Enhanced Power System Operation with Coordination and Forecasting Techniques

Fatemeh Najibi

CitAI Research Centre
Department of Computer Science
City, University of London

A thesis submitted in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

City, University of London September 2021

To My Love, Vali

To My Parents

To My Lovely Nieces and Nephews

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains less than 65,000 words including appendices, bibliography, footnotes, tables and equations and has less than 150 figures. I grant powers of discretion to the City, University of London librarian to allow the dissertation to be copied in whole or in part without further reference to myself (the author). This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgment.

Fatemeh Najibi

September 2021

Acknowledgements

This thesis would not have been possible without the constant encouragement, inspiration and support of a number of wonderful individuals. My very special thanks and appreciation to all of them for being part of my PhD journey and making this dissertation possible.

I would like to express my deep and sincere gratitude to my research supervisor, Dr Eduardo Alonso, and my advisor, Dr Dimitra Apostolopoulou, to let me this opportunity to put all ideas together, well above the level of simplicity and into something concrete. This thesis would have never been accomplished without their constant support, motivation, and encouragement, especially during pandemics. I am thankful to them for providing invaluable guidance throughout this thesis, which allowed me to grow as an independent researcher.

I would like to express my gratitude and appreciation for Dr Esther Mondragón, my Senior Tutor for Research, for her treasured support and continuous encouragement during my study.

I would like to thank my friends, lab mates, colleagues, and research team for a cherished time spent together in the lab and social settings. My time at City was enjoyable due to many marvellous friends and groups who have become part of my life. I am sincerely grateful for all my friends who made my PhD journey more pleasant.

I am incredibly grateful to my parents for their love, prayers, caring, and sacrifices for educating and preparing me for my future. I am deeply indebted to them for all their love, support, and help.

I am extending my thanks to my lovely family, who were always there for me—especially my sister Sarah and her brilliant husband, Alireza. Thanks to them, who constantly remind me that I have a family to count on when life is rough.

And finally, to the love of my life Vali, whose love and support for me has always been unconditional. Throughout all ups and downs of my PhD, he has been beside me, living every success and failure of my PhD, even when I was anxious and depressed. Thanks to him, who never loses his faith in me.

Abstract

With the integration of renewable energy into power systems, traditional power systems face new challenges. Due to their inherent fluctuations and variability, the introduction of renewable energy in power systems poses new challenges in modelling uncertainty. Controlling and optimising the operation cost by adjusting the output generation of renewable energy resources makes the operation more reliable and secure.

We first formulate the optimal power flow (OPF) problems for both the transmission and distribution systems and investigate the variables that greatly affect the outcome.

Solving the power system optimal operation problem, we realise the importance of uncertainties involved with renewable energy due to the inherent variability of weather data. Accurate forecasting mechanisms that address their inherent intermittency and variability enable the smooth integration of such resources in power system operations. To solve this problem, in the next step, we propose a novel probabilistic framework to predict short-term PV output taking into account the variability of weather data over different days and seasons. We go beyond existing prediction methods, building a pipeline of processes, i.e., feature selection, clustering and Gaussian Process Regression (GPR). We make use of datasets that comprise power output and meteorological data such as irradiance, temperature, zenith, and azimuth. First, a correlation study is performed to select the weather features which affect solar output to a greater extent. Next, we categorise the data into four groups based on solar output and time using k-means clustering. Finally, we determine a function that relates the selected features with solar output using GPR and Matérn 5/2 as a kernel function. We validate our method with five solar generation plants in different locations and compare with the existing methodologies. More specifically, to test the proposed model, two different methods are used: (i) 5-fold cross-validation; and (ii) holding out 30 random days as test data. To confirm the model accuracy, we apply our framework 30 independent times on each of the four clusters. The average error follows a normal distribution, and with a 95% confidence level, it takes values between -1.6% to 1.4%. The proposed framework

decreases the normalised root mean square error and mean absolute error by 54.6% and 55.5%, respectively, compared with other relevant works.

Although we address the integration of a Microgrid into the distribution power network in the first research question, we yet need to address the transmission system constraints, as the incorporation of renewable energy into power systems poses serious challenges to the transmission and distribution power system operators (TSOs and DSOs). To fully leverage these resources, there is a need for a new market design with improved coordination between TSOs and DSOs. To answer the last research question, we propose two coordination schemes between TSOs and DSOs: one centralised and another decentralised that facilitate the integration of distributed based generation; minimise operational cost; relieve congestion; promote a sustainable system. To this end, we approximate the power equations with linearised equations so that the resulting OPFs in both the TSO and DSO become convex optimisation problems. In the resulting decentralised scheme, the TSO and DSO collaborate to allocate all resources in the system optimally. In particular, we propose an iterative bi-level optimisation technique where the upper level is the TSO that solves its own OPF and determines the locational marginal prices at substations. We demonstrate numerically that the algorithm converges to a near-optimal solution. We study the interaction of TSOs and DSOs and the existence of any conflicting objectives with the centralised scheme. More specifically, we approximate the Pareto front of the multi-objective optimal power flow problem where the entire system, i.e., transmission and distribution systems, is modelled. The proposed ideas are illustrated through a five-bus transmission system connected with distribution systems, represented by the IEEE 33- and 69-bus feeders.

Table of contents

List of figures **xv**

List of tables **xix**

1 Introduction **1**

1.1 Motivation 1

1.2 Problems Overview and Research Objectives 2

1.3 Research Questions 3

1.4 Main Contributions 3

1.5 List of Publications 4

1.5.1 Peer-reviewed Journal 4

1.5.2 Conference 5

1.5.3 Poster Presentation 5

1.6 Dissertation Outline 5

2 Literature Review **7**

2.1 Power System Optimal Operation 7

2.2 Solar Output Forecast Model 12

2.3 Transmission-Distribution Coordination Schemes 14

3 Power System Optimal Operation **17**

3.1 Introduction 17

3.2 Methodology 18

3.2.1 Microgrid modeling 18
3.3 Optimal Power Flow Formulation 20
3.3.1 Transmission level 20
3.3.2 Distribution Level 22
3.4 Quadratic Programming 26
3.5 Results 27
3.6 Conclusion 33

4 Solar Output Forecast **35**
4.1 Introduction 35
4.2 Data Processing 38
4.2.1 Features' Selection 38
4.2.2 Data Clustering 42
4.2.3 Clustering Sensitivity Implementation 48
4.3 Proposed Forecasting Framework 51
4.3.1 Gaussian Process Regression 51
4.3.2 Framework Validation 56
4.4 Numerical results 57
4.4.1 Dataset Information 57
4.4.2 Framework Implementation on Site A 58
4.4.3 Summary of results of all sites 64
4.5 Conclusion 66

5 Transmission-Distribution Coordination Schemes **69**
5.1 Introduction 69
5.2 Decentralised Scheme 73
5.3 Centralised Scheme 78
5.4 Numerical Results 81
5.4.1 System Description 81
5.4.2 Decentralised Coordination Scheme 83

5.4.3 Centralised Coordination Scheme 92
5.5 Conclusion 99

6 Conclusion and Future Work 101
6.1 Conclusion 101
6.2 Future Work 105
6.2.1 More detailed coordination scheme 105
6.2.2 Probabilistic model 105
6.2.3 Electro-thermal model of PV 106
6.2.4 Model of pump hydro 106
6.2.5 Model of hydro plant 107
6.2.6 Hyperparameters Optimization 107
6.2.7 Hosting Capacity 107

References 109

List of figures

3.1 The diagram of the network used with the locations of the various resources 19
3.2 Estimated values of the load demand, market price and PV and WT power production. 20
3.3 Cost before and after optimization 27
3.4 Optimal cost divided into different cost parts 28
3.5 The total cost value for H=1,2,3,4,5 for the first hour 30
3.6 The total cost value for H=1,2,3,4,5 for 24 hour 30
3.7 Output generation for H=1 for 24 hours 32
3.8 Maximum Voltage deviation from the reference voltage for 24 hours 32
3.9 Average Voltage deviation from the reference voltage for 24 hours 33

4.1 The proposed process of building trained models. 36
4.2 The process of testing our trained models to predict solar output. 37
4.3 3D graph of four clusters. Different colours represent different clusters. . . 42
4.4 Elbow statistic test. 43
4.5 Gap statistic test. 44
4.6 Sensitivity study on the number of clusters by comparing different normalised error metrics values. 47
4.7 Box plot of dataset that shows non existence of outliers. 47
4.8 Proposed framework predictions of the training data set for one cluster. . . . 50
4.9 Proposed framework predictions of the training data set for four clusters. . . 51
4.10 Proposed framework predictions of the training data set for eight clusters. . 51

4.11 Diagram of k-fold cross-validation with k=5; the grey boxes refer to testing the data and the white to training. 56
4.12 The overall forecast system. 56
4.13 Proposed framework predictions of the training data set. 58
4.14 Proposed framework hourly predictions of the training data set. 59
4.15 hourly prediction for 30 random days. 59
4.16 30 random days 24-hour prediction with one hour intervals, Denver. 60
4.17 Probability distribution fitting of average error of one sample model. 61
4.18 Prediction interval with 95% confidence level and one hour look-ahead time for 17 June, 2006 at Site A. 62
4.19 Prediction interval with 95% confidence level and one hour look-ahead time for 17 March, 2006 at Site A. 62
4.20 Q-Q plot of absolute error between the observed value of sample test data points and their predicted value. 64

5.1 Decentralised iterative scheme flowchart. 78
5.2 Transmission and distribution system. 82
5.3 Modified IEEE 33-bus distribution feeder. 83
5.4 Modified IEEE 69-bus distribution feeder. 83
5.5 Transmission operation cost for methods (i) current practise and (ii) proposed decentralised TSO-DSO coordination scheme. 85
5.6 The total amount of distributed generation for methods (i) current practise and (ii) proposed decentralised TSO-DSO coordination scheme at nodes 3 and 4. 85
5.7 The cost for each feeder for methods (i) and (ii). 90
5.8 Netload at nodes 2,3 with using methods (i) and (ii). 90
5.9 TSO and DSOs operational cost using the proposed decentralised coordination scheme. 91
5.10 Evolution of the hourly cost for F_2 with respect to the iteration number. . . 91

5.11 Evolution of the hourly cost for the transmission system with respect to the iteration number. 92
5.12 Evolution of hourly cost for F_2 for different penetration levels of distributed generation. 92
5.13 Transmission operation cost for methods (i) current practise and (iii) proposed centralised TSO-DSO coordination scheme. 93
5.14 The total amount of distributed generation for methods (i) current practise and (iii) proposed centralised TSO-DSO coordination scheme at nodes 3 and 4. 93
5.15 Net load at nodes 2,3 with using methods (i) and (iii). 94
5.16 TSO and DSOs operational cost using the proposed centralised coordination scheme. 94
5.17 Pareto Front of the sum of all feeders DG and voltage regulation daily cost with respect to the TSO cost. 97
5.18 Pareto Front of the sum of all feeders daily cost with respect to the TSO cost. 98
5.19 Pareto Front of daily cost for F_i, $i = 1,\ldots,4$ with respect to the TSO cost. . 98

List of tables

3.1 The limit information and bid information for each generator 19
3.2 Total cost before and after optimization 29
3.3 Optimal output generation 31

4.1 The absolute value of correlation coefficients between solar output and different features. 40
4.2 Correlation value thresholds 41
4.3 Site description. 45
4.4 Clusters' details 47
4.5 Training set error metrics for various number of clusters. 49
4.6 Test set error metrics for various number of clusters. 49
4.7 Site A forecasts' error metrics. 60
4.8 Training set error metrics for all sites. 64
4.9 Test set error metrics for all sites. 65
4.10 The comparison of forecast error metrics for our framework against two other methods ([1, 2]), for site E. 65

5.1 Distributed resources' physical limits and bid information. 84
5.2 Locational marginal prices for method (i): current practise for TSO-DSO coordination in €/MW. 86
5.3 Locational marginal prices for method (ii): proposed decentralised TSO-DSO coordination in €/MW. 87

5.4 The power output in MW of generators at the transmission level for method (i): current practise for TSO-DSO coordination. 88
5.5 The power output in MW of generators at the transmission level for method (ii): proposed decentralised TSO-DSO coordination. 89
5.6 The power output in MW of generators at the transmission level for method (iii): proposed centralised TSO-DSO coordination. 95
5.7 Locational marginal prices for method (iii): proposed centralised TSO-DSO coordination in €/MW. 96

Glossary of abbreviations

AS Ancillary Service. 2

CI Confidence Interval. 61

DG Distributed Generators. 17

EV Electric Vehicles. 8

IEM Iowa Environmental Mesonet. 39

LD Load Ddeman. 10

MAE Mean Absolute Error. 37

MG MicroGrid. 18

MLP Multi-layered Perceptron. 65

MT MicroTurbine. 18

NREL National Renewable Energy Laboratory. 39

NSR National Solar Radiation. 39

PICP Prediction Interval Coverage Probability. 62

PINRW Prediction Interval Normalized Root-mean-square Width. 62

PV Photovoltaic. 3, 8, 58

RMSE Root Mean Square Error. 37

TSO Transmission System Operator. 7

Chapter 1

Introduction

1.1 Motivation

Over the past years, many countries have opted to integrate renewable energy resources such as solar energy in the grid to increase the penetration of environmental friendly resources (see, e.g., [3–5]) as well as to find cleaner methods of powering electric vehicles [6]. As an example, Japan, China, Germany, USA, and UK are able to meet 80% of their demand from solar generation; and the total installed solar energy capacity at the end of 2018 was more than 500 GW [4, 7]. However, the large-scale penetration of renewable energy resources into conventional power distribution systems presents different challenges to power system operation. Considering the connected power distribution grid, the optimal operation of these renewable energy resources provides a promising future for the power grid. The ratio of r/x, the ratio of the system resistance to the system reactance, makes the voltage of each bus sensitive to active power injections (e.g., [8]). Therefore, one way to control the voltage is to inject active power that storage-based resources and renewable energy resources can do through power system optimal operation management (e.g., [9]).

More specifically, the inherent uncertainty of these resources, e.g., irradiance, temperature and cloud conditions, makes its smooth integration into power system operations a formidable challenge. The intermittency of solar generation might cause issues in system stability, power

balance and frequency response, and reactive power generation (see, e.g., [10–12]). As a result, building accurate forecast models of solar generation is of vital importance.

Furthermore, incorporating renewable energy into the power system also challenges power system operators' transmission (TSO) and distribution (DSO). On the one hand, the transition of the power system into the new carbon-free system is inevitable; on the other hand, these resources might disturb the power system quality, voltage, and frequency [13]. The TSO can use these renewable energy resources (RES) to mitigate the power quality, improve the congestion management and the operation cost, and the DSO can use the local energy resources. However, despite all the techno-economic benefits of RESs, and also the new incentive policies which governments have introduced [14] to support the investor in the energy sector, the power network operators need to work in a more coordinated and collaborative mode to be able to maintain the reliable power system [15–17]. Specifically, the procurement of RESs as ancillary services (AS) in the power system needs the coordination between the distribution (DS) and the transmission system (TS) [18].

1.2 Problems Overview and Research Objectives

As mentioned in the previous section, there are difficulties in integrating renewable energy resources into traditional power systems. This dissertation aims to propose solutions to smooth the power system's transition into the new carbon-free system. To this end, the dissertation starts with modelling a distribution network into a linear optimal distribution network considering a Microgrid that comprises different renewable energy resources. The main objectives are to optimally operate a microgrid comprising different distributed generators comprising the upper-hand distribution power system and to minimise each bus's voltage deviation from the reference value. Thus, we approximate the AC load flow into a linearised AC load flow and solve the optimization problem in a deterministic framework using Quadratic programming.

Then, to model the PV output in a more realistic framework in the power system, we propose a forecast model to predict the solar output for five different sites. Due to the solar

output dependency on the weather type and its parameters, most recent works have focused on the seasonality trends or weather types rather than on the time of day. However, the solar output in one season does not necessarily follow the same pattern. Also, PV energy resources' inherent uncertainty due to, e.g., irradiance, temperature and cloud conditions, makes its smooth integration into power system operations a formidable challenge.

Considering a large-scale penetration of PV-Batteries into the power network, we aim to help the Transmission system operator (TSO) and distribution system operator (DSO) maintain and operate their system more reliably and collaboratively. To this end, we propose an iterative algorithm to solve a bi-level TSO-DSO coordination scheme in a decentralised manner where no sensitive information is being exchanged. Furthermore, we will analyse the interaction of TSOs and DSOs, i.e., how conflicting their objectives and priorities are, by formulating a common TSO-DSO OPF scheme where the Pareto front is determined.

1.3 Research Questions

The essential focus of this dissertation addresses the following research questions:

- How do we improve the solar output forecast model considering the inherent uncertainty of weather data?
- Considering the integration of the distributed generation into the power systems, how do we model the coordination schemes between the distribution and transmission power system?

1.4 Main Contributions

In this dissertation, we have focused on the impact of renewable energy on the power system. The main contribution and focus of our work are summarised as:

- The identification of highly correlated features with solar output through a preprocessing stage which will make the solar forecast model simpler, demanding less data with yet greater prediction accuracy.

- Development of clustering techniques based on time of day and categorization of data into four clusters that improve the accuracy of the GPR model beyond the state-of-the-art techniques.
- Choice of using Matérn 5/2 as kernel function that has not been exploited for solar output prediction previously despite its capability to solve probabilistic problems.
- Proposing the transmission-distribution coordination schemes by using a bi-level iterative approach and Quadratic optimization techniques.
- Constructing a centralised TSO-DSO framework which is used to quantify the operators' conflicting objectives and provide appropriate incentives for their coordination.
- Proposing a decentralised TSO-DSO scheme that reaches a near-least cost solution by respecting the privacy concerns of TSOs, DSOs; is computationally efficient; relieves congestion; and increases the level of DG resources' integration.

1.5 List of Publications

A number of articles have been published during this PhD. Part of this dissertation is based on the following peer-reviewed journal and conference papers.

1.5.1 Peer-reviewed Journal

- F. Najibi, D. Apostolopoulou, and E. Alonso, **"Gaussian process regression for probabilistic short-term solaroutput forecast**," International Journal of Electrical Power and Energy Systems, 2021 .
- F. Najibi, D. Apostolopoulou, and E. Alonso , **"TSO-DSO Coordination Schemes to Facilitate Distributed Resources Integration"**, Sustainability, 2021.

1.5.2 Conference

- F. Najibi, D. Apostolopoulou, and E. Alonso, **"Clustering Sensitivity Analysis for Gaussian Process Regression Based Solar Output Forecast"**, PowerTech IEEE PES, 2021.

- F. Najibi, D. Apostolopoulou, and E. Alonso, **"Optimal dispatch of pumped storage hydro cascade under uncertainty"**, in 2018 UKACC 12th International Conference on Control, pp. 187–192, IEEE, 2018.

1.5.3 Poster Presentation

- Poster/Talk on **'Microgrid Optimal Operation'** , Doctoral Symposium, City University of London, 2018.

1.6 Dissertation Outline

This dissertation includes six chapters and is organized as follows:

1. Chapter 1 discusses the motivations for the research conducted in this dissertation. Also, a short description of the related work is discussed. The contributions of this thesis and the list of publications are also presented in this chapter.

2. Chapter 2 provides the constructive literature review on the power system optimal operation, different forecast models, and different coordination schemes between the transmission and distribution power systems.

3. In Chapter 3, the power system optimal operation is discussed. The linear model that we developed for the deterministic optimal operation is formulated.

4. Chapter 4 presents an enhanced performance Gaussian Process Regression framework with the choice of a specific kernel for probabilistic short-term solar output forecast. This chapter starts with a brief introduction to the related work. Then, the data

processing as the correlation study and clustering analysis is described. Next, the formulation of the GPR with Matérn 5/2 as a kernel function related to the solar output with the input features is discussed as the framework validation methodologies. In the last section of this chapter, we illustrate the proposed methodology through five different datasets. In conclusion, we will summarise the results and make some concluding remarks.

5. In Chapter 5 proposes two different coordination schemes will be presented to help the TSO and DSO to operate the power system more smoothly. This section encompasses the power grid data, load data, and the mathematical formulation of linear approximation. Then we will go over implementation and the numerical results. And finally, the comparison between different schemes will be elaborated in detail as in the conclusion.

6. Chapter 6 is the final chapter where all results will be discussed, final remarks will be made, and reflections will be presented on the research aims and outcomes. Also, we will provide more suggestions for future research.

Chapter 2

Literature Review

The literature review is divided into the following sections: In the first section, a relevant history on the power system optimal operation considering the integration of renewable energy is discussed. Then, following the renewable energy challenges in the previous section, section 2.2 provides more information on the exigency of proposing solar output forecast model. More specifically, this section presents a brief history of different solar forecast models and the specific Gaussian Process Regression technique which we exploited in this thesis. After providing more information on the forecast model, we elaborate the impact of renewable energy, e.g., PV-battery systems on the power system in detail, and we review different coordination schemes provided to help the Transmission System Operator (TSO) and Distribution System Operator (DSO) collaborate, and operate the system in a more reliable way.

2.1 Power System Optimal Operation

In this part, we review the literature about various studies on the power network operation management approaches. The power system's optimal operation in Microgrids' presence can be either steady-state or dynamic. Moreover, both operation modes can be simulated in deterministic or probabilistic frameworks. In this section, firstly, we go through studies on

microgrid optimal operation problem, Followed by the research question that we solved in this dissertation.

By the advent of new technologies in power systems, the regulation of active power and voltage in distribution systems led to growing interest in injecting the active power by utilising renewable energy resources into the power system [19]. Moreover, with the increase in the number of electric vehicles (EV), the distributed generators outputs need to be optimally coordinated to avoid any undesirable voltage fluctuation. As a revolution in technology, interest in Microgrids (MG) is increasingly growing. Microgrids are electrical distribution systems including loads, distributed generators (DGs) such as microturbines, wind generators, photovoltaic modules (PV) within clearly defined generations that act as a single controllable system with respect to the grid [20]. Microgrids can be operated either as islanded or as connected to the power network. Some relevant studies on the optimal Microgrid operation are provided as follows. In [21], the authors proposed an optimal load frequency control that improves a load frequency control system's efficiency in all balancing areas. In this work, the authors suggest a distributed load frequency control system that imitates economic dispatch behaviour. In [22], one neural network control approach is presented to control a Microgrid and the DGs within the Microgrid. This controller is more potent than the conventional method, which is investigated in [23]. In [24], one approach is implemented to control the load in islanded MGs. In [25], the authors tried to review comprehensive methodologies implemented on the optimal allocation of renewable energy resources.

