data is applied—removing unused and sorting data columns. The irradiance data was captured in five-minute intervals but stored as UTC; therefore, it required convert_datetime_timezone(dt, tz1, tz2) before merging the two datasets. The return was to write a new file format as a parquet to apply consistency between the datasets.

Measurement

The monitoring data is in a parquet format as introduced in Chapter Three; due to the number of data files over the three years, data was sorted under the guide conditionals. It was established that the NMI was the most appropriate argument merge() on using a conditional pass-through. Removal of unused and sorting of data columns is the same process as irradiance but was defined by a function. At the same time, the merging had to be verified if it contained approximately 365- or 1095 days' worth of measurements. After the cleanse_df(df) and power_calc_df(df) were completed, the df was written into a file location before irradiance locals could be merged.

Merge

The merging process was simplified by the measurement and irradiance preprocessing and sorting. This allowed the link between the NMI's and substation connections with irradiance values to be organised through a desensitised processing, mitigating the ability to correlate without authorised accesses. After this link was processed, all available data were concatenated for ease of analysis and modelling.

Stage Two – Analysis and Modelling

Stage two commissions the cleansed and merged data from stage one to analyse different meaningful PV reactive power contribution scenarios. These include Q versus S_{inv} grouped, Q during low irradiance events and the aggregate Q over 24 hours. It was established while linking measurement to irradiance that only four of the 20 irradiance locations had a given amount of household connections. The impact over these four substation connections differs in reactive flow influence; it cannot be established at this project stage if the entire substation population has been covered. For this reason, this stage of the project sites will be examined as an aggregate rather than an isolated impact.

Programs such as sklearn statistical modelling were employed to compare the differing levels of contributions of Q applied to each scenario and contribute to the derivation of the analytical expressions. Two constraints were applied to the cleansed and merged dataset to confirm that the data being examined was mainly influenced by solar impacts. Where PV current I_{PV} is prominent compared to the Load I_{Load}, and three levels of S relative to the compliance of the latest AS/NZS 4777.2:2020 [16].

1. $(abs(I_{Load} + I_{PV}) > abs(I_{PV} \cdot 5\%)$
2. PV Output Ranges, where compliance correlates:
    a. Low: $|S_{PV}| \leq 20\%$, compliance for $Q_{PV} < \mp 44\%$
    b. Mid: $20\% < |S_{PV}| \leq 60\%$, compliance for $Q_{PV} < \mp 44\%$
    c. High: $60\% < |S_{PV}| \leq 80\%$, compliance for $Q_{PV} < \mp 60\%$

The inverter is a unifying practical element that links all aspects of power contributions that interact with grid operation—specifically, the AC-connected components or, as illustrated in Figure six, the output filter of the inverter. The significance of these components is essential to understanding the reactive power contributions of grid-connected PV's. The design philosophy of an inverter is to regulate the output current feeding to the grid while operating with high efficiency [26]. In contrast, grid-connected inverters must observe and comply with local standards like 4777.2 at installation. The operational fitment testing is focused on optimal operating times, with low irradiance event testing as optional or only when compliance circumstances occur [17]. The reactive contribution supplied from one inverter is of minimal significance, but the increasing growth in PV penetrations on the distribution network increases these contributions.

While current flows through an inductor, it generates a magnetic field around the core and stores this as electrical energy. The energy stored is proportional to the square of the current passing through the inductor's core [5]. However, suppose there is a reduction of current flow or zero current like during low irradiance conditions. Therefore, the circuit's inductance as in Figure six, is reduced until zero. Thus, similar considerations can be applied to the capacitance of the output filter component. While voltage is applied across a capacitor's terminals, electrical energy is stored between the two conducting plates [5]. The difference of potential electrical energy across the two open terminals or, in the case of a capacitor, the two conducting plates are not eliminated as there is always a voltage connection to and from the grid.

Q Vs S

Discussion

Using the PV-rated output as the operation's apparent base for per-unit calculations grants the comparison between different inverter ratings. Figures 21 to 23 illustrate the three corresponding output ranges, low, mid, and high, comparing the relevant compliance of supply or absorption of Q, relating to inverter ratings. There is an interesting trend in capacitive behaviour over all three ranges where all inverters operate within compliance until the mid-range of around S 0.4 p.u. Though between 0.4 and 0.6 does operate over the 44% (orange dashed line) reactive compliance, especially when relating to reactive power supply. However, the operation complies with standards when evaluating the higher range of 0.6 to 0.8 S of PV, at 60% (red dashed line) reactive compliance, with some outlier moments of oversupply or absorbing.

An interesting observation is illustrated in the regression modelling over the entire range, where $|S_{PV}| \leq 80\%$. The mid-range between 0.2 and 0.6 of S contributes, on average, 10% to near 13% Q. The relationship is the potential sizing difference between components of the filter elements of the inverter, causing a phase shift in pf. The filter elements are designed to the inverter's rated active or apparent power [27].

Metrics and Regression Equation

Regression Type		R^2	***MSE***
Simple Linear		0.19	0.00072
Polynomial	Quadratic	0.87	0.00011
	Cubic	0.91	0.00008
	Bi-Quadratic	0.91	0.00008
	Quintic	0.93	0.00006
	Sextic	0.93	0.00006

Table 7: Regression Metrics for Q vs S[inv]

Equation 19 represents the best-fitting chosen regression, including all ranges of S using the metrics R^2 of 0.91 from Table seven. The polynomial expressed using Equation 16 is illustrated in Figure 24, including the shading area showing the 95% confidence interval of predictions.

$$Q_{pv[S]} = (-0.012) + (-1.80e^{-2})X_i + (8.80e^{-4}){X_i}^2 + (-1.18e^{-5}){X_i}^3 \quad \text{Equation 19}$$

