Grid impact analysis and interval forecasting of solar PV generation

Verma, Jatin

2019

Verma, J. (2019). Grid impact analysis and interval forecasting of solar PV generation.
Master's thesis, Nanyang Technological University, Singapore

<https://hdl.handle.net/10356/84946>

<https://doi.org/10.32657/10220/49167>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**GRID IMPACT ANALYSIS AND INTERVAL FORECASTING
OF
SOLAR PV GENERATION**

**VERMA JATIN
SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING
2019**

Grid Impact Analysis and Interval Forecasting Of Solar PV generation

Verma Jatin

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University

In partial fulfillment of the requirement for the degree of

Master of Engineering

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

18/06/2019

.....
Date

.....
Verma Jatin

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

18/6/2019

.....

Date

.....

Asst. Prof. Xu Yan

Authorship Attribution Statement

This thesis contains material from 2 papers published in the following conferences where I was the first and/or corresponding author.

Chapter 2 is published as J. Verma, C. Dong, D. Shi, K. L. Hai, H. C. Ong, X. Yang, and Y. Xu, "Impact Study for PV and ESS integration in Jurong Port Distribution Grid." *2018 Asian Conference on Energy, Power and Transportation Electrification (ACEPT)*. IEEE, 2018.

The contributions of the co-authors are as follows:

- Dr. Koh served as the main project supervisor. Along with Mr. Ong, he served as the project manager who represented Energy Research Institute @ NTU while dealing with Jurong Port Pte. Ltd.
- Mr. Dong initially helped me with the eTAP software setup and rectifying various technical issues throughout the course of the project.
- I went through the official documentation of the Jurong Port distribution grid. With Mr. Ong's and Xia's help, I determined various given and necessary distribution grid parameters for the impact study.
- Mr. Shi offered various expert guidelines in regular meetings which helped in deciding the next steps for the project.
- Asst. Prof. Xu provided the scenario structure for the impact study of the PV and ESS integration and helped in writing the research paper.

Chapter 3 is published as J. Verma, Y. Xu, J. Zhao, & Z. Xu (2018, Nov.), "Short Term PV Power Forecasting using ELM and Probabilistic Prediction Interval Formation", *Paper presented at the 9th International Conference on Extreme Learning Machines (ELM2018), Singapore, 2018*.

The contributions of the co-authors are as follows:

- Asst. Prof. Xu provided the fundamental guidance with the Extreme Learning Machine (ELM) algorithm and other machine learning concepts.
- I searched the free dataset source for training the proposed approach. I presented my initial idea towards Prediction Interval formation to Prof. Xu who suggested necessary refining to use it with ELM.
- Under Prof. Xu's guidance and timely suggestions for improvement, I did all the experimentation with multiple approaches, optimization of model hyper-parameters and input feature selection.
- With the guidance and reviews from Mr. Junhua Zhao and Mr. Zhao Xu, I could complete the paper writing.

18/06/2019

.....
Date

.....
Verma Jatin

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to everyone whose direct or indirect support enabled me to complete this report.

First and foremost, I am utterly grateful to Professor Xu who saw the potential in me and hired me to pursue Masters in the prestigious NTU Singapore. The amount of effort he puts in supervising and consulting students is unbelievable and ever-inspiring. He never misses a chance to give the right advice, be it research or other career matters. I am grateful to him on providing me the funding required for my projects. Secondly, I express my thankfulness to all the colleagues and supervisors, namely, Dr. Marcus Koh, Mr. Ong Hueh Chuah, Dong Chaoyu, Shi Donghan, Xia Yang and all the members of my research group, who were a part of my projects. It was a great learning experience working with everyone. I am grateful to the National Research Foundation (NRF) Singapore, project #NRF-ENIC-SERTD-SMES-NTUJTCI3C-2016, SMES, for providing financial support to conduct the study 'Impact Study for PV and ESS integration in Jurong Port Distribution Grid'.

I would like to thank the School of Electrical and Electronic Engineering to provide me a platform to show my research potential as a Masters student. Particularly, the resources provided by the Electric Power Research Lab, managed by the Lab-in-charge Jennifer Tan, offered me a conducive environment to perform my studies and work on my algorithms. As a whole, I am thankful to Nanyang Technological University for offering a multi-disciplinary, diverse and active environment, not to mention all the housing and other prominent facilities, which made the entire research process easier to bear.

Last but not least, I would not miss the chance to genuinely show my gratitude to my parents who are always there for me, my brothers and my dear friends who supported me while I went through various twists and turns during my tenure as a Masters student at NTU.

Contents

Statement of Originality	iii
Supervisor Declaration Statement	v
Authorship Attribution Statement	vii
Acknowledgements	ix
Abstract	xiii
1 Introduction	1
1.1 Background	1
1.2 The Climate Action Plan - Singapore	1
1.2.1 Solar Potential	2
1.2.2 Challenges	2
1.3 Grid Impact Study of PV integration	3
1.4 Research on PV forecasting	3
1.4.1 Indirect Solar Forecasting	4
1.4.2 Interval Forecasting	4
2 Grid Impact Analysis on PV Integration	7
2.1 Grid Overview	7
2.2 Jurong Port Grid Model in eTAP	7
2.3 PV-ESS integration and Loading scenarios	8
2.4 Stability Criteria	8
2.5 eTAP Simulation Results and Analysis	9
2.5.1 Medium Load	9
2.5.2 Light Load	11
2.5.3 Heavy Load	14
3 Short Term PV Power Forecasting using ELM and Probabilistic Prediction	17
Interval Formation	17
3.1 ELM	17
3.2 Hyperparameter Tuning	18
3.2.1 Cross-Validation and Grid search	18
3.2.2 Differential Evolution	19
3.3 PI formation	20
3.4 PI evaluation metric	22
3.5 Data Selection	23
3.5.1 Monthly Solar Irradiance Analysis	23
3.5.2 Autocorrelation	24
3.5.3 Permutation Importance	25
3.6 Results and Discussion	26

4 Conclusion and Future Directions	31
4.1 Grid Impact Analysis for PV integration	31
4.2 PV Power Interval Forecasting	31
4.3 Future directions	32
Bibliography	33

NANYANG TECHNOLOGICAL UNIVERSITY

Abstract

Dr. Xu Yan

School of Electrical and Electronic Engineering

Master of Engineering

Grid Impact Analysis and Interval Forecasting of Solar PV generation

by Jatin VERMA

Solar PV is renowned all around the world, but its prevalence varies with a country's solar profile. Singapore is an island country which has solar energy as the only promising renewable energy option. On the UN's request to submit the Intended Nationally Determined Contribution (INDC) against carbon emissions, Singapore, as a promoter of efficient energy, has come up with an elaborate Climate Action Plan towards carbon emission reduction. An aspiring leap by Singapore towards 350MWp PV capacity by 2020 is on trial, which would involve numerous PV installation projects.

The thesis at one end demonstrates the quantitative impact of PV installation on long-length transmissions, while at the other end, it showcases the use of intelligent algorithms for forecasting future PV power output. Chapter 2 deals with a 3.3MWp PV installation in the low voltage side of the Jurong Port distribution grid, providing comprehensive steady-state analyses in different loading scenarios. Comparison of different configurations of PV and ESS is done and it is found that the PV installation improves the bus voltage profile, reduces line congestion and circumvents substantial transmission losses.

Accurate predictions of solar power are important to the grid operator for ensuring energy management from multiple sources without jeopardizing stability and to the PV plant owner for scheduling plant maintenance periods and avoiding penalties imposed by the grid operators due to power imbalance costs. It is evident that meteorological data like solar irradiance is more readily available than historical PV power output series with hourly samples. In this case, indirect forecasting can be utilized where solar irradiation is predicted first, which is followed by obtaining the PV power output using a PV performance model of the plant. Since PV series during the day experiences lots of variations, point forecasting can be quite uncertain. The inherent uncertainty in the point forecasts can be quantified by associating them with a probability distribution to form prediction intervals (PIs) which is a more interpretable representation of uncertainty. Chapter 3 presents a probabilistic forecasting approach using a nonparametric PI formation method based on Extreme Learning Machine. No prior assumption on the error distribution is required for the PI formation. Solar irradiance data from NUS geography weather station, Singapore, is analyzed and assembled into two separate sets for better model performance. Cross-validation and Grid search followed by Differential Evolution are utilized to tune the hyper-parameters of the proposed model. Coverage probability and interval scores are evaluated for the resulting PIs which show promising results.

List of Figures

2.1	Medium and Low Voltage JP distribution grid structure	8
2.2	PV and ESS addition on PDL – 6.6kV/400V Substation P1. ESS can be in charging or discharging mode	9
2.3	Bus Voltage profiles in the medium loading condition	10
2.4	Cable load factors for medium loading condition	11
2.5	Bus Voltage profiles for light loading condition	12
2.6	Cable load factors for light loading condition	13
2.7	Structure of Substation P1 showing bus locations	14
2.8	Bus Voltage profiles for heavy loading condition	15
2.9	Cable load factors for heavy loading condition	16
3.1	ELM architecture	17
3.2	An illustration of the error signal, e_{train}	20
3.3	PI formation from the training error, e_{train}	21
3.4	Average of 2015-2017 monthly solar irradiance in Singapore	23
3.5	The Autocorrelation values of solar irradiance series at time lags vary- ing from 0 to 10000	25
3.6	Permutation Importance evaluation for the input features of MayOct dataset	26
3.7	Permutation Importance evaluation for the input features of NovApr dataset	26
3.8	Prediction intervals for <i>NovApr</i> set with CI=0.90	28
3.9	Prediction intervals for <i>NovApr</i> set with CI=0.95	28
3.10	Prediction intervals for <i>MayOct</i> set with CI=0.90	29
3.11	Prediction intervals for <i>MayOct</i> set with CI=0.95	29

List of Tables

2.1	Buses in the JP Distribution Grid Model	7
2.2	Branches in the JP Distribution Grid Model	8
2.3	Instability Levels	9
2.4	Average Bus Voltages in Medium Loading Condition	10
2.5	Comparing Cable Load Factors of Base Case and PV Addition Case - Medium Loading	11
2.6	Average Bus Voltages in Light Loading Condition	12
2.7	Comparing Cable Load Factors of Base Case and PV Addition Cases - Light Loading	13
2.8	Load Flow Results after PV addition for Buses related to the cables in Table 2.7	14
2.9	Average Bus Voltages in Heavy Loading Condition	15
2.10	Comparing Cable Load Factors of Base Case and PV Addition Case - Heavy Loading	16
3.1	An illustration of Grid Search	19
3.2	Training and Testing Data	24
3.3	Input features determined by Autocorrelation	25
3.4	Final set of input features after Permutation Importance evaluation	27
3.5	Best set of hyper-parameters and their cross-validation scores for <i>No- vApr</i> and <i>MayOct</i> sets	27
3.6	Final set of hyper-parameters obtained from DE	27
3.7	PICP and interval scores for <i>NovApr</i> and <i>MayOct</i> sets	28

Dedicated to my parents...