In [26, 27] one intelligent method is used to manage the MG system to optimise operational cost. Also, one Fuzzy ARTMAP Neural Network is used to forecast hourly day-type productions based on which output generations can be predicted. The authors in [28, 29] presented one heuristic algorithm to optimise a Microgrid operation cost. This algorithm is GSA. In this control system, the load demand and generation are both foretasted by using fuzzy systems and ANN. In [30], one approach is presented to control the residential loads by predicting the market price in a real-time electricity pricing environment. In [31], the

authors proposed an IGDT method to investigate an operation strategy for CHP units in a competitive electricity market.

As a further analysis, probabilistic analyses are needed in MG problems to promote the system's performance besides optimising the operation cost. The renewable energy resources insert uncertainty into the power system, and the load demand and utility price are uncertain. Therefore, modelling uncertainties is necessary. Different methodologies can be exploited to model uncertainties. In [32], the authors tried to propose a hybrid stochastic/robust optimisation model to minimise the expected net cost, which equals operation cost minus the total benefit demand. In this article, the results with and without Microgrid bidding in the day ahead are compared. It is clearly demonstrated that if a Microgrid participated in the day-ahead market, it can anticipate the energy from the day-ahead market. In [33] an approximation of dynamic programming-based approach is proposed for economic dispatch of Microgrid comprising different distributed generators. In this work, the authors modelled the uncertainties by employing the Monte Carlo method for sampling training scenarios. Also, they used piecewise linear function approximation for their proposed strategy. This work showed that the optimal solution is very close to the optimal global solution. However, in this work, the upper hand power grid is not modelled, and the method is implemented on Microgrid. In [34], the research focuses on renewable-based AC/DC Microgrid system price and service quality. The authors tried to utilise renewable energy resources by maintaining the state of charges in batteries. The system working is categorised into four distinct operating manners to satisfy both DC and AC loads in the system without power transfer between the micro-grids. Eleven operational modes are defined with power conversation amongst the AC source and DC source grids. Although this work's results show the effectiveness of their method, this approach is more successful in isolated Microgrids. In [35], one stochastic framework is proposed for the Microgrid energy scheduling problem. The proposed approach aims to minimise the expected operation cost and power losses while considering renewable energy resources' uncertainties. The production cost for PV and Wind turbine units is assumed to be zero; however, these units are expensive, and their operation cost cannot be neglected. In [36] a model predictive control based on linear mixed-integer linear

programming is proposed and tested on an experimental Microgrid located in Athens, Greece. This method is using commercial solvers to solve the problem efficiently. However, in this work, uncertainties are not modelled. New energy management methodologies are introduced by many researchers all over the world day-to-day. One of the new methods emphasised utilising the whole energy system is proposed in [37]. The authors aimed to present an approach based on deep learning to optimise renewable energy resources utilisation and minimise power loss. They tested the results on a 100-routers test system to prove the effectiveness of their methods.

In [38], the authors address the problems of Microgrid energy management in Energy Internet (EI). Actually, EI's main benefit is autonomous systems such as solar panels and wind turbines (WT), which can be integrated into the traditional power system and supply energy to their customers. This work aims to maximise each market player's individual objective function while guaranteeing a reliable system operation and satisfying demand load. To model the uncertainties, they consider WT uncertainties by using a deep learning-based short-term forecasting algorithm for wind power by combining stacked auto-encoders (SAE), and a back-propagation genetic algorithm. The energy management problem is then modelled by a three-stage Stackelberg game to capture the dynamic interactions and interconnections among electricity users, the Microgrid, the utility company, and the energy storage company. Despite the model effectiveness, the stochastic behaviour of all sources is not modelled. In [39], the authors aimed to optimise the droop coefficients of dispatchable distributed energy resources for a Microgrid in the Energy Internet (EI) considering the volatility of renewable energy generation. Although renewable energy resources can be integrated into EI, they can not be penetrated on a large scale because of uncoordinated renewable generators into the Microgrid, especially the distribution network. This will cause a high level of volatility system disturbances. In [40], an intelligent management method is proposed to optimise a Microgrid operation cost and environmental impact. In this work, an artificial neural network is used to predict PV, WT output and load demand (LD). Also, a fuzzy logic expert system is used for battery scheduling. This work shows considerable improvement in minimising the

cost compared to previous methods based on opportunity charging and Heuristic Flowchart (HF) battery management.

In [41], the authors model one Microgrid considering the uncertainty of storage based on a 2m point estimate method. The evolutionary algorithm is also used to optimise the objective function. In [42], the authors try to consider the uncertainties of DGs, market price and load demand for one MG operation problem by considering one electro-thermal model of PV. In this research, the authors suggest a probabilistic framework to take into account the uncertainties. Also, one optimisation Back Tracking Search algorithm is used to optimise the objective operation cost. The optimisation algorithm, electro-thermal PV model and the probabilistic approach, which are implemented in [42] have better performances than the models represented in previous work [43]. In [43], the authors model the probabilistic behaviour of random parameters with one scenario-based methodology. One simple PV module is simulated as one of the Microgrid distributed generators, and an evolutionary algorithm is presented as a tool to optimise the objective function.

All in all, Microgrid contribution to the power distribution grid improves energy efficiency, power quality, and reliability. To achieve this, the optimal operation of the Microgrids is of the primary concern. The optimal Microgrid operation objectives are typically based on optimising the benefit, which can be minimising the real power loss, optimising the voltage profile or maximising the revenue (e.g. [20, 44–51]). To this end, we consider one Microgrid, which comprises different renewable energy resources to inject active power to the grid. Moreover, the operation cost is regarded as another objective goal that can be achieved by controlling the output generation of each renewable energy resource.

In our work, we aim to model the power network into a linear model to minimise cost as well as to optimise voltage deviation. We use Quadratic programming to solve the problem. We will provide the model in Chapter 3. It is worth noting that this chapter motivates what we propose in Chapter 5 as coordination management schemes between transmission and distribution power systems.

2.2 Solar Output Forecast Model

PV output forecasting may be classified into four groups based on the approach used to model solar panels and weather behaviour, namely: (i) statistical methods; (ii) Artificial Intelligence (AI); (iii) physical models; and (iv) hybrid approaches [52]. Statistical approaches are based on available historical measured meteorological and PV output data as well as numerical weather forecasts. AI methods use machine learning techniques such as Artificial Neural Networks (ANNs) to capture the non-linear relationship between weather data and solar output and construct a probabilistic model [53]. These methods may be classified in a group (i) above if their performance is judged by statistical metrics [54]. Physical models focus on numerical weather forecasts and the use of satellite images to predict weather parameters such as solar irradiation as input to a PV model to determine the solar generation output. Last, hybrid models combine the approaches mentioned above.

There are several advantages and disadvantages associated with each group of methods. For instance, in physical models where a detailed description of the panels based on the single diode model is used, weather data's stochastic nature is neglected (see, e.g., [55]). Moreover, the output is based on a PV datasheet; therefore, partial failure and downtime of a PV plant are not considered. As such, physical models usually have less accuracy in their forecasts compared to AI algorithms. Other studies pivot around statistical approaches or hybrid models that incorporate machine learning and statistical techniques. In [56] a probabilistic forecast model is proposed as a linear programming model. The authors used the Extreme Learning Machine (ELM) and quantile regression to efficiently develop a statistical approach to generate a confidence interval on the forecasted power generation. ELM was also used in [2] to forecast day-ahead solar output. In [1, 57], different distribution functions were combined to predict a 15-minute ahead probability distribution function of PV output based on a higher-order Markov chain. This method has been recently proved to improve generalisation in comparison to a standard multilayer perceptron (MLP) [58]. Although a plethora of contemporary studies have focused on ANNs and Support Vector Regression approaches in the context of forecasting [59], other machine learning techniques such as regression trees can also be used based on available historical data. According to [60], which

discusses the assessment of different forecasting techniques, most ANNs and persistence models disregard the uncertainty provoked by the random behaviour of meteorological data. On the other hand, regression techniques incorporate uncertainty and are able to build a probabilistic forecast model. For example, in [61, 62], the authors utilised a Support Vector Machine (SVM) to predict PV output based on different meteorological conditions.

In addition, physical models for PV forecasting need a large amount of accurate equipment data that are hard to obtain due to measurement and simulation errors [63]. On the other hand, AI techniques, such as SVM and ANNs, are solely based on historical statistical data for training [64]. Gaussian Process Regression (GPR) exploits the advantages of both methods in the sense that it uses both historical data and data fitting approaches to build a robust model [65]. It should be noted that mathematical modelling of the uncertainties of output as a function of uncertainties of input is outside the scope of this work (see, e.g., [66]).

Among all the approaches used to predict solar output, GPR is one of the most powerful due to its flexibility to be applied to a wide range of time-series data [67, 68]. GPR is a unique method for modelling uncertainty in a probabilistic framework setup [69]. In modelling weather forecasts, the uncertainty of input attributes is taken into account by using GPR, which treats input data as random points with an unknown distribution function. Therefore, the uncertainties are reflected in the output forecast with a specific confidence interval. GPR is based on Bayesian statistics, which help us to model and quantify uncertainty in the parameters. Moreover, the non-linear relationship between solar output and meteorological weather parameters can be explicitly modelled using an appropriate kernel function [70]. In comparison to other methods, GPR is more efficient for prediction in time-series events with a wide range of variation for each hour of a day over one year [71]. By using GPR, the uncertainty of the input data is reflected in the output forecasts since this technique assumes each input as a random variable with an unknown distribution. This is true due to the Bayesian nature of GPR. Details on the proposed method may be found in [72] Recent work has focused on using GPR [67] to forecast wind power output (see, e.g., [73–77]). For instance, the authors in [77] proposed GPR using the squared exponential kernel for wind output forecast. Their approach showed promising results. However, the focus of GPR

methods for forecasting solar output is more limited (see, e.g., [78, 79]). In [78], a weighted GPR approach is proposed so that data outliers have less importance in the forecast model. More specifically, the authors use outlier detection by integrating a weighted covariance function into GPR to identify different features and outliers that influence output forecast. In [79] a Grouped Gaussian Process (GPP) is used as a linear combination of different independent GPRs to forecast the solar output. The authors combined data from twelve Sydney sites and ten Adelaide sites in Australia. They exploited GGP with separable kernel functions as a multi-task framework for solar output forecast based on different data from different sites.

In Chapter 3, we elaborate on GPR in detail, and also we provide more information on the main contribution of our work and address the gap in the previous studies.

2.3 Transmission-Distribution Coordination Schemes

In recent years, power systems have undergone critical changes due to the penetration of renewable energy. In turn, incorporating renewable energy into power systems poses serious challenges to transmission and distribution system operators (TSOs and DSOs). The transition to a carbon-free power system is welcome; however, concerns about the quality, voltage and frequency of such systems have been raised [13]. The main objective is to use renewable energy sources (RESs) and guarantee efficient congestion management, reduction in operational costs, and increased flexibility while using local energy resources [80–82]. Working in this direction, governments have introduced incentives through policies that support the integration of RESs and encourage the collaboration and coordination of operators to maintain reliable and cost-efficient power systems [14, 17]. For instance, in [15] a hierarchical economic dispatch model was proposed to control the congestion in a power network and provide a unified bid function to network operators. In [16], the authors addressed issues about the intermittent nature of non-dispatchable resources, which requires the network operators to cooperate on new regulations, network designs, and congestion management solutions. Ancillary services are an example of the need for coordination

between TSOs and DSOs [18]. RESs can specifically provide distribution systems with ancillary services such as spinning reserves, voltage support and real-time frequency control. Currently, such services are commonly priced, and cleared in the wholesale markets. However, to fully leverage such services from these resources, it is paramount to create a new market design where new technologies such as microgrids become smoothly integrated into power systems [83, 84]. Existing centralised power market models lack appropriate mechanisms to insert more environmentally friendly resources into distributed grids.

In this dissertation, we assume that in the current practise, the TSO solves its own optimal power flow (OPF) and determines the locational marginal prices (LMPs) at the substations. Next, the DSOs dispatch distributed generation (DG) by optimising cost and considering the LMP at the substation as a fixed parameter. To facilitate the integration of RESs into power systems, the interaction between TSOs and DSOs, which are responsible for balancing the demand and supply, could be further improved (see, e.g., [85, 86]). Research has been focused on proposing methods that increase the level of coordination between TSOs and DSOs. These vary from centralised to totally decentralised methodologies. In centralised schemes, the TSO is responsible for satisfying the system demand in both the transmission and distribution systems using generators at both levels. In a more common market model, on the other hand, each operator is responsible for its own operation cost minimisation taking into account the RESs connected to each system respectively [87]. Such models are referred to as decentralised schemes where the TSO and DSO collaborate [88]. More specifically, in decentralised schemes, DSOs and TSOs need to agree on the point of common coupling (PCC) power flow interchange. The DSO operates its local system considering the TSO's bid to supply energy to the distribution system at the PCC; this is usually the LMP at the PCC. Before solving the DSO OPF, the TSO solves its own OPF representing the entire distribution system by its netload. Therefore, the DSO can operate its system with the knowledge of the supply function for the real power, i.e., the bid function, from the TSO. After the DSO solves the OPF considering the local constraints, the DSO can again participate in the TSO market and receive the payment for its energy supply sent back to the transmission system [89]. Decentralised TSO-DSO coordination approaches are categorised

as hierarchical or distributed [90]. In hierarchical TSO-DSO coordination schemes, the interaction between distributed resources in the distribution (lower level) system and the transmission (upper level) power system is like a leader-follower type, where the leader has fixed decision variables and leads the followers in making decisions [91]. In distributed TSO-DSO, all local RESs connected to the market communication graph can potentially be selected to meet the load. A detailed representation of the physical distribution system at a nodal basis as well as its market structure is necessary [92].

In Chapter 5, We will elaborate more on our proposed coordination schemes.

Chapter 3

Power System Optimal Operation

In order to solve the power system optimal operation problem, we started our research by addressing the issues associated with the integration of microgrids into the conventional power system. Microgrids are electrical distribution systems including loads, distributed generators(DGs) such as microturbines, wind generators, photovoltaic modules (PV) within clearly defined generations that act as a single controllable system with respect to the grid [20].This chapter lays the ground for the research undertaken and reported in Chapter 5.

3.1 Introduction

Integrating renewable energy resources into conventional power systems makes their operation more complex. Microgrids are gaining popularity due to their ability to satisfy local demand and sell energy to the grid. Moreover, they can inject active power into buses in power systems and improve bus voltages. Although the microgrid can supply the energy for local consumers, it is more beneficial for these systems to store energy in the storage devices and sell it to the grid in time intervals such that the energy is expensive. One way to integrate the storage devices into the microgrid is to allocate batteries in the system. The other way is coupling the resources that have complementary characteristics, such as a photovoltaic source, to a hydroelectric resource. In this case, the pumped hydro serves as storage when the solar generation exceeds the netload. For example, when the solar generation is high, the

potential power output of the other is lower, e.g., hydro generation (e.g. [93], [94]). This is true when the solar irradiance is high and there is no rainfall, thus reducing the water inflows to the hydroelectric system and the potential output of the hydroelectric resource. Moreover, when the solar generation is not sufficient to meet the net load, the pumped hydro has satisfactory ramping capabilities to quickly ramp up to meet the load. The other resources that can be integrated into MG include a wind turbine and microturbine.

This chapters aims to model the distribution power system AC power flow considering the microgrid optimal management problem to control the generators power outputs and improve the voltage of each bus. The uncertainty, intermittency and variability associated with renewable generators and the nonlinear AC load flow make the problem nonlinear. To linearize the problem, we consider the nonlinear nature of AC load flow and approximate the power system to a linear system. The objective function is thus linearized and can be solved by exploiting linear programming techniques.
Briefly, in this chapter, we linearize the load flow to make the optimization problem linear and convex. The case study is a 33-bus IEEE standard feeder. A microgrid comprising five renewable generators such as photovoltaic(PV), wind turbine (WT), hydroelectric plant, battery and microturbine (MT) is connected to this feeder. The main objectives include the cost of operation, generation, and the voltage deviation of each bus.

3.2 Methodology

In this part, the objectives are clarified, and the mathematical model is presented:

3.2.1 Microgrid modeling

Firstly the grid-connected microgrid is modelled. The microgrid is implemented on one 33 buses IEEE test feeder as a case study as shown in Fig. 3.1. The bus data for this feeder can be found in [95]. This microgrid comprises one photovoltaic (PV), wind turbine (WT), battery as storage, hydroelectric plant and microturbine (MT). The local load is met by microgrid; however, excessive generation can be sold to the grid. Also, the load is met by

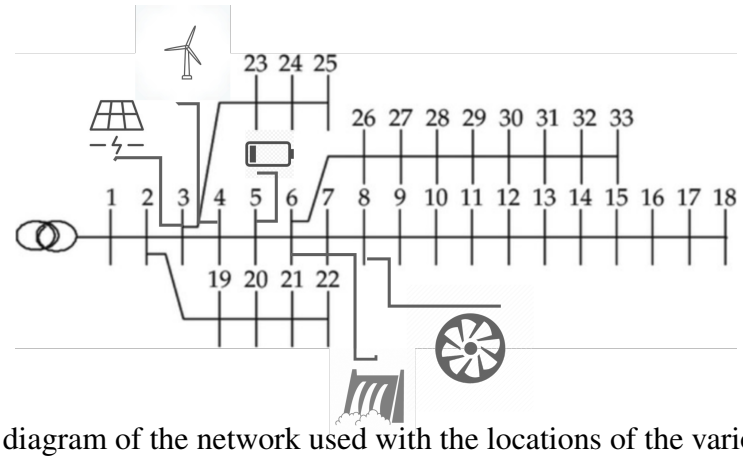

Fig. 3.1 The diagram of the network used with the locations of the various resources

the grid when local generation is not sufficient. These loads include different loads such as commercial, residential and domestic heat water load [43]. The bids information of each generators are presented in table 3.1 [43]. The tariff for different renewable technologies can be found in [96]. The values used for the load demand, market price and PV and WT power production can also be seen in figure 3.2. As demonstrated in table 3.1 the solar plant, hydro plant and WT are the most expensive units but have an environmental advantage since they don't use any fossil fuel. Thus, it is more beneficial to use them in the new power system. The other reason to use such green energies would be finishing fossil fuels shortly. Using the battery as a storage device, we can store energy to sell to the grid at an appropriate time to gain revenue.

Type	Node	Min power(KW)	Max power(KW)	Bid (€ct/kwh)
PV	3	0	30	2.584
WT	4	7	30	1.073
Battery	5	-25	25	.38
Hydro plant	6	5	30	2.294
MT	8	20	50	.457
External grid	1	-75	75	-

Table 3.1 The limit information and bid information for each generator

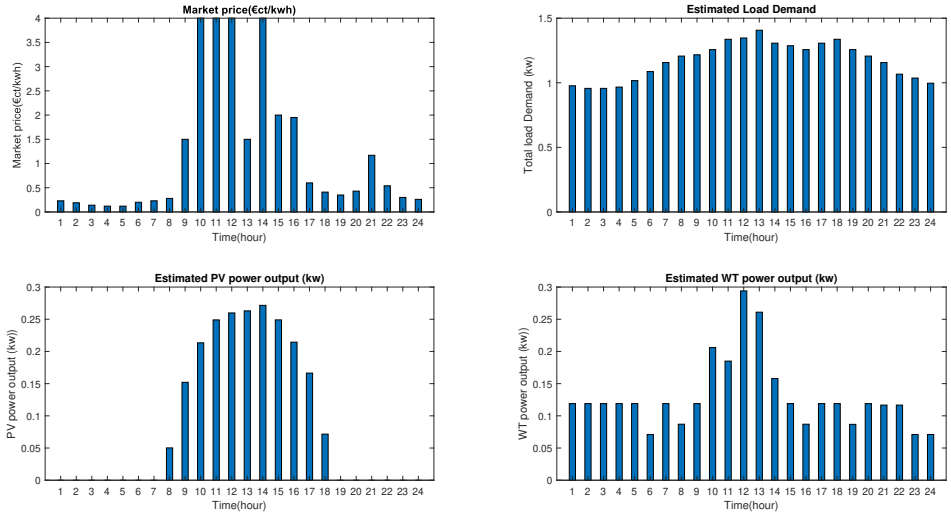

Fig. 3.2 Estimated values of the load demand, market price and PV and WT power production.

3.3 Optimal Power Flow Formulation

In this section, we formulate the linearized OPF models for transmission and distribution systems. More specifically, we formulate the augmented DC OPF for the transmission system by defining its objective and constraints. Next, we present the linearized model for the network representation of the distribution system and other constraints and determine the objective of the DSO; these are used as input to the DSO OPF.

This chapter mainly focuses on the distribution network constraints, and we assume that the transmission system constraints are satisfied. However, this chapter motivates us to address the interactions between the transmission and distribution systems in Chapter 5 as in different coordination schemes.

3.3.1 Transmission level

The AC OPF at the transmission level is a nonlinear non-convex problem as it has nonlinear equality constraints, e.g., the power balance. Using a DC formulation of the power flow, we obtain a convex problem known as the DC OPF. The objective function at the transmission DC OPF usually comprises the generators' cost. In this paper, we augment the objective function with a soft penalty function on the sum of the squared voltage angle differences, as

suggested in [97]. This augmentation has both physical and mathematical benefits. From a physical perspective, it provides a way to conduct sensitivity experiments on the size of the voltage angle differences that could be informative for estimating the size and pattern of AC-DC approximation errors. From a mathematical perspective, augmentation could help to improve the numerical stability and convergence properties of any applied solution method. The resulting augmented DCOPF is a strictly convex quadratic problem that can be solved through quadratic programming. The constraints of the OPF refer to the nodal power balance, whose dual variables are the LMPs, the line flow limits, and the generation limits.