Graphics

Low

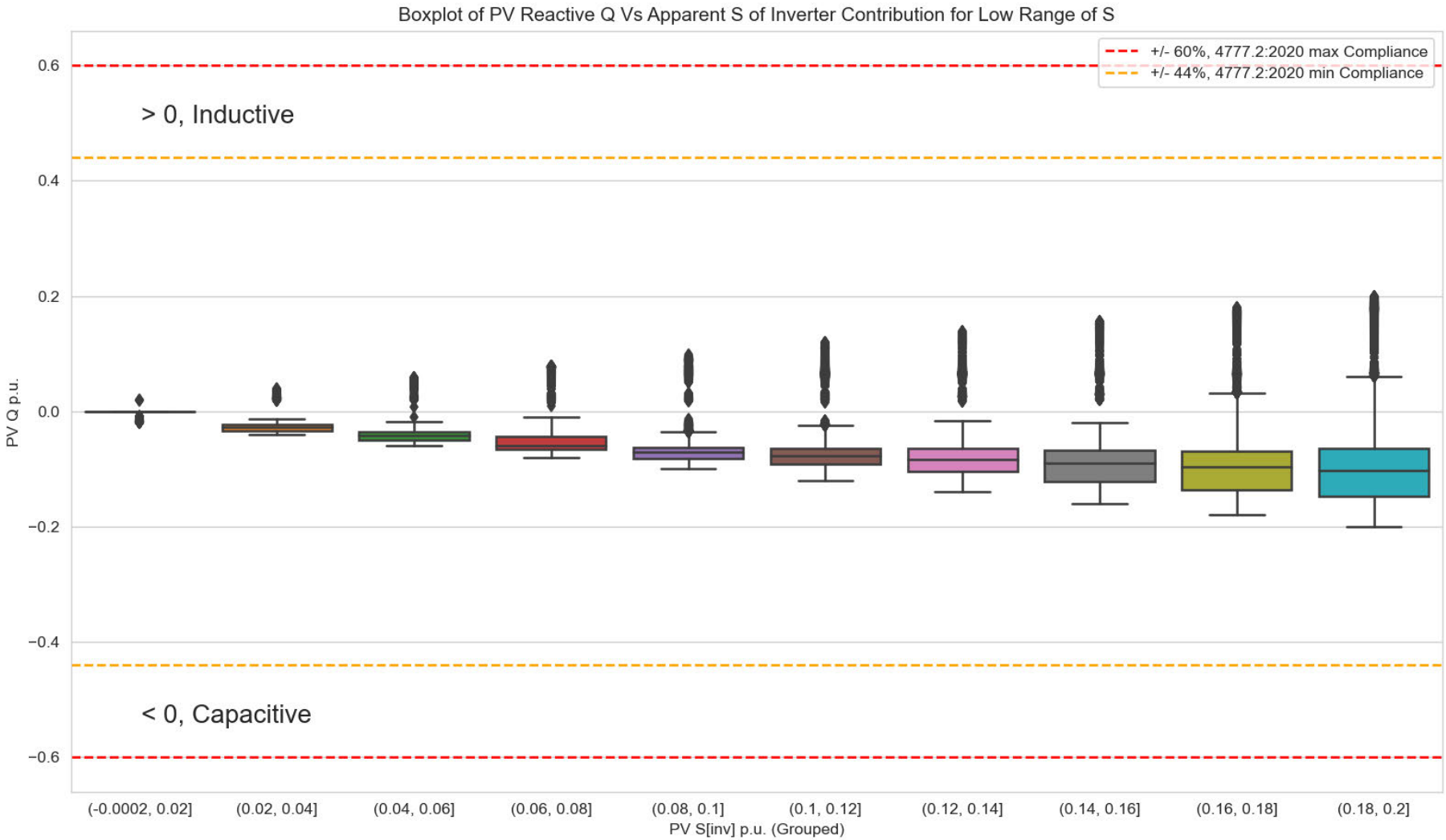


Figure 21: Boxplot Q vs S for Low-Range of S

Mid

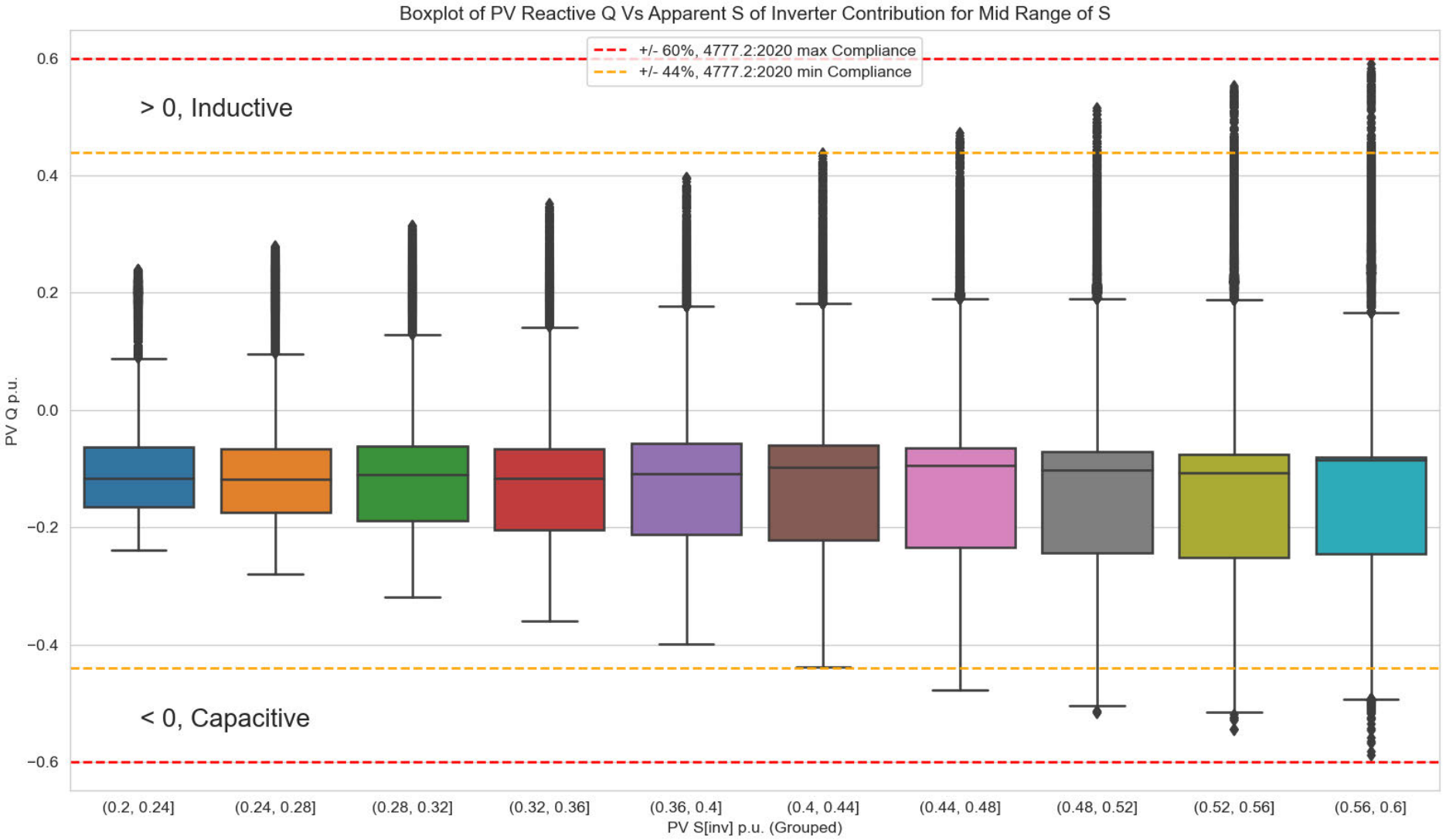


Figure 22: Boxplot Q vs S for Mid-Range of S

High

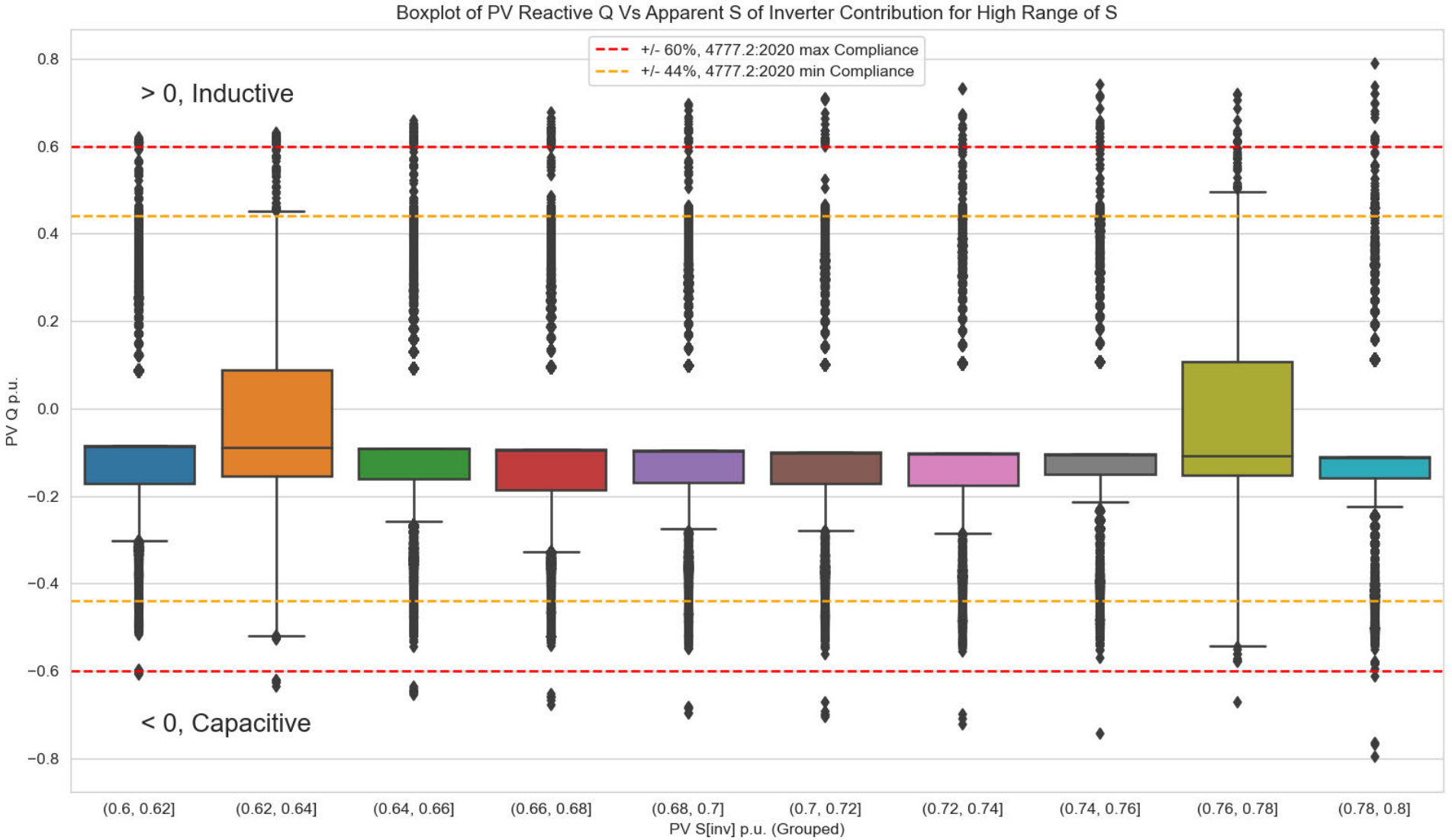


Figure 23: Boxplot Q vs S for High-Range of S

Regression Modelling

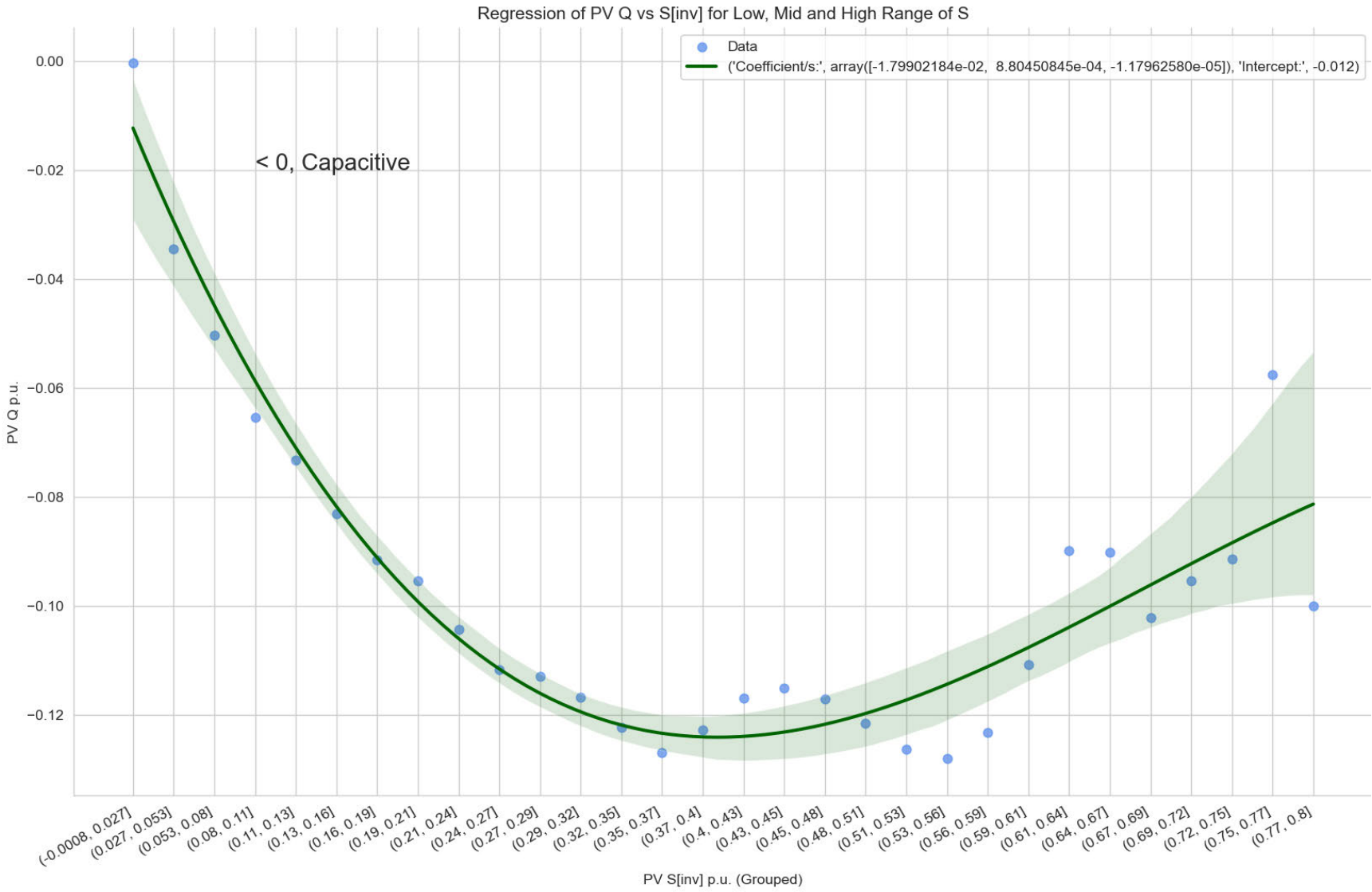


Figure 24: Regression Modelling of Q vs S[inverter]

Q Vs Irradiance

Discussion

This section presents the results of the reactive power contributions compared over irradiance. Low irradiance events can drastically cause changes in PV output over both real and reactive power contributions. A low irradiance condition is defined in AS/NZS IEC 60904.1:2023 and measured as a condition less than 200 watts per metre squared after extra-terrestrial and atmospheric loss have been accounted [17]. Figure 14 is an illustration of seasonal variations of Ghi over 24 hours; the distributions used in this section are of a similar representation. However, the actual instantaneous conditions acting on the PV panels will differ as the irradiance measurements were accumulated within a 10km radius. An analysis by Patsalides et al. [36] researched PV power quality behaviours during different irradiance conditions [36]. Figure 24 compares similar findings for active P and pf behaviour over the full irradiation dataset available to validate this project model outcome. Though this provided insight into the generation of active power, the pf analysis differed considerably. Where this dissertation seeks to differentiate between inductive and capacitive flows of Q, Patsalides et al. [36] applied an absolute on the pf; therefore, any vars supplied or absorbed were absolute.

Figures 26 to 28 capture the reactive power contributions range during low irradiance events. These illustrations have a similar trending of Q contributions that are normally distributed with the trend to capacitive flow. It is interesting to notice that while the low range is within 4777.2:2020 compliance, once greater than 20% of S PV, there are instances of oversupply and absorption of reactive contribution throughout the whole range of irradiance. The over factor to correlate is that irradiance occurs during daytime hours (which will be examined throughout the next section), which are influenced by load behaviours of the connection point.

The regression modelling illustrated in Figure 29 reinforces a PV trend towards a capacitive Q influence. Irradiance conditions below 200 have average contributions that approach 10%, Whereas above 200, these average increases to some moments near 15%. As introduced earlier, the influence of load behaviours cannot be eliminated unless isolated system monitoring is involved. However, this section demonstrates the influence weather factors can have on the shift in phase between current and voltage and, therefore, overall power quality, increasing the contributions of Q during these events.

Metrics and Regression Equation

Regression Type		R^2	***MSE***
Simple Linear		0.10	0.00065
Polynomial	Quadratic	0.43	0.00042
	Cubic	0.43	0.00041
	Bi-Quadratic	0.54	0.00033
	Quintic	0.57	0.00031
	Sextic	0.57	0.00031

Table 8: Regression Metrics for Q vs Irradiance

Equation 20 represents the best-fitting regression model with an R^2 value of 0.57, as demonstrated in Table eight. Using the polynomial expression from Equation 16, the model is illustrated in Figure 28, including the shaded area representing the 95% confidence interval of predictions.

$$Q_{pv[Irr]} = (-0.038) + (-2.55e^{-4})X_i + (-4.76e^{-7})X_i{}^2 + (3.32e^{-9})X_i{}^3 + (-4.83e^{-12})X_i{}^4 + (2.17e^{-15})X_i{}^5 \quad \text{Equation 20}$$

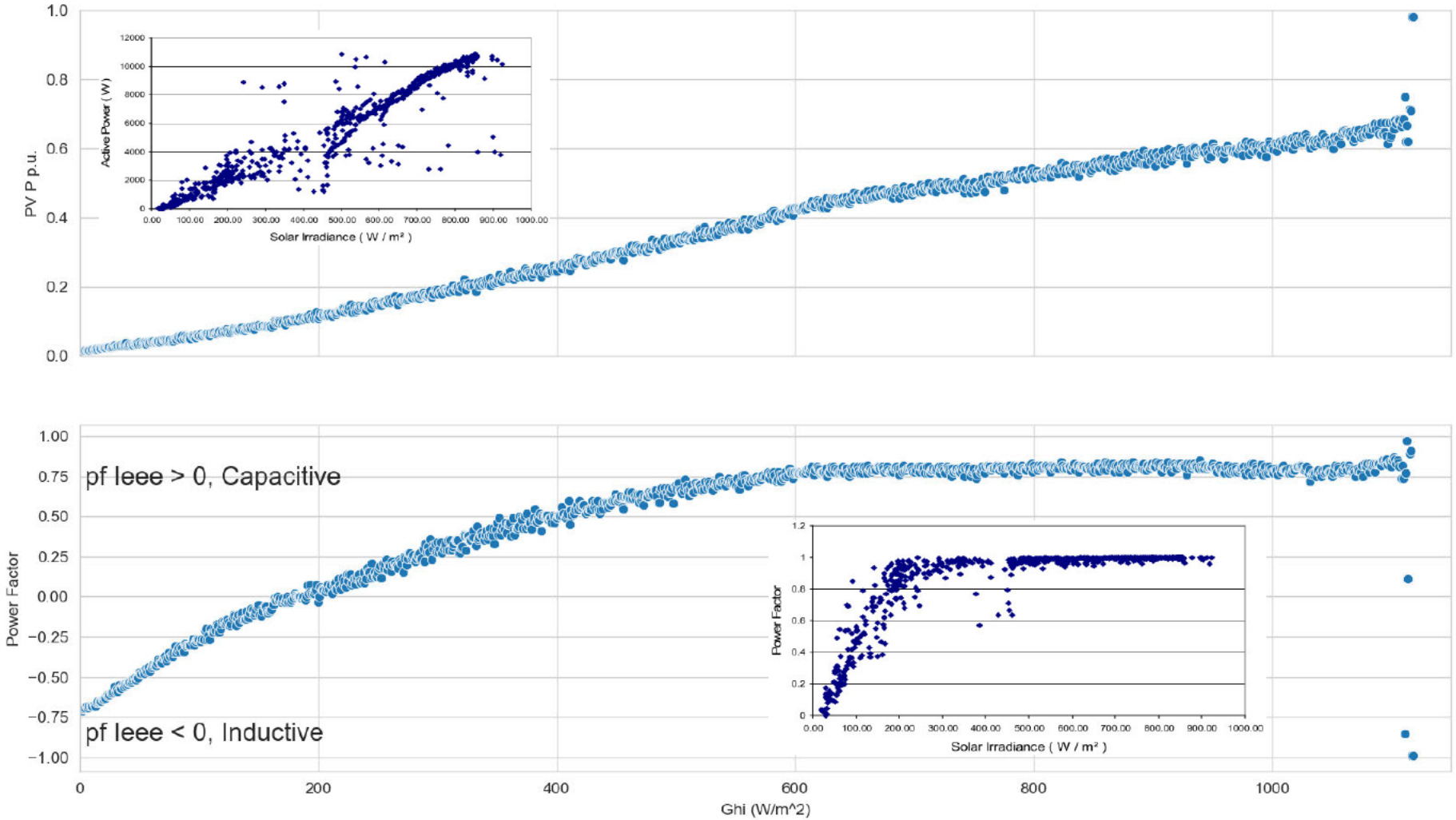


Figure 25: P (mean) and pf (median) vs entire Ghi Range [32]