Chapter 1

Introduction

1.1 Background

With every year, the urgency to reduce carbon emissions is globally increasing. The issue regarding the Greenhouse Gas (GHG) emissions, which includes CO₂, oxides of sulphur and nitrogen, is seeking the world's attention and calls for their action. The Kyoto Protocol (1997) and the Paris Convention (2016) are some chief measures against GHG emissions organized by the United Nations Framework Convention on Climate Change (UNFCCC). In March 2015, Singapore submitted its Intended Nationally Determined Contribution (INDC) to the UNFCCC as an extended plan towards the nationwide mitigation of carbon emission [1]. Singapore, as a promoter of efficient energy, has taken up various measures to improve electricity generation efficiency and promote smart and efficient energy usage in its infrastructure. According to the National Climate Change Secretariat (NCCS), Singapore, the carbon emissions will peak at 65 million tonnes by 2030, after which it is supposed to stabilize [2]. In other words, the increase in the amount of yearly emissions is consistently decreasing, and after 2030, the yearly emissions themselves will start decreasing. This would require a coordinated and comprehensive effort from the government development boards, private agencies and the community.

Singapore, as advanced as it is, already began taking measures against carbon emission in the early 2000s by switching to electricity production using natural gas as fuel [3]. Natural gas has considerably less carbon emissions than fuel oil. Consequently, Singapore has achieved 95.3 % of its total electricity generation from natural gas [4]. With a land area of 719km² and low wind profile, wind turbine plantation becomes infeasible. Additionally, the pre-feasibility studies conducted by the government discourage the possibility of nuclear power generation in the city-state due to high risks [5]. Due to this, the scope of further reduction in carbon emission is limited. However, NCCS, with the support from Prime Minister's Office, has come up with an elaborate Climate Action Plan towards carbon emission reduction [6].

1.2 The Climate Action Plan - Singapore

The strategy for a carbon-efficient Singapore is based on four broad measures – improving energy efficiency, reducing carbon emissions from power generation, developing low carbon technologies and encouraging a coordinated and collective action as a nation. Increasing the energy efficiency of the industrial and housing sector is of immense importance to Singapore's action plan as it would create a considerable difference in the required energy fuel quantity. The country is investing in energy-efficient buildings as the number of green building projects has increased from 17

(2005) to 2,700 (2016) under the Green Mark Scheme, a rating system by the country's Building and Construction Authority [7].

To make a remarkable difference in reducing the emissions, adoption of alternative sources of energy is an indispensable step that Singapore must take. Among all other options of renewable sources of energy, solar PV is the most promising one particularly for Singapore. Although the land constraint would not allow, the growth from 2MWp (2009) to 60MWp (2015) has been achieved by leveraging the rooftop areas and limited open land areas. A promising leap towards 350MWp PV capacity by 2020 is on trial, which would be 5% of the total peak demand [6]. The descent in the cost of PV-produced electricity over the last 5 years is a driving factor in the increasing popularity of solar PV [8]. Singapore is determined towards investing in research and development for PV forecasting and ESS to overcome the intermittency of solar PV. A PV system installation at Choa Chu Kang Waterworks, and other upcoming installations at Changi Water Reclamation Plant, and Bedok Waterworks & WaterHub represent some major achievements by the PUB Singapore [9]. Another innovative effort of introducing the concept of Floating PV that allows the solar PV to be installed in waterbodies, like reservoirs, is an attempt to overcome the land constraints and increase the solar potential of the country [10]. Testbeds and experimental analysis are being conducted by PUB which is a critical part of the installation process. Many corporate giants like Apple are coming forward to promote solar PV and conduct comprehensive techno-economic analyses [11]. Comprehensive simulation studies with analytically derived results and observations favorable for PV and ESS implementation are essential before the actual implantation.

1.2.1 Solar Potential

In the past few decades, there has been a significant increase in the prevalence of solar power all around the world. However, the prevalence of solar power might vary with a country's solar profile. The Middle East is blessed with solar potential thereby hosting around 1.3GW of installed solar power (2018) while expecting a growth by 4.8GW more [12]. Meanwhile, countries like Singapore experience limited overall renewable energy potential [6]. The growth in attention towards solar power is attributed to some global factors, like, the worldwide attempts towards the reduction of carbon emissions, advanced photovoltaic technology with reduced cost of production and increased efficiency, and the slow but fateful depletion of fossil fuels [8].

1.2.2 Challenges

Before the installation of solar PV in any region, proper economic and technical feasibility studies should be performed, such that a new addition of PV to the grid does not jeopardize the grid's stability and is worth the installation and operational costs. The newly added PV plant can have the capacity large enough to affect the power flow pattern in the grid, which might call for violations in the power flow constraints. Such undesirable possibilities should be foreseen by proper grid impact studies using various power flow simulation environments. Larger the PV capacity, larger is the magnitude of variation in the PV output power. Dealing with uncertainties is a major concern when it comes to solar power, given its dependence on the meteorological variables, primarily the solar irradiance pattern. Any deviation from the expected PV output power would affect the entire energy management planning done by the grid operators. It could result in the shortage or excess of power,

both of which are system stability concerns. Thus, reliable estimation of the future PV power using intelligent PV power forecasting approaches is necessary for the smooth operation of the grid.

1.3 Grid Impact Study of PV integration

The assurance regarding solar PV usage comes from different research studies and simulations conducted around the world. Patil et al. [13] studied the effect of PV penetration at different levels on bus voltages as well as on the life of tap-change transformers. Denser is the PV penetration in the distribution grid, more is the number of tap changes required for voltage regulation, which in turn, reduces the life time of the transformer. However, the research was performed in a small grid with 2.5 MW load and generation. It is important to find out a higher level impact on the distribution lines of a realistic power grid. Lazzeroni et al. [14] recommend the proposition of certain regulations on feed-in PV power to avoid contingent critical grid operations, including line congestion, after conducting load-flow studies in a realistic grid of Hebron, Palestine. However, the case of Middle-East countries is a bit different from Singapore in a way that the former has enormous solar potential. Ruiz-Romero et al. [15] suggest that depending on the ratio of the conductor resistance and reactance, (R/X) , an active power injection would proportionally affect the bus voltage. In medium voltage (MV) lines, (R/X) is greater than 1. Thus, the bus voltage increases with the increase in active power injection. The active power injection from a Distributed Generator, like PV, would affect the most when connected at the end of the cable connecting the MV and low voltage (LV) substations because the reduction in the cable loading would affect many more MV line sections and cause increase in bus voltages at the low voltage end. Keeping this in mind, the PV is positioned accordingly for grid impact analysis in Chapter 2. In this thesis, an existing Singapore MV/LV distribution grid is simulated to perform a load-flow analysis to evaluate the impact of PV and ESS addition at a node in the grid where it would help reduce line congestion and improve bus voltage profiles. The results of the simulation can be used for further analysis to realize PV integration in the Singapore grid.

1.4 Research on PV forecasting

The large-scale penetration of PVs brings in operational and managerial challenges in power systems due to their inherent variability and weather dependence. Grid operators need to rely on PV power forecasts to plan for real-time market clearing mechanisms, reserve allocations, optimum power flow and maintaining system stability [16]. Inaccurate forecasts provided by the plant owners might cause them considerable power imbalance costs imposed by the grid operator if the imbalance is more than certain tolerance level. The impact of inaccurate forecasting also increases with the PV penetration [17, 18]. These penalties persuade the plant owners to look for reliable forecasting techniques with higher accuracy to maximize their profit.

Marquez et al. [19] worked on the US National Weather Service's dataset and employed an ANN model to predict the medium-scale solar irradiance using relevant meteorological variables, like sky cover, precipitation and temperature. Da Silva Fonseca Jr et al. [20] demonstrated the advantage of using the cloudiness parameter in the power prediction of a 1MW PV plant using support vector machines. Without the cloudiness input, the mean absolute error goes up by the difference of

42%. The cloudiness information was received from the weather forecast system of the Japan Meteorology Agency. In general, the cloudiness information requires sky imaging and would provide great improvements in PV power forecasting models only in the case if it is available. Zhang et al. [21] designed a series ensemble of models of a novel learning technique, known as Extreme Learning Machine (ELM) for predicting load consumption using the historical data from Australian National Electricity Market. Since it was designed in 2006 [22], it has found a plethora of applications in data-driven applications because of its thousand-folds faster learning speed and generalization ability over the training data. In this way, it overcomes overfitting, high computational costs and local minima convergence which are the chief drawbacks of the conventional back-propagation based neural network algorithms. For this reason, the research presented in the Chapter 3 focuses on utilizing ELM for PV power forecasting.