We consider a time period of interest $\mathscr{T} = \{1,\ldots,T\}$ with time increments denoted by Δt and a power system consisting of the set of K nodes $\mathscr{K} = \{1,\ldots,K\}$, with the slack bus at node 1. We denote the set of I generators by $\mathscr{I} = \{1,\ldots,I\}$, the set of J loads by $\mathscr{J} = \{1,\ldots,J\}$, the set of generators connected to bus k by $\mathscr{I}_k$, i.e., $\mathscr{I} = \cup_{k\in\mathscr{K}}\mathscr{I}_k$; the set of loads connected to bus k by $\mathscr{J}_k$, i.e., $\mathscr{J} = \cup_{k\in\mathscr{K}}\mathscr{J}_k$; and the set of L lines by $\mathscr{L} = \{\ell_1,\ldots,\ell_L\}$. Each line is denoted by the ordered pair $\ell = (n,m)$ where n is the *from* node, and m is the *to* node with $n,m \in \mathscr{K}$, with the real power flow $f_\ell \geq 0$ whenever the flow is from n to m and $f_\ell < 0$ otherwise. We assume that each bus is connected to at least one other bus. We consider a lossless network with the diagonal branch susceptance matrix $B_d \in \mathbb{R}^{L\times L}$. Let $A \in \mathbb{R}^{L\times K}$ be the reduced branch-to-node incidence matrix for the subset of nodes $\mathscr{K}/\{1\}$ and $B \in \mathbb{R}^{K\times K}$ be the corresponding nodal susceptance matrix. We assume that the network contains no phase shifting devices and so $B^\top = B$. We denote the slack bus nodal susceptance vector by $b_1 = [b_{11},\ldots,b_{1K}]^\top$, with $b_1 + B\mathbb{1}^K = 0$, where $\mathbb{1}^K$ is the unit K-dimensional vector. We denote by P_{G_i} the power injection of generator $i \in \mathscr{I}$, by P_{L_j} the power withdrawal at load $j \in \mathscr{J}$, and by θ_k the angle at node k, as node 1 is the slack bus $\theta_1 = 0$.

The mathematical formulation of the augmented DC OPF at the transmission level at hour $t \in \mathscr{T}$ is presented as follows:

$$\begin{aligned}
\min_{P_{G_i}(t), i\in\mathscr{I}, \theta_k(t), k\in\mathscr{K}} & \sum_{i\in\mathscr{I}} c_i(t) + \pi \sum_{\ell=(m,n)\in\mathscr{L}} (\theta_n(t) - \theta_m(t))^2 \\
\text{subject to } & \sum_{i\in\mathscr{I}_k} P_{G_i}(t) - \sum_{\ell\in\mathscr{L}} B_{d_\ell} A\theta(t) = \sum_{j\in J_k} P_{L_j}(t), k \in \mathscr{K}, \longleftrightarrow \quad \lambda_k(t), \\
& f^m \le f(t) = B_d A\theta(t) \le f^M, \\
& P_G^m \le P_G(t) \le P_G^M, \qquad (3.1)
\end{aligned}$$

where B_{d_ℓ} is the ℓ^{th} row of the B_d matrix; f^M and f^m are the values of the maximum real power flow allowed through the lines in $\mathscr{L}$ in the same direction and in the opposite direction of line ℓ, respectively; π is a coefficient for the generator's angles as part of our objective. and P_G^m (P_G^M) is the vector of lower (upper) generation limits. Usually, the cost of generator $i \in \mathscr{I}$ is a quadratic function in the form of $c_i(t) = \alpha_i P_{G_i}(t) + \beta_i P_{G_i}^2(t) + \gamma_i$. The LMPs are the dual variables of the nodal power balance denoted by $\lambda(t) = [\lambda_1(t), \ldots, \lambda_K(t)]^\top$.

3.3.2 Distribution Level

We assume a radial distribution feeder with a set of N buses denoted by $\mathscr{N}$ and a set of $N-1$ lines denoted by $\mathscr{L}'$. Bus 1 denotes the PCC with the TSO and is considered to be the slack bus. For each bus, i, V_i stands for the bus voltage magnitude, while p_i and q_i represent the injected active and reactive power, respectively. For each line segment in $\mathscr{L}'$ that connects bus i to bus j, r_{ij} and x_{ij} stand for its resistance and reactance, respectively, and P_{ij} and Q_{ij} for the real and reactive power from bus i to j respectively. In addition, the set $\mathscr{N}_j \subset \mathscr{N}$ denotes bus j's neighbouring buses, which are further downstream. The linear equations that

model the distribution feeder for each line (i, j) are as follows (see, e.g., [8]):

$$P_{ij} - \sum_{k\in\mathscr{N}_j} P_{jk} = -p_i + r_{ij}\frac{P_{ij}^2+Q_{ij}^2}{V_i^2}, \tag{3.2}$$

$$Q_{ij} - \sum_{k\in\mathscr{N}_j} Q_{jk} = -q_i + r_{ij}\frac{P_{ij}^2+Q_{ij}^2}{V_i^2}, \tag{3.3}$$

$$V_i^2 - V_j^2 = 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) - (r_{ij}^2 + x_{ij}^2)\frac{P_{ij}^2+Q_{ij}^2}{V_i^2}. \tag{3.4}$$

The nonlinear part in the equations above, i.e., $\frac{P_{ij}^2+Q_{ij}^2}{V_i^2}$, corresponds to the power losses in the system, which are assumed to be zero in our work. Also, under relatively flat voltage profile, i.e., $V_i \approx 1$, we have $V_i^2 - V_j^2 \approx 2(V_i - V_j)$. Thus, we have:

$$M_0^\top \left[V_1 V^\top\right]^\top = m_0 + M^\top V = D_r P + D_x Q, \tag{3.5}$$

where $M^0 \in \mathbb{R}^{N\times(N-1)}$. More specifically, its l^{th} column corresponds to one line segment $(i, j) \in \mathscr{L}'$, the entries of which are all zero except for the i^{th} and j^{th} ones, where $M_{il}^0 = 1$ and $M_{jl}^0 = -1$ when $j \in N_i$, i.e., bus i is closer to the feeder head. m_0^T corresponds to the first row of M^0 and denotes the slack bus while the rest of the matrix is shown by M with the size of $(N-1)\times(N-1)$ [98]. We assume $V_1 = 1$ and define the vectors $[V_i : \forall i \in \{\mathscr{N} / 1\}]$, $P = [P_{ij} : \forall (i, j) \in \mathscr{L}']$, $Q = [Q_{ij} : \forall (i, j) \in \mathscr{L}']$. We define D_r and D_x as $(N-1)\times(N-1)$ diagonal matrices with the l^{th} column and row entry that corresponds to one line segment $(i, j) \in \mathscr{L}'$ equal to r_{ij} and x_{ij} respectively. Thus, (3.2)-(3.4) can be written in the form of matrices as:

$$-MP = -p, \tag{3.6}$$

$$-MQ = -q, \tag{3.7}$$

$$V = Rp + Xq - {M^{-1}}^\top m_0, \tag{3.8}$$

with $p = [p_i : \forall i \in \{\mathscr{N} / 1\}]$, $q = [q_i : \forall i \in \{\mathscr{N} / 1\}]$, $R = {M^{-1}}^\top D_r M^{-1}$ and $X = {M^{-1}}^\top D_x M^{-1}$. As can be seen in (3.8), the relationship between the voltage and real power is now linear.

Let us assume a set of D distribution systems denoted by $\mathscr{D} = \{1, \ldots, D\}$ connected to the transmission system. For each $d \in \mathscr{D}$ we know the PCC, which is denoted by k_d. The OPF at

each distribution system $d \in \mathscr{D}$ has a goal to minimise the cost of electricity purchased from the transmission system, the cost of distributed resources and the voltage deviation from the reference value. The cost of electricity at the substation for the time period $\mathscr{T}$ is a function of the LMP at the PCC at time t denoted by $\lambda_{k_d}(t)$ and the amount of power purchased from the transmission system at time t, i.e., $P^d_{\text{grid}}(t)$, and is defined as follows:

$$\sum_{t\in\mathscr{T}} \left(\lambda_{k_d}(t) P^d_{\text{grid}}(t)\Delta t\right). \tag{3.9}$$

We denote by $\mathscr{N}^d_{PV}$ the set of PVs connected to distribution system d. The cost of PV generation resource is formulated as:

$$\sum_{t\in\mathscr{T}} \sum_{i\in\mathscr{N}^d_{PV}} B_{PV_i} P_{PV_i}(t)\Delta t, \tag{3.10}$$

We denote by $\mathscr{N}^d_{DG}$ the set of DGs connected to distribution system d. The cost of DGs are formulated as:

$$\sum_{t\in\mathscr{T}} \sum_{i\in\mathscr{N}^d_{DG}} B_{DG_i} P_{DG_i}(t)\Delta t, \tag{3.11}$$

where B_{PV_i} is the cost of PV generation at node i, and $B_{DG_i}$is the cost of DG generation at node i. We denote by $\mathscr{N}^d_B$ the set of battery systems connected to the distribution system d. The cost of battery systems is equal to:

$$\sum_{t\in\mathscr{T}} \sum_{i\in\mathscr{N}^d_B} B_{B_i}(P^{\text{ch}}_{B_i}(t) + P^{\text{dis}}_{B_i}(t))\Delta t, \tag{3.12}$$

where B_{B_i} is the cost of the battery system at node i. We denote by $P^{\text{ch}}_{B_i}(t)$ the charging power of the battery system at node i at time t and by $P^{\text{dis}}_{B_i}$ the discharging power of the battery system at node i at time t. The voltage deviation from some reference value is defined as follows:

$$\sum_{i\in\mathscr{N}} \sum_{t\in\mathscr{T}} \alpha(V_i(t) - V_{\text{ref}})^2, \tag{3.13}$$

where α is the voltage regulation cost and V_{ref} is the voltage reference value. The constraints of the distribution system OPF include the maximum and minimum limits for the decision variables:

$$P_{PV,i}^{\min} \leq P_{PV_i}(t) \leq P_{PV,i}^{\max}, i \in \mathscr{N}_{PV}, t \in \mathscr{T}, \tag{3.14}$$

$$P_{DG,i}^{\min} \leq P_{DG_i}(t) \leq P_{DG,i}^{\max}, i \in \mathscr{N}_{DG}, t \in \mathscr{T}, \tag{3.15}$$

$$P_{B,i}^{\text{ch,min}} \leq P_{B_i}^{\text{ch}}(t) \leq P_{B,i}^{\text{ch,max}}, i \in \mathscr{N}_B, t \in \mathscr{T}, \tag{3.16}$$

$$P_{B,i}^{\text{dis,min}} \leq P_{B_i}^{\text{dis}}(t) \leq P_{B,i}^{\text{dis,max}}, i \in \mathscr{N}_B, t \in \mathscr{T}, \tag{3.17}$$

$$V_i^{\min} \leq V_i(t) \leq V_i^{\max}, i \in \mathscr{N}, t \in \mathscr{T}, \tag{3.18}$$

$$P_{\text{grid}}^{d,\min} \leq P_{\text{grid}}^{d}(t) \leq \textstyle\sum_{i \in \mathscr{I}_k} P_{G_i}(t), t \in \mathscr{T}, \tag{3.19}$$

where $P_{\text{grid}}^{d,\min}$ is defined by the interchange flow limit between the distribution system d and the transmission system. We model the battery system i as follows (see, e.g., [99])

$$E_{\min,i} \leq \sum_{t \in \mathscr{T}} \left(\eta_{\text{ch},i} P_{B_i}^{\text{ch}}(t) - \frac{1}{\eta_{\text{dis},i}} P_{B_i}^{\text{dis}}(t) \right) \Delta t + E_{0,i} \leq E_{\max,i}, \forall i \in \mathscr{N}_B, \tag{3.20}$$

where $E_{0,i}$ is the initial value of the energy stored, $E_{\max,i}$ and $E_{\min,i}$ are the maximum and minimum energy that can be stored in the battery. The network constraints from (3.6)-(3.8) for every time step $t \in \mathscr{T}$ are defined as follows:

$$V(t) = R p_i(t) + X q_i(t) - {M^{-1}}^{\top} m_0, \tag{3.21}$$

$$p_i(t) = P_{PV_i}(t) + P_{DG_i}(t) + P_{B_i}^{\text{dis}}(t) - P_{B_i}^{\text{ch}}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_{PV} \cap \mathscr{N}_B, \tag{3.22}$$

$$p_i(t) = P_{PV_i}(t) + P_{DG_i}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_{PV} \setminus \mathscr{N}_B, \tag{3.23}$$

$$p_i(t) = P_{B_i}^{\text{dis}}(t) - P_{B_i}^{\text{ch}}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_B \setminus \mathscr{N}_{PV}, \tag{3.24}$$

$$p_i(t) = -P_{\text{load}_i}(t), \forall i \in \mathscr{N} \setminus \mathscr{N}_{PV} \cap \mathscr{N}_B, \tag{3.25}$$

$$q_i(t) = -Q_{\text{load}_i}(t), \forall i \in \mathscr{N}, \tag{3.26}$$

where $P_{\text{load}_i}(t)$ is the real load at bus i at time t and $Q_{\text{load}_i}(t)$ is the reactive load at bus i at time t.

The OPF at the distribution system $d \in \mathscr{D}$ is formulated as follows:

$$\min_{P_{PV_i}(t),P^{\text{ch}}_{B_i}(t),P^{\text{dis}}_{B_i}(t),V_i(t),P^{d}_{\text{grid}}(t)} (3.9)+(3.11)+(3.12)+(3.13)$$
$$\text{subject to } (3.15)-(3.26). \tag{3.27}$$

3.4 Quadratic Programming

After we linearise the problem, we can solve the problem using quadratic programming [100] as follows:

$$\begin{aligned} \min_x f(x) &= \frac{1}{2} X^T H\, X + c^T X \\ \text{subject to } A_{in}\, X &\leq B_{in}, \\ A_{eq}\, X &= B_{eq}, \end{aligned} \tag{3.28}$$

where $f(x)$ in our problem is the objective function of the distribution OPF, as described in Equation 3.27. X^T denotes the vector transpose of all decision variables. $A_{in}\, X \leq B_{in}$ and $A_{in}\, X \leq B_{in}$ represent all inequality and equality constraints respectively. H stands for the Hessian matrix of the objective function, wherein our problem represents a coefficient for the voltage regulation cost. c denotes the generation cost in our problem. When we optimize the cost to a reference point x_0, H is set to its Hessian matrix $H(f(x_0))$ and c is set to its gradient $\nabla f(x_0)$.

3.5 Results

In this part, the results presented are based on the case study of one 33 bus IEEE standard feeder which has been defined in previous parts. The cost before the optimization is calculated using the estimated value in Fig. 3.2. After we run the optimization, the cost is the summation of generation cost and voltage regulation as explained in Section 3.3.2. In Figure 3.3 the cost before and after optimization is shown. As can be seen, the cost after optimization has decreased. However, this cost is calculated for $H = 1$ including the utility cost, the generation cost, and the value of voltage optimization (sum of square of the voltage deviation).

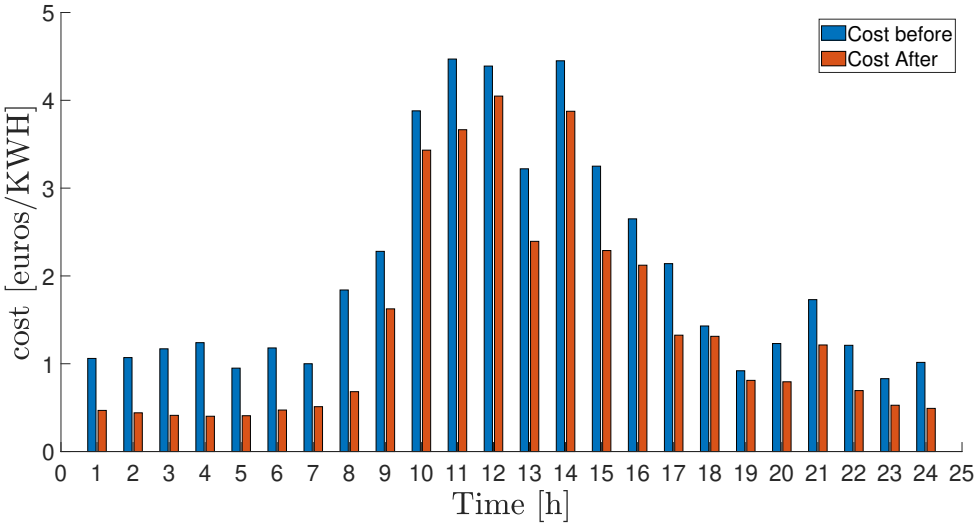

Fig. 3.3 Cost before and after optimization

In Figure 3.4, the cost is divided into different parts and as seen the generation cost is the most significant part of the objective function which is beneficial for the operators because the utility cost decreased as can be seen in Figure 3.4.

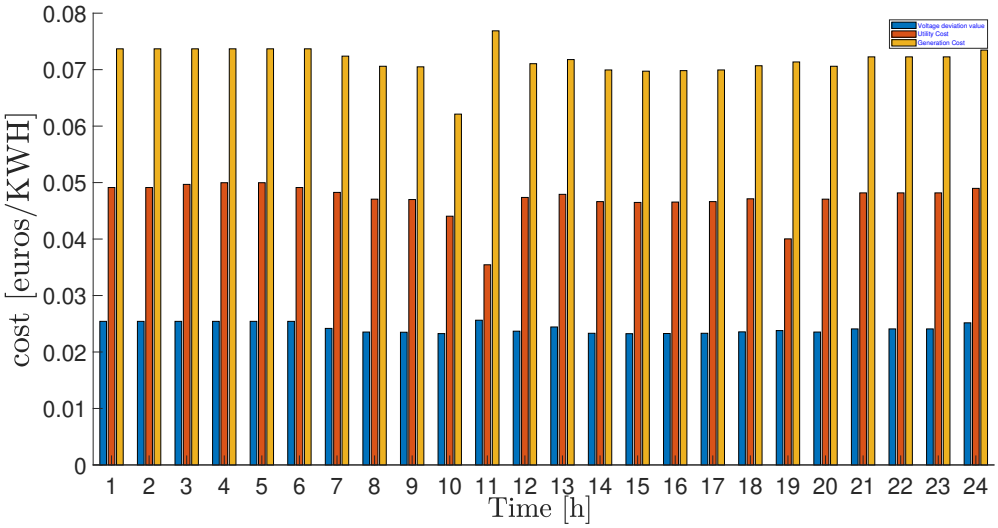

Fig. 3.4 Optimal cost divided into different cost parts

In table 3.2 the cost value before and after optimization is demonstrated. As seen, the cost is significantly reduced after the optimization. More specifically, buying energy from renewable energy resources is more beneficial for the grid than supplying energy as standalone system to all customers.

Value	Total cost before (€ct/kwh)	Total cost after(€ct/kwh)
1	1.06	0.47
2	1.07	0.44
3	1.17	0.41
4	1.24	0.40
5	0.95	0.41
6	1.18	0.47
7	1.00	0.51
8	1.84	0.68
9	2.28	1.62
10	3.88	3.43
11	4.47	3.67
12	4.39	4.05
13	3.22	2.39
14	4.45	3.88
315	3.25	2.29
16	2.65	2.12
17	2.14	1.33
18	1.43	1.31
19	0.92	0.81
20	1.23	0.79
21	1.73	1.21
22	1.21	0.69
23	0.83	0.53
24	1.02	0.49

Table 3.2 Total cost before and after optimization

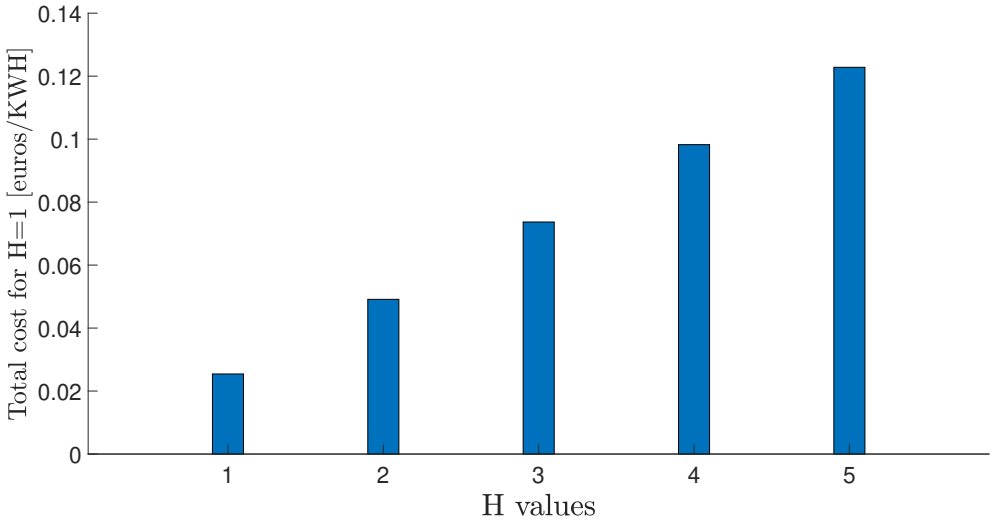

Fig. 3.5 The total cost value for H=1,2,3,4,5 for the first hour

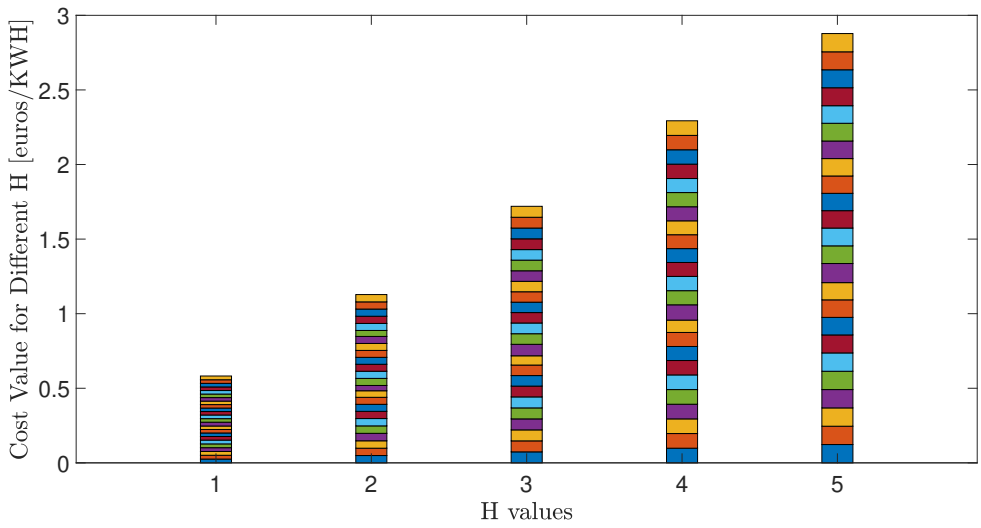

Fig. 3.6 The total cost value for H=1,2,3,4,5 for 24 hour

The output generation of different generators is presented in Figure. 3.7. It can be seen that the PV, WT, and hydroelectric plants are the most expensive units but have the benefits of not requiring fuel as well as being green energy. As expected, the summation of output generation and utility production meets the load. The values of the generators' output are presented in table3.3. As can be seen in the table in the intervals that the utility bid is low, the battery starts to charge and in the time intervals that the utility cost is more, the battery

starts to discharge. The hydro plant is producing the minimum amount of energy since they are expensive units. The PV just produce energy in the intervals that the irradiance of the sun is present, however as seen in table 3.3, the PV and the battery can act as complementary to each other as well.