Graphics

Low

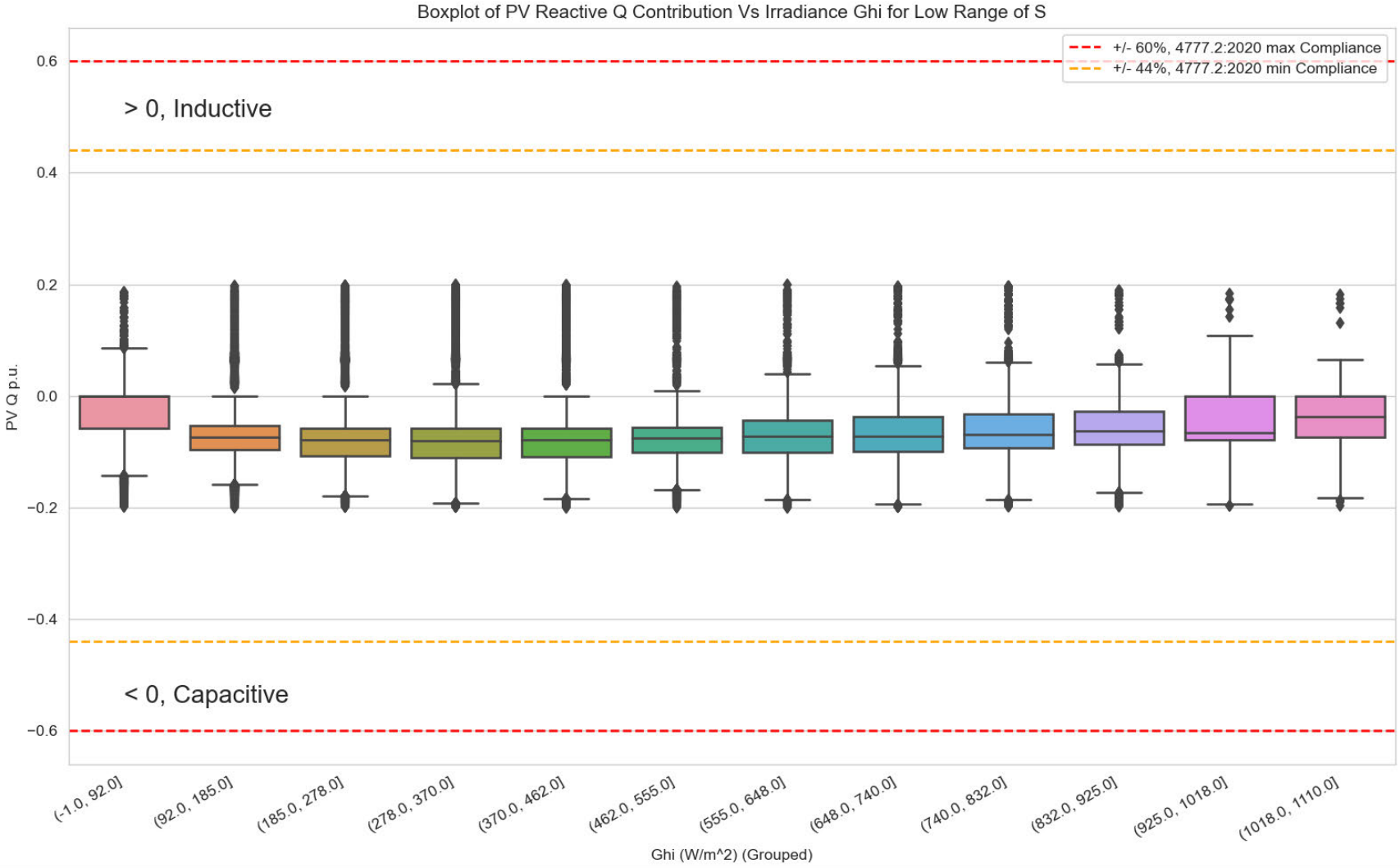


Figure 26: Boxplot Q vs Irradiance for Low-Range of S

Mid

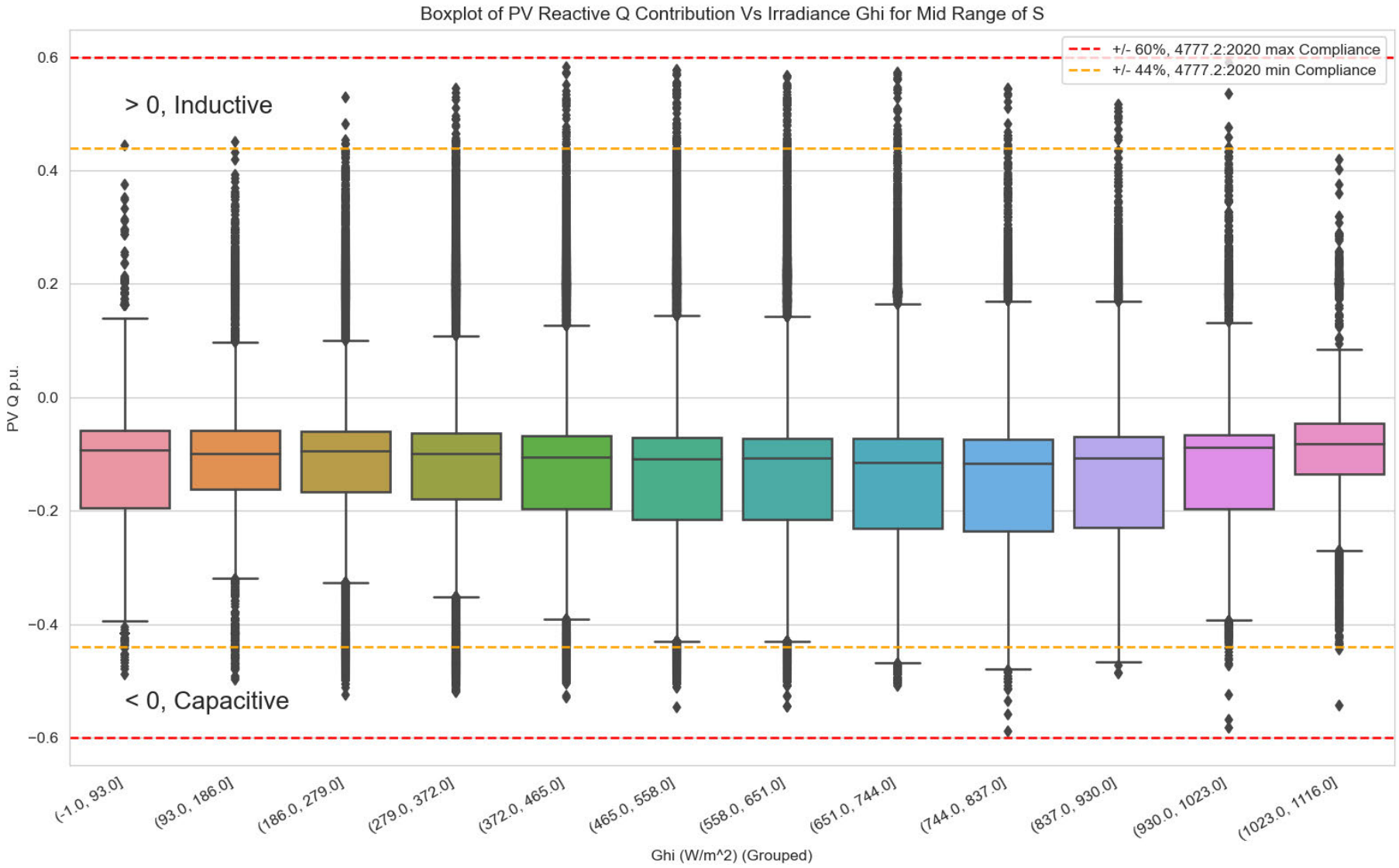


Figure 27: Boxplot Q vs Irradiance for Mid-Range of S

High

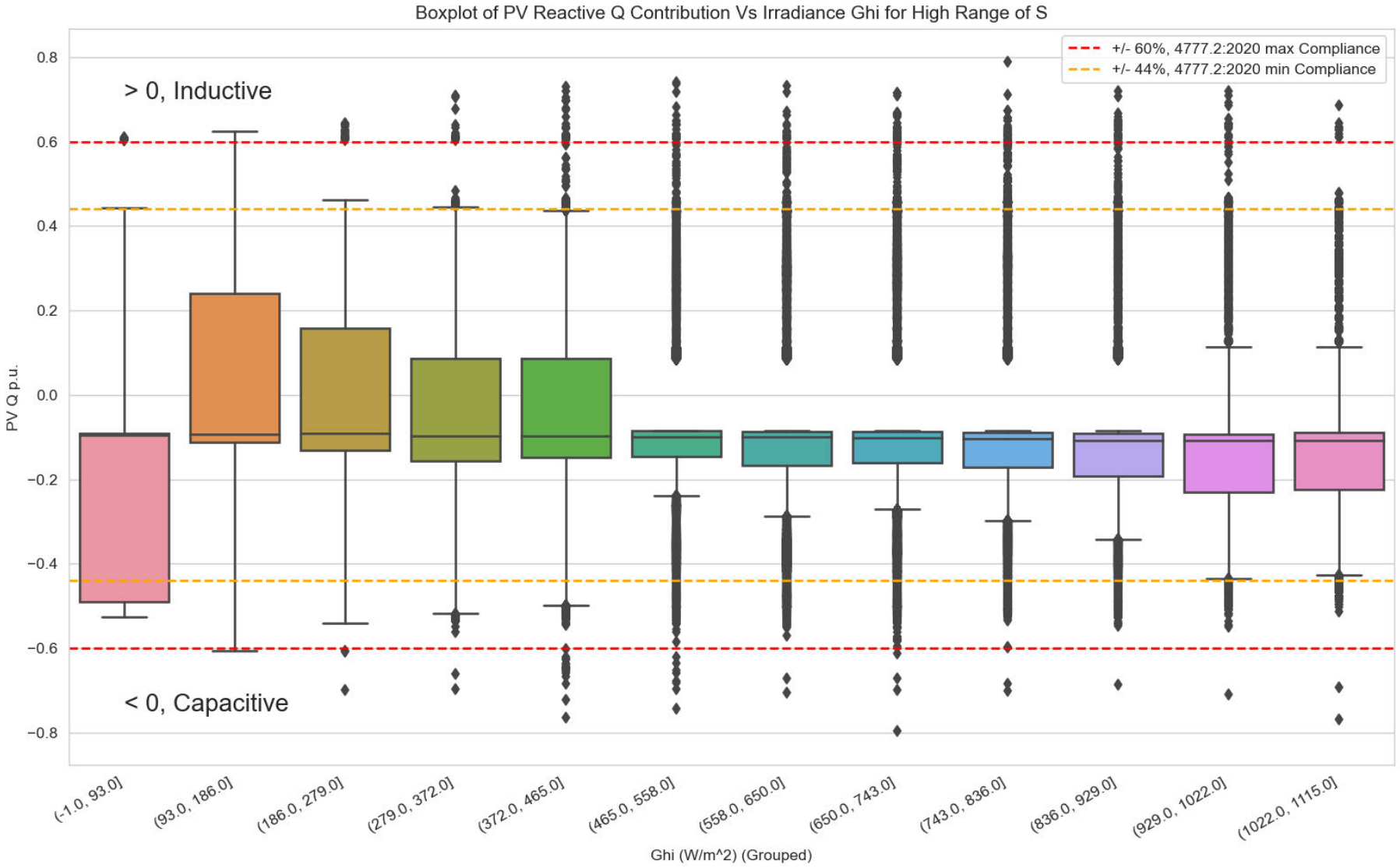


Figure 28: Boxplot Q vs Irradiance for High-Range of S

Regression Modelling

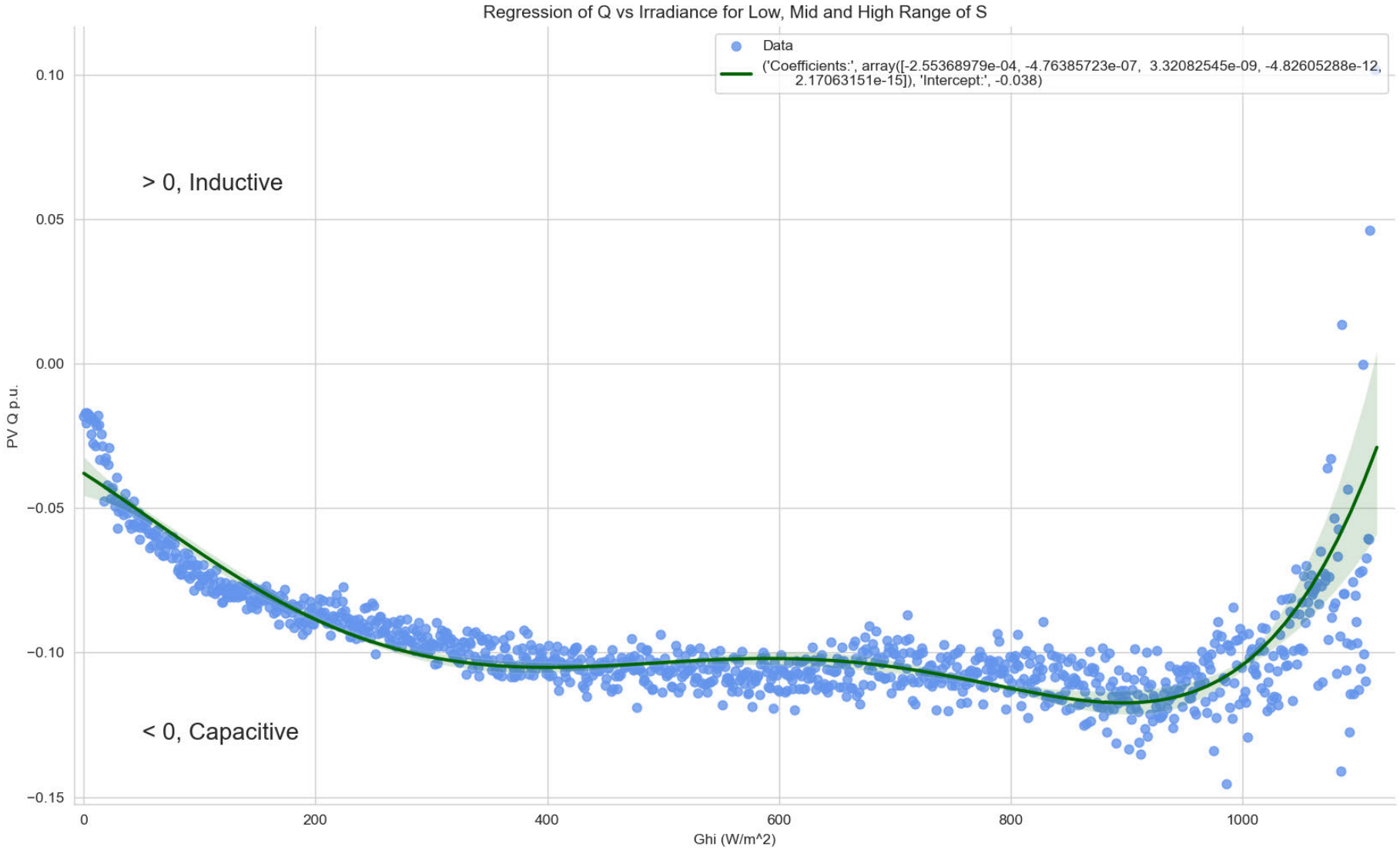


Figure 29: Regression Modelling of Q vs Irradiance

Q Vs Time

Discussion

This section for reactive power contributions allows an equivalence between times of the day and household behaviours that are common during those time frames. A similar capacitive nature can be established in the initial observation when comparing these results to the other two scenarios. Figure 30 illustrates that during low demand of S, contributively, Q is minimal and operates well within the 44% compliance. Figures 30 to 32 also show a similar oval pattern between the morning at 0630 and the afternoon around 1730. Though the outliers of mid and high contribute outside of compliance, this indicates the influence of the load can vary the pf at the point of connection. The noteworthy contribution of this section happens during the late hours between 1800 and 0600 when there is no sun for solar PV generation capability. It demonstrates that though mid and high are operating with 4777.2:2020 compliance, the contributions of Q are dominated capacitive during periods where the grid is in low demand. This correlates with PLQ's assessment of reducing reactive supply during minimum demand periods, as illustrated in Figure five [8]. It is linked to the inverter filter capacitance phase, shifting periods of low active demand to a capacitive reactive power flow.

Figure 33 simulates the average reactive Q contributions over 24 hours. The significance of this regression holds the most significant consideration. Allowing comparisons to be drawn between trends of end-user behaviours. Figures 31 and 32 illustrate that the contributions are most considerable when light is absent, yet the cumulative averages, regardless of apparent demands, are around one to five percent.