1.4.1 Indirect Solar Forecasting

PV power can either be forecasted directly from the given time series PV power data (known as direct forecasting) or calculated using appropriate models whose inputs are the forecasts of variables which are the factors of PV power production, for example, solar irradiance [20]. It is evident that meteorological data like solar irradiance is more readily available than the historical PV power output series with hourly samples. Therefore, in this report, the historical solar irradiance data from NUS Geography Weather Station, Singapore, is taken as input which is described in the later sections [23].

1.4.2 Interval Forecasting

In practice, the operational decisions in power systems solely based on point forecasts can be risky considering the sharp variations in solar irradiance patterns. In the case of solar energy, point forecasts cannot be relied upon because of the degree of uncertainty associated with the irradiance pattern. A means to quantify this degree of uncertainty becomes necessary to avoid power imbalance risks. The solar irradiance profile of Singapore is different from other countries which attributes to its tropical rainforest climate where there is no true distinction between seasons. One of the research works involving solar irradiance interval forecasting using Singapore irradiance data is from Dong et al. [24]. The data is from a meteorological station of Solar Energy Research Institute of Singapore (SERIS). They make use of a Fourier trend model to extract residual signals from the solar irradiance data. These residual signals are tested for weak stationarity after which they are passed into a definite set of exponential smoothing state space (ESSS) models. Through a model selection process, one model is finalized. The model offers 0.90 target coverage for 0.95 confidence level. Evaluation in terms of sharpness of the PIs is not performed. The model is claimed to perform better than ARIMA, Linear Exponential Smoothing, Simple Exponential Smoothing and Random Walk. However, there are some issues with nonlinear exponential smoothing models [25]. The models, especially purely multiplicative models, which make an assumption on the error process to be Gaussian, generate prediction distribution having infinite variance beyond a certain forecasting horizon. In some cases, the PIs become increasingly skewed for longer-term forecasts, while in the cases where there is a Gaussian distribution assumption, the resulting PIs have a progressively larger negative component.

Another way of representing uncertainty is to associate a probability distribution function (PDF) with the given forecasts to form prediction intervals (PIs), which is termed as probabilistic forecasting [16]. This probability distribution can either be parametric (for example, assuming the PDF shape as Gaussian) or nonparametric. Fixating to a specific type of distribution could result in unrealistic PI formation. Thus, in nonparametric probabilistic forecasting, PDFs are evaluated from the finite sample points obtained as the inputs to the PI formulation approach.

Various recent researches have tested different probabilistic forecasting approaches for PI formation. Chai et al. [26] use the nonparametric Kernel Density Estimation method for PV power forecasting based on a PV power data source from Belgium. They also make use of a tool named *copula* to model the interdependence of the point forecasts and the actual PV measurements. For a successful machine learning approach, input feature selection is as much important as optimizing model parameters. A relevant set of input features would increase the probability of achieving accurate forecasts. The most popular tool for input feature selection involving time-series data is Autocorrelation [27]. Further, a recent technique named Permutation Importance has been discussed on Kaggle to find out individual feature importance [28]. This technique is further elaborated and applied later in Chapter 3.

A renowned PI formation approach called *adapted resampling* has been elaborated by Pinson et al. [29]. Chapter 3 presents a similar approach to use the probability distribution of training errors to derive PIs with associated confidence levels without any prior assumption on error distribution. Another important aspect of PI based forecasting is to produce sharp PIs while maintaining accurate target coverage. Wan et al. [30] introduce a hybrid intelligent optimization approach which considers the coverage probability and the sharpness of the PIs to obtain optimal PIs for Wind Power forecasting from an ELM based model. For quantifying the sharpness of the PI, a metric called interval score is utilised later in the section 3.4. In future, the research would be taken forward to focus on producing more precise and accurate PIs.

Chapter 2 deals with the installation of an ESS-backed PV capacity in the low voltage distribution side in Jurong Port (JP) distribution grid, Singapore. The JP grid structure, grid modeling in eTAP [31] and PV-ESS integration are described. The load flow results and analysis are presented in different loading scenarios concluding overall benefits of the PV-ESS integration. Chapter 3 describes the PV interval forecasting using the proposed approach which is composed of the ELM model and the nonparametric PI formation. It is also shown how ELM model is tuned for the hyper-parameters with better training accuracy. Finally, the PI evaluation metric is explained. Chapter 3 also elaborates the solar irradiance historical data used for the proposed approach, is followed by the implementation of the approach and ends with the discussion of the results. Chapter 4 provides some ways for improving the current research and concludes the thesis.

Chapter 2

Grid Impact Analysis on PV Integration

A realistic grid scenario is modeled based on the JP distribution grid, Singapore, for understanding the impact of the PV addition on the distribution grid.

2.1 Grid Overview

The JP distribution grid is spread into two locations – Mainland (ML) and Pulau Damar Laut (PDL) island. Fig. 2.1 represents the grid structure. The Singapore grid supplies the ML – 22kV Main Substation which has its own low voltage LT 22kV/400V substation. The ML – 22kV Main Substation provides supply to ML – 22kV/400V Substation 1, 2 and 4. The PDL substations have a 1800m line connection with the ML – 22kV Main Substation through PDL – 22kV/6.6kV Substation P1 which is also connected to PDL – 22kV/400V Substation P5. The PDL grid has a connection with Singapore grid through PDL – 22kV/400V Substation P5. Further, the PDL – 22kV/6.6kV Substation P1 steps down to 6.6 kV and supplies PDL – 6.6kV/400V Substations P1, P2 and P3. This study also incorporates the Warehouse B13 which is under construction and aims to establish a hybrid AC-DC microgrid within itself. This warehouse is supplied by the PDL – 6.6kV/400V Substation P1.

2.2 Jurong Port Grid Model in eTAP

The JP distribution grid modeling and simulation is done in eTAP 16.0.0 [31] and the load flow method is Adaptive Newton-Raphson. The system frequency is 50Hz. The two Singapore grid connections are represented by 'Power Grid' components in eTAP acting as 22kV swing buses. Table 2.1 and 2.2 show the specifications of the grid.

Total Static Load connected at various buses sums up to 21.11 MVA. The information required to build the single line diagrams is provided by Jurong Port Pte. Ltd, Singapore.

TABLE 2.1: Buses in the JP Distribution Grid Model

No. of Buses	Swing	V-Control	Load	Total
	2	0	99	101

TABLE 2.2: Branches in the JP Distribution Grid Model

No. of Branches	XFMR	Line/Cable	Tie PD	Total
	21	75	5	101

FIGURE 2.1: Medium and Low Voltage JP distribution grid structure

2.3 PV-ESS integration and Loading scenarios

1. PV integration on PDL – 6.6kV/400V Substation P1: A proposal of implanting a 3.3 MWp PV plant is considered for load flow study. Integrating a distribution PV plant could help fulfill the load demand on the distribution side and strengthen the bus stability, and in turn, the grid stability. For realizing this proposal, a PV plant with the following specifications is connected to 415 V Bus10 and Bus11 (see Fig. 2.2):

SUNTECH Monocryst. 72-cell (180W) panel.
 $3.3\text{MWp PV plant} = 132 \times 140 \text{ panels (5.082 kV)}$
Interlinking Inverter = 10 MVA, 5kV/415V, 90% η

2. Energy Storage System (ESS) connection: For backing up the mercurial PV plant, a 2.5 MW battery backup is considered, and impact study is done. The ESS is connected on the same bus as the PV plant. It can basically operate in two modes: charging and discharging. For simplifying study, in charging mode, the ESS is treated as a 2.5 MW load, and in discharging mode, it is treated as a DC supply providing 2.5 MW power.

The simulation is done for three loading scenarios: medium load, light load and heavy load. In each scenario, four cases are dealt with: base case, PV addition, PV with ESS in charging mode (ESS-C), and PV with ESS in discharging mode (ESS-D). Base case denotes the scenario described in Section 2.1 and 2.2.

2.4 Stability Criteria

The steady-state AC load flow results would help in understanding the steady state stability of the buses and cables given their ratings and ampacity. In eTAP, for a stable system, the bus voltages should be close to 1 p.u. (100% voltage level), and the power flow through the cables should be considerably lower than the cable ampacity [32]. A bus undergoes overvoltage condition if the voltage level is beyond 5% from

FIGURE 2.2: PV and ESS addition on PDL – 6.6kV/400V Substation P1. ESS can be in charging or discharging mode

the rated voltage. If the deviation is within 5%, then it is in the safest range (Table 2.3). Cable loading factor is given by:

$$\text{Cable L.F.(\%)} = \frac{\text{Current carried by the cable}}{\text{Cable Ampacity}} \times 100 \quad (2.1)$$

TABLE 2.3: Instability Levels

Electrical Component	Unstable voltage/Overload Limits
Bus Voltage	Beyond 95% - 105% range
Cable	100% loading

2.5 eTAP Simulation Results and Analysis

This section describes the results after simulating all the cases. The results include bus voltage report and the percentage load factors for all the cables which would tell whether the cable ampacity is enough for the respective power flow. On comparing the results for different cases, one could understand the impact of integration of a PV plant and Energy Storage System (ESS).

2.5.1 Medium Load

In the medium loading condition, all the loads in the network draw rated power. Fig. 2.3 shows the bus voltages for the four scenarios - *base case*, *PV addition*, *PV with ESS in charging mode* (ESS-C), and *PV with ESS in discharging mode* (ESS-D). Base case denotes the scenario without PV and ESS integration as described in the sections 2.1 and 2.2.