	Grid	PV	WT	Battery	Hydro Plant	MT
1	59.42	0.00	7.00	4.18	5.00	22.07
2	57.42	0.00	7.00	4.18	5.00	22.07
3	57.42	0.00	7.00	4.18	5.00	22.07
4	58.42	0.00	7.00	4.18	5.00	22.07
5	63.42	0.00	7.00	4.18	5.00	22.07
6	70.42	0.00	7.00	4.18	5.00	22.07
7	75.00	0.00	7.00	8.48	5.00	20.19
8	75.00	0.00	23.63	-2.96	5.00	20.00
9	75.00	0.00	27.13	-5.46	5.00	20.00
10	72.23	0.00	29.83	-1.67	5.00	20.00
11	76.47	0.00	33.14	4.94	5.00	20.00
12	75.00	21.72	30.00	-17.05	5.00	20.00
13	75.19	30.33	30.34	-19.35	5.00	20.00
14	75.00	14.58	30.00	-13.92	5.00	20.00
15	75.00	11.02	30.00	-12.35	5.00	20.00
16	75.00	5.67	30.00	-10.00	5.00	20.00
17	75.00	14.58	30.00	-13.92	5.00	20.00
18	75.00	19.93	30.00	-16.27	5.00	20.00
19	75.47	0.00	30.90	-3.93	5.00	20.00
20	75.00	0.00	23.63	-2.96	5.00	20.00
21	74.76	0.00	7.00	8.91	5.00	20.00
22	65.76	0.00	7.00	8.91	5.00	20.00
23	62.76	0.00	7.00	8.91	5.00	20.00
24	61.00	0.00	7.00	4.93	5.00	21.74

Table 3.3 Optimal output generation

The voltage deviation in this work is between .95pu to 1.05pu which is the standard range in the distribution power system. However, the maximum deviation after and before optimization is .07pu which is in the range of acceptable voltage deviation, e.g., %10. The mean value of deviation is .03pu. These results can be seen in Figure 3.8 and Figure 3.9 respectively.

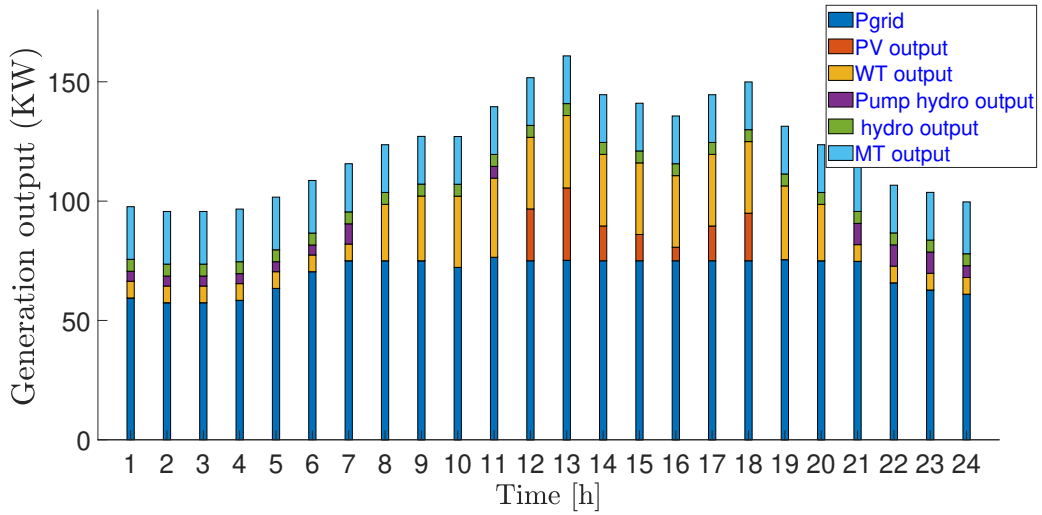

Fig. 3.7 Output generation for H=1 for 24 hours

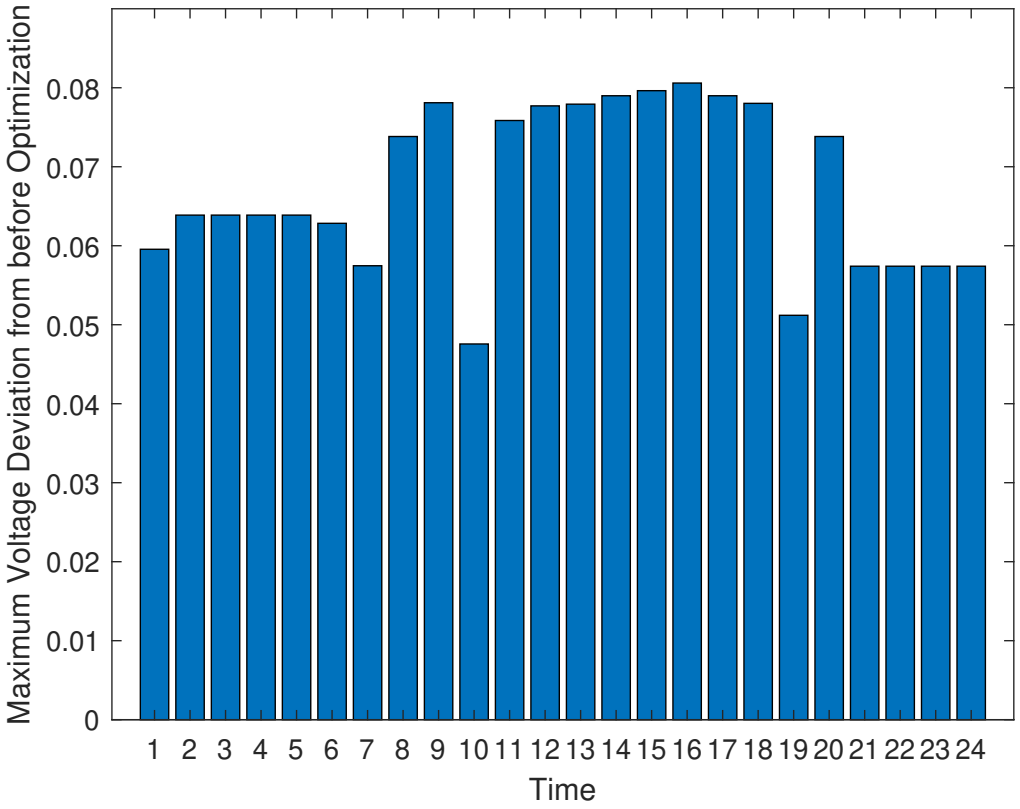

Fig. 3.8 Maximum Voltage deviation from the reference voltage for 24 hours

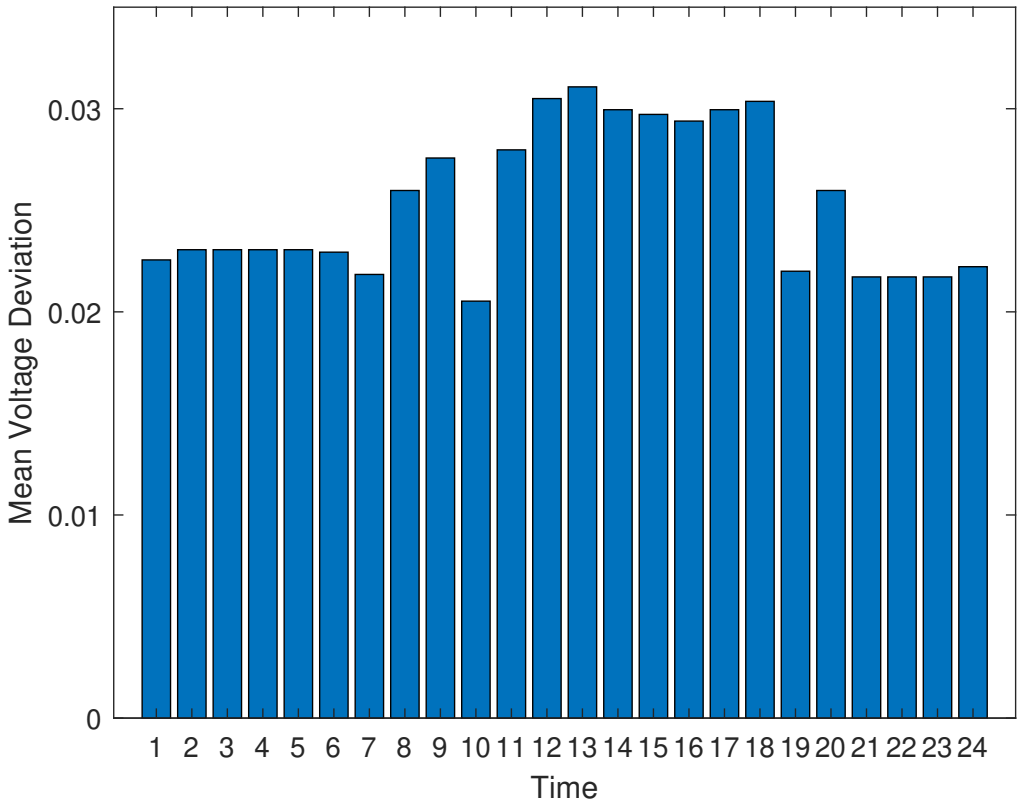

Fig. 3.9 Average Voltage deviation from the reference voltage for 24 hours

3.6 Conclusion

This chapter solved the first problem we aimed to solve as an optimal microgrid operation to address the first research question. The power system is moving fast towards a more carbon-free power energy system. However, the high penetration of renewable energy resources into the power system forces the conventional power to adapt to these new changes resulting in a smarter power system. To maintain the power system stability, frequency, voltage and optimize the operation cost, in the chapter, we proposed a deterministic linear microgrid operation framework to optimize the operation cost and maintain the busses' voltages at a nominal range. The microgrid included different resources such as photovoltaic(PV), wind turbine (WT), hydroelectric plant, battery and microturbine (MT). The microgrid is modelled in a 33 buses IEEE standard feeder. The results illustrated improvement in the cost before

and after we ran the optimization algorithm. Also, to quantify the influence of parameter H in the optimization, we have run a sensitivity analysis as by increasing the value of this parameter, the objective function increased.

Chapter 4

Solar Output Forecast

After addressing the issues with the operation of distribution power network, we realised the importance of an effective forecast model for the renewable energy output. Thus, we move into the second research question in this chapter by providing solutions on an enhanced forecast model for the solar output.

4.1 Introduction

Based on the relevant literature on solar output forecast models, despite the appealing features of GPR methods, there are still significant challenges in the implementation of GPR to ensure satisfactory performance and avoid ad-hoc implementations. The new smart grid paradigm involves the availability of numerous input parameters and introduces the idea of working with high dimensional datasets. However, without an effective input feature selection mechanism, such data-driven models become computationally expensive. Also, the models become more prone to overfitting and their prediction accuracy decreases with the increase of redundant features [101, 102]. Data clustering helps identify and group datapoints with similar patterns and distributions together. This way, we transform the unstructured data into more homogeneous clusters and empower the machine learning algorithm to analyze and extract important information much more efficiently [103–105]. In this regard, we enhance GPR performance by integrating feature selection and pre-clustering techniques;

and selecting the appropriate kernel function. Our choice of kernel Matérn 5/2 has shown promising results for wind power output forecast and for fault location prediction in power systems [106, 107]. Using Matérn 5/2 as a kernel function increases the accuracy of forecast due to the capability of this kernel in solving stochastic problems [108]. The process of training and testing models are depicted in Figs. 4.1, 4.2. As shown in Fig. 4.1, we build a dataset which comprises of solar output and meteorological weather data. We perform a correlation study to identify the features that have a high impact on solar output in order to improve the accuracy of the forecast and reduce computational complexity. The features selected are: direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, temperature, zenith, and azimuth. Next, we partition the data into four groups based on time and solar output using k-means clustering. Each hour of the day is considered to belong in one specific cluster. We use GPR to relate solar output generation with the selected features and train each cluster using Matérn 5/2 as a kernel function for the forecasting model. We apply it to different datasets from different sites, i.e., Denver, New York, Dallas, San Francisco, and St. Lucia, to validate the proposed framework. We build a dataset combining meteorological weather data and solar output for four geographical locations which is publicly available to foster research in the area[1]. We validate the results by utilising both k-fold cross validation and holding-out data techniques.

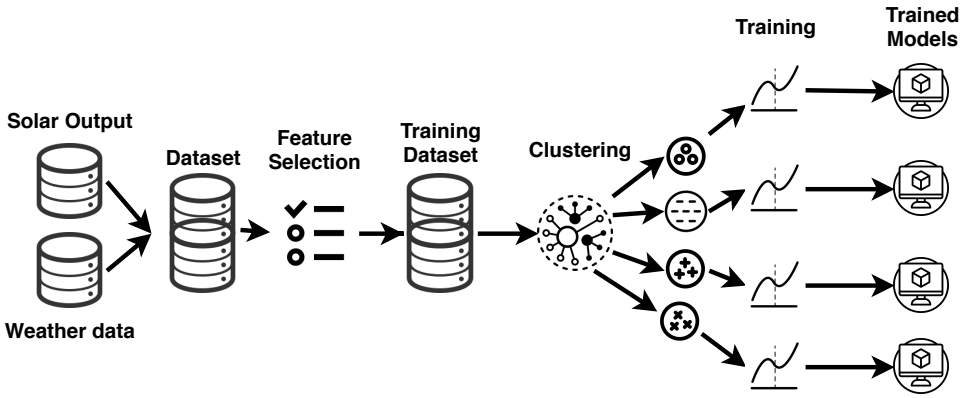

Fig. 4.1 The proposed process of building trained models.

[1]https://figshare.com/s/3b114c8fa4574dc0dd69

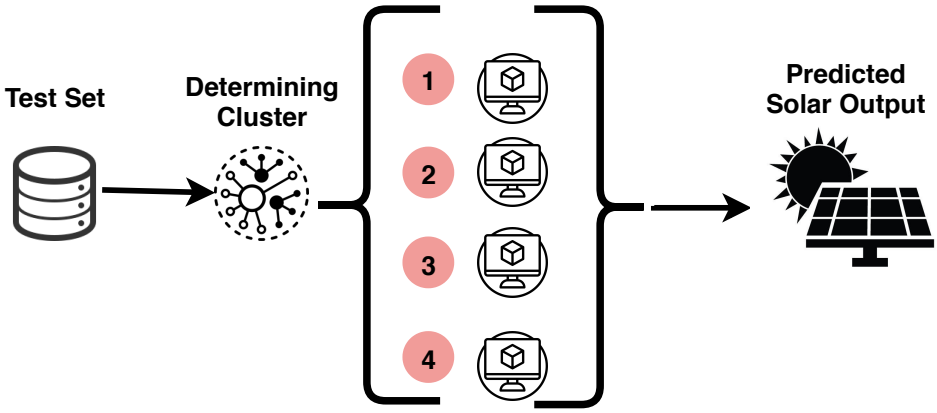

Fig. 4.2 The process of testing our trained models to predict solar output.

We choose 30 random days from the dataset as hold-out test set to represent different weather conditions. To test the model, we first use the clustering algorithm to determine the cluster each test data point belongs to. Then, we use the corresponding trained GPR model to predict the solar output for each test data point, as shown in Fig. 4.2. The numerical results' section shows that the average forecast error follows a normal distribution, and with 95% confidence level, it takes values between -1.6% to 1.4%. The proposed framework decreases the normalised root mean square error (RMSE) and mean absolute error (MAE) by 54.6% and 55.5%, respectively, when compared with other relevant works. It should be noted that clustering and feature selection were implemented only on the training dataset, i.e., excluding the testing dataset, to avoid data leakage. The main contribution of this chapter is summarised as the combination of all following steps, resulting in one coherent novel forecast model: i) the identification of highly correlated features with solar output through a preprocessing stage which makes the model simpler, demanding less data while prediction accuracy is still high; ii) development of clustering techniques based on time of day and categorisation of data into four clusters that improves the accuracy of the GPR model beyond state of the art techniques; and (iii) choice of using Matérn 5/2 as kernel function which has not been exploited for solar output prediction previously in spite of its capability in solving probabilistic problems.

The remainder of this chapter is organised as follows. In Section 4.2 the data processing is described. More specifically, in Section 4.2.1 the correlation study to identify the features

that have a high impact on solar output is presented and in Section 4.2.2 the clustering of the dataset is described. In Section 4.3 the proposed framework is developed. In particular, in Section 4.3.1 the GPR with Matérn 5/2 as a kernel function that relates the solar output with the input features is discussed and in Section 4.3.2 the framework validation methodologies are presented. In Section 4.4, we illustrate the proposed methodology through five different datasets. In Section 4.5, we summarise the results and make some concluding remarks.

4.2 Data Processing

In this section, we present the processing that needs to be performed to the data in order to formulate the proposed framework. In particular, we describe the feature selection and clustering methodologies.

4.2.1 Features' Selection

Before clustering the data into four clusters, we carry out a correlation study to identify the highly related features to the historical output power. In fact, studies have found that meteorological data such as temperature and solar irradiance are the main features which affect the solar output [109], however there are still other features which affect the solar output significantly. We choose to perform feature selection to make the proposed model less sensitive to changes and not prone to overfitting. If a very simple model is developed, it has a high bias and is likely to fail to predict complex relationships. As such, by selecting the appropriate number of features a balance between complexity and bias is achieved [110, 111]. Another reason to select fewer features is that less data needs to be collected to apply the framework.

There are different feature selection approaches, such as filter (e.g., Pearson correlation, Spearman correlation), wrapper (e.g., cross-validation) and embedded techniques [112]. Although, the Pearson correlation coefficient might not be the most robust one, it is nevertheless one of the standard coefficients as shown in the recent studies (see, e.g., [113–117]). The main benefit of using Pearson coefficient is its time-effectiveness and small compu-

tational complexity [118]. To this end, we calculate Pearson and Spearman correlation coefficients of solar generation and meteorological features of different datasets from Denver, New York, Dallas, and San Francisco that can be found in National Solar Radiation (NSR), Iowa Environmental Mesonet (IEM), and National Renewable Energy Laboratory (NREL) databases.

Let us assume, we have a collection of data over a period of T hours. We use the Pearson correlation, which linearly measures the relation between solar output and each feature, defined as a correlation coefficient. The Pearson correlation coefficient between vectors $a \in \mathbb{R}^T$ and $b \in \mathbb{R}^T$ is calculated by the following formulation [119]:

$$\rho(a,b) = \frac{1}{T-1} \sum_{t=1}^{T} \frac{a_t - \mu_a}{\sigma_a} \cdot \frac{b_t - \mu_b}{\sigma_b}, \tag{4.1}$$

where a_t (b_t) is the value of a (b) at time t, μ_a and σ_a (μ_b and σ_b) are respectively the mean and the standard deviation of a (b). The Spearman coefficient is equivalent to calculating the Pearson correlation coefficient on the ranked, i.e., ordered data and used to determines if there is a nonlinear relationship between the selected features and solar generation [120].

We assume that we have data for different features for T time intervals denoted as $X_i \in \mathbb{R}^T$ where $i = 1,\ldots,M$ is the index of each feature we perform the correlation study on and $Y \in \mathbb{R}^T$ is the time-series solar output. More specifically, meteorological weather data refer to direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, temperature, sky cloud covering, zenith (angle between sun and zenith), azimuth (angle between sun and the North), albedo, visibility, ozone (cm), humidity, and wind speed (knots) i.e., $M = 12$. We calculate the correlation coefficients $\rho(X_i, Y)$ for $i = 1,\ldots,M$ to determine which features affect in a greater extent the solar output.

In Table 4.1, the correlation coefficient values for all attributes are presented. As seen in this table, we obtain the similar results by using either Pearson or Spearman coefficients. The value of correlation coefficients for sky cloud covering, albedo, visibility, ozone, humidity and wind speed are very small in comparison to other features, therefore the aforementioned features are not included in the selected feature set.

Feature	Pearson coeff	Spearman coeff
Horizontal solar irradiance	0.81	0.86
Zenith	0.81	0.86
Direct solar irradiance	0.71	0.77
Diffused solar irradiance	0.64	0.81
Azimuth	0.49	0.52
Temperature	0.27	0.28
Albedo	0.07	0.09
Humidity	0.06	0.05
Sky cloud covering	0.03	0.02
Visibility	0.02	0.02
Ozone	0.01	0.01
Wind speed	0.003	0.003

Table 4.1 The absolute value of correlation coefficients between solar output and different features.

It is worthy to note that although albedo is a good measure to show how well a surface is reflecting the light (for instance, green grass and white sand snow have albedo values of 0.24 and 0.67 respectively [121]), representation of albedo as a constant value produces unsatisfactory and unreliable results for radiation exchange [122]. Considering that our datasets manifest a constant value of albedo and taking into account the small correlation value, we do not include this feature in our dataset. Moreover, cloud covering increases the interpretability of the model, i.e., the effect of the amount of clouds in the sky is more understandable for a human observer than other measurements, e.g., zenith. However, as shown in Table 4.1 cloud covering has a very small correlation with solar output generation. It has been shown in [123] that the effect of cloud covering on solar panels, without the knowledge of thickness, density and type of clouds, cannot be quantified. As the correlation coefficient also shows, one can ignore this feature and expect the model's accuracy not to deteriorate.