Metrics and Regression Equation

Regression Type		R^2	***MSE***
Simple Linear		0.03	0.00137
Polynomial	Quadratic	0.79	0.00029
	Cubic	0.79	0.00029
	Bi-Quadratic	0.92	0.00011
	Quintic	0.95	0.00007
	Sextic	0.96	0.00006

Table 9: Regression Metrics for Q vs Time (30-min Periods)

Equation 21 represents the best-fitting regression model with an R^2 value of 0.92, as demonstrated in Table nine. Using the polynomial expression from Equation 16, the model is illustrated in Figure 33, including the shaded area displaying the 95% confidence interval of predictions.

$$Q_{pv[t]} = (-0.012) + (5.85e^{-3})X_i + (-1.31e^{-3})X_i{}^2 + (5.01e^{-5})X_i{}^3 + (-5.33e^{-7})X_i{}^4 \quad \text{Equation 21}$$

Graphics

Low

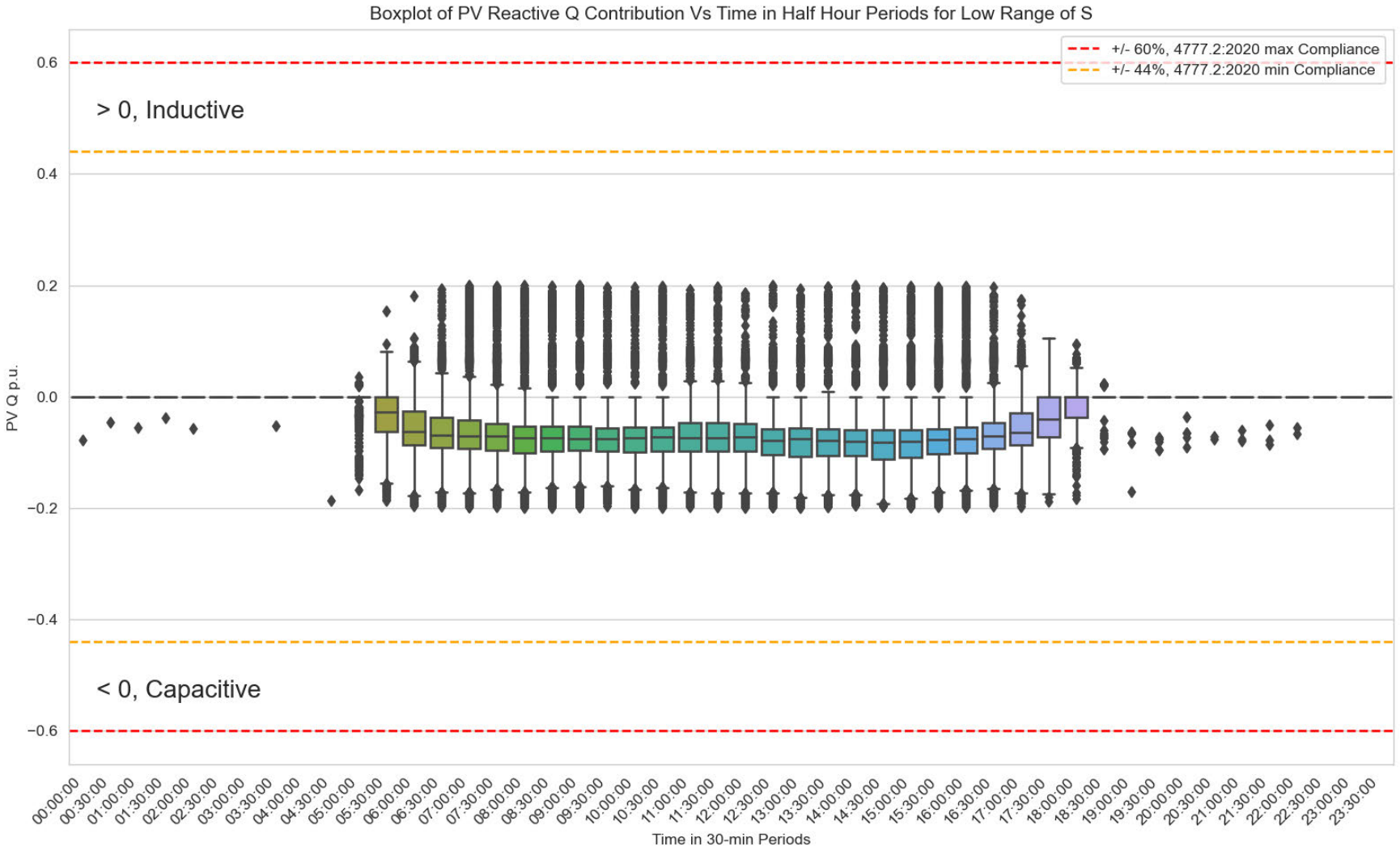


Figure 30: Boxplot Q vs Time (30-min Periods) for Low-Range of S

Mid

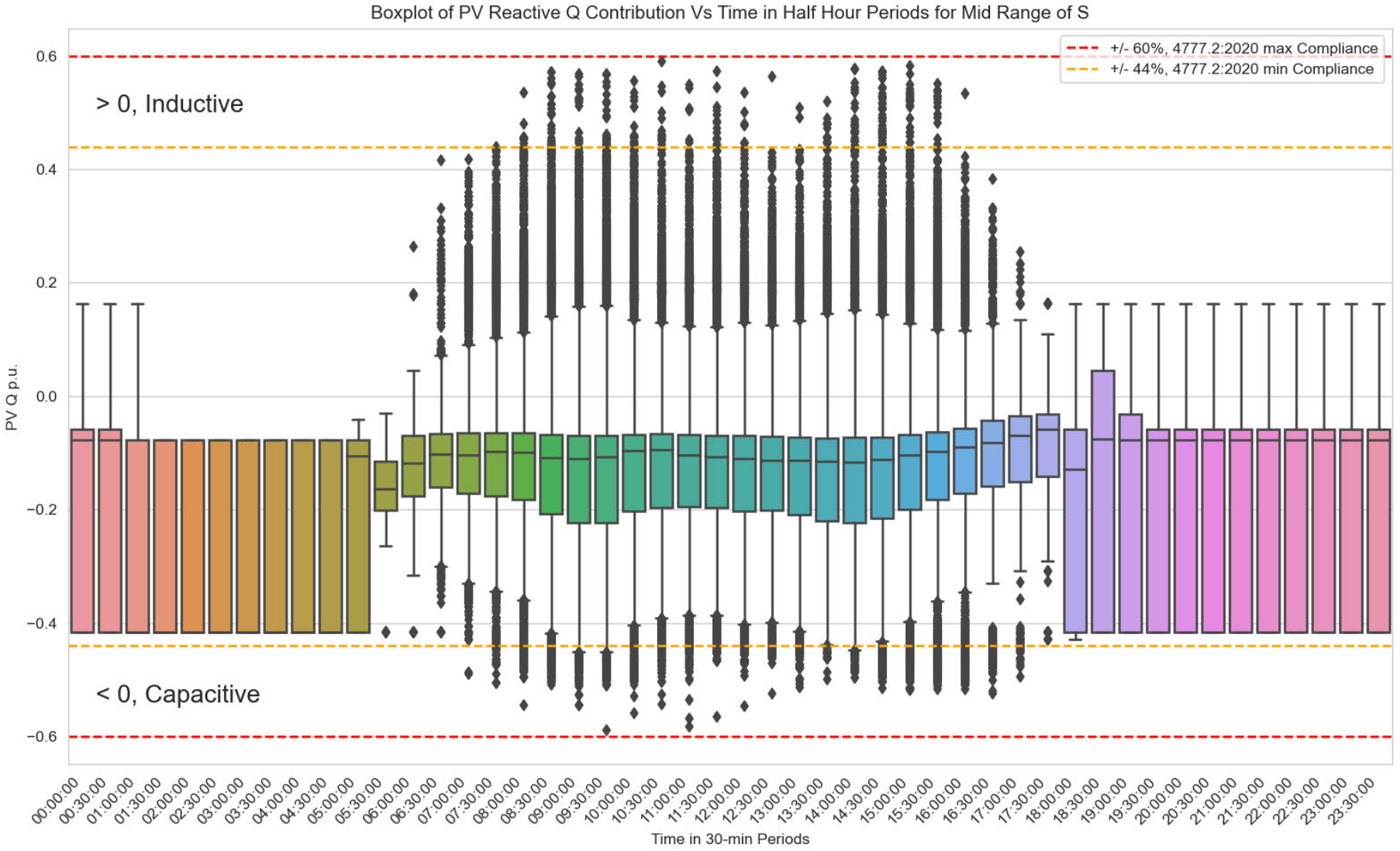


Figure 31: Boxplot Q vs Time (30-min Periods) for Mid-Range of S

High

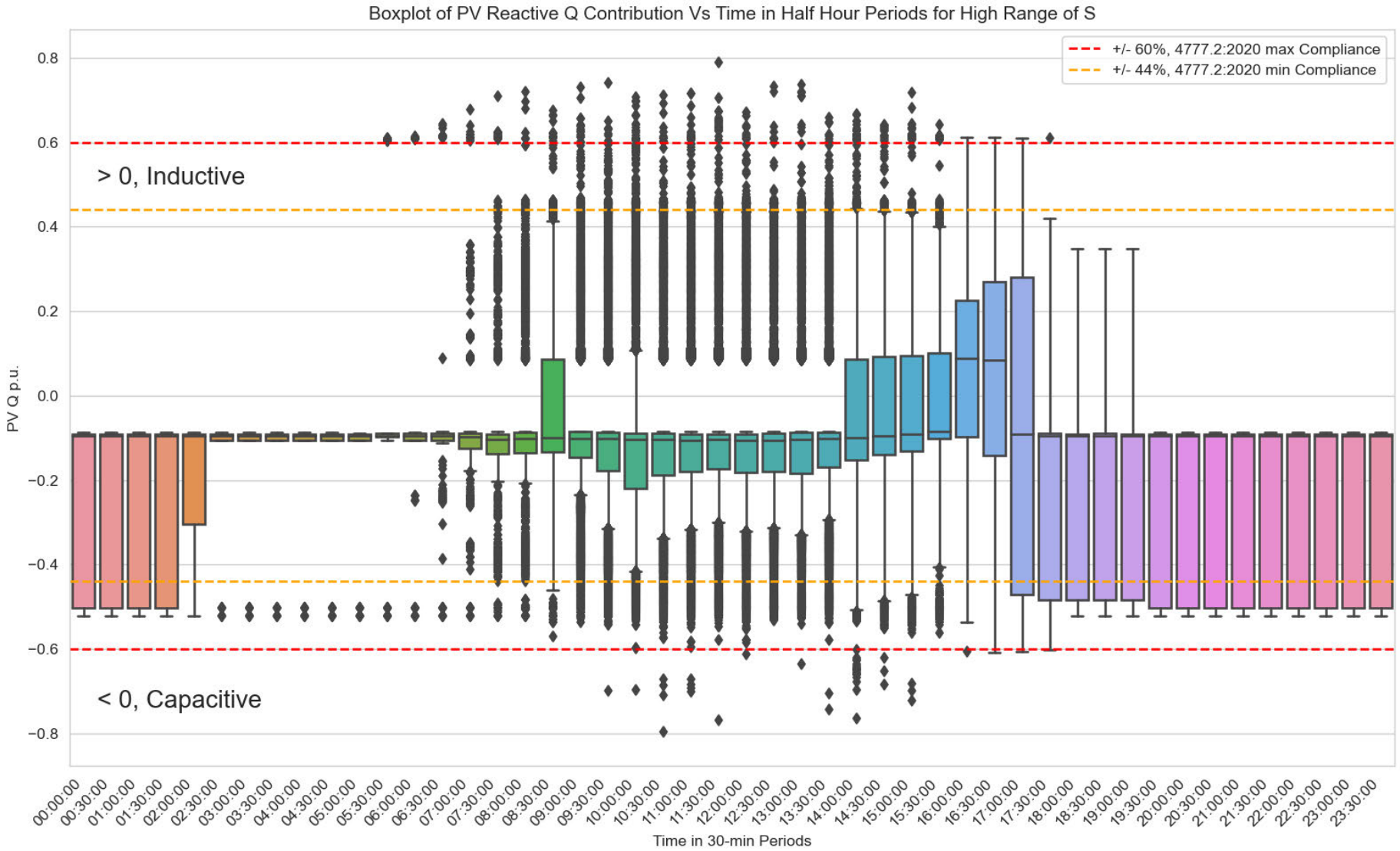


Figure 32: Boxplot Q vs Time (30-min Periods) for High-Range of S

Regression Modelling

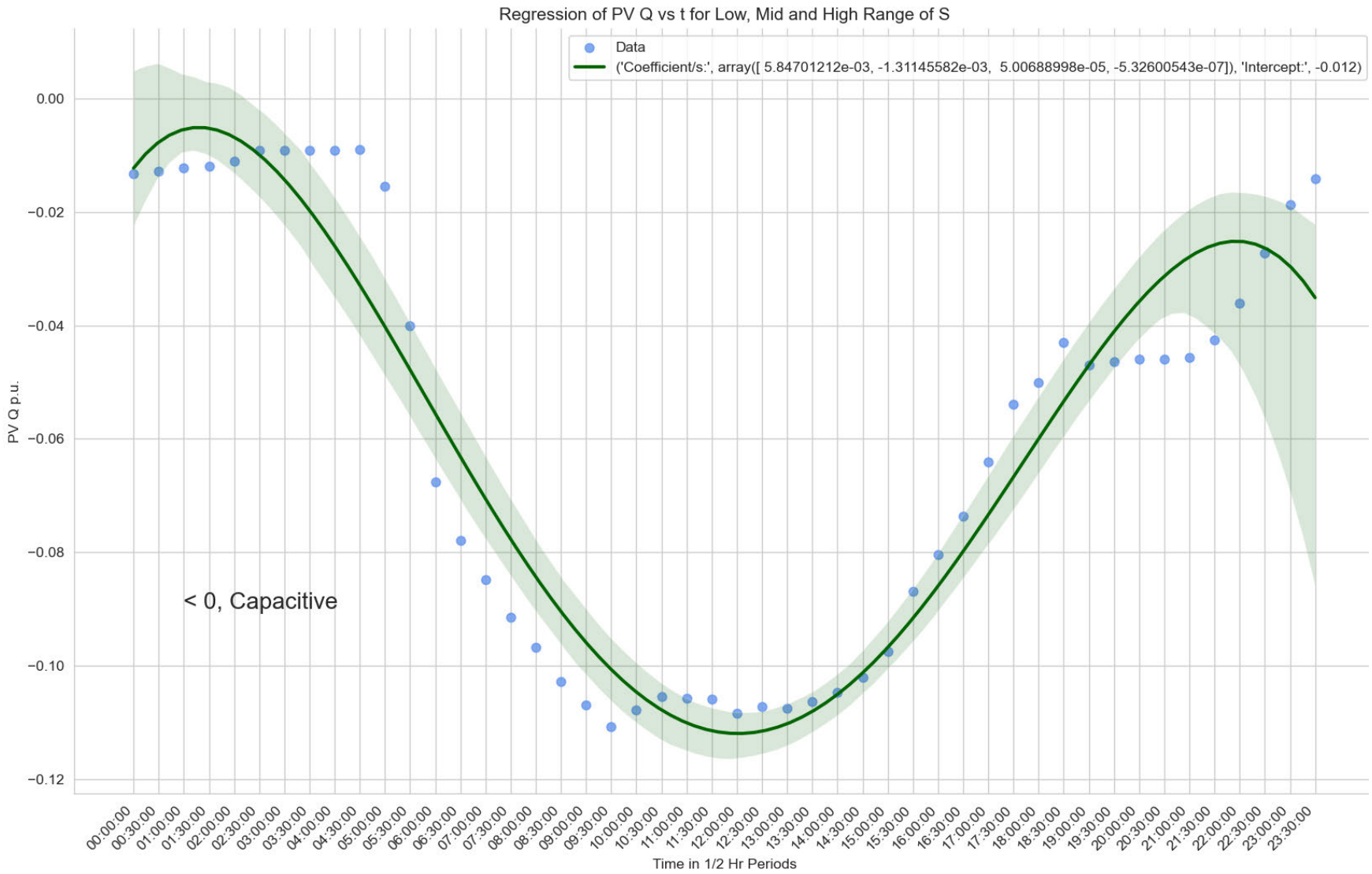


Figure 33: Regression Modelling of Q vs Time (30-min Period)

Chapter Five – Discussions and Conclusions

The dissertation aimed to recognise the reactive power contributions of grid-connected solar photovoltaic systems. The evaluation of the current literature established considerable knowledge linking PV to optimal operating conditions and less regarding sub-optimal operating conditions like low irradiance events. However, there is a significance towards the electricity utilities experiencing a rise in capacitive reactive power flows. With the growth in home and small business solar-PV systems reaching three million as of early 2022, nearly 26% of those existing in Queensland [3]. A positive side effect of grid-connected solar PV is that it reduces or offsets consumption BTM. Yet, the electrical landscape is continually in temporal flux for utility providers. This ever-changing nature of reactive power flow results in the growth of capacitive contribution from various areas, like end-user household characteristics and increases in underground electric development (URD), especially within the distribution network.