From the Fig 2.3, the bus voltages in the buses from 7 to 14, the buses from 17 to 24, and the buses from 83 to 86 go as low as 0.955 p.u. The buses closer to the Singapore grid connection have voltages close to 1 p.u. whereas those buses located

FIGURE 2.3: Bus Voltage profiles in the medium loading condition

closer to the loads have lower voltages. By the addition of PV at PDL – 6.6kV/400V Substation P1, the improvements in bus voltage profile in the buses from 7 to 14, and buses from 80 to 86 show that these buses are the most affected buses when the PV is integrated. On addition of an ESS in charging mode, some bus voltages drop down. This is because most of the PV power is used for charging the ESS. In the ESS-D scenario, the bus voltage values are approximately same as in the ‘PV addition’ scenario. Table 2.4 shows the average operating voltage for the four cases. In the Base Case, the average bus voltage is 0.97124 p.u. After PV addition, due to the increase in bus voltages in some buses, the average bus voltage reaches 0.985241 p.u., a 1.44% increase. In the ESS-C scenario, the average voltage falls down to 0.9832 p.u., which is still better than the Base Case. In the ESS-D scenario, the average bus voltage rises up to 0.985242 p.u. Overall, the addition of PV and ESS shows improvement in the buses’ steady state stability.

TABLE 2.4: Average Bus Voltages in Medium Loading Condition

Case	Average of Bus Voltage (p.u.)
Base Case	0.97124
PV	0.985241
ESS-C	0.98320
ESS-D	0.985242

In the medium loading scenario, the average cable load factor for the base case is 21.05%. From Fig. 2.4, the addition of PV at the low voltage end has some impact on some cable load factors. It can be seen that the impact is mainly visible in the cables numbered Cable5, Cable6, Cable13, Cable14, Cable15, Cable16, Cable48 and Cable49. The loading of these cables which are also the primary connections to the Singapore grid is relaxed a bit. The reason behind this is that a significant portion of load is now supplied by PV at the low voltage distribution grid. The net power requirement from the Singapore grid is reduced. Substantial transmission loss from the Singapore grid is avoided. In the ESS-C scenario, the cable loading is again increased in the aforementioned cables as most of the PV power is used for charging the ESS. In the ESS-D scenario, the cable loading is similar to that of the PV addition case.

Table 2.5 finds out the amount of decrease in load factors after PV addition in the Base Case. The cables which have experienced major decrease in the load factor are those which connect the PDL – 6.6kV/400V Substation P1 to the Singapore grid

FIGURE 2.4: Cable load factors for medium loading condition

at PDL - 22kV/400V Substation P5. Some cables, like Cable5, experience substantial decrease (5.6%) in the cable load factor, whereas some cables, like Cable15, show an increase (-2.2%) in the loading. The equivalent decrease in the amount of Ampere is calculated for all cables. There has been considerable decrease in Cable loading and equivalent amount of transmission losses have been avoided because of PV addition.

TABLE 2.5: Comparing Cable Load Factors of Base Case and PV Addition Case - Medium Loading

Cable	Physical Location	Difference of Load Factor (%)	Net Amp diverted by PV
Cable5	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	5.6	24.230
Cable6	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	9.2	39.806
Cable13	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	17.1	80.154
Cable14	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	28.1	131.716
Cable15	within PDL – 6.6kV/400V S/S P1	-2.2	-7.337
Cable16	within PDL – 6.6kV/400V S/S P1	8.6	28.684
Cable48	within PDL – 6.6kV/400V S/S P1	-3.1	-119.046
Cable49	within PDL – 6.6kV/400V S/S P1	11.9	456.983

* S/S refers to Substation

2.5.2 Light Load

In the light loading condition, all the loads in the network draw 50% of the rated power. Fig. 2.5 shows the bus voltages for the four scenarios - base case, PV addition, PV with ESS in charging mode (ESS-C), and PV with ESS in discharging mode (ESS-D).

FIGURE 2.5: Bus Voltage profiles for light loading condition

Base Case denotes the scenario without PV and ESS integration as described in the sections 2.1 and 2.2.

From the Fig 2.5, the bus voltages in the buses from 7 to 14, buses from 17 to 24, and buses from 83 to 86 go as low as 0.97 p.u. The buses closer to the Singapore grid connection have voltages close to 1 p.u. whereas those buses located closer to the loads have lower voltages. By the addition of PV at PDL – 6.6kV/400V Substation P1, the drops in bus voltage profile in the buses from 7 to 14, and buses from 80 to 86 improves, showing that these buses are the most affected buses when PV is integrated. On addition of an ESS in charging mode, some bus voltages drop down. This is because most of the PV power is used for charging the ESS. In the ESS-D scenario, the bus voltage values are approximately same as in the ‘PV addition’ scenario. Table 2.6 shows the average operating voltage for the four cases. In the Base Case, the average bus voltage is 0.97945 p.u. After PV addition, due to the increase in bus voltages in some buses, the average bus voltage reaches 0.99337 p.u., a 1.42% increase. In the ESS-C scenario, the average voltage falls down to 0.99145 p.u., which is still better than the Base Case. In the ESS-D scenario, the average bus voltage rises up to 0.99337 p.u. Overall, the addition of PV and ESS shows improvement in the buses’ steady state stability.

TABLE 2.6: Average Bus Voltages in Light Loading Condition

Case	Average of Bus Voltages (p.u.)
Base Case	0.97945
PV	0.99337
ESS-C	0.99145
ESS-D	0.99337

In the light loading scenario, the average cable load factor for the base case is 10.83%. From Fig. 2.6, the addition of PV at the low voltage end has some impact on some cable load factors. It can be seen that the impact is mainly visible in the cables numbered Cable5, Cable6, Cable13, Cable14, Cable15, Cable16, Cable48 and Cable49. The cable load factors of Cable5, Cable6, Cable13 and Cable14 have been relaxed a bit. However, the cable load factors of Cable15, Cable16, Cable48 and Cable49 have unexpectedly increased.

Table 2.7 finds out the amount of change in load factors of the cables, namely Cable5, Cable6, Cable13, Cable14, Cable15, Cable16, Cable48 and Cable49. These cables connect the PDL – 6.6kV/400V Substation P1 to the Singapore grid at PDL – 22kV/400V Substation P5. In the light loading scenario, the negative flows of

FIGURE 2.6: Cable load factors for light loading condition

currents show an increase in the cable loading which looks undesirable. For understanding the reason behind the increase in the cable loading despite the PV addition, we must take a look at the Load Flow report involving the respective buses so that we can get an idea about the direction of current as well. From Table 2.8, it is shown

TABLE 2.7: Comparing Cable Load Factors of Base Case and PV Addition Cases - Light Loading

Cable	Physical Location	Difference of Load Factor (%)	Net Amp diverted by PV
Cable5	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	3.8	16.4418
Cable6	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	5.4	23.3647
Cable13	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	10.0	46.8740
Cable14	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	14.1	66.0923
Cable15	within PDL – 6.6kV/400V S/S P1	-15.3	-51.0301
Cable16	within PDL – 6.6kV/400V S/S P1	-10.4	-34.6871
Cable48	within PDL – 6.6kV/400V S/S P1	-21.2	-814.1200
Cable49	within PDL – 6.6kV/400V S/S P1	-14.3	-549.1470

* S/S refers to Substation

that the power flow from Bus29 to Bus35 is negative which means that after PV addition in the light load condition the surplus power generated by the PV is sent back towards higher voltage substation loads. The physical location of the buses in the substation P1 is shown in Fig. 2.7. Similarly, for the remaining buses, the power flow is in the opposite direction as it was in the Base Case. Thus, an extra amount of power is supplied by the 3.3 MWp PV to the higher substation loads as compared to Base Case. Again, this is a benefit from the PV addition.

TABLE 2.8: Load Flow Results after PV addition for Buses related to the cables in Table 2.7

From	To	Connecting Cable	P (MW)	Q(MVAR)	Amp (A)	PF
Bus29 (6.6kV)	Bus35 (6.6kV)	Cable15	-0.844	0.169	75.5	-98.0
Bus30 (6.6kV)	Bus36 (6.6kV)	Cable16	-1.091	0.248	98.3	-97.5
Bus35	Bus78 (415V)	XFMR T3	-0.844	0.169	75.5	-98.0
Bus36	Bus79 (415V)	XFMR T3	-1.092	0.248	98.3	-97.5
Bus78	Bus10 (415V)	Cable48	-0.854	0.134	1201.1	-98.8
Bus79	Bus11 (415V)	Cable49	-1.103	0.207	1562.7	-98.8

FIGURE 2.7: Structure of Substation P1 showing bus locations

In the ESS-C scenario, the surplus PV power which was available to be sent back towards the higher voltage substation loads is used for charging the ESS. Therefore, the cable loading is decreased for Cable15, Cable16, Cable48 and Cable49. In the ESS-D scenario, the cable loading is similar to that of the PV addition case.

2.5.3 Heavy Load

In the heavy loading condition, all the loads in the network draw 150% of the rated power. Fig. 2.8 shows the bus voltages for the four scenarios - *base case*, *PV addition*, *PV with ESS in charging mode* (ESS-C), and *PV with ESS in discharging mode* (ESS-D). Base Case denotes the scenario without PV and ESS integration as described in the sections 2.1 and 2.2.

From the Fig 2.8, the bus voltages in the buses from 7 to 14, buses from 17 to 24, and buses from 83 to 86 go as low as 0.93 p.u. The buses closer to the Singapore grid connection have voltages close to 1 p.u. whereas those buses located closer to the loads have lower voltages. By the addition of PV at PDL – 6.6kV/400V Substation P1, the drops in bus voltage profile in the buses from 7 to 14, and buses from 80 to 86 improves, showing that these buses are the most affected buses when PV is integrated. On addition of an ESS in charging mode, some bus voltages drop down. This is because most of the PV power is used for charging the ESS. In the ESS-D scenario, the bus voltage values are approximately same as in the ‘PV addition’ scenario. Table 2.9 shows the average operating voltage for the four cases. In the Base

FIGURE 2.8: Bus Voltage profiles for heavy loading condition

Case, the average bus voltage is 0.96295 p.u. After PV addition, due to the increase in bus voltages in some buses, the average bus voltage reaches 0.977005 p.u., a 1.46% increase. In the ESS-C scenario, the average voltage falls down to 0.97485 p.u., which is still better than the Base Case. In the ESS-D scenario, the average bus voltage rises up to 0.977007 p.u. Overall, the addition of PV and ESS shows improvement in the buses' steady state stability.