To sum up, based on the analysis above and the criterion we define using Table 4.2 [124], we eliminate the features with small and no correlation from the dataset. In our model, temperature represents the medium correlation, and direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, zenith, and azimuth represent the strong correlation

value. Thus, six features were identified as the parameters which affect more prominently PV output generation, namely, direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, temperature, zenith, and azimuth. The number of features whose relationship with solar generation was originally studied was twelve. The accuracy of the forecasts was better and the computational complexity was reduced in the case of six selected features compared to twelve. To estimate computation time, on average a run with all twelve features takes 17.9 minutes. With our selected six features it takes 10.3 minutes in a Windows machine which is equipped with AMD® FX-9830P RADEON R7 CPU with 4 Cores at 3.00 GHz and 16 GB of RAM. The lead time is hourly. However, if weather data for the selected features are available for a greater amount of time, the one hour lead time can be extended. Moreover, in terms of forecast errors introduced with training models with six compared to twelve feature data we compare the effect in the accuracy against the increase in the number of features. More specifically, we select solar data from Denver International Airport PV, i.e., Site A (see Table 4.3 for more details) and calculated the following error metrics when six and twelve features were selected respectively. Let us denote by $y_\star^{(t)}$, the forecasted value for solar generation at time t, and by $\tilde{y}^{(t)}$ the actual value at time t; the error metrics are calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{T_\star}\sum_{t=1}^{T_\star}\left(\tilde{y}^{(t)} - y_\star^{(t)}\right)^2}, \tag{4.2}$$

$$\text{MAE} = \frac{1}{T_\star}\sum_{t=1}^{T_\star}\left|\tilde{y}^{(t)} - y_\star^{(t)}\right|, \tag{4.3}$$

where $T_\star$ is the number of hourly intervals we are predicting the solar output. We also compute normalised values of the above metrics. Such metrics are widely-used to test the performance of both deterministic and probabilistic forecast models [59, 125–129]. In terms

Strength	Minimum value	Maximum value
No correlation	0	0.09
Small correlation	0.10	0.24
Medium correlation	0.25	0.40
Strong correlation	0.41	1.00

Table 4.2 Correlation value thresholds

of error for Site A the RMSE and MAE are equal to 1.48 and 0.69 MW, respectively, when all twelve features are taken into account and 1.23 and 0.56 MW when six features were used. Thus, we can see that selecting a large number of features can lead to worse results than a small number due to overfitting issues.

4.2.2 Data Clustering

Prior to clustering, we perform a correlation study to identify the features that highly impact solar output. Several papers that are dedicated to the development of solar forecasting techniques focus on the seasonality trends or weather types rather than on the time of day [1, 78]. The latter affects the solar output to a great extent, as shown in Fig. 4.3, where hours in one season belong in different clusters. The figure shows that solar output in one season does not necessarily follow the same pattern.

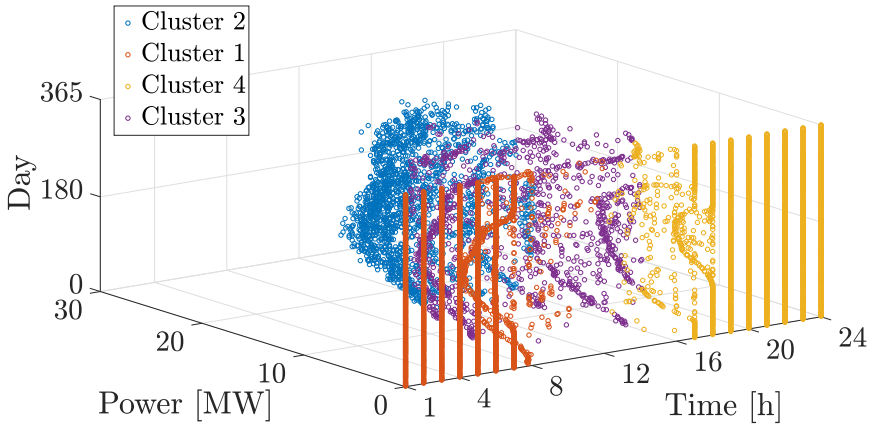

Fig. 4.3 3D graph of four clusters. Different colours represent different clusters.

Clustering is an unsupervised pattern classification learning technique used to partition data with high similarity into different groups based on a distance or dissimilarity function [130–132]. The key concepts and different clustering algorithms are discussed in [133]. k-means is a very popular clustering algorithm which is used to cluster data into different groups while each point belongs to a cluster with the least Euclidean distance to the centroid [134, 135].

In previous studies, k-means is not employed to cluster output solar energy based on time [136], while in our proposal the dataset is clustered based on output and time. PV output has a huge amount of scattering across both the time of day and the day of year. Also, another advantage of using k-means clustering with solar output and time of day is that no data need to be excluded, e.g., night hours or seasonal effects, making our model flexible for all geographical locations on earth. k-means aims to partition the data into K categories in a way that the sum of squares from points to the assigned cluster centres is minimised. In each cluster, all cluster centres are at the mean of the data points which belong to the corresponding cluster. Consider a set $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ with N elements, where $x_i \in \mathbb{R}^n$ for all of $i = 1, \ldots, N$; the data point cluster number $C(i) \in \{1, \ldots, K\}$, $i \in \{1, \ldots, N\}$; the cluster centroid for cluster k $c_k \in \mathbb{R}^n$, $k = 1, \ldots, K$; and the Euclidean distance $d(x_i, c_k) = ||x_i - c_k||$, which is the distance between x_i and cluster centroid c_k. Then k-means clustering tries to minimise the following squared error function:

$$\underset{\{c_k\}_{k=1}^{K}}{\text{minimize}} \sum_{k=1}^{K} N_k \sum_{C(i)=k} d^2(x_i, c_k), \tag{4.4}$$

where N_k is the number of points assigned to cluster k.

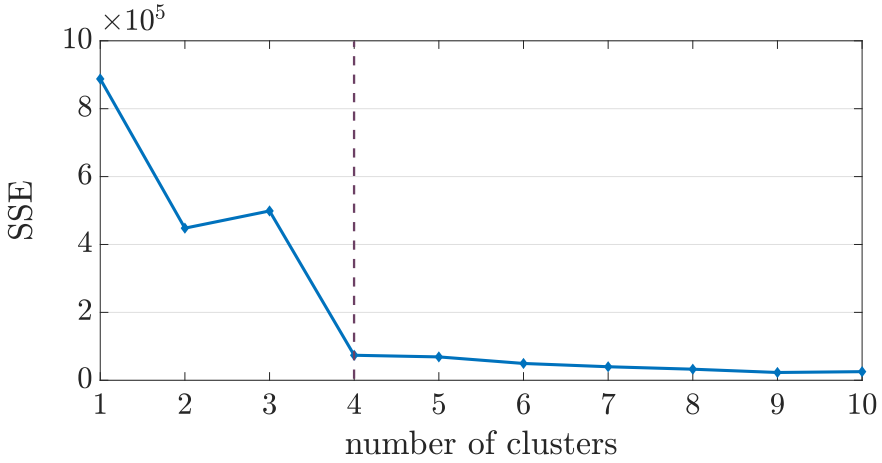

Fig. 4.4 Elbow statistic test.

To determine the number of clusters we use the Elbow and Gap statistic methods. The Elbow technique uses the sum of squared errors (SSE), which is the sum of the distances

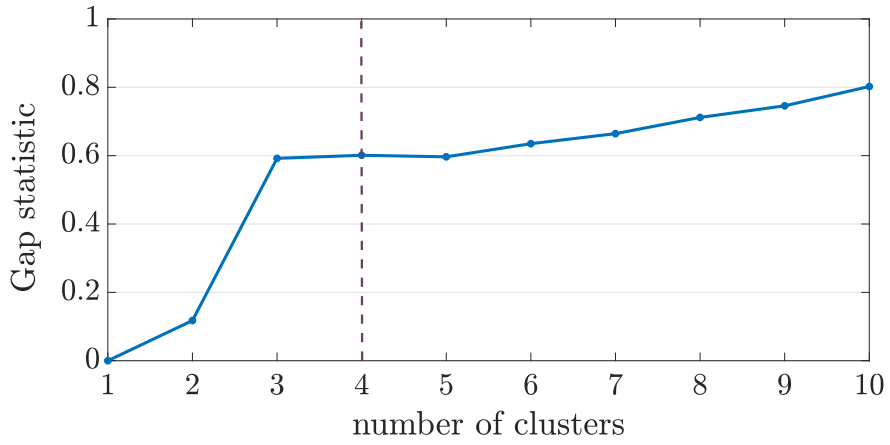

Fig. 4.5 Gap statistic test.

between the sample points in each cluster and the centroid of the cluster as a performance indicator for a set number of clusters [137]. More specifically, the SSE is calculated over a series number of clusters. If small SSE values are obtained then that is an indication that each cluster is more convergent. When the number of clusters is set to approach the optimal number of clusters K, SSE shows a rapid decline. When the number of clusters exceeds K, SSE continues to decline but with a slower rate. Usually the optimal number of clusters K is obtained graphically at the point that looks like an "elbow", i.e., at the largest inflection point down. Once K is determined, if the selected number of clusters is less than K, the SSE will be greatly reduced for every 1 increase of the number of clusters. On the other hand, when the selected number of clusters is greater than K the change of the SSE will not be so obvious for every 1 increase of the selected number of clusters. The basic idea of Gap Statistic is to introduce reference datasets, which are generated with independent Monte Carlo simulations sampling from an empirical distribution and to calculate the sum of the squares of the Euclidean distance between two measurements in each cluster. To describe the Gap methodology we define the summation of all pairwise Euclidean distances for all datapoints in cluster k to be $D_k = \sum_{i,i' \in C_k} d(x_i, x_i')$ and the normalized sum of intra-cluster

distances to be $W_k = \sum_{k=1}^{K} \frac{1}{2N_k} D_i$. Then, we use the following function to measure the Gap value [138]:

$$\text{Gap}_n(k) = \mathbb{E}_n\left[\log(W_k)\right] - \log(W_k), \tag{4.5}$$

where $\mathbb{E}_n[\cdot]$ denotes the expectation operator under a sample of size n from the empirical distribution of the data. The optimal number of clusters based on the Gap statistic is the smallest number k that satisfies the following expression:

$$\text{Gap}_n(k) \geq \text{Gap}_n(k+1) - s_{k+1}, \tag{4.6}$$

where $s_k = \sqrt{1+1/B}s_d(k)$ is the simulation error that is calculated using the standard deviation $s_d(k)$ of B Monte Carlo replicates, in this study $B = 500$, drawn from the empirical distribution.

In our framework the outcome of both Elbow and Gap methods for the optimal number of clusters is four (see Figs. 4.4 and 4.5). We use R version 4.0.2 with factoextra (v. 1.0.7) and NbClust (v. 3.0) libraries to analyse Elbow and Gap statistics. It means that if the selected number of clusters is less than four, the cost value (i.e., the total within sum of square) will be greatly reduced for every one increase in the number of clusters. On the other hand, when the selected number of clusters is greater than four the change of the cost value will not be so obvious for every one increase of the selected number of clusters. By using the Gap method we determined that the value of four clusters corresponds to the minimum value of k such that $\text{Gap}_k \geq \text{Gap}_{k+1} - S_{k+1}$. Specifically, the Gap value for the optimal k as the optimal number of clusters is within one standard deviation away from the Gap value at $k+1$.

Site	Location	Size [MW]	Latitude [°]	Longitude [°]
A	Denver Intl Airport	30	39.8561 N	104.6737 W
B	John F. Kennedy IntlAirport	30	40.6413 N	73.7781 W
C	Dallas Executive Airport	35	32.6807 N	96.8672 W
D	San Francisco Intl Airport	30	37.6213 N	122.3790 W
E	St Lucia	0.433	27.498 S	153.013 E

Table 4.3 Site description.

To further study the optimal number of clusters, we perform a sensitivity study and compare the increase in the accuracy against the increase in the number of clusters. More specifically, we select solar data from Denver International Airport PV, i.e., Site A (see Table 4.3 for more details) and cluster the data into one to eight clusters. For each of the clusters we train a GPR model, as discussed in Section 4.3, and depict the error between the forecasted and the actual values in Fig. 4.6. We use the RMSE and MAE error metrics as defined in Section 4.2.1 along with

$$\text{MSE} = \frac{1}{T_\star} \sum_{t=1}^{T_\star} \left(\tilde{y}^{(t)} - y_\star^{(t)} \right)^2. \tag{4.7}$$

These metrics compare how accurate the prediction of PV is with respect to the number of clusters. Each bar shows the average prediction error for each experiment with a different number of clusters. A choice of a large number of clusters increases the computational complexity of the forecasting algorithm since one GPR model needs to be trained for each distinct cluster. From the graph, we can see that there is a big decrease in all error metrics when the number of clusters is four. However, after increasing the number of clusters from four to eight, we notice that there is a marginal decrease in the error metrics. Taking into account Elbow and Gap results along with Fig. 4.6, where the sensitivity study on the number of clusters by comparing different normalised error metrics values is presented, the optimal number of clusters is set to four. This number provides a trade-off between two different objectives of minimum forecast error and minimum number of clusters. Thus, the data are grouped into four clusters, which are depicted in Fig. 4.3. Clusters 2 and 3 represent early morning and night times. Clusters 1 and 4 represent seasonal variations. More information about the clusters' characteristics may be found in Table 4.4.

When applying the k-means algorithm on the datasets, one should bear in mind that the algorithm (i) is sensitive to the outliers and the number of clusters; and works better with spherical data [139, 140]. As such, we test the presence of outliers in our dataset using R's Box plot (see, e.g., [141]). No outliers are detected, as shown in the Fig. 4.7 where no data points are observed outside the whiskers of the box plot. The y-axis is showing the output

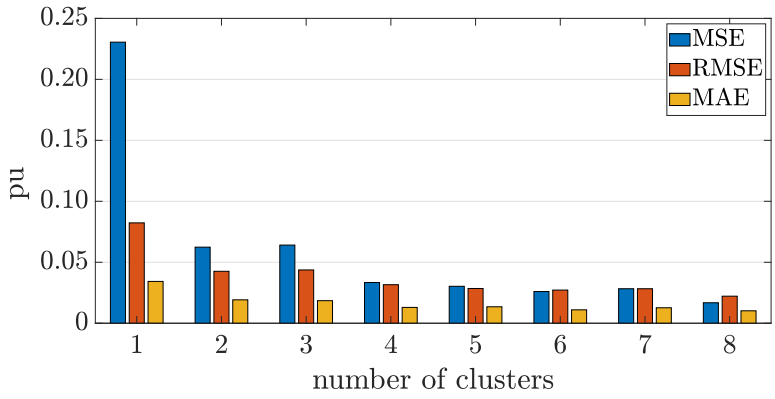

Fig. 4.6 Sensitivity study on the number of clusters by comparing different normalised error metrics values.

	Time			PV output [MW]		
Cluster number	min	centroid	max	min	centroid	max
Cluster 1	8	11.25	15	14.1	18.78	25.70
Cluster 2	0	3.37	11	0.0	0.35	5.40
Cluster 3	12	19.40	23	0.0	0.24	5.30
Cluster 4	7	11.43	16	1.8	9.35	14.10

Table 4.4 Clusters' details

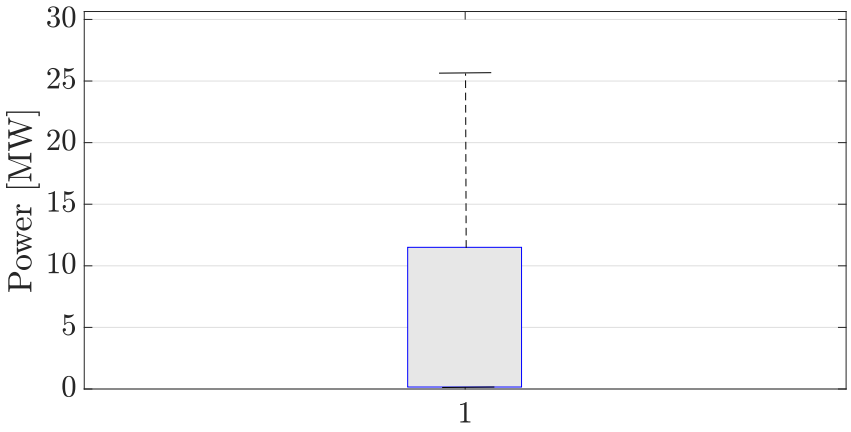

Fig. 4.7 Box plot of dataset that shows non existence of outliers.

power, which is between zero and 25.7 MW. In order to determine the optimal number of clusters, we use the Elbow and Gap statistic methods [138]. Although k-means works better on globular clusters [142], in the 3D graph of Fig. 4.3 we can see that k-means distinguished

four clusters. We have also applied the k-medoids method on our datasets, which is more robust to noise and outliers than k-means [143]. The results on the RMSE and MAE on the forecast errors for Site A (see Table 4.7 for more details) when using k-medoids instead of k-means are 1.20 and 0.55 MW, respectively. These values are very close to the ones obtained with k-means clustering, which are 1.23 and 0.56 MW, respectively.

4.2.3 Clustering Sensitivity Implementation

To demonstrate the effect of different numbers of clusters in the performance of the proposed forecasting methodology, we present the results for one to eight clusters. Wee use the input weather datafor each cluster to train our GPR model and determine the nonlinear relationship between the solar output and the weather data. To validate our methodology, we use two datasets based on two locations at Denver and St. Lucia, and the 5-fold cross-validation and holding-out data techniques. More specifically, to take into account different days in different seasons, we choose 30 random days as a hold-out test dataset.

To show how the solar power forecast is affected by the number of clusters, we apply the proposed framework on Site A for one to eight clusters. More specifically, in the case of one cluster, we only train one GPR model for the entire dataset; in the case of two clusters, we train two GPR models, one for each cluster; and so on until we have eight clusters and eight GPR models. We use the 5-fold cross-validation and hold out validation techniques to obtain the forecast errors and be able to analyse the clustering effect on the accuracy of the solar power forecasting.

For Site A, the available historical data comprise of hourly input weather data: diffused solar irradiance, horizontal solar irradiance, direct solar irradiance, temperature, zenith, and azimuth from 2006, i.e., we have 6×8760 data points for weather input data and the solar generation output. We implement the proposed framework in one to eight number of clusters and select 30 random days as hold-out data as representative of different days of the year during different seasons. Each cluster is trained by using Matérn 5/2 GPR and tested by 5-fold cross-validation and hold-out techniques. The error metrics used are defined in equations 4.2 and 4.7.

no. of clusters	RMSE [MW]	MAE [MW]	RMSE [%]	MAE [%]
8	0.90	0.34	2.72	1.02
7	0.80	0.26	2.43	0.78
6	1.02	0.40	3.10	1.20
5	1.18	0.49	3.58	1.48
4	1.29	0.36	3.91	1.08
3	1.53	0.47	4.63	1.43
2	1.64	0.66	4.98	2.00
1	2.94	0.58	8.91	1.77

Table 4.5 Training set error metrics for various number of clusters.

The results of the forecast error metrics for the training and test sets for one to eight number of clusters are given in Tables 4.5, 4.6. The error metrics of the training data between the actual and the predicted values are based on the average error of all 5 folds for the training set. It should be noted that the test results are expected to be different from the training set results, since 30 hold-out days are not shown to the model during the training process. However, the results with any test set should be approximately the same as those obtained with the training set, as it may be seen in Tables 4.5, 4.6. We notice that the error metrics are usually improved as we increase the number of clusters. However, at the same time a choice of a large number of clusters increases the computational complexity of the model since for each cluster, we build a GPR model. The number of clusters needs to balance the

no. of clusters	RMSE [MW]	MAE [MW]	RMSE [%]	MAE [%]
8	0.80	0.38	2.41	1.14
7	1.01	0.48	3.05	1.45
6	0.95	0.41	2.87	1.25
5	1.00	0.47	3.02	1.43
4	1.08	0.50	3.26	1.52
3	1.46	0.65	4.43	1.97
2	1.44	0.68	4.36	2.05
1	2.75	1.16	8.35	3.53

Table 4.6 Test set error metrics for various number of clusters.

trade-off between two different objectives of minimum forecast error and minimum number of clusters due to the computational complexity.

In this regard, we further study the effect of the number of clusters in the forecast error and depict in Figs. 4.8, 4.13, 4.10 the forecasts for the training set along with the actual values. As seen in these figures, the different patterns of solar generation are better captured and modelled in the case of eight clusters. However, partitioning the data into four clusters also leads to good results in comparison to eight based on the results, we can see in Fig. 4.6, where the sensitivity on the number of clusters to different normalised error metrics values is depicted. As such, we partition the data into four clusters and as seen in Fig. 4.3, clusters two and three, represent the seasonal variations while clusters one and four represent early morning and night times.In Fig. 4.8, we trained only one cluster by using Matérn 5/2 GPR. This figure shows different patterns of solar output which are all combined together.

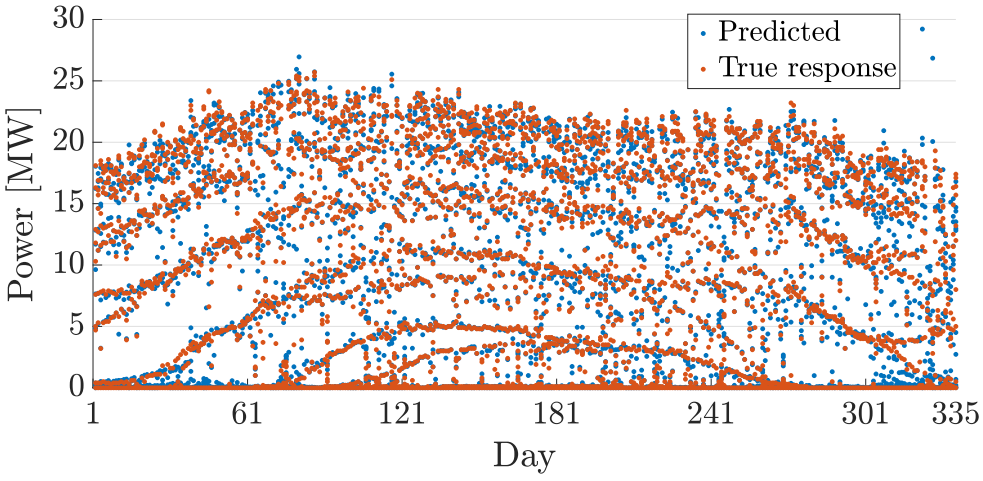

Fig. 4.8 Proposed framework predictions of the training data set for one cluster.