With those two significant concepts at the forefront, both an increase in grid-connected solar and the changing reactive power flow environment. Analytical ideas were leveraged to research the reactive contributions during the less explored low irradiance events. The dissertation found that the average cumulative capacitive contribution could be as high as 10 to 13% during these low irradiance conditions. The investigation also discovered that most reactive contributions comply with the latest AS/NZS 4777.2 standard regardless of system rating or legacy [16]. However, resultant instances absorb and supply reactive flow outside of the 2020 operating standard.

Stage Three – Forecasting and Direction for Future Research

The project's inception had a clear projection towards the next steps in advancing the project's potential. Derived from the research within this dissertation and available literature, some obstacles must be overcome.

Initial stages for refining research included in this dissertation:

- Development of differing methodologies to further or challenge research conclusions.
- Algorithm refinement to streamline analysis of standardised data retrieval for reactive contributions of grid-connected PV.
- Voltage management analysis and modelling to determine relevant compliance with standards.

Future research that can extend the concepts developed from the dissertation:

- Apply research from this dissertation and machine learning to produce or integrate predictive models to reactive power contributions.
- Forecasting potential reactive power contributions for growing penetrations of PV using AEMO projections.
- Grid planning optimisation frameworks for the expected supply or absorption concerning PV projections.
- Practical feeder analysis and modelling regarding capacitive nature concerning increased PV penetrations.

Bibliography

[1] Lin, Y, Eto, JH, Johnson, BB, Flicker, JD, Lasseter, RH, Villegas Pico, HN, Seo, G-S, Pierre, BJ & Ellis, A 2020, *Research roadmap on grid-forming inverters*, National Renewable Energy Lab.(NREL), Golden, CO (United States), viewed 02/07/2023, <https://www.nrel.gov/docs/fy21osti/73476.pdf>.

[2] Department of Climate Change, E, the Environment and Water 2022, *Australia's emissions projections 2022*, E Department of Climate Change, the Environment and Water, Online, viewed 25/04/2023, <https://www.dcceew.gov.au/sites/default/files/documents/australias-emissions-projections-2022.pdf>.

[3] Department of Climate Change, E, the Environment and Water 2022, *Solar PV and batteries*, Australian Government, Online, viewed 23/04/2023, <https://www.energy.gov.au/households/solar-pv-and-batteries>.

[4] Widén, J, Wäckelgård, E, Paatero, J & Lund, P 2010, 'Impacts of distributed photovoltaics on network voltages: Stochastic simulations of three Swedish low-voltage distribution grids', *Electric power systems research*, vol. 80, no. 12, pp. 1562-71.

[5] Hughes, E 2016, *Hughes Electrical and Electronic Technology*, Pearson, [N.p.].

[6] Zobaa, A 2004, 'Comparing power factor and displacement factor corrections based on IEEE Std. 18-2002', *2004 11th International Conference on Harmonics and Quality of Power (IEEE Cat. No. 04EX951)*, IEEE, pp. 1-5, viewed 10/07/2023, <https://ieeexplore.ieee.org/abstract/document/1409319?casa_token=leTyhA3SoywAAAAA:cVHhNKFLd-HB8_qdI4ySPpi1gG9EkDokuwlP65RzptVqaOBFB-ChbeJKRBE7BTqf5FgauvoWYw&signout=success>.

[7] Chen, Q, Kuang, Z, Liu, X & Zhang, T 2022, 'Transforming a solar-rich county to an electricity producer: Solutions to the mismatch between demand and generation', *Journal of Cleaner Production*, vol. 336, p. 130418.

[8] PLQ 2021, *Managing Southeast Queensland Voltages*, Powerlink, AER, viewed 04/04/2023, <https://www.aer.gov.au/system/files/Powerlink%20-%20CP.02814%20Managing%20Southeast%20Queensland%20Voltages%20-%20January%202021.pdf>.

[9] Pukhrem, S, Basu, M, Conlon, MF & Sunderland, K 2016, 'Enhanced network voltage management techniques under the proliferation of rooftop solar PV installation in low-voltage distribution network', *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 5, no. 2, pp. 681-94.

[10] Kaloudas, CG, Ochoa, LF, Fletcher, I, Marshall, B & Majithia, S 2015, 'Investigating the declining reactive power demand of UK distribution networks', *2015 IEEE Power & Energy Society General Meeting*, IEEE, pp. 1-5, viewed 18/04/2023, <https://ieeexplore.ieee.org/abstract/document/7286464>.

[11] Tazky, M, Regula, M & Otcenasova, A 2021, 'Impact of changes in a distribution network nature on the capacitive reactive power flow into the transmission network in Slovakia', *Energies*, vol. 14, no. 17, p. 5321.

[12] Katiraei, F & Agüero, JR 2011, 'Solar PV integration challenges', *IEEE power and energy magazine*, vol. 9, no. 3, pp. 62-71.

[13] Murillo-Yarce, D, Alarcón-Alarcón, J, Rivera, M, Restrepo, C, Muñoz, J, Baier, C & Wheeler, P 2020, 'A review of control techniques in photovoltaic systems', *Sustainability*, vol. 12, no. 24, p. 10598.

[14] AS 2005, *Grid connection of energy system via inverters*, Part2: Inverter requirements, 47777.2:2005, Standards Australia, Online, viewed 14/05/2023, <www.standards.org.au>.

[15] AS/NZS 2015, *Grid connection of energy system via inverters*, Part2: Inverter requirements, 4777.2:2015, Australian/New Zealand Standard, Online, viewed 14/05/2023, <www.standards.org.au>.

[16] AS/NZS 2020, *Grid connection of energy system via inverters*, Part2: Inverter requirements, 4777.2:2020, Australian/New Zealand Standard, Online, viewed 14/05/2023, <www.standards.org.au>.

[17] AS/NZS:IEC 2023, *Photovoltalic devices*, Part 1: Measurement of photovoltalic current-voltage charteristics, 60904.1:2023, Australian/New Zealand Standard, Online, viewed 16/08/2023, <www.standards.org.au>.

[18] Lubośny, Z, Klucznik, J & Dobrzyński, K 2015, 'The Issues of Reactive Power Compensation in High-voltage Transmission Lines', *Acta Energetica*, no. 2, pp. 102-13.

[19] Karki, U, Gunasekaran, D & Peng, FZ 2015, 'Reactive compensation of overhead AC transmission lines using underground power cables', *2015 IEEE Power & Energy Society*

General Meeting, IEEE, pp. 1-5, viewed 13/05/2023, <https://ieeexplore.ieee.org/abstract/document/7285628>.

[20] AEMO 2023, *DISTRIBUTION LOSS FACTORS FOR THE 2023/24 FINANCIAL YEAR*, Australian Energy Market Operator, Online, viewed 17/05/2023, <https://aemo.com.au/-/media/files/electricity/nem/security_and_reliability/loss_factors_and_regional_boundaries/2023-24-financial-year/distribution-loss-factors-for-the-2023-24-financial-year.pdf?la=en>.

[21] Niitsoo, J & Palu, I 2012, 'Undesirable usage of energy in residential house', *11th International Symposium PÄRNU*, pp. 16-20, viewed 25/04/2023, <https://www.researchgate.net/profile/Jaan_Niitsoo/publication/265942346_Undesirable_Usage_of_Energy_in_Residential_House/links/54c7516b0cf22d626a36402d.pdf>.

[22] Wlas, M & Galla, S 2018, 'The influence of LED lighting sources on the nature of power factor', *Energies*, vol. 11, no. 6, p. 1479.

[23] Lakervi, E & Holmes, EJ 1995, *Electricity distribution network design*, IET.

[24] Xu, T, Gao, W, Qian, F & Li, Y 2022, 'The implementation limitation of variable renewable energies and its impacts on the public power grid', *Energy*, vol. 239, p. 121992.

[25] EQL 2021, *ANNUAL REPORT 2021-22*, Energy Queensland Limited, Online, viewed 26/04/2023, <https://www.energyq.com.au/__data/assets/pdf_file/0019/1030834/EQL-Annual-Report-2021-22_DIGITAL.pdf>.

[26] Ti 2020, *Grid Connected Inverter Reference Design*, Texas Instuments, Online, viewed 13/10/2023, <https://www.ti.com/lit/ug/tidub21d/tidub21d.pdf?ts=1697121495679&ref_url=https%253A%252F%252Fdev.ti.com%252F>.

[27] Dursun, M & DÖŞOĞLU, MK 2018, 'LCL filter design for grid connected three-phase inverter', *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, IEEE, pp. 1-4, viewed 13/10/2023, <ieeexplore.ieee.org/abstract/document/8567054?casa_token=iO-ynuKxnIQAAAAA:jCeymYom18i3ihK9ISE2aqtj1pFPK5GgDoCklAXdyb-qkBdJsPjz_04EN-cDgyKSNprEhCV9FNk>.

[28] Kim, H & Sul, S-K 2011, 'A novel filter design for output LC filters of PWM inverters', *Journal of Power Electronics*, vol. 11, no. 1, pp. 74-81.

[29] Johnson, BC, Dunn, DG & Hulett, R 2002, 'A comparison of the IEEE and IEC standards processes', *Record of Conference Papers. Industry Applications Society. Forty-Ninth Annual Conference. 2002 Petroleum and Chemical Industry Technical Conference*, IEEE, pp. 1-12, viewed 25/08/2023, <https://ieeexplore-ieee-org.ezproxy.usq.edu.au/stamp/stamp.jsp?tp=&arnumber=1044979>.

[30] Electric, S 2023, *What is the difference between IEEE and IEC standards in terms of kW, kVAR, and PF (power factor)?*, Schneider Electric, Online, viewed 16/05/2023, <https://www.se.com/au/en/faqs/FA212521/>.

[31] Hahn, H, Meyer-Nieberg, S & Pickl, S 2009, 'Electric load forecasting methods: Tools for decision making', *European journal of operational research*, vol. 199, no. 3, pp. 902-7.

[32] Xiong, T, Bao, Y & Hu, Z 2014, 'Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework', *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 353-62.

[33] Mousavi, OA & Cherkaoui, R 2014, 'Investigation of P–V and V–Q based optimization methods for voltage and reactive power analysis', *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 769-78.

[34] Grid, N 2013, *National Electricity Transmission System Performance Report 2012–13*, Report to the Gas and Electricity Markets Authority, Great Britian, viewed 07/04/2023, <https://www.nationalgrideso.com/document/59021/download>.

[35] Kaloudas, CG, Ochoa, LF, Marshall, B, Majithia, S & Fletcher, I 2017, 'Assessing the future trends of reactive power demand of distribution networks', *IEEE Transactions on power systems*, vol. 32, no. 6, pp. 4278-88.

[36] Patsalides, M, Evagorou, D, Makrides, G, Achillides, Z, Georghiou, GE, Stavrou, A, Efthimiou, V, Zinsser, B, Schmitt, W & Werner, JH 2007, 'The effect of solar irradiance on the power quality behaviour of grid connected photovoltaic systems', *International Conference on Renewable Energy and Power Quality*, pp. 1-7, viewed 29/04/2023, <https://www.icrepq.com/icrepq07/284-patsalides.pdf>.

[37] Cangi, H 2022, 'Quality Problems in Low İrradiance at Grid Connected PV Solar System'.

[38] Bird, RE & Riordan, C 1986, 'Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres', *Journal of Applied Meteorology and Climatology*, vol. 25, no. 1, pp. 87-97.

[39] BOM 2016, *Gridded Hourly Solar Direct Normal Irradiance Metadata*, Online, <http://www.bom.gov.au/climate/how/IDCJAD0111.shtml>.

[40] BOM 2016, *Gridded Hourly Solar Global Horizontal Irradiance Metadata*, Online, <http://www.bom.gov.au/climate/how/IDCJAD0111.shtml>.

[41] Parquet, A 2023, *Apache Parquet Documentation*, Online, viewed 10/07/2023, <https://parquet.apache.org/docs/>.

[42] Melnik, S, Gubarev, A, Long, JJ, Romer, G, Shivakumar, S, Tolton, M & Vassilakis, T 2010, 'Dremel: interactive analysis of web-scale datasets', *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 330-9.

[43] AEMO 2017, *MSATS PROCEDURES National Metering Identifier Procedure,* A Markets, https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Retail-and-metering/-/media/EBA9363B984841079712B3AAD374A859.ashx>.

[44] Energy, E 2023, *What is the NMI?*, Energy Queensland, Online, viewed 08/07/2023, <https://www.ergon.com.au/network/connections/metering/what-is-the-nmi>.

[45] AER 2012, *Confidentiality requirements for energy, metering and NMI standing data*, AE Regulator, Commonwealth of Australia, Online, https://www.aer.gov.au/system/files/D12%2065032%5BV2%5D%20%2020120629%20-%20Compliance%20Bulletin%208%20-%20Confidentiality%20requirements%20for%20energy%2C%20metering%20and%20NMI%20standing%20data%20-%20for%20web%20upload_0.pdf>.

[46] Kushida, CA, Nichols, DA, Jadrnicek, R, Miller, R, Walsh, JK & Griffin, K 2012, 'Strategies for de-identification and anonymization of electronic health record data for use in multicenter research studies', *Medical care*, vol. 50, no. Suppl, p. S82.

[47] Nunez, E, Steyerberg, EW & Nunez, J 2011, 'Regression modeling strategies', *Revista Española de Cardiología (English Edition)*, vol. 64, no. 6, pp. 501-7.