TABLE 2.9: Average Bus Voltages in Heavy Loading Condition

Case	Average Bus Voltage (p.u.)
Base Case	0.96295
PV	0.977005
ESS-C	0.97485
ESS-D	0.977007

In the heavy loading scenario, the average cable load factor for the base case is 30.925%. From Fig. 2.9, the addition of PV at the low voltage end has some impact on some cable load factors. It can be seen that the impact is mainly visible in the cables numbered Cable5, Cable6, Cable13, Cable14, Cable15, Cable16, Cable48 and Cable49. The loading of these cables which are also the primary connections to the Singapore grid is relaxed a bit. The reason behind this is that a significant portion of load is now supplied by PV at the low voltage distribution grid. The net power requirement from the Singapore grid is reduced. Substantial transmission loss from the Singapore grid is avoided. In the ESS-C scenario, the cable loading is again increased in the aforementioned cables as most of the PV power is used for charging the ESS. In the ESS-D scenario, the cable loading is similar to that of the PV addition case.

Table 2.10 finds out the amount of change in load factors after PV addition in the Base Case. The cables which experienced change in load factor are those which connect the PDL – 6.6kV/400V Substation P1 to the Singapore grid at PDL – 22kV/400V Substation P5. A considerable amount of cable loading and equivalent amount of transmission losses have been avoided.

FIGURE 2.9: Cable load factors for heavy loading condition

TABLE 2.10: Comparing Cable Load Factors of Base Case and PV Addition Case - Heavy Loading

Cable	Physical Location	Difference of Load Factor (%)	Net Amp diverted by PV
Cable5	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	5.7	24.6628
Cable6	PDL – 22kV/400V S/S P5 to 22kV/6.6kV S/S P1	9.7	41.9700
Cable13	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	17.9	83.9045
Cable14	PDL – 22kV/6.6kV S/S P1 to 6.6kV/400V S/S P1	30.4	142.4970
Cable15	within PDL – 6.6kV/400V S/S P1	9.0	30.0177
Cable16	within PDL – 6.6kV/400V S/S P1	24.5	81.7148
Cable48	within PDL – 6.6kV/400V S/S P1	12.5	480.0238
Cable49	within PDL – 6.6kV/400V S/S P1	33.9	1301.8240

* S/S refers to Substation

Chapter 3

Short Term PV Power Forecasting using ELM and Probabilistic Prediction Interval Formation

Consider a dataset $D = \{x', y'\}$, out of which a certain amount of training data is picked. Following are the respective training inputs and output targets,

$$x_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\} \quad (3.1)$$

$$y_i = \{y_{i1}, y_{i2}, \dots, y_{iq}\} \quad (3.2)$$

where $i = 1, 2, \dots, N$, N is the number of training samples, q is the number of outputs which equals the number of output neurons, and p is the number of feature inputs which is same as the number of input neurons of the ELM architecture (as shown in Fig. 3.1). Before inputting it into the model, the data is first normalized using Z-scores normalization as follows:

$$D_{norm} = \frac{D - \text{mean}(D)}{\text{std}(D)} \quad (3.3)$$

where, $\text{std}(\cdot)$ is standard deviation function.

3.1 ELM

ELM [22] is a Single Layer Feed-forward Neural network (SLFN) which doesn't require iterative learning for training its hidden layer neurons. The input weights and biases for the hidden-layer neurons are randomly chosen, and the output weights

FIGURE 3.1: ELM architecture

are analytically determined. This increases the learning rate of the algorithm an unprecedented times than the Backpropagation in ANNs. Consider an ELM model with n hidden layer neurons. The outputs of the hidden layer neurons are given by:

$$\begin{aligned} h_{ij} &= \psi(w_j \cdot x_i + b_j) \\ i &\in 1, 2, \dots, N \\ j &\in 1, 2, \dots, n \end{aligned} \quad (3.4)$$

Here, $\psi(\cdot)$ denotes the activation function of the neurons, $w_j \in \mathbb{R}^p$ and $b_j \in \mathbb{R}$ are the randomly generated weight vectors and biases of the neurons, and $w_j \cdot x_i$ is the dot product of the two vectors. Using equation (3.4), the hidden layer output matrix can be formed,

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \dots & h_{Nn} \end{bmatrix} \quad (3.5)$$

Further, the target matrix, $\mathbf{y}$, is formed as follows:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (3.6)$$

After obtaining $\mathbf{H}$, and the target matrix, $\mathbf{y}$, the output weight matrix $\beta \in \mathbb{R}^{n \times q}$ of ELM is analytically determined as follows,

$$\beta = \mathbf{H}^\dagger \mathbf{y} \quad (3.7)$$

$\mathbf{H}^\dagger$ refers to the Moore-Penrose psuedo-inverse [33] of the $\mathbf{H}$ matrix. Training output, Q_{train} , is then obtained as follows:

$$Q_{train} = \mathbf{H}\beta \quad (3.8)$$

Mean Absolute Error (MAE) is calculated for the training output,

$$MAE_{train} = \text{mean}(\text{abs}(Q_{train} - \mathbf{y})) \quad (3.9)$$

3.2 Hyperparameter Tuning

Besides training the weights and biases, an ELM model requires the tuning of the number of neurons and the activation function parameters, which can be termed as the hyper-parameters.

3.2.1 Cross-Validation and Grid search

Grid search [34] is an approach of evaluating a range of possible hyperparameter values and choosing the best parameter set with the least prediction error. In popular practice, k -fold cross validation [35] is used, in which the training set is divided further into k 'folds' or subparts. One of the folds is taken out to be the validation set, and the remaining $k - 1$ folds are used as the training set. Based on a given scoring method, the validation scores are calculated. This is repeated k times until

all the folds are selected for validation. The final validation score is the average of the individual scores. The idea of applying Grid search and Cross-Validation is to repeat the whole process for different sets of hyper-parameters of the model. Suppose, the model, $ELM(u, v)$, has two hyper-parameters, each of which could have a range of possible values. Then, the model can be cross-validated for each possible combination of those values, and their validation scores can be stored in the form of a grid, as shown in the Table 3.1. In the end, the hyper-parameter set with the best score, which can be denoted by (u^*, v^*) , is selected for the calculation of training error and PI formation.

TABLE 3.1: An illustration of Grid Search

$\begin{matrix} v \\ u \end{matrix}$	v_1	v_2	v_3	v_4
u_1	$\text{score}(u_1, v_1)$	$\text{score}(u_1, v_2)$	$\text{score}(u_1, v_3)$	$\text{score}(u_1, v_4)$
u_2	$\text{score}(u_2, v_1)$	$\text{score}(u_2, v_2)$	$\text{score}(u_2, v_3)$	$\text{score}(u_2, v_4)$
u_3	$\text{score}(u_3, v_1)$	$\text{score}(u_3, v_2)$	$\text{score}(u_3, v_3)$	$\text{score}(u_3, v_4)$

3.2.2 Differential Evolution

Differential Evolution (DE) has been popularly used for tuning the hyperparameters of the prediction models [36]. It maintains a population of candidate solutions which undergo fitness evaluation, selection and crossover operations to create newer candidates at a given crossover rate [37]. It works for optimizing nonlinear, non-differentiable continuous functions. The optimization algorithm finds out the best hyperparameter set within the given boundaries for which the objective function is minimum. The objective function can be an error measurement criteria for the forecasting model. In our approach, the error measurement criteria is the mean absolute error (MAE).

For ELM, the hyper-parameter set (n^*, σ_{width}^*) is obtained as the optimal hyper-parameter set with the best cross-validation score using Grid search. DE is employed to further explore the search space around (n^*, σ_{width}^*) . Hence, the new search space for DE would be $n \in [n^* - k_n, n^* + k_n]$, and $\sigma \in [\sigma_{width}^* - k_\sigma, \sigma_{width}^* + k_\sigma]$, where $k_n, k_\sigma \in \mathbb{R}$. k_n and k_σ can be any user-defined numbers set to define a search space for DE. The advantage of using DE following the Grid search is that it reduces the search space to a great extent. It is computationally expensive to look for the optimal hyper-parameter set over a large continuous search space with DE alone. Therefore, first Grid search is used to evaluate a definite search space as illustrated in Table 3.1. After obtaining the Grid search optimized hyper-parameters, DE, a metaheuristic algorithm, can further explore the hyper-parameter space around them.

DE usually works for continuous variables, whereas, n is an integer in this problem. To enable the smooth operation of DE, n is treated as a continuous variable during the DE procedure. In each iteration of DE, during fitness evaluation, n is rounded off to the closest integer for inputting it into ELM as a hyper-parameter.