The sensitivity study demonstrated the improved framework performance, when four clusters are chosen in terms of balancing model complexity and accuracy.

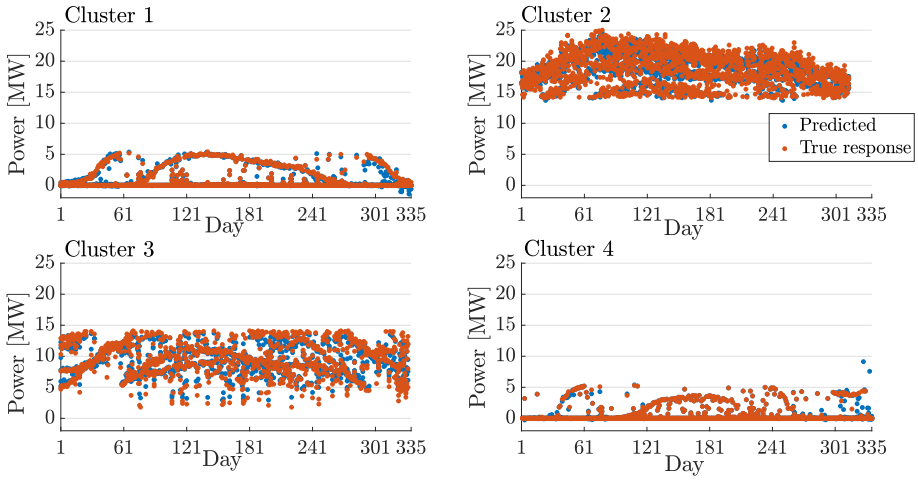

Fig. 4.9 Proposed framework predictions of the training data set for four clusters.

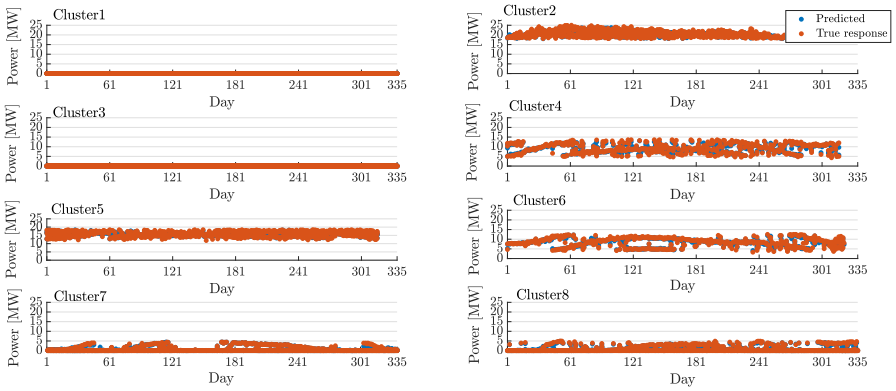

Fig. 4.10 Proposed framework predictions of the training data set for eight clusters.

4.3 Proposed Forecasting Framework

In this section, the stochastic framework for the short-term forecast of PV output is presented. More specifically, the formulation of the GPR is described and the validation methodologies are discussed.

4.3.1 Gaussian Process Regression

In this dissertation, a model is trained for each cluster using a GPR model, which is a supervised learning technique. We aim to learn a mapping function that relates the input feature set data to the output data in supervised learning. In fact, GPR is a kernel-based

nonlinear nonparametric regression technique in which the covariance function plays a crucial role in defining the relation between input data and the responses.

Let the training set $\mathscr{S} = \{(x^{(t)}, y^{(t)})\}_{t=1}^{T}$ be a set of i.i.d. samples from some unknown distribution, where T is the period of available data with one hour resolution; q stands for the number of selected features, i.e., $q = 6$; $x^{(t)} \in \mathbb{R}^q$ is the vector containing all selected features at time t; and $y^{(t)} \in \mathbb{R}$ the solar output at observation t. With the use of a Gaussian model we may relate the input with the output terms by:

$$y^{(t)} = f(x^{(t)}) + h(x^{(t)})^{\top}\beta + \varepsilon^{(t)}, \text{ for } t = 1, \ldots, T, \tag{4.8}$$

where $\varepsilon^{(t)}$ are i.i.d. "noise" variables with independent $\mathscr{N}(0, \sigma^2)$ distributions, $f(x^{(t)})$ is the mapping function $\mathbb{R}^q \to \mathbb{R}$ and $h(x^{(t)})$ is a set of a fixed basis function. The explicit use of basis functions is a way to specify a non-zero mean over $f(x^{(t)})$. In this work we assume that $h(x^{(t)})$ is a $q \times 1$ vector whose all entries are equal to the constant value of one, and β is the basis function coefficient $q \times 1$ vector and is evaluated by maximising a likelihood function as described below. For notational convenience, we define:

$$X = \begin{bmatrix} (x^{(1)}) \\ \vdots \\ (x^{(T)}) \end{bmatrix} \in \mathbb{R}^{T \times q}, y = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(T)} \end{bmatrix} \in \mathbb{R}^{T}, \varepsilon = \begin{bmatrix} \varepsilon^{(1)} \\ \vdots \\ \varepsilon^{(T)} \end{bmatrix} \in \mathbb{R}^{T},$$

$$f = \begin{bmatrix} f(x^{(1)}) \\ \vdots \\ f(x^{(T)}) \end{bmatrix} \in \mathbb{R}^{T}, H = \left[h(x^{(1)}), \ldots, h(x^{(T)}) \right] = \mathbb{1}_{q \times T},$$

where $\mathbb{1}_{q \times T}$ is a q by T matrix whose all elements are one. In matrix form we may rewrite (4.8) as

$$y = f(X) + H^{\top}\beta + \varepsilon. \tag{4.9}$$

We assume a prior distribution over functions $f(X)$ as

$$f(X) \sim \mathcal{N}(0, K(X,X)), \tag{4.10}$$

where 0 is the mean value; $K(X,X)$ is the covariance matrix:

$$K(X,X) = \begin{bmatrix} k(x^{(1)},x^{(1)}) & \dots & k(x^{(1)},x^{(T)}) \\ \vdots & \ddots & \vdots \\ k(x^{(T)},x^{(1)}) & \dots & k(x^{(T)},x^{(T)}) \end{bmatrix},$$

where $k(\cdot,\cdot)$ is the covariance or kernel function. By using the kernel function we aim to actively model the unknown relationship between the input and the output variables. The kernel function is defined based on the likely pattern that we can observe in the data. One assumption to model the kernel may be that the correlation between any two points in our input set, i.e., $x^{(t)}, x^{(t')} \in \mathcal{S}$, with $t, t' = 1, \dots, T, t \neq t'$, decreases with increasing the euclidean distance between them. This means that points with similar features behave similarly. Under this assumption, in this work we use the Matérn 5/2 as a kernel function, which is parameterised as follows:

$$\begin{aligned} k(x^{(t)},x^{(t')}) \quad &= \sigma_f^2 \left(1 + \frac{\sqrt{5}d(x^{(t)},x^{(t')})}{\sigma_l} + \frac{5d^2(x^{(t)},x^{(t')})}{3\sigma_l^2}\right) \\ &\quad e^{-\frac{\sqrt{5}d(x^{(t)},x^{(t')})}{\sigma_l}}, \end{aligned} \tag{4.11}$$

where $d(x^{(t)},x^{(t')})$ is the euclidean distance between any two input observations $x^{(t)}, x^{(t')} \in \mathcal{S}$ as defined in Section 4.2.2; σ_l and σ_f, are two other kernel parameters which show respectively the characteristic length scale and the signal standard deviation that both belong in $\mathbb{R}^q$. The characteristic length scale σ_l defines how far the response variable $y^{(t)}$ needs to be away from the predictor $x^{(t)}$ to become uncorrelated. These two parameters are greater than zero and are formulated as follows:

$$\sigma_l = 10^{\theta_l}, \sigma_f = 10^{\theta_f}. \tag{4.12}$$

We now define a new parameter θ to be:

$$\theta = \begin{bmatrix} \theta_l \\ \theta_f \end{bmatrix} = \begin{bmatrix} \log(\sigma_l) \\ \log(\sigma_f) \end{bmatrix} \in \mathbb{R}^{q\times 2}. \tag{4.13}$$

From (4.9) we may write that

$$y|f(X),X \sim \mathcal{N}(H^\top \beta, \sigma^2 I + K(X,X)), \tag{4.14}$$

since both $f(X)$ and ε have zero means. In order to determine the distribution that y follows, we need to determine three parameters, i.e., β, σ^2 and θ. $K(X,X)$ is a function of θ as may be seen in (4.11)-(4.13). β, σ^2, and θ are also known as the hyperparameters of the kernel function. In order to estimate the parameters we maximise the following marginal log-likelihood function

$$\log P(y|f(X),X) = \log P(y|X,\beta,\theta,\sigma^2). \tag{4.15}$$

Thus, the estimates of β, θ, and σ^2 denoted by $\hat{\beta}$, $\hat{\theta}$ and $\hat{\sigma}^2$ are given by

$$\hat{\beta},\hat{\theta},\hat{\sigma}^2 = \underset{\beta,\theta,\sigma^2}{\operatorname{argmax}} \log P(y|X,\beta,\theta,\sigma^2). \tag{4.16}$$

We may write from (4.14) and (4.15) that

$$P(y|X) = P(y|X,\beta,\theta,\sigma^2) = \mathcal{N}(H^T\beta, K(X,X)+\sigma^2 I). \tag{4.17}$$

Thus, the marginal log-likelihood function is

$$\begin{aligned} \log P(y|X,\beta,\theta,\sigma^2) = -\frac{1}{2}(y-H^\top\beta)^T[K(X,X)+\sigma^2 I]^{-1} \\ (y-H^\top\beta) - \frac{1}{2}\log 2\pi - \frac{1}{2}\log|K(X,X)+\sigma^2 I|. \end{aligned} \tag{4.18}$$

We concentrate the likelihood function for the subset of parameters, σ^2 and θ, by expressing β as a function of the parameters of interest and replacing them in the likelihood function. Thus, we have that the estimate of β for given θ and σ^2 is:

$$\begin{aligned}\hat{\beta}(\theta,\sigma^2) = [H^\top [K(X,X|\theta)+\sigma^2 I]^{-1} H]^{-1} \\ H^\top [K(X,X|\theta)+\sigma^2 I]^{-1} y.\end{aligned} \tag{4.19}$$

By substituting (4.19) in (4.18) we have

$$\begin{aligned}\log P(y|X,\hat{\beta}(\theta,\sigma^2),\theta,\sigma^2) = -\frac{1}{2}(y-H\hat{\beta}(\theta,\sigma^2))^T \\ [K(X,X|\theta)+\sigma^2 I]^{-1}(y-H\hat{\beta}(\theta,\sigma^2)) \\ -\frac{1}{2}\log 2\pi - \frac{1}{2}\log|K(X,X|\theta)+\sigma^2 I|.\end{aligned} \tag{4.20}$$

We now determine the hyperparameters as the output of the above optimisation problem.

Once the hyperparameters are evaluated, we may use (4.14) to predict the output of solar generation based on the input parameters. More specifically, $\{x_*^{(t)}\}_{t=1}^{T_*}$ be a set of i.i.d. input points of the features drawn from the same unknown distribution; we will plug these values in (4.14) and the unknown $\{y_*^{(t)}\}_{t=1}^{T_*}$ can be calculated as the predicted solar output value for the time period T_*. More details on GPR model may be found in [66, 144].

After training our model and estimating the kernel parameters for each of the four clusters, we can use the proposed framework for solar generation forecasting. As also noted in [1, 2], training the models is time-consuming. Therefore, it will not be cost-efficient to build a model each time we intend to predict the solar output. As such, it is more efficient to train the model based on historical data and then predict the solar output online. The overall forecast block-scheme in Fig 4.12 illustrates offline and online processes. The training process (pictured on the left) is performed offline, while the trained GPRs are used in the online process (depicted on the right) to forecast solar output.

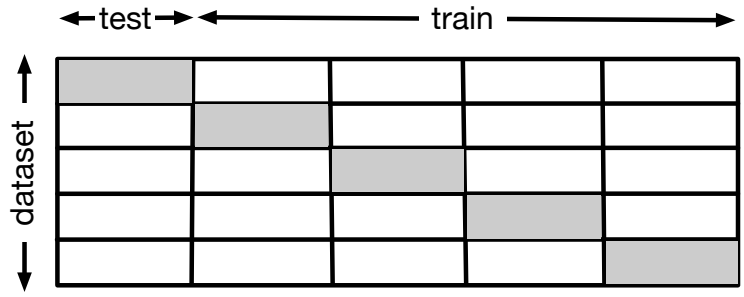

Fig. 4.11 Diagram of k-fold cross-validation with k=5; the grey boxes refer to testing the data and the white to training.

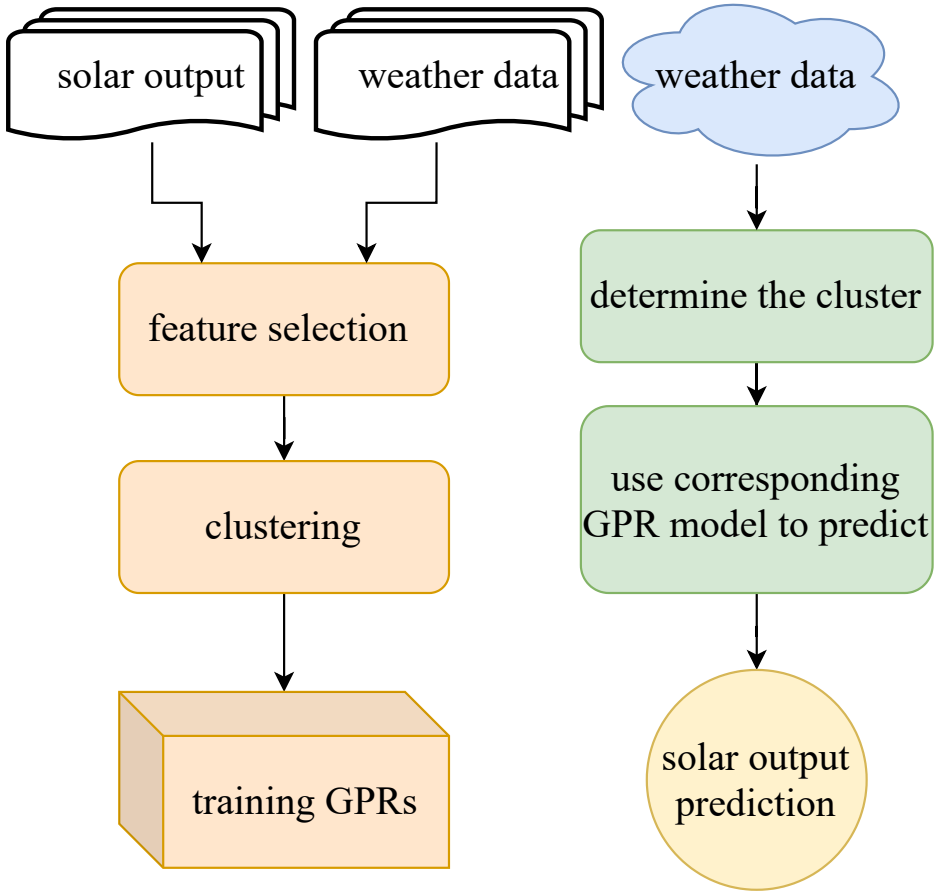

Fig. 4.12 The overall forecast system.

4.3.2 Framework Validation

To test the accuracy of the proposed method for solar output forecasting, different tests and validation methods are exploited. K-fold cross-validation and hold out validation are the

most prevalent validation methods used in recent studies [145]. As depicted in Fig. 4.11, in k-fold cross-validation, the whole data set is split into k folds: at each time, k-1 folds are used as a training set and a one-fold as a testing set, until all folds used to build the forecast model, typical values for k range between 3 to 10 [146]. In addition, hold-out is used to avoid overfitting [147]. In this work both methods are used for test and validation.

In our implementation, 30 days of a year are randomly selected as hold-out data, while the remaining data are used for training and testing using 5-fold cross-validation.

4.4 Numerical results

In this section, we provide the proposed framework results at five sites and its comparisons with existing forecasting methodologies from the literature.

In Section 4.4.1, the five sites' information is given; in Section 4.4.2 detailed results and analysis of site A are given so that the reader better understands the proposed framework. In Section 4.4.3, summarised results for all sites are provided as well as comparisons with other methods to prove the efficiency of the proposed framework.

4.4.1 Dataset Information

Different datasets from five sites are used to test the model's efficiency based on available historical data from National Solar Radiation, Iowa Environmental Mesonet (IEM) and National Renewable Energy Laboratory. The five sites' details are given in Table 4.3. The data for sites A-D are lagged hourly data for the year 2006 from National Solar Radiation, Iowa Environmental Mesonet (IEM) and National Renewable Energy Laboratory. The data for site E is also lagged observed data for years 2012 and 2013. For site E, the temporal resolution is one-minute data for the year 2012, which are used as training set and year 2013, which are used as test set; however, to implement our model on site E and compare it with state of the art, we used hourly resolution data.

4.4.2 Framework Implementation on Site A

The model is tested on the Denver International Airport PV plant, i.e., site A. The dataset comprises of hourly attributes' values from 2006, i.e., of 8760 data points for each feature and the solar generation output. The experiments are run 30 independent times for each cluster; 30 random days are selected as hold-out data representing different days of the year during different seasons. As described in Sections 4.2 and 4.3, the training set is partitioned into four clusters, and all clusters are trained using GPR Matérn 5/2 independently and validated using 5-fold cross-validation and hold-out methods. It is also worth mentioning that in our model, the lead time is hourly. However, if weather data for the selected features are available for a greater amount of time, the one hour lead time can be extended. For example, if we have an accurate weather prediction, e.g., for the next 36 hours [148], then we can forecast the hourly solar output for the next 36 hours.

We first train the GPR model with the available hourly dataset of 335 (365-30=335) days. In Fig. 4.13 the box plots of the predicted and actual solar output for the training data set for the four different clusters are depicted. We use 5-fold cross-validation as a test and validation method for our training set, which comprises hourly data points of 335 days. Since clusters are partitioned based on similarities between the points, as shown, each cluster follows specific patterns with specific ranges which prove the similarity of the data in them.

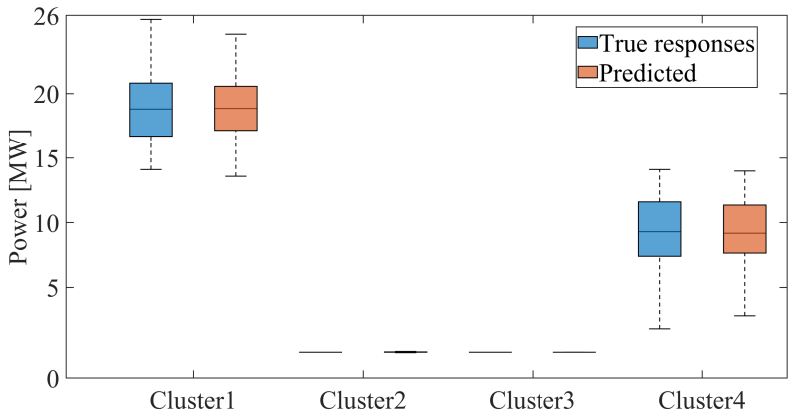

Fig. 4.13 Proposed framework predictions of the training data set.

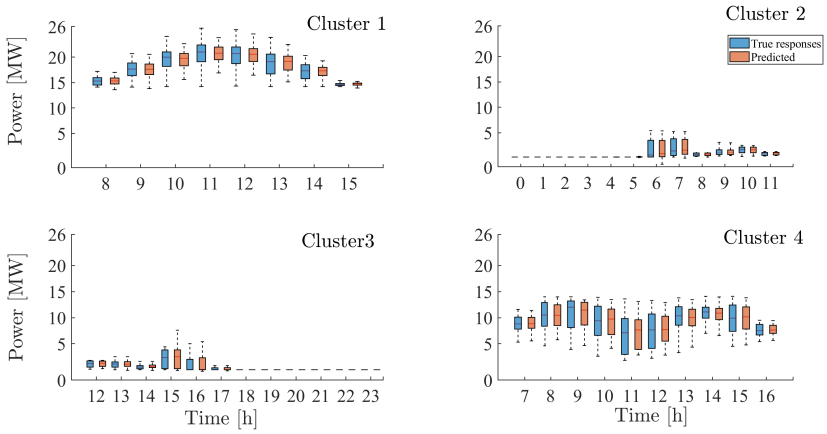

Fig. 4.14 Proposed framework hourly predictions of the training data set.

In order to further understand the value of clustering in Fig. 4.14, the box plots of the hourly predictions and actual values of the testing data are depicted. It can be seen that there are hours that belong to more than one cluster, which means that even the same hour patterns may be different based on which cluster they are identified to be in. In other words, different hours on different days, even if the days are in the same season, may behave completely different. It can also be seen in Fig. 4.14 that Clusters 2 and 3 represent early morning and night times and Clusters 1 and 4 represent seasonal variations, as also mentioned in Section 4.2.2.

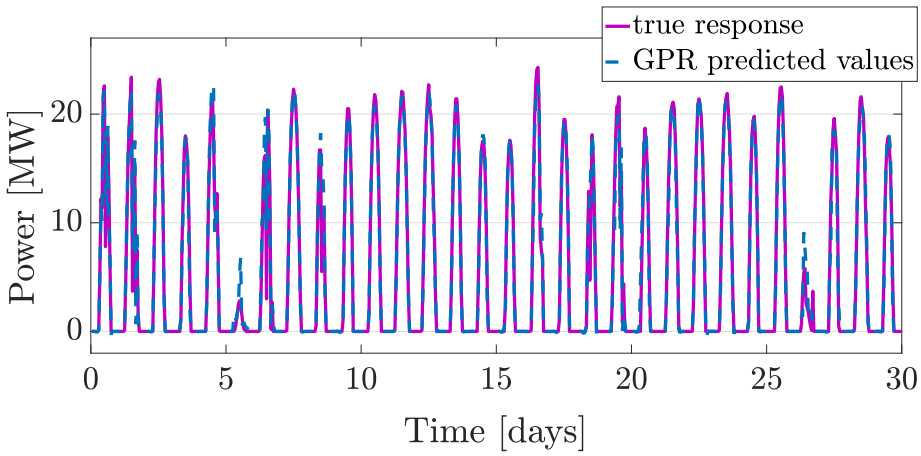

Fig. 4.15 hourly prediction for 30 random days.