[48] Harrell, FE 2017, 'Regression modeling strategies', *Bios*, vol. 330, no. 2018, p. 14.

[49] Freund, RJ, Wilson, WJ & Sa, P 2006, *Regression analysis*, Elsevier.

Appendices

Appendix A – Research Project Specification for ENG4111 & ENG4112

ENG4111/4112 Research Project

Project Specification

For: Dylan Perrett

Title: Reactive Power Contribution of Grid Connected Solar

Major: Electrical & Electronic Engineering

Supervisors: Dr Jason Brown, University of Southern Queensland
Dr Joel Kennedy, Energy Queensland Ltd

Confidentiality: The data being accessed contains energy usage and identifiers of individual customers. There will be de-identification of any specific linking information that will distinguish individuals or small groups of customers before publishing.

Enrollment: ENG4111 – EXT S1, 2023
ENG4112 – EXT S2, 2023

Project Aim: To identify and understand reactive power performance of grid-connected solar photovoltaic (PV) systems during a range of operating conditions while comparing performance with requirements of AS4777.2. Furthering, an examination of reactive power performance on power quality in relation to voltage magnitude disturbance, comparing this performance in contrast with irradiance conditions.

Programme**: Version 1, 15th March 2023**

1. Conduct initial background research linking grid-connected solar PV and voltage management, include financial challenges that over/under voltage imposes from a general and refined Energy Queensland Limited (EQL) perspective. Examine applicable Australian standard (AS4777.2:2005, 2015 or 2020) and policies that are associated with theoretical and practical operation of grid-connected solar PV systems and power flow metering. Review existing literature related to grid-connected solar PV system operation during all reasonably encountered irradiance conditions.
2. Analyse EQL provided irradiance data confirming appropriate site locations. Relate site selection/s with de-identified corresponding customer consumption and PV export data. Identify a suitable case study to support potential findings relating to voltage management.
3. Understanded and relate the reactive power contributions of an average solar PV unit in Queensland under differing conditions determining the voltage contributions are within accordance with Australian Standard AS4777.2 of operation.
4. Review and categorise this over/under voltage contribution while relating any financial or practical responsibilities required by all significant stakeholders including at a minimum customer and distribution network operator (DNO).

5. Apply regression modelling techniques to reactive power output of solar PV capacity under different irradiance conditions.

If time and resources permit:

6. Refine and develop the modeled data in extrapolating a potential future reactive power contribution forecasting algorithm, depending on previous achievements.
7. Choose a feeder that has potential from an overvoltage perspective and apply forecasting algorithm with future demand solar PV forecasts published by Australian energy market operator (AEMO) and Commonwealth scientific and industrial research organisation (CSIRO). (This option may require access to specialised software, potential to exhaust time constraints)

Appendix B – Project Resources

These resources are already readily available and can either be acquired throughout previous studying efforts, personal acquisition or supplied via UniSQ as a part of student licencing accessibility.

Item	Source	Cost ($AUS)	Comments
Computer Operating on Windows	Personally accessed. UniSQ	Nil	Personal PC and/or laptop UniSQ networked student computers as redundancy
Internet Access	Personal/university	Various	Personal household cost. UniSQ student access as redundancy
Microsoft Office 365	Student licence	Nil	Word: Dissertation processing Excel: additional data analysis tool PowerPoint: Presentation processing EndNote: UniSQ recommended research referencing software
Grammarly	Personal purchase	~200	Identify and improve writing. Some plagiarism detection tooling
MATLAB	Student licence University supplied	Nil	Data analysis and modelling tool
Visual Studio Code & Associated Components	Open Sourced	Nil	Analysis, data management, modelling and machine learning
Sundries (pens, paper, printing, etc)	Personal/UniSQ printing budget	0-100	N/A

Table 10: Resources required for project.

Appendix C – Risk Mitigation Plan

	Consequence				
Probability	**A**. Insignificant	**B.** Minor	**C.** Serious	**D.** Major	**E.** Catastrophic
1. Almost Certain	MEDIUM	HIGH	EXTREME	EXTREME	EXTREME
2. Likely	MEDIUM	HIGH	HIGH	EXTREME	EXTREME
3. Possible	LOW	MEDIUM	HIGH	HIGH	HIGH
4. Unlikely	LOW	LOW	MEDIUM	MEDIUM	MEDIUM
5. Rare	LOW	LOW	LOW	LOW	LOW

Table 11: Five-by-Five Consequence Matrix

Risk score before	**Phase**	**Hazzard**	**Minimisation**	**New risk score**
	Phase One	Project Preparation		
3E	1A	Unsuccessful project approval	Contact with potential supervisor early to confirm and define scope	4B
4C	1B	Resource licencing	Project is using open-source software	5A
3D	1C	Confidentiality	Agreement with organisations and deidentification of data	4A
	Phase Two	Characterisation of data gathering		
3C	2A	Incorrect population sizing	Review and adjust	4B
2D	2B	Unable to achieve real-time data processing for configuration	Review and adjust profiling and management	3A
2C	2C	Minimal Literature	Clarify and confirm methods with supervisor define understanding from other sources	3A
	Phase Three	Refinement of data Characterisation in phase two		
2C	3A	Data will not fit towards model	Move to 3B	3A
3C	3B	Incorrect calculations and/or model	Re-evaluate characterisation with existing models and review with supervisor	3A
	Phase Four	Regression Modelling		
2C	All 4	Model not adequate	Re-evaluate definition	3A
	Phase Five	Data analysis		
3B	All 5	Project does not achieve performance	Review project with the potential for a failed simulation incorporate into results discussions	4B
	Phase Six	Dissertation write-up and presentation of results		
3D	All 6	Computer failure/loss of records Insufficient time due to unforeseeable events	Multi-platform backup of work, implement version control for each phase. Keep communication channels open with supervisor possible small extension application/s	2A

Table 12: Project risk profile

Appendix D –Plan (Gantt Chart)

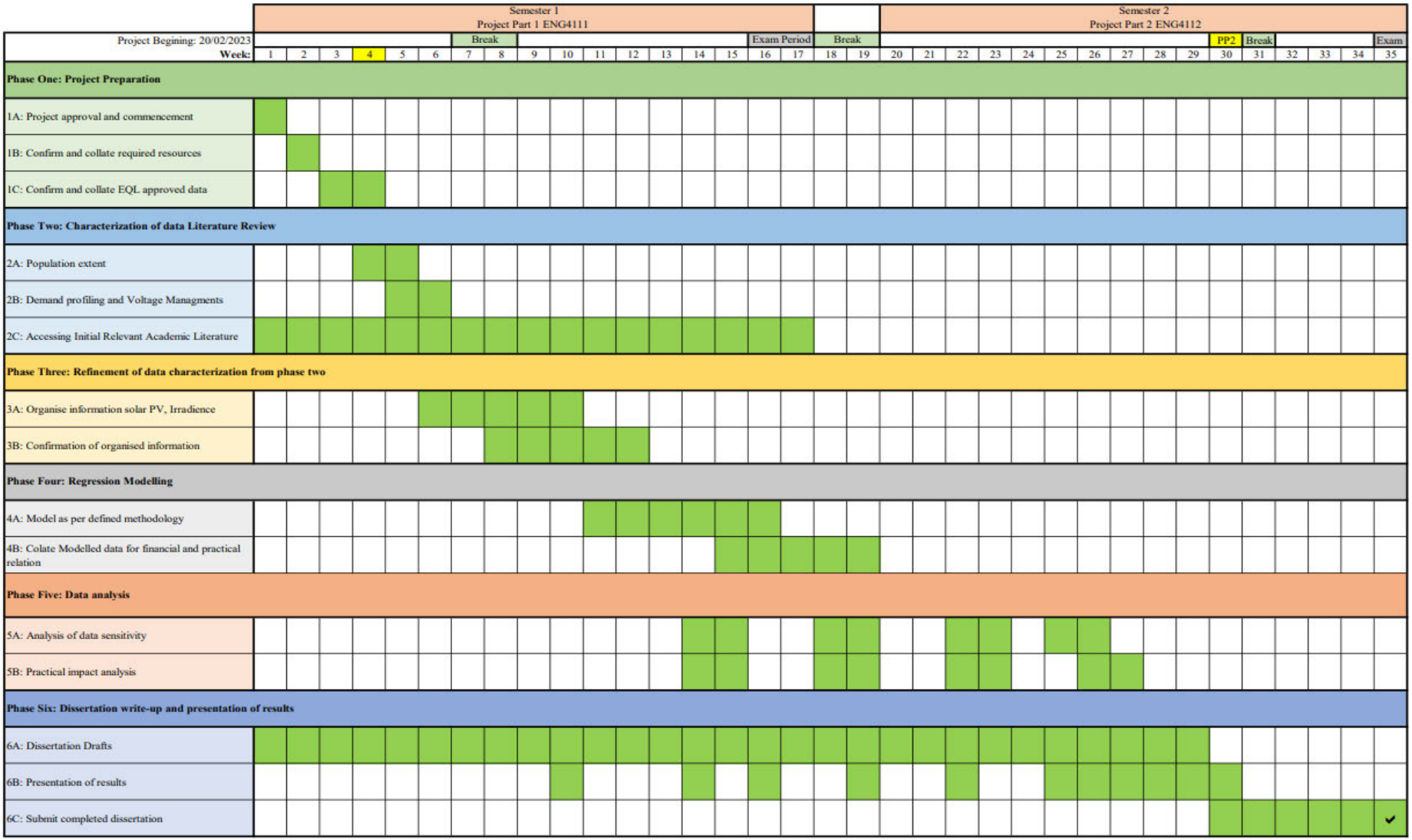


Figure 34: Research Project Plan (Gant Chart)

Appendix E – MATLAB Source Code

Introduction graph

```
%% Thesis Introduction Plot

% Example of In-Phase, Lagging, and Leading Currents on Power
% Produced for course ENG4111 & ENG4112
% student name: Dylan Perrett
% student number: [redacted]
% Draft Due Date: 24/05/2023

clc, clear, close all

%----------------------------------------
%% Wave Set-up

% Random Definitions
Fs = 8000;
dt = 1/Fs;
f = 50; % frequency in (hertz)
T = 1/f;
tt = (0:dt:T); % One cycle of 50Hz

omega = 2*pi*f;
Vamp =10*sqrt(2); % Voltage Peak
Iamp = 2*sqrt(2); % Current Peak

Phi = 0; % Phase in Radians
Alt_Phi = pi/3; % Phi change in current
V = Vamp*sin(omega*tt+Phi); % Voltage wave

% Current Waves
i = Iamp*sin(omega*tt+Phi); % i in phase
i_Lag = Iamp*sin(omega*tt-Alt_Phi); % i Lagging Inductive Load
i_Lead = Iamp*sin(omega*tt+Alt_Phi); % i Leading Capacitive Load

% Power
P = V.*i; % inPhase
P_Lag = V.*i_Lag; % Lagging
P_Lead = V.*i_Lead; % Leading

%----------------------------------------
%% Plot the Output
t = tiledlayout(2,2);
t.Padding = 'compact';
t.TileSpacing = 'compact';

% Display In Phase
nexttile([1 2])
P_plot = hatchfill2(area(tt,P), 'cross', 'HatchAngle', 45,
'HatchDensity', 40, 'HatchColor', 'yellow');
hold on
V_plot = plot(tt,V,'b','LineWidth',2);
hold on
i_plot = plot(tt,i,'r','LineWidth',2);
hold off
```

```
title({'\fontsize{14}\bfIn Phase'});
ylim([(min(V)*(1.05)) (max(P)*1.05)]); % 5% over Amplitudes
grid on;
xlabel('\fontsize{12}\bftime (s)');
ylabel('\fontsize{12}\bfAmplitude');
yticklabels({'','0'})
legend([V_plot i_plot P_plot],{'Voltage', 'Current',' Power'})

% Display Lagging
nexttile([1 1])
PLag_plot = hatchfill2(area(tt,P_Lag), 'single', 'HatchAngle', 45 ,
'HatchDensity', 60, 'HatchColor', 'yellow');
hold on
V_plot = plot(tt,V,'b','LineWidth',2);
hold on
iLag_plot = plot(tt,i_Lag,'r','LineWidth',2);
hold off
title({'\fontsize{14}\bfLagging (Inductive Load)'});
ylim([(min(V)*(1.05)) (max(P_Lag)*1.05)]); % 5% over Amplitudes
grid on;
xlabel('\fontsize{12}\bftime (s)');
ylabel('\fontsize{12}\bfAmplitude');
yticklabels({'','0'})

% Display Leading
nexttile([1 1])
PLead_plot = hatchfill2(area(tt,P_Lead), 'single', 'HatchAngle',
45+90 , 'HatchDensity', 60, 'HatchColor', 'yellow');
hold on
V_plot = plot(tt,V,'b','LineWidth',2);
hold on
iLead_plot = plot(tt,i_Lead,'r','LineWidth',2);
hold off
title({'\fontsize{14}\bfLeading (Capactive Load)'});
ylim([(min(V)*(1.05)) (max(P_Lead)*1.05)]); % 5% over Amplitudes
grid on;
xlabel('\fontsize{12}\bftime (s)');
ylabel('\fontsize{12}\bfAmplitude');
yticklabels({'','0'})
```

Appendix F – Python Source Code

Initial Draft Code

```
'''
Thesis Data Interpolation & Organisation

Convert erratic data increments to 5-minute intervals
Produced for course ENG4111 & ENG4112
student name: Dylan Perrett
student number: [redacted]
Draft Due Date: 24/05/2023
'''

import pandas as pd
import glob
from pathlib import Path


path = 'Path to original file location'

# Generate an array of files to evaluate (note: path)
all_files = glob.glob(path + '*.parquet')
# print(all_files) # <-- Evaluation point

'''
Clean data and Interpolate to achieve 5-min intervals.
For each original file Pad data with NAN/s to 1-min intervals and interpolate
for increased accuracy
then gather 5-min intervals and place in new file ready for analysis stage.
'''

for filename in all_files:
    df = pd.read_parquet(filename)
    file_name = Path(filename).stem # isolate filename
    #print(file_name) # <-- Evaluation point
    df = df.reset_index()
    df = df.iloc[:,1:] # locks columns from index 1('ts') onwards using an int
    # Can define specific column headings (but one data-set had different
column designations)
    #df = df.loc[:,['ts', 'point_web_id_Current_A',
'point_web_id_Current_Solar',
                #'point_web_id_Impedance', 'point_web_id_Power_Factor',
                #'point_web_id_Voltage_A', 'point_web_id_Voltage_Variation']]
    #print(df) # <-- Evaluation point
    df_1min = df.groupby(pd.Grouper(key='ts', freq='1min')).first() # Pad data
with NAN/s
    df_1min = df_1min.reset_index()
    df_interpol_1min = df_1min.interpolate() # Interpolate 1-min interval
frame (note: I found errors when grouping 5-min straight away)
```

```
    df_5min = df_interpol_1min.groupby(pd.Grouper(key='ts',
freq='5min')).first().round(2) # rounding to 2 as this may be useful for Power
Factor Analysis
    #df_5min = df_5min.reset_index() # Can reset index but may not need to.
    newFilePath = 'Create New Path to store interpolated data'
    df_5min.to_parquet(newFilePath + '{name}.parquet'.format(name =
file_name))
```