FIGURE 3.2: An illustration of the error signal, e_{train}

3.3 PI formation

The prediction error from the training set is used for the PI formation. The training error is given by the equation (3.10). Fig. 3.2 illustrates an instance of the error signal.

$$e_{train} = \|vect(Q_{train} - \mathbf{y})\| \quad (3.10)$$

Here, $vect(\cdot)$ is used to vectorize the matrix into a row/column vector. The range of training errors, $[0, rng_e]$, can be calculated as:

$$rng_e = \text{ceil}(\max(e_{train})) \quad (3.11)$$

where, $\text{ceil}(\cdot)$ is the ceiling function. The idea is to divide the series $[0, rng_e]$ into n_{int} intervals (see Fig. 3.3a) and form the probability distribution of the errors, which could then be used to set confidence intervals for the PIs. To obtain the PI width, ϵ_0 for a confidence interval, CI , following steps should be followed:

1. Form the interval profile array, int_{prof} by dividing the interval $[0, rng_e]$ into n_{int} subintervals. Each subinterval has the interval width, w_{int} , which is given by equation (3.12).

$$w_{int} = \frac{rng_e}{n_{int}} \quad (3.12)$$

Each subinterval can be represented by two ways - the center value or the left and right bounds.

$$int_{prof} = \{i^1(\alpha), i^2(\alpha), \dots, i^{n_{int}}(\alpha)\} \quad (3.13)$$

$$\alpha = \{l, r \text{ or } c\}$$

(A) Dividing $[0, rng_e]$ into n_{int} subintervals to form error frequency distribution

(B) Using the frequency of errors in each subinterval, $prob_{dist}$ is formed

(C) The center value of that subinterval is chosen to be ϵ_0 for which the $cdf_e(k)$ is just greater than the given CI

FIGURE 3.3: PI formation from the training error, e_{train}

2. Iterate through e_{train} and count the number of times an error falls into k^{th} subinterval of int_{prof} . f_{ik} gives the frequency of k^{th} subinterval. And then, the probability distribution, $prob_{dist}$, is formed using the following equation,

$$prob_{dist}(k) = \frac{f_{ik}}{\text{total error samples in } e_{train}} \quad (3.14)$$

3. The probability distribution is used to form the cumulative distribution function, $cdf(k)$, as shown in the Fig. 3.3c.
4. Choose that subinterval's center value, $i^k(c)$, as the final PI width, ϵ_0 , for which equation (3.15) satisfies. In other words, choose the smallest $cdf(k)$ which is greater than the confidence interval, CI . Then, choose the subinterval center value, $i^k(c)$, as the final PI width, ϵ_0 .

$$\begin{aligned} \epsilon_0 &= i^k(c) \\ \text{where, } cdf(k-1) &< CI \leq cdf(k) \end{aligned} \quad (3.15)$$

The so obtained PI width, ϵ_0 , is used to split the testing output into PIs. Testing output, Q_{test} can be obtained from equation (3.8) by inputting the testing inputs to form the hidden layer output matrix, $\mathbf{H}$. The PIs can then be formed for the testing input, x_t , by using the following equation,

$$\{L^{(CI)}(x_t), U^{(CI)}(x_t)\} = Q_{test} \pm \epsilon_0 \quad (3.16)$$

3.4 PI evaluation metric

The quality of PIs can be judged on two chief parameters: *reliability* and *sharpness*. A PI is reliable if it successfully encloses a target value within itself with a given confidence level. The reliability of PIs with a nominal confidence level of $100 \times CI\%$ is measured by evaluating the PI coverage probability (PICP) [30] which is as follows:

$$PICP = \frac{1}{N_t} \sum \kappa_i^{(CI)} \quad (3.17)$$

where, N_t is the number of testing samples, and $\kappa_i^{(CI)}$ is the indicator of PICP, given by

$$\kappa_i^{(CI)} = \begin{cases} 1, & y_t^i \in [L^{(CI)}(x_t), U^{(CI)}(x_t)] \\ 0, & y_t^i \notin [L^{(CI)}(x_t), U^{(CI)}(x_t)] \end{cases} \quad (3.18)$$

A PI with 100% reliability but poor precision is of very less use. PI's sharpness gives an amount of credibility to the results. A sharp PI would ease the planning operation in any field because it involves lesser uncertainty. The interval scores, $S^{(CI)}(x_t)$, incorporate the sharpness factor of the PIs and quantify the PI quality [30], expressed as,

$$\begin{aligned} S^{(CI)}(x_t^i) &= \begin{cases} -2CI\omega - 4[L^{(CI)}(x_t^i) - y_t^i], & \text{if } y_t^i < L^{(CI)}(x_t^i) \\ -2CI\omega, & \text{if } y_t^i \in [L^{(CI)}(x_t^i), U^{(CI)}(x_t^i)] \\ -2CI\omega - 4[y_t^i - U^{(CI)}(x_t^i)], & \text{if } y_t^i > U^{(CI)}(x_t^i) \end{cases} \\ \omega &= U^{(CI)}(x_t^i) - L^{(CI)}(x_t^i) \end{aligned} \quad (3.19)$$

where, (x_t^i, y_t^i) are the testing data samples. The final score is given by,

$$S^{(CI)} = \frac{1}{N_t} \sum S^{(CI)}(x_t^i) \quad (3.20)$$

A narrower PI does not mean that it would have less confidence level in target coverage. For the same confidence level, CI , we can have different sharpness or precision. If the point forecasting model exhibits higher training accuracy, then we could have sharper PIs for the same CI . If the model shows bad accuracy, then PIs would need to be wider to have the same target coverage. So, for the same target coverage accuracy, we can have different precision. This explains the difference between accuracy and precision. On the other hand, for the same training accuracy of a point forecasting model, PIs associated with higher CI would be wider. A PI associated with 0.95 CI would be wider than the one associated with 0.90 CI . In this case, the target coverage accuracy, PICP, can be traded off for the PI's precision or sharpness.

3.5 Data Selection

The data for training and testing the proposed algorithm is taken from the NUS Geography Weather Station, Singapore [23]. The station is located at the $1^\circ 18' N$ & $103^\circ 46' E$ at the height of 90m from the sea level. The solar irradiance data (unit: W/m^2) from Oct 26th, 2014 to Apr 1st, 2018 is taken. It is sampled in 15 min intervals.

3.5.1 Monthly Solar Irradiance Analysis

In order to get a better idea from the solar irradiance data series, the monthly averages of years 2015-2017 are calculated as shown in Fig. 3.4. It is notable that the amount of sunlight falling in the months May-Oct is considerably greater than in the months Nov-Apr. This motivates to segregate the dataset into two parts- May-Oct for the years 2015, 2016 and 2017 (26,830 samples), and Nov-Apr for the years 2014-15, 2015-16, 2016-17 and 2017-18 (34,800 samples). These parts are then used to train two separate models of the proposed approach. The night hours are removed from the dataset because there is no point in forecasting them. Table 3.2 showcases the training data and the testing data used in this paper.

FIGURE 3.4: Average of 2015-2017 monthly solar irradiance in Singapore

TABLE 3.2: Training and Testing Data

<i>model1</i>	Training	Nov-Apr	2014-15 2015-16 2016-17 2017-18	
<i>model1</i>	Testing	Nov-Apr	2014-15 2015-16 2016-17 2017-18	15 th , 16 th , 17 th of Jan & Mar 15 th , 16 th , 17 th of Jan & Mar 15 th , 16 th , 17 th of Jan & Mar 15 th , 16 th , 17 th of Jan & Mar
<i>model2</i>	Training	May-Oct	2015 2016 2017	
<i>model2</i>	Testing	May-Oct	2015 2016 2017	15 th , 16 th , 17 th of Jul & Sep 15 th , 16 th , 17 th of Jul & Sep 15 th , 16 th , 17 th of Jul & Sep

3.5.2 Autocorrelation

For a time-series signal, like the given solar irradiance data, there exists a relation between values at different time samples. However, not all the values can be taken as the input to our forecasting approach. Autocorrelation is used to define the correlation between the values at different time lags. If a value at a certain time lag shows strong correlation with the value at the current time sample, then it can be used as an input to the forecasting approach [27]. The obtained datasets - *NovApr* and *MayOct*, are passed through the Autocorrelation function, $A_X(\lambda)$, to find the correlation of the discrete signal, $x(k)$, with itself at different time sample lags, as follows:

$$A_X(\lambda) = \frac{1}{N} \sum_{k=0}^N \frac{x(k)x(k-\lambda)}{\text{variance}(x)} \quad (3.21)$$

Here, λ is the time sample lag, and N is the number of samples in X . Fig. 3.5 shows the autocorrelation plot for the *NovApr* dataset for λ varying from 0 to 10000. Autocorrelation is the largest at zero time lag and decreases in value as the time lag increases. The autocorrelation curve is approximately sinusoidal in nature as it is alternating after a time period of 48 time samples (1 day from 7:00 to 18:45 hours at 15 minutes interval). This corroborates the fact that solar irradiance, being periodic in nature, at any time would be highly correlated with the solar irradiance 24 hours before. A portion of past values can be chosen whose correlation values are greater than a certain set-threshold. Setting a threshold of 0.45, Table 3.3 shows the feature positions that have Autocorrelation value more than 0.45. The current hour and minute are also considered to be a useful information for irradiance forecasting. *NovApr* dataset has 11 input features whereas *MayOct* dataset has 40 features. *NovApr* dataset has features with time lags varying from two hours to one day, and to two days. Similarly, *MayOct* has features with time lags including one hour, two hours, one day, two days, and so on up to eight days. This tells us that in the months *May-Oct*, the current day's solar irradiance patterns are more correlated to the irradiance patterns from the last week as compared to *NovApr*.

FIGURE 3.5: The Autocorrelation values of solar irradiance series at time lags varying from 0 to 10000

TABLE 3.3: Input features determined by Autocorrelation

NovApr	Hour, Min, t-1, t-2, t-3, t-4, t-5, t-47, t-48, t-49, t-96
MayOct	Hour, Min, t-1, t-2, t-3, t-4, t-5, t-6, t-46, t-47, t-48, t-49, t-50, t-94, t-95, t-96, t-97, t-98, t-142, t-143, t-144, t-145, t-146, t-190, t-191, t-192, t-193, t-238, t-239, t-240, t-241, t-286, t-287, t-288, t-289, t-382, t-383, t-384, t-385, t-386

3.5.3 Permutation Importance

The feature space can be further reduced by opting out features which hardly affect the forecasting accuracy. Data scientists come up with newer approaches to evaluate the effect of features on model accuracy. Permutation importance [28] is one of the methods which involves the following steps:

- The prediction model is first trained on the training data containing all the input feature values arranged in columns.
- The testing data is used to predict the output value and the testing accuracy is noted.
- Now, the values in one of the testing feature columns are shuffled and again testing is performed.
- The decrease in accuracy is calculated due to shuffling of that particular feature column.
- If there is a significant fall in accuracy, then that feature is an important input to the forecasting model. The decrease in accuracy is termed as the Permutation Importance of that feature.