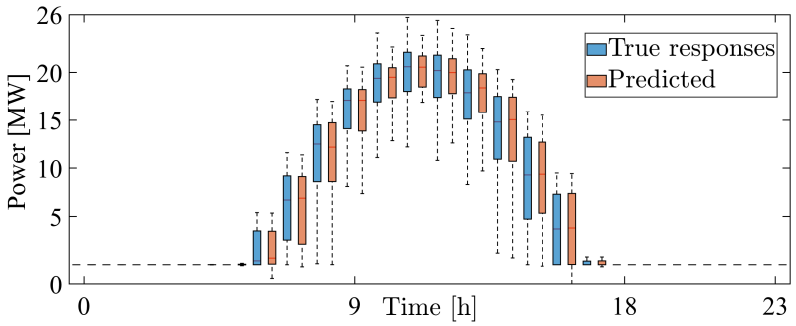

Fig. 4.16 30 random days 24-hour prediction with one hour intervals, Denver.

In Fig. 4.15 the forecasts and actual values of the 30 hold-out selected days that are representative of different seasons are depicted. The x-axis of the figure has 30 days, which correspond to 24-hour intervals for each day. As it may be seen, the two values are very close to each other. Another visual representation of the same result, i.e., the daily hourly forecast of the 30 days, seen in Fig. 4.16, where we notice that the predicted and the actual values (true response) follow the same pattern.

In order to test the accuracy of the forecasts, we use the following statistical metrics: RMSE, and MAE, as defined in (4.2) and (4.3). The statistical results for the training set and the test set are summarized in Table 4.7. The error metrics of the testing data between the actual and the predicted values are based on the average error of all five folds for the training set. To interpret these values, notice that the higher RMSE and MAE values, the less predictive the model is. In addition, it should be noted that the test results are expected to be different from the training set results since 30 hold-out days are not shown to the model during the training process. However, the results with any test set should be approximately the same as those obtained with the training set, as it may be seen in Table 4.7.

Site A	RMSE [MW]	MAE [MW]	RMSE[%]	MAE[%]
Training set	1.24	0.36	4.18	1.22
Test set	1.23	0.56	4.12	1.89

Table 4.7 Site A forecasts' error metrics.

The error between the actual and the forecasted value for the 30 hold-out days is depicted in Fig. 4.17. The average prediction error of the hold-out days for one cluster is fitted into

a normal distribution. In this figure, the y-axis represents the percentage of hours. It can be seen that for 54% of the hours, i.e., 382 hours, the prediction error is less than 0.03 MW. In order to provide a confidence level to the forecast, we use the confidence interval (CI) [149]. The selection of a confidence level for an interval determines the probability that the confidence interval produced will contain the true parameter value. Common choices for the CI are 0.90, 0.95, and 0.99. The CI is defined as follows:

$$\mathrm{CI} = \left(\bar{\varepsilon} - z^{\star} \frac{\sigma_{\varepsilon}}{\sqrt{T_{\star}}}, \bar{\varepsilon} + z^{\star} \frac{\sigma_{\varepsilon}}{\sqrt{T_{\star}}} \right), \tag{4.21}$$

Where $\bar{\varepsilon}$ is the mean value of the errors, σ_{ε} is the standard deviation, and $T_{\star}$ is the sample size of the errors. The value $z^{\star}$ represents the point on the standard normal density curve, such that the probability of observing a value greater than $z^{\star}$ is equal to p. The relationship between CI and p is $p = (1 - \mathrm{CI})/2$. Thus, if we wish to have a CI of 95% then $p = 0.025$. The value $z^{\star}$ such that $P(Z > z^{\star}) = 0.025$, or $P(Z < z^{\star}) = 0.975$, is equal to 1.96 as we may find in a standard normal distribution table. As the size of the width of the corresponding interval decrease, the confidence level decreases. By fitting a normal distribution in Fig. 4.17 we have mean value $\bar{\varepsilon} = 0.03$ and a standard deviation $\sigma_{\varepsilon} = 0.50$. Now, we may calculate the CI for various confidence levels; for instance, with 95% confidence level, the difference between the actual data and the prediction value of each point ranges between $[-0.47, 0.43]$ MW or $[-1.6\%, 1.4\%]$. On the other hand, our model produces a prediction interval (PI) for each

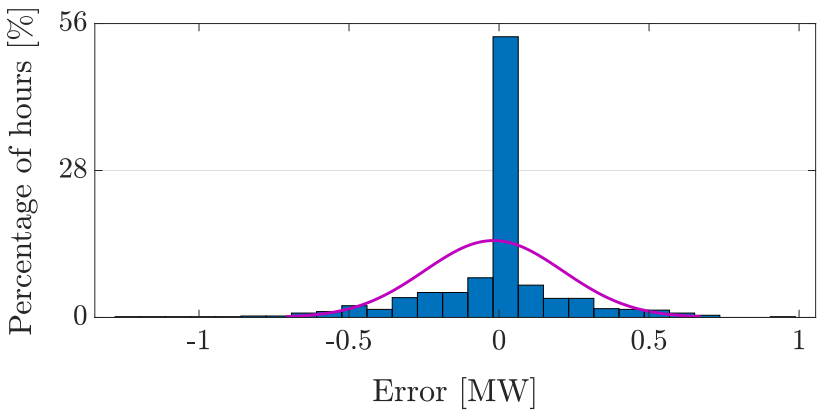

Fig. 4.17 Probability distribution fitting of average error of one sample model.

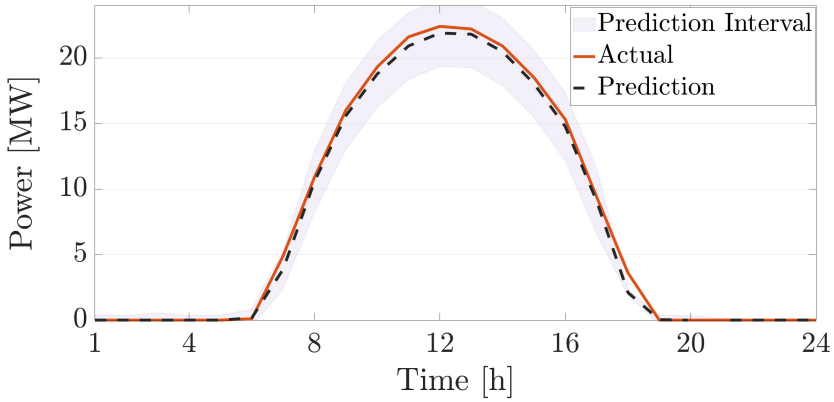

Fig. 4.18 Prediction interval with 95% confidence level and one hour look-ahead time for 17 June, 2006 at Site A.

point estimate, giving the lower and upper bounds within which the predicted solar output is expected to fall with a 95% confidence level. As seen in Figs. 4.18, 4.19, the hourly forecast is bounded with the PI lower and upper bounds produced for each hour. As we can see, the predicted PV line, which is constructed by connecting the central points of PIs, is close to the actual PV output. In order to provide more statistical metrics appropriate for probabilistic forecasts (see, e.g., [129]), we evaluate the Prediction Interval Coverage Probability (PICP) and the Prediction Interval Normalized Root-mean-square Width (PINRW), which are two standard performance metrics used to evaluate model accuracy [150]. PICP measures the

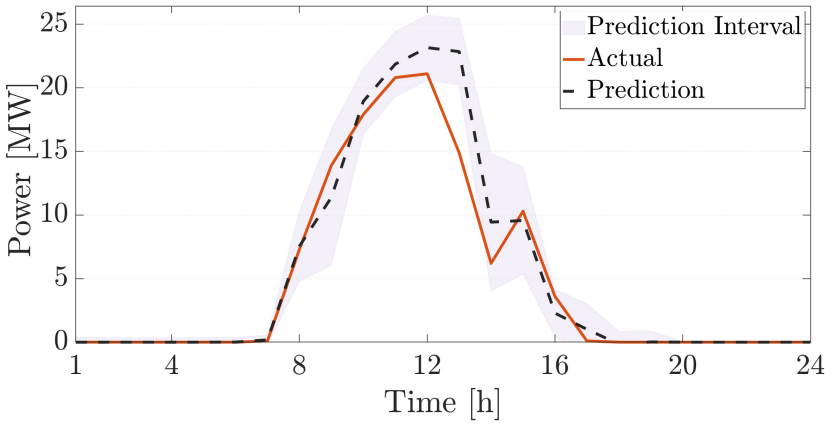

Fig. 4.19 Prediction interval with 95% confidence level and one hour look-ahead time for 17 March, 2006 at Site A.

number of observed values within the predicted intervals and the reliability of the Prediction Interval (PI). These are defined as follows:

$$\text{PICP} = \left(\frac{1}{T_*} \sum_{t=1}^{T_*} C^{(t)} \right) \cdot 100\%, \tag{4.22}$$

where

$$C^{(t)} = \begin{cases} 1 & \tilde{y}^{(t)} \in [L^{(t)}, U^{(t)}] \\ 0 & \tilde{y}^{(t)} \notin [L^{(t)}, U^{(t)}] \end{cases},$$

where $L^{(t)} = \bar{y} - z^{\star} \frac{\sigma_y}{\sqrt{T_\star}}$ and $U^{(t)} = \bar{y} + z^{\star} \frac{\sigma_y}{\sqrt{T_\star}}$ are the lower and upper values of the prediction interval and $\tilde{y}^{(t)}$ is the observed value at time t with $\bar{y}$ and σ_y representing the mean and standard deviation of the solar output predictions respectively.

$$\text{PINRW} = \frac{1}{A} \sqrt{\frac{1}{T_*} \sum_{t=1}^{T_*} (L^{(t)} - U^{(t)})^2}, \tag{4.23}$$

where A is the range of target data. It can be seen that larger PICP and narrower PINRW indicate higher quality PI [125]. In our work for Site A, the PICP score is 93.89%, which implies that the reliability of the PI is very high. Also, PINRW is 5.68% which shows a narrow prediction interval. Thus, the PICP and PINRW results show a high-quality prediction interval that has a high value of PICP and a low value of PINRW. It is worthy to note that these values are averaged over 30 independent runs.

As a further investigation, to validate the rationality of the proposed approach, we use a Q-Q plot as a graphical tool to illustrate the reliability of the forecast [151]. If the forecast system is statistically reliable, the forecast falls inside the prediction interval [152]. Furthermore, as depicted in Fig. 4.20, the error is falling around the straight line, which shows that the error between the observations and the estimations follows a normal distribution and thus indicates the suitability of our approach.

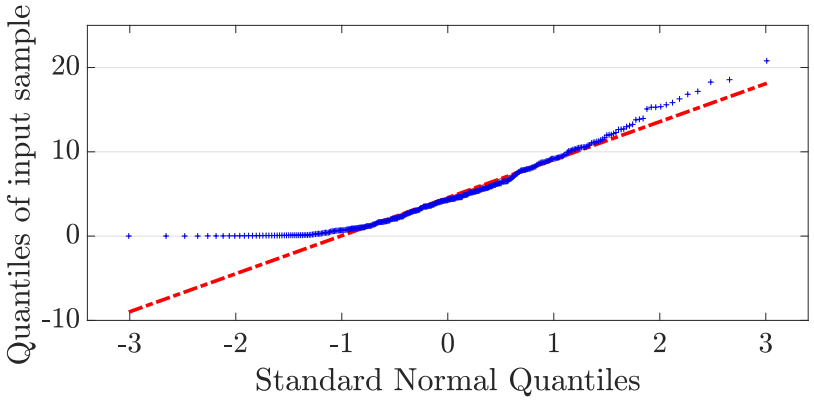

Fig. 4.20 Q-Q plot of absolute error between the observed value of sample test data points and their predicted value.

4.4.3 Summary of results of all sites

In order to further validate our framework, we have applied it to the remaining four sites as given in Table 4.3. Following the same procedure, as described in more details in Section 4.4.2, we have four clusters per site; from each dataset, we hold-out 30 representative days and train a GPR model for each cluster with the remaining data. The results for each site are summarised in Table 4.8 for the training dataset and in Table 4.9 for the test dataset. As seen above, the results for all datasets are approximately in the same range, which means that the model may be applied in any site under the assumption that the data of the selected features are available.

To further prove the effectiveness of the proposed framework, we compare our results with other recent studies. In order to make the comparison meaningful, we need to have access to the same set of data. The work presented in [1, 2] use the same data for site E,

Site	RMSE [MW]	MAE [MW]	RMSE [%]	MAE [%]
A	1.25	0.36	4.18	1.22
B	1.39	0.63	4.18	1.91
C	1.51	0.59	4.33	1.69
D	1.39	0.30	4.21	0.92
E	0.02	0.01	4.48	1.96

Table 4.8 Training set error metrics for all sites.

Site	RMSE [MW]	MAE [MW]	RMSE [%]	MAE [%]
A	1.23	0.56	4.12	1.89
B	1.51	0.66	4.58	2.00
C	1.61	0.72	4.60	2.06
D	1.44	0.65	4.38	1.98
E	0.02	0.01	3.48	1.85

Table 4.9 Test set error metrics for all sites.

		RMSE[%]	MAE[%]
Proposed framework		**3.48**	**1.85**
[1]	Fall	13.85	8.48
	Winter	7.67	4.16
	Spring	13.60	8.08
	Summer	16.43	10.73
[2]	ELM	12.84	6.68
	MLP	13.33	7.53

Table 4.10 The comparison of forecast error metrics for our framework against two other methods ([1, 2]), for site E.

which are available from the University of Queensland. The temporal resolution of the data in [2] is 1-minute; however, since we are interested in hourly values, we select historical data with hourly resolution. We used 2012 data for training and 2013 data for testing. The authors in [1] calculate predictions for each of the four seasons, i.e., fall, winter, spring and summer. In [2] the authors calculate hourly forecasts using their proposed ELM method and a multilayer perceptron (MLP) (referred to as FFBPG in their work). Yearly results are better than each season prediction in [1, 2], as shown in Table 4.10.

The errors of the proposed framework are small since the variation of solar output over different times of day and year is taken into account with the use of k-means clustering. The use of a clustering algorithm results in similar points that belong in the same cluster being trained with a GPR model. More specifically, k-means divides similar data in one group, which follows a distribution with specific characteristics, making each cluster's training more efficient with lower errors. Moreover, using an appropriate kernel function that relates the input features to output improves the forecast. In this work, using Matérn 5/2 as a kernel

function increase the accuracy of prediction due to the capability of the kernel in solving stochastic problems [108].

4.5 Conclusion

In this Chapter, we proposed a probabilistic framework for short-term photovoltaic forecasting to answer the second research question we defined in the first Chapter. We carried out a correlation study to identify the features which are highly related to solar output power. The six selected features that affected the PV output generation more prominently were: direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, temperature, zenith, and azimuth.

After identifying the highly correlated features, since solar output relies on solar irradiance, we clustered our data into four groups based on day-time. Two clusters appeared to represent early morning and night times, and the remaining two represent seasonal variations. We then trained a model for each of the four clusters using GPR in order to learn the relationship between the six input features and the PV generation. GPR is a kernel-based nonlinear nonparametric regression technique in which the covariance function plays a crucial role. In this work, we selected the Matérn $5/2$ as a covariance or kernel function. This function was selected under the assumption that the correlation between any two points in the input feature set decreases with increasing the euclidean distance between them. To test the accuracy of the proposed method for solar output forecasting, different tests and validation methods were exploited, i.e., k-fold cross-validation and hold-out validation methods.

In the case studies, we demonstrated the framework implementation in five different sites. For each site, the experiments were run 30 independent times for each cluster. Then, 30 random days were selected as hold-out data that were representative of different days of the year during different seasons. As shown in Fig. 4.3, selecting an appropriate number of clusters significantly improves the results, i.e., RMSE for eight clusters is 2.4%, while this metric value is 8.34% without clustering and for four clusters, the choice of our study, this value is 4.12%. The largest RMSE and MAE were 4.60 % and 2.06 %, respectively, showing

the efficacy of the proposed framework. Furthermore, the comparison of our proposed framework with the existing state of the art in Table 4.10 showed an improvement on the same accuracy metrics they used for point prediction. In the clustering, we used groups of similar data points that are more likely to produce similar PV outputs and use them for training a single model, reducing the variance of training data and increasing the accuracy. The rationale behind this is that similar data points in different seasons can be grouped in the same cluster in our model, while in the seasonal model, they happen to be grouped in different seasons.

Chapter 5

Transmission-Distribution Coordination Schemes

As we addressed the first two research questions in the previous Chapters, we now are aware of the role of an effective forecast model in the massive integration of renewable energy into the conventional power system. Also, we know how the constraints associated with the power network need to be appropriately addressed in the optimal power flow problem. The network operators play a significant role in the transition towards the net-zero emission system. So, now we are moving into the final research question. We are modelling the interaction between the network operators in two different frameworks considering the penetration of PV-Batteries into the distribution network.

5.1 Introduction

Several coordination schemes that can precisely model the system taking into account Several coordination schemes that can precisely model the system taking into account nonlinear bidirectional AC power flow constraints present in transmission and distribution systems have been recently proposed. In [153], the authors propose five coordination schemes to evaluate the recent proposals of the SmartNet project consortium. In order to do so, they model the optimisation problem considering the AC load flow and the topology of the grid in

each scheme. The main objective of this work was to quantify the proximity of the optimal solution to a physically compatible solution in different coordination schemes. In [154], the study aims at minimising the deviation from the real-time dispatch and maximising the share contribution of renewable energy while addressing uncertainty using Dynamic AC Optimal Power Flow. In [155], distribution locational marginal pricing is designed through quadratic programming. The case studies include a high number of electric vehicles and heat pumps to address issues associated with these resources in the distribution system. In [156], the authors summarise the main challenges proposed in the SmartNet project in three different countries (Denmark, Italy, and Spain) by providing techno-economic analysis on various coordination schemes in 2030 scenarios.

Alternative approaches are based on approximations of the AC power flow and represent the distribution and transmission systems with linearised power equations to overcome the challenges associated with nonlinearities (see, e.g., [97]. Approximations of AC power flow have been used in various problems in power systems that can also be applied in this particular setting. For instance, to control the reactive power at every bus, a method that approximates the distribution network into a linear distribution load flow was proposed in [8]. The results show that by linearising the load flow, the error on the voltage mismatch error is minimised. The authors in [157] address the power loss optimisation in smart power distribution by linearising the distribution power flow. This work demonstrates that the results of quadratic programming are better than conventional power flow in both robustness and computational complexity. In [158], a linear optimal load flow has been introduced using quadratic programming to cope with the increase in the number of DC microgrids.

How the network is represented is one of the main aspects to consider in TSO-DSO coordination. For instance, as the integration of RESs affects the voltage levels and the line thermal limits, network constraints need to be considered to ensure that these resources do not adversely disturb the power system operations [159]. In [160] the authors propose a coordination scheme that does not explicitly represent the grid topology but incorporates some information concerning, e.g., bus voltages. In [161], three market designs are proposed to mitigate coordination between the TSO and the DSO that provide a flexible, competitive

market design for retailers. In the model, the main focus is on the market rather than on the operation and topology of the grid. A control framework that provides the DSO with information on the contribution of each smart home, the unbalanced power flow and network voltage constraints are given in [162]. In this way, DG participates in the electricity market while ensuring that the upstream constraints are satisfied. In [86], three TSO-DSO coordination models are discussed. First, a TSO-managed model is presented, where the TSO is responsible for the optimal operation of the system by considering DG and transmission system constraints. Next, a TSO-DSO hybrid-managed model is introduced, where the TSO operates the system regarding the transmission network constraints and the DG that submits bids to demonstrate its willingness to participate in the market. Last, a DSO-managed model is mentioned where the DSO is responsible for operating its own system taking into account the distributed energy sources and sending back the outcomes to the TSO [153]. Centralised TSO-managed schemes make the coordination model simpler to implement (see, e.g., [13, 163]). By using a centralised scheme, we utilise the traditional SCADA system to monitor, measure and collect the data from different assets of the grid [163].

However, they might fail to fully utilise DG resources at the distribution system as the DSO has less visibility of their usage. TSO-DSO hybrid systems are an improvement of the latter since DG resources indicate by their bids to the TSO and DSO their willingness to participate, and both operators, based on their priorities, can decide whether they accept the offer or not [164]. Thus, a DSO-managed scheme has the potential to reach to the highest level of efficient use of distributed resources. However, it incorporates the risk that there might be a conflict between the TSO and DSO requirements and needs. Thus, making a real-time exchange of information between both operators necessary to ensure a reliable operation. Notwithstanding the merits of the solutions mentioned above, there are still gaps to assist operators with practical solutions to smoothly adapt to the large-scale integration of renewable energy resources and reliably transition into carbon-free power systems. Different solutions have been proposed to model coordination schemes. However, privacy concerns of the entities involved and their individual priorities need to be further analysed.

The aforementioned centralised schemes face a variety of regulatory challenges that make their actual implementation difficult. However, centralised schemes can still be used to provide insights into the desired coordination between TSOs and DSOs. As such, in practice, decentralised schemes need to be further investigated. These schemes need to respect the privacy concerns of the entities involved, be computationally efficient, depend on realistic communication infrastructure, achieve an optimal with some objective outcome, relieve congestion, and facilitate the integration of renewable-based generation. As discussed in the previous section, the methods present in the literature fail to meet at least one of the points mentioned above.