Final Source Code

```
'''
Data Organisation & Master file

Convert erratic data increments to 5-minute intervals
Produced for course ENG4111 & ENG4112
student name: Dylan Perrett
student number: [redacted]
Final Due Date: 15/10/2023
Ext Due Date: 18/10/2023

'''


# Library Import/s
import pandas as pd
import numpy as np
import os
import glob
from pathlib import Path
from datetime import datetime
import pytz
import matplotlib.pyplot as plt


# Variable Setup
zerothStep = 0
backTwoStep = -2

unity = 1
squared = 2
half = 1/2

thirtyMins = 30
roundingtoOne = 1
roundingtoTwo = 2
roundingtoThree = 3
roundingtoFour = 4
roundingtoFive = 5
roundingtoSix = 6
hoursInDay = 24
monthsInYear = 12
```

```
tenDigits = 10
elevenDigits = 11

# Initialisation of lists, arrays or DataFrame storage
unique_all_NMIs = []
unique_nov_NMIs = []
unique_mar_NMIs = []
unique_jan_NMIs = []

# File or Folder Paths into Virtual Environment
path_to_Virtual_Loc = r[“DESENSITISED”]

# Original data file Locations
path_Nov_Data = r[“DESENSITISED”]
path_Mar_Data = r[“DESENSITISED”]
path_Jan_Data = r[“DESENSITISED”]
irr_Path = r[“DESENSITISED”] # Irradiance

# Sorted Data file Locations
merged_path_A = r[“DESENSITISED”]
merged_path_B = r[“DESENSITISED”]
merged_path_C = r[“DESENSITISED”]
sorted_IRR_Data = r[“DESENSITISED”] # Irradiance
joined_PQ_to_IRR = r[“DESENSITISED”]

# Array generation of files paths to evaluate/organise
files_Nov = glob.glob(os.path.join(path_to_Virtual_Loc, path_Nov_Data,
'*.parquet'))
files_Mar = glob.glob(os.path.join(path_to_Virtual_Loc, path_Mar_Data,
'*.parquet'))
files_Jan = glob.glob(os.path.join(path_to_Virtual_Loc, path_Jan_Data,
'*.parquet'))
all_files =  files_Mar + files_Jan + files_Nov
all_IRR_Files = glob.glob(os.path.join(path_to_Virtual_Loc, irr_Path, '*',
'*.csv')) # Irradiance

# Generate Series with Unique NMIs for sorting
for path in all_files:
    filename = Path(path).stem # isolate filename
    sections = (filename.split('.'))
    if sections[zerothStep] not in unique_all_NMIs:
        if (len(sections[zerothStep]) < elevenDigits) &
(len(sections[zerothStep]) >= tenDigits):
            unique_all_NMIs.append(sections[0])

for path in files_Nov:
    filename = Path(path).stem # isolate filename
    sections = (filename.split('.')) # isolate NMI
```

```
    if sections[zerothStep] not in unique_nov_NMIs:
        if (len(sections[zerothStep]) < elevenDigits) &
(len(sections[zerothStep]) >= tenDigits):
            unique_nov_NMIs.append(sections[zerothStep])

for path in files_Mar:
    filename = Path(path).stem # isolate filename
    sections = (filename.split('.')) # isolate NMI
    if sections[zerothStep] not in unique_mar_NMIs:
        if (len(sections[zerothStep]) < elevenDigits) &
(len(sections[zerothStep]) >= tenDigits):
            unique_mar_NMIs.append(sections[zerothStep])

for path in files_Jan:
    filename = Path(path).stem # isolate filename
    sections = (filename.split('.')) # isolate NMI
    if sections[zerothStep] not in unique_jan_NMIs:
        if (len(sections[zerothStep]) < elevenDigits) &
(len(sections[zerothStep]) >= tenDigits):
            unique_jan_NMIs.append(sections[zerothStep])

# Functions

def convert_datetime_timezone(dt, tz1, tz2):
    tz1 = pytz.timezone(tz1)
    tz2 = pytz.timezone(tz2)

    dt = datetime.strptime(dt,"%Y-%m-%dT%H:%M:%SZ") #  T (meaning: time or
used as a spacer) & Z (meaning: zulu) are litterals
    dt = tz1.localize(dt)
    dt = dt.astimezone(tz2)
    dt = dt.strftime("%Y-%m-%d %H:%M:%S")

    return dt

def get_half_interval(minutes):
    minutes = minutes/thirtyMins
    if roundingtoOne <= minutes < roundingtoTwo:
        return '30'
    else:
        return '00'

#Season function
def get_season(month):
    x = month%monthsInYear // roundingtoThree + roundingtoOne
    if x == roundingtoOne:
        season = 'Summer'
    if x == roundingtoTwo:
```

```
        season = 'Autumn'
    if x == roundingtoThree:
        season = 'Winter'
    if x == roundingtoFour:
        season = 'Spring'
    return season

def get_Time_Ref(hour):
    x = (hour%hoursInDay + roundingtoFour) // roundingtoFour
    if x == roundingtoOne:
        definedPeriod = 'Late Night'
    if x == roundingtoTwo:
        definedPeriod = 'Early Morning'
    if x == roundingtoThree:
        definedPeriod = 'Morning'
    if x == roundingtoFour:
        definedPeriod = 'Noon'
    if x == roundingtoFive:
        definedPeriod = 'Evening'
    if x == roundingtoSix:
        definedPeriod = 'Night'
    return definedPeriod

def pf_assumption_fix(pf):
    if (pf >= unity or pf <= (-unity)):
        pf = unity
    else:
        pf = pf

    return pf

def solar_current_assumption_fix(df):
    df.loc[df['Solar_Current'] < 0 , 'Solar_Current'] *= -1

    return df

def negate(number):
    return -number

def sort_dfs(df, NMI, columns_to_drop, column_arrangement):
    df = df.reset_index()
    df = df.drop(columns = columns_to_drop)
    df = df.rename(columns = column_arrangement)
    df['NMI'] = NMI
    reorder_columns = ['ts', 'NMI', 'Voltage', 'Agg_Current', 'Solar_Current',
'Power_Factor']
    df = df[reorder_columns]
```

```
    return df

def cleanse_df(df, NMI):
    df = solar_current_assumption_fix(df)
    df['Power_Factor'] = [pf_assumption_fix(count) for count in
df['Power_Factor']]
    df = df.groupby(pd.Grouper(key = 'ts', freq =
'5min')).first().round(roundingtoTwo) # rounding to 2 decimal points as this
may be useful for PowerFactor Analysis
    df = df.reset_index()


    return df

def power_calc_df(df):
    apparentPower = pd.DataFrame({'Agg_Apparent_Power (S)':
df['Voltage']*(df['Agg_Current'].abs())})
    realPower = pd.DataFrame({'Agg_Real_Power (P)':
negate(apparentPower['Agg_Apparent_Power (S)'].abs() *
(np.copysign(df['Power_Factor'].abs(),df['Agg_Current'])))}).round(roundingtoT
wo)
    reactivePower = pd.DataFrame({'Agg_Reactive_Power (Q)':
negate(np.copysign((apparentPower['Agg_Apparent_Power (S)'].abs().pow(squared)
- realPower['Agg_Real_Power
(P)'].pow(squared)).pow(half),df['Power_Factor']))}).round(roundingtoTwo)

    apparentSolar = pd.DataFrame({'Solar_Apparent_Power (Spv)':
df['Voltage'].abs()*df['Solar_Current'].abs()})
    realPowerSolar = pd.DataFrame({'Solar_Real_Power (Ppv)':
apparentSolar['Solar_Apparent_Power (Spv)'].abs() *
(np.copysign(df['Power_Factor'].abs(),df['Solar_Current']))}).round(roundingto
Two)
    reactivePowerSolar = pd.DataFrame({'Solar_Reactive_Power (Qpv)':
negate(np.copysign((apparentSolar['Solar_Apparent_Power
(Spv)'].abs().pow(squared) - realPowerSolar['Solar_Real_Power
(Ppv)'].pow(squared)).pow(half),df['Power_Factor']))}).round(roundingtoTwo)

    halfHrPeriods = pd.DataFrame({'Half Hour Period':
df['ts'].dt.hour.astype(str) + ':' +
df['ts'].dt.minute.apply(get_half_interval)})
    halfHrPeriods['Half Hour Period'] = pd.to_datetime(halfHrPeriods['Half
Hour Period'], format='%H:%M').dt.time
    year = pd.DataFrame({'Year': df['ts'].dt.year})
    season = pd.DataFrame({'Season': df['ts'].dt.month.apply(get_season)})
    periodOfDay = pd.DataFrame({'Defined 24hr Period':
df['ts'].dt.hour.apply(get_Time_Ref)})
    df =
df.join(apparentPower).join(realPower).join(reactivePower).join(apparentSolar)
```

```
.join(realPowerSolar).join(reactivePowerSolar).join(halfHrPeriods).join(year).
join(season).join(periodOfDay)


    return df


# Main Code

'''
# IRR Tidying
for path in all_IRR_Files:
    filenumber = ((path.split('\\'))[backTwoStep].split('_'))[zerothStep] #
isolate IRR N.o. 0-19?
    print(filenumber)
    irr_data = pd.read_csv(path)
    irr_columns_to_drop =
['PeriodEnd','Period','CloudOpacity','Dhi','Dni','GtiFixedTilt','GtiTracking',
'Zenith']
    irr_data = irr_data.drop(columns = irr_columns_to_drop)
    irr_data = irr_data.rename(columns={'PeriodStart': 'ts'})
    irr_data['ts'] = pd.Series(pd.to_datetime([convert_datetime_timezone(ts,
'UTC', 'Australia/Queensland') for ts in irr_data['ts']]),
dtype='datetime64[ns]').values

    NewFilePath = os.path.join(path_to_Virtual_Loc, sorted_IRR_Data)
    filename = '{name}.IRRData.parquet'.format(name = filenumber)
    irr_data.to_parquet(os.path.join(NewFilePath, filename))
    print(filenumber + ' Tick parquet')
'''

# Measurement Data Tidying
for NMI in unique_all_NMIs:
    print(NMI)
    columns_to_merge = ['ts', 'NMI', 'Agg_Current', 'Solar_Current',
'Power_Factor', 'Voltage']
    if NMI in unique_nov_NMIs:
        path_nov = os.path.join(path_to_Virtual_Loc, path_Nov_Data)
        path = os.path.join(path_nov, ('{NMI}.PhaseA.parquet'.format(NMI =
NMI)))
        CA = {'asset':'NMI', 'point_web_id_Grid_Current':'Agg_Current',
'point_web_id_Solar_Current':'Solar_Current','point_web_id_Power_Factor':'Powe
r_Factor','point_web_id_Grid_Voltage':'Voltage'}
        df0 = pd.read_parquet(path)
        if 'point_web_id_Impedance' not in df0.columns:
            CtD = ['asset', 'point_web_id_Grid_Voltage_Variation']
        elif 'point_web_id_Neutral_Current' in df0.columns:
             CtD = ['asset', 'point_web_id_Grid_Voltage_Variation',
'point_web_id_Impedance', 'point_web_id_Neutral_Current']
```

```
        else:
            CtD = ['asset', 'point_web_id_Grid_Voltage_Variation',
'point_web_id_Impedance']
        df0 = sort_dfs(df0, NMI, CtD, CA)
    if NMI in unique_mar_NMIs:
        path_mar = os.path.join(path_to_Virtual_Loc, path_Mar_Data)
        path = os.path.join(path_mar, ('{NMI}.PhaseA.parquet'.format(NMI =
NMI)))
        CA = {'asset':'NMI', 'point_web_id_Current_A':'Agg_Current',
'point_web_id_Current_Solar':'Solar_Current',
'point_web_id_Power_Factor':'Power_Factor',
'point_web_id_Voltage_A':'Voltage'}
        df1 = pd.read_parquet(path)
        if 'point_web_id_Impedance' not in df1.columns:
            CtD = ['asset', 'point_web_id_Voltage_Variation']
        else:
            CtD = ['asset', 'point_web_id_Voltage_Variation',
'point_web_id_Impedance']
        df1 = sort_dfs(df1, NMI, CtD, CA)
    if NMI in unique_jan_NMIs:
        path_jan = os.path.join(path_to_Virtual_Loc, path_Jan_Data)
        path = os.path.join(path_jan, ('{NMI}.PhaseA.parquet'.format(NMI =
NMI)))
        CA = {'asset':'NMI', 'point_web_id_Current_A':'Agg_Current',
'point_web_id_Current_Solar':'Solar_Current',
'point_web_id_Power_Factor':'Power_Factor',
'point_web_id_Voltage_A':'Voltage'}
        df2 = pd.read_parquet(path)
        if 'point_web_id_Impedance' not in df2.columns:
            CtD = ['asset', 'point_web_id_Voltage_Variation']
        else:
            CtD = ['asset', 'point_web_id_Impedance',
'point_web_id_Voltage_Variation']
        df2 = sort_dfs(df2, NMI, CtD, CA)

    if (NMI in unique_nov_NMIs) and (NMI in unique_mar_NMIs) and (NMI in
unique_jan_NMIs):
        merge1 = pd.merge(df0,df1, how='outer', on = columns_to_merge)
        merge2 = pd.merge(merge1,df2, how='outer', on = columns_to_merge)
        df = merge2
    elif (NMI in unique_nov_NMIs) and (NMI in unique_mar_NMIs) and (NMI not in
unique_jan_NMIs):
        merge1 = pd.merge(df0,df1, how='outer', on = columns_to_merge)
        df = merge1
    elif (NMI not in unique_nov_NMIs) and (NMI in unique_mar_NMIs) and (NMI in
unique_jan_NMIs):
        merge1 = pd.merge(df1,df2, how='outer', on = columns_to_merge)
        df = merge1
```

```
    elif (NMI in unique_nov_NMIs) and (NMI not in unique_mar_NMIs) and (NMI in
unique_jan_NMIs):
        merge1 = pd.merge(df0,df2, how='outer', on = columns_to_merge)
        df = merge1
    elif (NMI in unique_nov_NMIs) and (NMI not in unique_mar_NMIs) and (NMI
not in unique_jan_NMIs):
        df = df0
    elif (NMI not in unique_nov_NMIs) and (NMI in unique_mar_NMIs) and (NMI
not in unique_jan_NMIs):
        df = df1
    elif (NMI not in unique_nov_NMIs) and (NMI not in unique_mar_NMIs) and
(NMI in unique_jan_NMIs):
        df = df2
    else:
        print('Error has occured during merge process!')
        break

    df = cleanse_df(df,NMI)
    df = power_calc_df(df)

    NewFilePath = os.path.join(path_to_Virtual_Loc, merged_path_A)
    filename = '{name}.parquet'.format(name = NMI)
    df.to_parquet(os.path.join(NewFilePath, filename))
    print(NMI + ' Tick parquet cleansed')


# Join Sorted Measurement data with respective IRR location
dfunique =
pd.read_parquet(os.path.join(path_to_Virtual_Loc,'unique_NMIs.parquet'))
NMI_to_IRR =
pd.read_parquet(os.path.join(path_to_Virtual_Loc,'NMI_to_IRR.parquet'))
IRR_to_Data = pd.merge(NMI_to_IRR,dfunique, how='inner', on='NMI')

# During merge process found some NMIs that had been excluded from analysis.
errorNMIs = [“DESENSITISED”]

for idx in IRR_to_Data.index:
    NMI = IRR_to_Data['NMI'][idx]
    if NMI in errorNMIs:
        print('No data or different Phase B/C was present, move through -->
Unused')
    else:
        PQfilename = '{name}.parquet'.format(name = NMI)
        dfPQ = pd.read_parquet(os.path.join(path_to_Virtual_Loc,
merged_path_A, PQfilename))
        fileID = IRR_to_Data['FILE_ID'][idx]
        IRRfilename = '{number}.IRRData.parquet'.format(number = fileID)
        dfIRR = pd.read_parquet(os.path.join(path_to_Virtual_Loc,
sorted_IRR_Data, IRRfilename))
```

```
        merged = dfPQ.merge(dfIRR, how='left', on='ts')
        merged['Site'] = fileID

        NewFilePath = os.path.join(path_to_Virtual_Loc, joined_PQ_to_IRR)
        filename = '{name}.{number}.parquet'.format(name = NMI, number =
fileID)
        merged.to_parquet(os.path.join(NewFilePath, filename))
        print(NMI, ' ', fileID, ' Tick joined IRR')


IRR_files_to_concat = glob.glob(os.path.join(path_to_Virtual_Loc,
joined_PQ_to_IRR, '*.parquet'))

df_from_each_file = (pd.read_parquet(f) for f in IRR_files_to_concat)
Master_df_inc_IRR = pd.concat(df_from_each_file, ignore_index = True)

NewFilePath = os.path.join(path_to_Virtual_Loc, 'MergedData')
filename = '{name}.parquet'.format(name = 'Master_df_inc_IRR')
Master_df_inc_IRR.to_parquet(os.path.join(NewFilePath, filename))

'''
Analysis & Regression Modelling

Produced for course ENG4111 & ENG4112
student name: Dylan Perrett
student number: [redacted]
Final Due Date: 15/10/2023
Ext Due Date: 18/10/2023
'''

# Library Import/s
import pandas as pd
import glob
from pathlib import Path
from datetime import datetime
import os

import seaborn as sns
import matplotlib.pyplot as plt
from sklearn import linear_model
from sklearn import preprocessing
from sklearn.preprocessing import PolynomialFeatures

from sklearn.metrics import mean_squared_error, r2_score
from sklearn.impute import KNNImputer


n = 0
```

```
perhr = 60/12
val = 0.5

path_to_Virtual_Loc = r[“DESENSITISED”]

merged_path_A = r[“DESENSITISED”]

joined_PQ_to_IRR = r[“DESENSITISED”]


# No Inverter?
no_Inverter = []
siteNumber = []
NMIs = []
Numbers = []

all_files = glob.glob(os.path.join(path_to_Virtual_Loc, joined_PQ_to_IRR) +'/'
'*.parquet')

for path in all_files:
    filename = Path(path).stem # isolate filename
    NMIs.append((filename.split('.'))[0])
    N = (filename.split('.'))[1]
    siteNumber.append(N)
    if N not in Numbers:
        Numbers.append(N)

path = os.path.join(path_to_Virtual_Loc, 'MergedData',
'Master_df_inc_IRR.parquet')
df = pd.read_parquet(path)

invRatings =
pd.read_parquet(os.path.join(path_to_Virtual_Loc,'NMIs_with_InvRating.parquet'
))
invRatings = invRatings.astype({'INV_RATING_KW':'float'})
df = pd.merge(df,invRatings , how='inner', on='NMI')

def per_unit_df(df):
    df['Sagg per Unit'] = df['Agg_Apparent_Power
(S)']/df['INV_RATING_KW'].multiply(1000) # P per unit of Inverter on site
    df['Pagg per Unit'] = df['Agg_Real_Power
(P)']/df['INV_RATING_KW'].multiply(1000) # P per unit of Inverter on site
    df['Qagg per Unit'] = df['Agg_Reactive_Power
(Q)']/df['INV_RATING_KW'].multiply(1000) # Q per unit of Inverter on site

    df['Spv per Unit'] = df['Solar_Apparent_Power
(Spv)']/df['INV_RATING_KW'].multiply(1000) # P per unit of Inverter on site
    df['Ppv per Unit'] = df['Solar_Real_Power
(Ppv)']/df['INV_RATING_KW'].multiply(1000) # P per unit of Inverter on site
```

```
    df['Qpv per Unit'] = df['Solar_Reactive_Power
(Qpv)']/df['INV_RATING_KW'].multiply(1000) # Q per unit of Inverter on site

    return df



sns.set_style('whitegrid')

twentyPercent = 0.2
sixtyPercent = 0.6
eightyPercent = 0.8


def get_range(value):
    if value <= twentyPercent:
        return 'Low'
    elif twentyPercent < value <= sixtyPercent:
        return 'Mid'
    elif sixtyPercent < value <= eightyPercent:
        return 'High'
    else:
        return 'Out of Range'


df = per_unit_df(df)
df['S Range'] = df['Spv per Unit'].apply(get_range)

# Boundary Conditions
#Low
#df = df[df['Spv per Unit'] <= twentyPercent]

#Mid
#df = df[df['Spv per Unit'] <= sixtyPercent]
#df = df[df['Spv per Unit'] > twentyPercent]

#High
df = df[df['Spv per Unit'] <= eightyPercent]
#df = df[df['Spv per Unit'] > sixtyPercent]


df = df[abs(df['Solar_Current'] + df['Agg_Current']) <=
df['Solar_Current'].abs().multiply(0.05)]
#df = df[df['Ghi'] <= 200]

#print(df['Agg_Reactive_Power (Q)'])
df['Ppv bins'] = pd.cut(df['Ppv per Unit'], 7,precision=0)

df['Spv bins'] = pd.cut(df['Spv per Unit'], 30,precision=0)
```

```
df['Ghi bins'] = pd.cut(df['Ghi'], 12, precision=0)

unitOfD = 3

first = 1
second = 2
third = 3
forth = 4

SpuVals = 'Spv per Unit'
QpuVals = 'Qpv per Unit'
PpuVals = 'Ppv per Unit'
PBins = 'Ppv bins'
SBins = 'Spv bins'
pfVals = 'Power_Factor'
PVRatings = 'INV_RATING_KW'
hhrVals = 'Half Hour Period'
irrVals = 'Ghi'
volts = 'Voltage'
# Sorted data by half hours slots

dfhhr = df.sort_values(by = [hhrVals], ascending = True)
#dfGhi = df.sort_values(by = [irrVals], ascending = True)


n = n+1
plt.figure(n)
sns.boxplot(data = df, x = SBins, y = pfVals).set(title = 'Boxplot of PV
Reactive Q Vs Apparent S of Inverter Contribution for High Range of S', xlabel
= 'PV S[inv] p.u. (Grouped)', ylabel = 'PV Q p.u.')
plt.xticks(rotation=45, ha='right')
plt.annotate('> 0, Inductive',xy = (0.01,0.7), fontsize=16)
plt.annotate('< 0, Capacitive',xy = (0.01,-0.74), fontsize=16)
plt.axhline(y = 0.6, color = 'r', linestyle = '--', label =  '+/- 60%,
4777.2:2020 max Compliance')
plt.axhline(y = 0.44, color = 'orange', linestyle = '--', label = '+/- 44%,
4777.2:2020 min Compliance')
plt.axhline(y = -0.44, color = 'orange', linestyle = '--')
plt.axhline(y = -0.6, color = 'r', linestyle = '--')
plt.legend(loc = 'upper center')


dfGhi = df.groupby(irrVals)[QpuVals].mean().reset_index()


# Create linear regression object
regr = linear_model.LinearRegression()
ransac = linear_model.RANSACRegressor()
```

```
QvS = df.groupby(SBins)[QpuVals].mean().reset_index()
dfHH = df.groupby(hhrVals)[QpuVals].mean().reset_index()

dfHH['Index Rank'] = (dfHH.index - dfHH.index[0])
QvS['Index Rank'] = (QvS.index - QvS.index[0])


imputer = KNNImputer()
#Qvt
#X = dfHH['Index Rank'].values.reshape(-1, 1)
#y = dfHH['Qpv per Unit'].values.reshape(-1, 1)
#QvS
X = QvS['Index Rank'].values.reshape(-1, 1)
y = QvS['Qpv per Unit'].values.reshape(-1, 1)
# & QvP
#X = QvP['Index Rank'].values.reshape(-1, 1)
#y = QvP['Qpv per Unit'].values.reshape(-1, 1)
#QvIrr
#X = dfGhi['Ghi'].values.reshape(-1, 1)
#y = dfGhi['Qpv per Unit'].values.reshape(-1, 1)
#y = imputer.fit_transform(y)
poly = PolynomialFeatures(degree=unitOfD, include_bias=False)

poly_features = poly.fit_transform(X)
lab_enc = preprocessing.LabelEncoder()

#regr.fit(X, y)
regr.fit(poly_features, y)
ransac.fit(X, y)
y_ransac = ransac.predict(X)
y_predicted = regr.predict(poly_features)

# The coefficients
print("Estimated coefficients (linear regression, RANSAC):")
print(regr.coef_, ransac.estimator_.coef_)
print("Mean squared error: %.5f" % mean_squared_error(y, y_ransac))
print("Mean squared error: %.5f" % mean_squared_error(y, y_predicted))
print("Coefficient of determination: %.2f" % r2_score(y, y_ransac))
print("Coefficient of determination: %.2f" % r2_score(y, y_predicted))
print(f"intercept: {regr.intercept_}")

n = n+1
plt.figure(n)
# Plot outputs
plt.scatter(X, y, color="black")
```

```
plt.plot(X, y_ransac, color="cornflowerblue", linewidth=3, label="RANSAC
regressor")
plt.plot(X, y_predicted, color='red', linewidth=3, label="Poly regressor
Degree = 4")
plt.text(15, 0.08,'> 0, Inductive', fontsize=16)
plt.text(4, -0.07,'< 0, Capacitive', fontsize=16)

'''
# Regression Qvt
Qvtplot = sns.lmplot(data = dfHH, x = 'Index Rank', y = 'Qpv per Unit', order
= unitOfD, scatter_kws={"color": "cornflowerblue"}, line_kws={"color":
"darkgreen"})
plt.annotate('< 0, Capacitive',xy = (2,-0.09), fontsize=16)
plt.xticks(range(len(dfHH)), labels = dfHH['Half Hour Period'],rotation=45,
ha='right')
plt.legend(loc='upper right', labels=['Data', ('Coefficient/s:',
regr.coef_[0], 'Intercept:', regr.intercept_[0].round(3))])
plt.title('Regression of PV Q vs t for Low, Mid and High Range of S')
plt.xlabel('Time in 1/2 Hr Periods')
plt.ylabel('PV Q p.u.')
#plt.xlim([0,len(dfHH)])
'''
# Regression QvS
QvSplot = sns.lmplot(data = QvS, x = 'Index Rank', y = 'Qpv per Unit', order =
unitOfD, scatter_kws={"color": "cornflowerblue"}, line_kws={"color":
"darkgreen"})
plt.annotate('< 0, Capacitive',xy = (3,-0.02), fontsize=16)
plt.annotate("R-Squared: %.2f" % r2_score(y, y_predicted),xy = (0.2,-0.25),
fontsize=16)
plt.xticks(range(len(QvS)), labels = QvS['Spv bins'], rotation=30, ha='right')
plt.legend(loc='upper right', labels=['Data', ('Coefficient/s:',
regr.coef_[0], 'Intercept:', regr.intercept_[0].round(3))])
plt.title('Regression of PV Q vs S[inv] for Low, Mid and High Range of S')
plt.xlabel('PV S[inv] p.u. (Grouped)')
plt.ylabel('PV Q p.u.')

'''
# Regresion QvIRR
QvIrr = sns.lmplot(data = dfGhi, x = irrVals, y = QpuVals, order = unitOfD,
scatter_kws={"color": "cornflowerblue"}, line_kws={"color": "darkgreen"})
plt.annotate('> 0, Inductive',xy = (50,0.06), fontsize=16)
plt.annotate('< 0, Capacitive',xy = (50,-0.13), fontsize=16)
plt.legend(loc='upper right', labels=['Data', ('Coefficients:', regr.coef_[0],
'Intercept:', regr.intercept_[0].round(3))])
plt.title('Regression of Q vs Irradiance for Low, Mid and High Range of S')
plt.xlabel('Ghi (W/m^2)')
plt.ylabel('PV Q p.u.')
'''
```