For each of the input features of the datasets *MayOct* and *NovApr* selected by Autocorrelation (as shown in Table 3.3), the permutation importance is calculated. For the *MayOct* dataset, the procedure for permutation importance calculation is performed for each of the 40 features. The resulting values are shown in the Figure 3.6. Unlike the definition, here the importance is being calculated in terms of decrease in prediction error (MAE). If the decrease in MAE is negative (in other words, increase

FIGURE 3.6: Permutation Importance evaluation for the input features of MayOct dataset

FIGURE 3.7: Permutation Importance evaluation for the input features of NovApr dataset

in MAE), then the feature column which was shuffled is important. In case, the decrease in MAE is close to zero, then the corresponding feature does not effect the prediction. From Figure 3.6, *MayOct* dataset has features showing negative decrease in MAE, while there are some features which show zero decrease in MAE. $(t - 1)$ feature holds the most importance followed by $(t - 287)$ (six days time lag), $(t - 2)$, $(t - 5)$ and so on. Some features like $(t - 286)$, $(t - 48)$, $(t - 142)$, $(t - 191)$ and $(t - 240)$, show positive decrease in MAE, which means by shuffling the feature column coincidentally the error decreased and the actual feature column is insignificant and should actually be removed. After setting a threshold of -0.1 on the importance value, a subset of the input features are selected (as shown in Table 3.4).

Similarly, for the *NovApr* dataset, the permutation importance is calculated for each of the 11 features. The resulting values are shown in the Figure 3.7. For *NovApr* dataset, $(t - 1)$ feature holds the most importance followed by $(t - 47)$, $(t - 3)$, $(t - 2)$, $(t - 96)$ and *Hour*. All other features have permutation values less than the set-threshold -0.1, and hence are removed. Table 3.4 shows the final set of input features selected for the proposed approach.

3.6 Results and Discussion

As discussed before, the dataset has been categorized into *NovApr* and *MayOct* sets. Two models of the proposed approach, *model1* and *model2*, are trained on the datasets,

TABLE 3.4: Final set of input features after Permutation Importance evaluation

NovApr	Hour, t-1, t-2, t-3, t-47, t-96
MayOct	Hour, t-1, t-2, t-4, t-5, t-6, t-46, t-47, t-50, t-95, t-143, t-144, t-145, t-146, t-192, t-193, t-238, t-239, t-241, t-287, t-288, t-289, t-383, t-384, t-385

TABLE 3.5: Best set of hyper-parameters and their cross-validation scores for *NovApr* and *MayOct* sets

Datasets	$\{n, \sigma_{width}\}$	Best validation score (normalized MAE)
<i>NovApr</i>	{500, 0.05}	0.336544914
<i>MayOct</i>	{100, 0.005}	0.3062656659

NovApr and *MayOct* sets, respectively. Both the datasets are respectively normalized and then the test sets are separated from the training sets. The activation function chosen for the ELM model is Gaussian activation function. Grid search is performed on a range of parameter values to tune the hyper-parameters of the model $ELM(n, \sigma_{width})$. The search for the minimum error is performed for out of all combinations of $n = \{10, 50, 100, 500\}$ and $\sigma_{width} = \{0.005, 0.05, 0.5\}$. For each combination, 10-fold Cross-Validation is performed. Table 3.5 shows the best parameters with minimum cross-validation MAE for different datasets.

With a good idea about the rough location of optimal hyperparameter set, DE is implemented with *mutation rate*= 0.8, *crossover rate*= 0.7, *population size*= 20 and *iterations*= 1000. For *NovApr* dataset, the search space is chosen to be: $n \in [480, 520]$ and $\sigma_{width} \in [0.01, 0.1]$. For *MayOct* dataset, the search space is chosen to be: $n \in [80, 120]$ and $\sigma_{width} \in [0.001, 0.01]$. Table 3.6 shows the best candidate hyperparameter sets for the ELM model. The training errors are roughly around 8.3% and have a scope for improvement.

Training is done for the selected hyper-parameters, which is then followed by the PI formation for two confidence levels - 0.90 and 0.95. For the probability distribution of errors, the value of n_{int} is taken to be 500 which is high enough to form smooth distributions. Table 3.7 shows the final PICP and Interval scores for *NovApr* and *MayOct* sets. From the results, it can be seen that, the proposed approach is almost able to achieve the CIs. The *model1* results into 0.889 coverage probability with -1.5574 interval score for $CI = 0.90$. The 0.90 CI is fairly achieved by 0.889 PICP, while an interval score of -1.5574 shows that the model was able to maintain fair target coverage along with maintaining considerable sharpness. The idea is to maintain a trade-off between PI coverage probability and sharpness, while pushing the PI to achieve better PICP and sharpness. This means the interval score needs to be reduced from -1.5574 in magnitude while maintaining PICP close to 0.90. For 0.95 CI, *model1* offers 0.940 PICP with -1.5637 interval score. Similarly, *model2* results into

TABLE 3.6: Final set of hyper-parameters obtained from DE

Datasets	$\{n, \sigma_{width}\}$	$MAE_{train}/\max(\text{Error}) (W/m^2)$
<i>NovApr</i>	{495, 0.0368122717}	66.448/795.66
<i>MayOct</i>	{120, 0.0084154851}	66.68/797

TABLE 3.7: PICP and interval scores for *NovApr* and *MayOct* sets

Datasets	CI	PICP	Interval Score
<i>NovApr</i>	0.90	0.8897569444	-1.5573931799
	0.95	0.9401041667	-1.5637002833
<i>MayOct</i>	0.90	0.8912037037	-1.3136729998
	0.95	0.9363425926	-1.310009537

FIGURE 3.8: Prediction intervals for *NovApr* set with CI=0.90

0.891 and 0.936 PICPs with -1.3136 and -1.3100 interval scores for 0.90 and 0.95 CIs respectively. It can be seen that the interval scores obtained on *MayOct* dataset are better than that of *NovApr*. The interval scores represent PI width which implies that PI width achieved for the *MayOct* dataset is narrower. It was more difficult to provide narrower PIs for *NovApr*. This must be because of the north-east monsoon with relatively higher rainfall, frequent cloud cover and randomly occurring winds in the months from December to March [38]. However, the proposed approach is still providing interval scores with magnitudes as less as 1.5 maintaining PICPs close to the CI levels.

Figures 3.8, 3.9, 3.10 and 3.11 display the resulting PIs for the two datasets with CIs as 0.90 and 0.95. Figure 3.8 shows the PIs for *NovApr* test data (as defined in 3.5.1) for $CI = 0.90$. The resulting PIs are successful in enclosing the highly fluctuating solar irradiance pattern maintaining 0.90 CI. However, the PIs have to be quite wide in order to achieve 0.90 CI. Figure 3.9 shows the PIs for *NovApr* test data for $CI = 0.95$ which are successful in maintaining $PICP = 0.94$. However, the PI widths have noticeably increased slightly in order to achieve the PICP. Similarly, Figures 3.10 and 3.11 show the PIs for *MayOct* test data (as defined in 3.5.1) with CI as 0.90 and 0.95 respectively. The resulting PIs are successful in enclosing the highly fluctuating solar irradiance pattern maintaining CI levels.

As can be seen from the Figures 3.8, 3.9, 3.10 and 3.11, the PI widths need to

FIGURE 3.9: Prediction intervals for *NovApr* set with CI=0.95

FIGURE 3.10: Prediction intervals for *MayOct* set with CI=0.90 FIGURE 3.11: Prediction intervals for *MayOct* set with CI=0.95

be reduced to improve the sharpness factor, while maintaining the reliability. The wide intervals are useful in enclosing the highly fluctuating portions of the irradiance pattern. However, the PI width should be narrower in the portions where there is less fluctuation, providing variable-width PIs. The PI widths should adapt to the degree of fluctuations and change accordingly. Another aspect that can improve the PI widths is better training accuracy as the final PI width, ϵ_0 , depends on the magnitudes of error samples (as defined in 3.3). Lower the magnitudes of error samples, lesser would be ϵ_0 , and narrower would be the PIs. The current search space of the hyper-parameters during Grid search can be varied beyond $n = \{10, 50, 100, 500\}$ and $\sigma_{width} = \{0.005, 0.05, 0.5\}$ to explore hyper-parameters which could give even better training accuracy and eventually narrower PI widths. Eventually, the interval scores also can be improved.

Chapter 4

Conclusion and Future Directions

4.1 Grid Impact Analysis for PV integration

The 3.3 MWp PV addition at the low voltage end, i.e. 415V, of the distribution side helps reduce the critical instability cases of bus voltages in Medium, Light and Heavy loading scenarios. It also mitigates the electrical load on the cables by directly supplying the local load and avoiding possible transmission losses. The ESS acts as a backup for the PV in case the PV is not available, i.e. in the night hours or cloudy days. In the Light load scenario, the combination of PV and ESS could send an extra amperes back to the Singapore grid. Overall, PV and ESS addition to the low voltage end of the grid is beneficial to the grid. This research can be moved forward by evaluating the impact on the Singapore grid at different levels of PV penetration and finding out the points of critical operation to avoid violations of grid stability limits similar to how Cheng et al. [39] performed the steady-state load flow analysis on the power grids in the cities in California.

4.2 PV Power Interval Forecasting

The prominence of solar power makes its forecasting an important agenda for power system planning. When the actual PV output data is unavailable, the indirect PV forecasting approach can be followed, where PV output predictions are obtained using the solar irradiance forecasts. This report used the solar irradiance series as input to perform interval forecasting using a probabilistic PI formation approach on ELM. Interval forecasting enables the decision-makers to quantify the availability of PV power in a broader sense. Prediction intervals are essential when dealing with PV generation series which could have sharp variations. Parametric interval formation methods are based on error distribution assumption whereas non-parametric probabilistic forecasting approach does not required any prior assumption. Using the proposed approach, the deterministic forecasters, like ELM, can be transformed into interval forecasting by making use of the training error distribution.