Appendix G – Extra Resultant Metrics

Q v Irradiance

Low

Regression Type		R^2	***MSE***
Simple Linear		0.05	0.00030
Polynomial	Quadratic	0.32	0.00021
	Cubic	0.41	0.00018
	Bi-Quadratic	0.44	0.00017
	Degree 5	0.46	0.00017
	Degree 6	0.38	0.00019

Mid

Regression Type		R^2	***MSE***
Simple Linear		0.01	0.00141
Polynomial	Quadratic	0.23	0.00110
	Cubic	0.38	0.00088
	Bi-Quadratic	0.44	0.00080
	Degree 5	0.46	0.00077
	Degree 6	0.47	0.00076

High

Regression Type		R^2	***MSE***
Simple Linear		0.05	0.00485
Polynomial	Quadratic	0.07	0.00477
	Cubic	0.15	0.00436
	Bi-Quadratic	0.21	0.00403
	Degree 5	0.28	0.00371
	Degree 6	0.19	0.00414

Q v t

Low

Regression Type		R^2	***MSE***
Simple Linear		0.00	0.00119
Polynomial	Quadratic	0.77	0.00027
	Cubic	0.77	0.00027
	Bi-Quadratic	0.90	0.00011
	Degree 5	0.90	0.00011
	Degree 6	0.91	0.00011

Mid

Regression Type		R^2	***MSE***
Simple Linear		0.24	0.00332
Polynomial	Quadratic	0.55	0.00194
	Cubic	0.55	0.00194
	Bi-Quadratic	0.68	0.00139
	Degree 5	0.79	0.00091
	Degree 6	0.89	0.00047

All S < 60%

Regression Type		R^2	***MSE***
Simple Linear		0.02	0.00637
Polynomial	Quadratic	0.59	0.00264
	Cubic	0.60	0.00258
	Bi-Quadratic	0.65	0.00227
	Degree 5	0.67	0.00216
	Degree 6	0.68	0.00211

High

Regression Type		R^2	***MSE***
Simple Linear		0.01	0.00528
Polynomial	Quadratic	0.62	0.00203
	Cubic	0.62	0.00203
	Bi-Quadratic	0.63	0.00198
	Degree 5	0.65	0.00186
	Degree 6	0.70	0.00160