The proposed algorithm works well in achieving the given confidence levels. However, the interval scores, which incorporate the sharpness factor, show a scope for improvement. The motif of using Grid search is to obtain an initial idea about the performance of the proposed approach over a set of hyper-parameter values. It is followed by a comprehensive optimization using Differential Evolution over the model hyperparameter search space. The approach tries to reduce the optimal hyperparameter search time by narrowing down to the most possible optimal search space. Further study is required to achieve sharper PIs without affecting their coverage probability. The sharpness of the PIs is inversely dependent on the training

error. Recently, researchers are deploying deep learning approaches for deterministic forecasting followed by Quantile Regression for estimating quantiles of the output variable [40]. The research claims to be able to extract deep nonlinear features from the PV output series. More research is needed on improving the accuracy of the deterministic forecasters to implement efficient probabilistic forecasting. The modularity of the proposed approach makes it flexible in testing more advanced intelligent algorithms replacing ELM. The future study will also focus on the inclusion of exogenous features like the cloudiness parameter or sky cover which seems very promising in solar irradiance forecasting.

4.3 Future directions

Before the actual implantation of the 3.3MW PV plant, grid resilience against the variability of the PV output should be evaluated. The scenario should be tested with approximate local consumer load profile to evaluate any possible critical operation points. Also, techno-economical cost studies need to be done to quantify the exact benefit from PV and ESS addition. This could be included in the future studies.

For PV power interval forecasting, the future work would orient towards multiple directions. Firstly, the current approach of probabilistic prediction interval formation can be taken forward to improve the PI sharpness. For this, the PI needs to be optimized by considering a trade-off between PICP and the sharpness factor. Making ensembles of a number of ELM algorithms will also be considered. Secondly, considering the popularity of deep learning and its success in achieving greater accuracy, certain algorithms like the Long-Short Term Memory (LSTM) models can be aimed for implementing on the solar irradiance dataset.

Bibliography

- [1] UNFCCC, "The united nations framework convention on climate change," 2019. [Online]. Available: <https://unfccc.int/> Accessed: 20-May-2019.
- [2] V. Shah, "Singapore pledges to peak emissions by 2030," 2015. [Online]. Available: <http://www.eco-business.com/news/singapore-pledges-to-peak-emissions-by-2030/> Accessed: 17-Oct-2018.
- [3] GENI, "Grid summary, national energy grid singapore," 2016. [Online]. Available: http://www.geni.org/globalenergy/library/national_energy_grid/singapore/index.shtml Accessed: 17-Oct-2018.
- [4] EMA, "Singapore energy statistics 2016," 2016. [Online]. Available: https://www.ema.gov.sg/cmsmedia/Publications_and_Statistics/Publications/SES%202016/Publication_Singapore_Energy_Statistics_2016.pdf Accessed: 17-Oct-2018.
- [5] H. Ali, S. Sanjaya, B. Suryadi, and S. Weller, "Analysing co2 emissions from singapore's electricity generation sector: Strategies for 2020 and beyond," *Energy*, vol. 124, pp. 553–564, 2017.
- [6] NCCS, "Singapore's climate action plan, take action to-day for a climate-efficient singapore," 2016. [Online]. Available: https://sustainabledevelopment.un.org/content/documents/1545Climate_Action_Plan_Publication_Part_1.pdf Accessed: 17-Oct-2018.
- [7] Y. Deng, Z. Li, and J. M. Quigley, "Economic returns to energy-efficient investments in the housing market: evidence from singapore," *Regional Science and Urban Economics*, vol. 42, no. 3, pp. 506–515, 2012.
- [8] L. T. Lam, L. Branstetter, and I. L. Azevedo, "A sunny future: expert elicitation of china's solar photovoltaic technologies," *Environmental Research Letters*, vol. 13, no. 3, p. 034038, 2018.
- [9] PUB, "Pub singapore's national water agency, brochure 2017," 2017. [Online]. Available: <https://www.pub.gov.sg/AnnualReports/corporatebrochure2017.pdf> Accessed: 17-Oct-2018.
- [10] K. Trapani and M. Redón Santafé, "A review of floating photovoltaic installations: 2007–2013," *Progress in Photovoltaics: Research and Applications*, vol. 23, no. 4, pp. 524–532, 2015.
- [11] L. Bird, J. Heeter, E. O'Shaughnessy, B. Speer, O. Cook, T. Jones, M. Taylor, P. Ralon, and E. Nilson, "Policies for enabling corporate sourcing of renewable energy internationally: A 21st century power partnership report," National Renewable Energy Lab.(NREL), Golden, CO (United States), NREL/TP-6A50-68149, 2017.

- [12] MESIA, "Solar outlook report 2018," 2018. [Online]. Available: <https://www.mesia.com/wp-content/uploads/2018/03/MESIA-OUTLOOK-2018-Report-7March2018.pdf> Accessed: 03-Aug-2018.
- [13] A. Patil, R. Girgaonkar, and S. K. Musunuri, "Impacts of increasing photovoltaic penetration on distribution grid—voltage rise case study," in *International Conference on Advances in Green Energy (ICAGE)*. IEEE, 2014, pp. 100–105.
- [14] P. Lazzeroni, S. Olivero, F. Stirano, and M. Repetto, "Impact of pv penetration in a distribution grid: A middle-east study case," in *IEEE 1st International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*. IEEE, 2015, pp. 353–358.
- [15] S. Ruiz-Romero, A. Colmenar-Santos, F. Mur-Pérez, and Á. López-Rey, "Integration of distributed generation in the power distribution network: The need for smart grid control systems, communication and equipment for a smart city—use cases," *Renewable and sustainable energy reviews*, vol. 38, pp. 223–234, 2014.
- [16] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.
- [17] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. Martinez-de Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016.
- [18] A. Fabbri, T. G. S. Roman, J. R. Abbad, and V. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1440–1446, 2005.
- [19] R. Marquez and C. F. Coimbra, "Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the nws database," *Solar Energy*, vol. 85, no. 5, pp. 746–756, 2011.
- [20] J. G. da Silva Fonseca Jr, T. Oozeki, T. Takashima, G. Koshimizu, Y. Uchida, and K. Ogimoto, "Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in kitakyushu, japan," *Progress in photovoltaics: Research and applications*, vol. 20, no. 7, pp. 874–882, 2012.
- [21] R. Zhang, Z. Y. Dong, Y. Xu, K. Meng, and K. P. Wong, "Short-term load forecasting of australian national electricity market by an ensemble model of extreme learning machine," *IET Generation, Transmission & Distribution*, vol. 7, no. 4, pp. 391–397, 2013.
- [22] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [23] NUS-GWS, "Nus geography weather station," 2019. [Online]. Available: <https://inetapps.nus.edu.sg/fas/geog/ajxdirList.aspx> Accessed: 20-May-2019.
- [24] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, "Short-term solar irradiance forecasting using exponential smoothing state space model," *Energy*, vol. 55, pp. 1104–1113, 2013.

- [25] M. Akram, R. J. Hyndman, J. K. Ord *et al.*, "Non-linear exponential smoothing and positive data," Monash University, Department of Econometrics and Business Statistics, Tech. Rep., 2007.
- [26] S. Chai, M. Niu, Z. Xu, L. L. Lai, and K. P. Wong, "Nonparametric conditional interval forecasts for pv power generation considering the temporal dependence," in *Power and Energy Society General Meeting (PESGM)*, 2016. IEEE, 2016, pp. 1–5.
- [27] H. T. Pedro and C. F. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Solar Energy*, vol. 86, no. 7, pp. 2017–2028, 2012.
- [28] D. Becker, "Permutation importance," 2018. [Online]. Available: <https://www.kaggle.com/dansbecker/permutation-importance> Accessed: 30-Dec-2018.
- [29] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1845–1856, 2010.
- [30] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1166–1174, 2014.
- [31] ETAP, "Etap releases version 16.0," 2016. [Online]. Available: <https://etap.com/newsdetail/2016/11/23/etap-releases-version-16.0> Accessed: 18-May-2019.
- [32] ETAP, "Running a load flow analysis," 2009. [Online]. Available: <https://etap.com/docs/default-source/faqs-tutorials/load-flow-analysis.pdf?sfvrsn=10> Accessed: 17-Oct-2018.
- [33] D. A. Harville, *The Moore-Penrose Inverse*. New York, NY: Springer New York, 1997, pp. 497–519.
- [34] Scikit-learn, "Tuning the hyper-parameters of an estimator," 2019. [Online]. Available: http://scikit-learn.org/stable/modules/grid_search.html Accessed: 20-May-2019.
- [35] Scikit-learn, "Cross validation: evaluating estimator performance," 2019. [Online]. Available: http://scikit-learn.org/stable/modules/cross_validation.html Accessed: 20-May-2019.
- [36] G. S. dos Santos, L. G. J. Luvizotto, V. C. Mariani, and L. dos Santos Coelho, "Least squares support vector machines with tuning based on chaotic differential evolution approach applied to the identification of a thermal process," *Expert Systems with Applications*, vol. 39, no. 5, pp. 4805–4812, 2012.
- [37] J. Brownlee, *Clever algorithms: nature-inspired programming recipes*. Jason Brownlee, 2011.
- [38] Meteorological-Service-Singapore, "Climate of singapore," 2019. [Online]. Available: <http://www.weather.gov.sg/climate-climate-of-singapore/> Accessed: 18-May-2019.

- [39] D. Cheng, B. Mather, R. Seguin, J. Hambrick, and R. P. Broadwater, "Pv impact assessment for very high penetration levels," in *IEEE 42nd Photovoltaic Specialist Conference (PVSC)*. IEEE, 2015, pp. 1–6.
- [40] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, H. Jiang, and W. Liu, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Conversion and Management*, vol. 153, pp. 409–422, 2017.