This thesis proposes a linear transmission-distribution system coordination framework considering large-scale integration of distributed resources, e.g., photovoltaic (PV) and storage. More specifically, we approximated the power equations with linearised equations so that the resulting optimal power flows performed by both the TSO and DSO are convex optimisation problems (see, e.g., [8, 97]) in Chapter 3. Next, we propose two different coordination schemes, decentralised and centralised. In the decentralised scheme, the TSO and the DSO collaborate to allocate all resources in the system optimally. In particular, we develop an iterative bi-level optimisation technique where the upper level is the TSO. The TSO solves its own OPF and determines the LMPs at substations. Then, the LMPs are passed on to the lower level, a collection of DSOs, each of which solves its own OPF. The new demand of the distribution system is aggregated at the substation levels and sent back to the TSO. We iterate between the two levels until some stopping criterion, e.g., that the infinity norm of the vector containing the differences of LMPs at current and previous iterations does not change by some tolerance is met. We demonstrate numerically that this process converges to a point near the optimal solution. Moreover, in the numerical results' section, it is shown that the proposed decentralised scheme provides a balance between the TSO and DSO objectives in terms of cost. It is worthy to note that the only information used in the iterative decentralised scheme is the customers' net load at the PCC; thus, there is no issue associated with privacy concerns of individual entities. The transmission system acts as the entire system operator in the proposed centralised scheme and has all the necessary

information about the distribution system. In such a case, the objective function consists of the distribution system voltage deviation from reference, the distributed resources cost, and the transmission system operating cost, aggregated as one objective with some weighting coefficients. We modify the weighting coefficients to approximate the Pareto front of the TSO and DSO objectives and study their interaction. In particular, we quantify the conflicting objectives of TSOs and DSOs, which may be used by DSOs to submit bids to the TSO or by the TSO to appropriately incentivise DSOs to provide their services. The proposed framework is validated by constructing a transmission distribution system using the 33- and 69-bus IEEE distribution feeders and a five-node transmission system. More specifically, the main contributions of our work are: (i) to propose an iterative algorithm to solve a bi-level TSO-DSO coordination scheme in a decentralised manner where no sensitive information is being exchanged, (ii) to analyse the interaction of TSOs and DSOs, i.e., how conflicting their objectives and priorities are, by formulating a common TSO-DSO OPF scheme where the Pareto front is determined. To this end, we approximate both transmission and distribution networks with linearised formulations, construct a transmission-distribution power system, analyse the impact of DG integration in terms of cost and congestion, and validate the proposed framework by comparing the results against the benchmark of current practice.

In Chapter 3, Section 3.3 we model the augmented DC OPF for the transmission system and a linear OPF for the distribution system. Thus, we can use that model in this chapter. The remainder of the chapter is organised as follows. In Section 5.2 and 5.3, we formulate the proposed decentralised and centralised schemes. In Section 5.4, we illustrate the proposed framework through the constructed transmission-distribution system. Finally, in Section 5.5, we summarise the results and make some concluding remarks.

5.2 Decentralised Scheme

We define for each distribution system d the set of decision variables y_d and the vector $y = \cup_{d \in \mathscr{D}} y_d$ representing all distribution systems connected to the transmission system. The

proposed decentralised scheme is based on solving the following optimisation problem:

$$\begin{aligned} &\min_{x} f_1(x,y) \\ \text{subject to } & g_1(x,y) \leq 0, \\ & h_1(x,y) = 0, \\ & y_d \in \arg\min_{y_d}\{(f_2(x,y_d) : g_2(x,y_d) \leq 0, h_2(x,y_d) = 0\}, \forall d \in \mathscr{D}, \end{aligned} \tag{5.1}$$

where $f_1(x,y)$ in our problem is the objective function of the TSO OPF, i.e., $\sum_{i\in\mathscr{I}} c_i(t) + \pi\sum_{\ell\in\mathscr{L}}(\theta_n(t)-\theta_m(t))^2$ as described in Section 3.3.1. Similarly, $g_1(x,y)$ and $h_1(x,y)=0$ are the equality and inequality constraints of (3.1) evaluated at y. In the lower-level parametric optimisation problem for each distribution system d, $f_2(x,y_d)$, $g_2(x,y_d)$ and $h_2(x,y_d)$ are the collection of distribution level objective functions, equality and inequality constraints,

respectively, as defined in (3.27). More specifically, the functions $f_1, f_2, g_1, g_2, h_1,$ and h_2 can be easily mapped to the functions below:

$$
\begin{aligned}
\min_{P_{G_i}(t), i\in\mathscr{I}, \theta_k(t), k\in\mathscr{K}} \quad & \sum_{t\in\mathscr{T}} \left(\sum_{i\in\mathscr{I}} c_i(t) + \pi \sum_{\ell=(n,m),\ell\in\mathscr{L}} (\theta_n(t) - \theta_m(t))^2 \right) \Delta t \\
\text{subject to} \quad & f^m \leq f(t) = B_d A \theta(t) \leq f^M, t \in \mathscr{T}, \\
& P_{G_i}^m \leq P_{G_i}(t) \leq P_{G_i}^M, t \in \mathscr{T}, \\
& \sum_{i\in\mathscr{I}_k} P_{G_i}(t) - \sum_{\ell\in\mathscr{L}} B_{d_\ell} A \theta(t) = P_{\text{grid}}^d(t), k \in \mathscr{K}, t \in \mathscr{T}, d \in \mathscr{D} \\
& \forall d \in \mathscr{D}, P_{\text{grid}}^d(t) \in \arg \min_{\substack{P_{PV_i}(t), P_{B_i}^{\text{ch}}(t), \\ P_{B_i}^{\text{dis}}(t), V_i(t), \\ P_{\text{grid}}^d(t)}} \quad \sum_{t\in\mathscr{T}} \left(\lambda_{k_d}(t) P_{\text{grid}}^d(t) + \sum_{i\in\mathscr{N}_{PV}^d} B_{PV_i} P_{PV_i}(t) + \sum_{i\in\mathscr{N}_B^d} B_{B_i} (P_{B_i}^{\text{ch}}(t) + P_{B_i}^{\text{dis}}(t)) + \sum_{i\in\mathscr{N}} \alpha (V_i(t) - V_{\text{ref}})^2 \right) \Delta t \\
& \qquad \text{subject to} \quad P_{PV,i}^{\min} \leq P_{PV_i}(t) \leq P_{PV,i}^{\max}, i \in \mathscr{N}_{PV}, t \in \mathscr{T}, \\
& \qquad P_{B,i}^{\text{ch,min}} \leq P_{B_i}^{\text{ch}}(t) \leq P_{B,i}^{\text{ch,max}}, i \in \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad P_{B,i}^{\text{dis,min}} \leq P_{B_i}^{\text{dis}}(t) \leq P_{B,i}^{\text{dis,max}}, i \in \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad V_i^{\min} \leq V_i(t) \leq V_i^{\max}, i \in \mathscr{N}, t \in \mathscr{T}, \\
& \qquad P_{\text{grid}}^{d,\min} \leq P_{\text{grid}}^d(t) \leq \sum_{i\in\mathscr{I}_k} P_{G_i}(t), t \in \mathscr{T}, \\
& \qquad E_{\min,i} \leq \sum_{t\in\mathscr{T}} \left(\eta_{\text{ch},i} P_{B_i}^{\text{ch}}(t) - \frac{1}{\eta_{\text{dis},i}} P_{B_i}^{\text{dis}}(t) \right) \Delta t + E_{0,i} \leq E_{\max,i}, \forall i \in \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad V(t) = Rp(t) + Xq(t) - {M^{-1}}^\top m_0, t \in \mathscr{T}, \\
& \qquad p_i(t) = P_{PV_i}(t) + P_{B_i}^{\text{dis}}(t) - P_{B_i}^{\text{ch}}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_{PV} \cap \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad p_i(t) = P_{PV_i}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_{PV} \setminus \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad p_i(t) = P_{B_i}^{\text{dis}}(t) - P_{B_i}^{\text{ch}}(t) - P_{\text{load}_i}(t), \forall i \in \mathscr{N}_B \setminus \mathscr{N}_{PV}, t \in \mathscr{T}, \\
& \qquad p_i(t) = -P_{\text{load}_i}(t), \forall i \in \mathscr{N} \setminus \mathscr{N}_{PV} \cap \mathscr{N}_B, t \in \mathscr{T}, \\
& \qquad q_i(t) = -Q_{\text{load}_i}(t), \forall i \in \mathscr{N}, t \in \mathscr{T},
\end{aligned}
\tag{5.2}
$$

where the objective of the upper level problem is the TSO cost minimisation and angle deviation; its constraints refer to power flow and generator limits and power balance. The lower level optimisation problem has as an objective the DSO cost and voltage regulation cost minimisation; its constraints refer to voltage, power, energy storage limits, and power

balance. It is worth noting that P^d_{grid} is the net power imported into the distribution network from the transmission system.

This problem is a bi-level optimisation [165]. Such problems were introduced when Stackelberg (see, e.g., [166]) formulated a strategic game in 1934 where a leader and a follower make sequential moves, starting with the leader. Thus, the upper level and lower level can be considered as leader and follower. More specifically, bi-level optimisation problems are defined where one or some of the decision variables are constrained to the solutions of another optimisation problem. Then, the problem is formulated as in (5.1) in two levels of optimisation. Solving bi-level optimisation problems has been known to be NP-hard [167]. There are basically two main techniques for solving bi-level optimisation problems. The first one keeps the bi-level structure and treats the lower level (LL) problem as a parametric optimisation problem that is solved whenever the solution algorithm for the upper level (UL) problem requires it. The second technique is based on the formulation of first-order necessary optimality conditions for the lower level problem. The lower level problem is then replaced by its necessary conditions, which are considered as constraints in the upper-level problem. This reduces the bi-level problem to a single level nonlinear optimisation problem. The drawback of this method is that, in general, necessary conditions are not sufficient for optimality and thus information is lost in the single level formulation, which, in turn, may result in non-optimal solutions for the bi-level optimisation problem. In particular, the Karush-Kuhn-Tucker (KKT) conditions that should be satisfied in this approach are only guaranteed if the optimisation problem is convex [168].

In this work, we propose an approach that resembles the first one discussed above. Still, we treat the two levels as coupled optimisation problems while iteratively solving one after the other; that is, LL optimisation problem is treated as interdependent parametric optimisation problems that are solved whenever the solution algorithm for the UL requires it. In particular, the TSO and DSO collaborate to operate the power network optimally. Initially, the TSO optimises the transmission system, considering a feasible distribution system initial load solution. The transmission system's resources meet the distribution system's entire load, i.e., the distribution system does not use its distributed resources to meet the load. Next, the TSO

solves its own augmented DC OPF and announces the locational marginal price of the PCC to the DSO. Next, the DSO solves its own LL problem taking into account the capabilities of the distributed resources. The DSO net load is different in the next iteration, and the amount of energy that DSO buys from the TSO may be reduced, depending on cost. We iterate between these two levels until a convergence criterion is met, e.g., that the infinity norm of the vector containing the LMP differences between the current iteration and the previous iteration does not change by some tolerance. The proposed algorithm is described as follows:

Algorithm 1: Iterative algorithm for solving (5.1)

1: **Initialization**
2: Set $\nu = 0$.
3: Consider $y_d[0]$ so that it is a feasible solution of the LL optimisation $\forall d \in \mathscr{D}$.
4: **Repeat until convergence**
5: Solve the UL optimisation problem using $y_d[\nu]$; let the solution be $x[\nu]$ and $\lambda_{k_d}[\nu]$.
6: Solve the LL optimisation for $x[\nu]$ using $\lambda_{k_d}[\nu]$. Let the solution be $y_d[\nu+1], \forall d \in \mathscr{D}$.
7: Set $\nu \leftarrow \nu + 1$ and go to step (4).

Considering this iterative procedure, the LL and UL optimisation problems are solved the same number of times and the levels are treated as uncoupled problems, just coupled at the interface by the procedure. There is no formal proof of convergence for such an iterative scheme, however convergence has been experimentally shown [169]. We further demonstrate that the proposed algorithm converges to a near-optimal solution. The flowchart of the algorithm is given in Fig. 5.1.

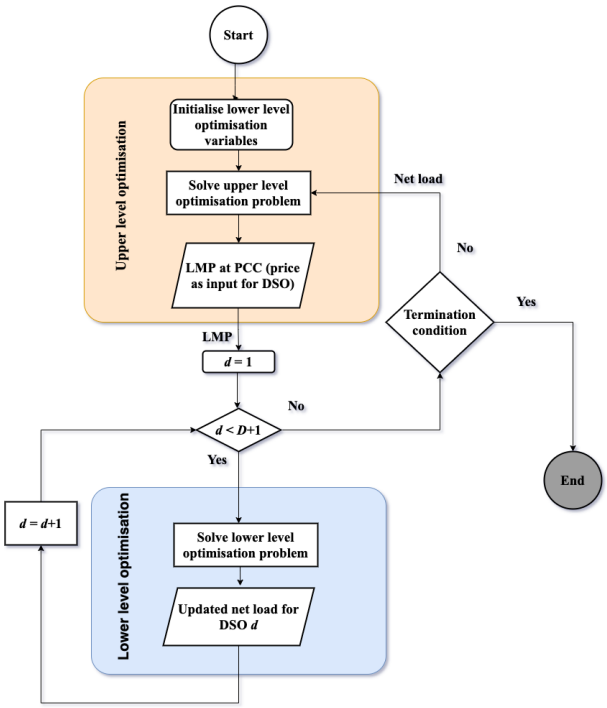

Fig. 5.1 Decentralised iterative scheme flowchart.

5.3 Centralised Scheme

This coordination scheme introduces the TSO as a leader who operates the transmission and distribution systems as one entire power network. In this case, the TSO solves a multi-objective optimisation (MOO) problem which can be formulated as follows:

$$\begin{aligned} \min_{x,y} \quad & (f_1(x,y), f_2(x,y)) \\ \text{subject to } & g_1(x,y) \leq 0, \\ & g_2(x,y) \leq 0, \\ & h_1(x,y) = 0, \\ & h_2(x,y) = 0, \end{aligned} \tag{5.3}$$

where x represents the decision variables for the transmission system and y the decision variables for all distribution systems. The first objective, $f_1(x,y)$, incorporates the TSO

objective functions, and $f_2(x,y)$ the objective functions of all the distribution systems in $\mathscr{D}$, that is, (3.11) + (3.12) + (3.13) as described in (3.1) and (3.27) respectively. The inequality and equality constraints are denoted as $g_1(x,y), g_2(x,y)$ and $h_1(x,y), h_2(x,y)$ respectively.

The notion of "optimality" in solving MOO problems is known as Pareto optimal. A solution is said to be Pareto optimal if there is no way to improve one objective without worsening the other, i.e., the feasible point $(x^\star, y^\star)$ is Pareto optimal if there is no other feasible point (x,y) such that for all i, j with $i \neq j, f_i(x,y) = f_i(x^\star, y^\star)$ with strict inequality in at least one objective, $f_j(x,y) < f_j(x^\star, y^\star)$. However, given their conflicting nature, it is difficult to minimise the objective functions simultaneously, and thus the Pareto solutions usually appear scattered. In solving the optimisation problem (5.3) we obtain the Pareto front. In general, identifying the set of all Pareto optimality points is not a tractable problem. A common approach for solving MOO is to find many evenly distributed efficient points, and use points to approximate the Pareto front. In this paper, we use the weighted sum method (see, e.g., [170, 171]) to convert the MOO into a single objective optimisation problem by using a convex combination of objectives. More formally, the weighted sum method solves the following scalar optimisation problem:

$$\begin{aligned} \min_{x,y} \quad & w_1 f_1(x,y) + w_2 f_2(x,y) \\ \text{subject to } & g_1(x,y) \leq 0, \\ & g_2(x,y) \leq 0, \\ & h_1(x,y) = 0, \\ & h_2(x,y) = 0 \\ & w_1 + w_2 = 1, \\ & w_1, w_2 \geq 0. \end{aligned} \tag{5.4}$$

By appropriately changing the weight vector $w = [w_1, w_2]^\top$ we can approximate the Pareto front. The weight w_2 corresponds to all $d \in \mathscr{D}$ distribution systems. We assign equal weights to each distribution system, i.e., $w_2 = \sum_{d \in \mathscr{D}} w_{2d}$, where $w_{2d} = \frac{w_2}{|\mathscr{D}|}, \forall d \in \mathscr{D}$ with $|\mathscr{D}|$ the

cardinality of the set $\mathscr{D}$. To make the formulation more clear we present here its detailed representation. The functions $f_1, f_2, g_1, g_2, h_1,$ and h_2 can be easily mapped to the functions below.

$$
\begin{aligned}
\min_{\substack{P_{G_i}(t), i\in\mathscr{I},\\ \theta_k(t), k\in\mathscr{K},\\ P_{PV_i}(t), P_{B_i}^{\mathrm{ch}}(t),\\ P_{B_i}^{\mathrm{dis}}(t), V_i(t)}} \quad & \sum_{t\in\mathscr{T}} \left(w_1 \left(\sum_{i\in\mathscr{I}} c_i(t) + \pi \sum_{\ell=(m,n)\in\mathscr{L}} (\theta_n(t)-\theta_m(t))^2 \right) + w_2 \sum_{d\in\mathscr{D}} \left(\sum_{i\in\mathscr{N}_{PV}^d} B_{PV_i} P_{PV_i}(t) + \sum_{i\in\mathscr{N}_B^d} B_{B_i}(P_{B_i}^{\mathrm{ch}}(t) + P_{B_i}^{\mathrm{dis}}(t)) + \sum_{i\in\mathscr{N}} \alpha(V_i(t)-V_{\mathrm{ref}})^2 \right) \right) \Delta t \\
\text{subject to} \quad & f^m \le f(t) = B_d A\theta(t) \le f^M, \\
& P_G^m \le P_G(t) \le P_G^M, \\
& \sum_{i\in\mathscr{I}_k} P_{G_i}(t) - \sum_{\ell\in\mathscr{L}} B_{d_i} A\theta(t) = p_1^d(t), k\in\mathscr{K}, d\in\mathscr{D}, t\in\mathscr{T}, \\
& P_{PV,i}^{\min} \le P_{PV_i}(t) \le P_{PV,i}^{\max}, i\in\mathscr{N}_{PV}, t\in\mathscr{T}, d\in\mathscr{D}, \\
& P_{B,i}^{\mathrm{ch,min}} \le P_{B_i}^{\mathrm{ch}}(t) \le P_{B,i}^{\mathrm{ch,max}}, i\in\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& P_{B,i}^{\mathrm{dis,min}} \le P_{B_i}^{\mathrm{dis}}(t) \le P_{B,i}^{\mathrm{dis,max}}, i\in\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& V_i^{\min} \le V_i(t) \le V_i^{\max}, i\in\mathscr{N}, t\in\mathscr{T}, d\in\mathscr{D}, \\
& E_{\min,i} \le \sum_{t\in\mathscr{T}} \left(\eta_{\mathrm{ch},i} P_{B_i}^{\mathrm{ch}}(t) - \frac{1}{\eta_{\mathrm{dis},i}} P_{B_i}^{\mathrm{dis}}(t) \right) \Delta t + E_{0,i} \le E_{\max,i}, \forall i\in\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& V(t) = Rp^d(t) + Xq^d(t) - M^{-1^\top} m_0, t\in\mathscr{T}, d\in\mathscr{D} \\
& p_i^d(t) = P_{PV_i}(t) + P_{B_i}^{\mathrm{dis}}(t) - P_{B_i}^{\mathrm{ch}}(t) - P_{\mathrm{load}_i}(t), \forall i\in\mathscr{N}_{PV}\cap\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& p_i^d(t) = P_{PV_i}(t) - P_{\mathrm{load}_i}(t), \forall i\in\mathscr{N}_{PV}\setminus\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& p_i^d(t) = P_{B_i}^{\mathrm{dis}}(t) - P_{B_i}^{\mathrm{ch}}(t) - P_{\mathrm{load}_i}(t), \forall i\in\mathscr{N}_B\setminus\mathscr{N}_{PV}, t\in\mathscr{T}, d\in\mathscr{D}, \\
& p_i^d(t) = -P_{\mathrm{load}_i}(t), \forall i\in\mathscr{N}\setminus\mathscr{N}_{PV}\cap\mathscr{N}_B, t\in\mathscr{T}, d\in\mathscr{D}, \\
& q_i^d(t) = -Q_{\mathrm{load}_i}(t), \forall i\in\mathscr{N}, t\in\mathscr{T}, d\in\mathscr{D},
\end{aligned}
\tag{5.5}
$$

where the objective of the centralised optimisation is the TSO cost, angle deviation, the DG cost and voltage regulation cost minimisation; its constraints refer to power flow and generator limits and power balance. The power balance in this case is modified to directly incorporate the real power injection/withdrawal at the PCC of each DSO.

Our problem has a convex Pareto front; thus we can generate all points of the Pareto front. Using the proposed method, we investigate how the objectives of TSO and DSOs interact with each other, and the TSO directly manages the entire system and purchases power from distributed energy sources in the distribution system; as for bidirectional power flows, if distributed energy sources generate excess energy needed at the distribution system level is fed into the transmission system.

5.4 Numerical Results

We present several numerical examples to demonstrate the capabilities of the proposed framework. We discuss the properties of the proposed decentralised coordination scheme in terms of convergence with some sensitivity studies. Insights are provided into both proposed coordination schemes. Furthermore, we demonstrate the interaction of TSOs and DSOs with the determination of the Pareto front of the centralised optimisation problem. Thus, in 5.4.1, the case study information is provided, followed by the numerical results of decentralised and centralised schemes in 5.4.2 and 5.4.3 respectively.

5.4.1 System Description

To validate the proposed framework we need to construct a power system with many voltage levels that will represent the transmission and distribution systems. As such, we select a five-node transmission system on which four distribution system feeders are connected to different nodes as depicted in Fig. 5.2.

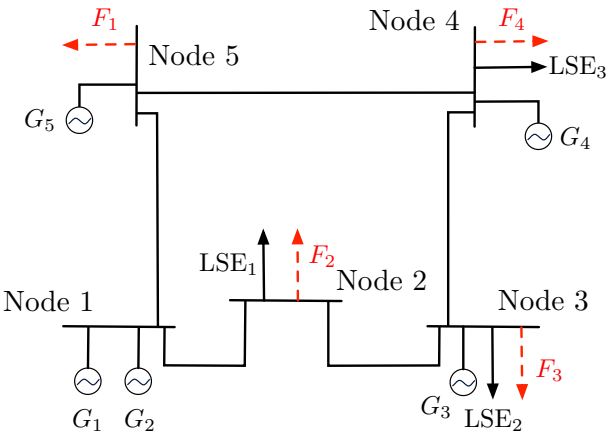

Fig. 5.2 Transmission and distribution system.

We denote by F_i the i_{th} feeder connected to the transmission system. More specifically, F_1 and F_3 correspond to the IEEE standard 33-bus feeder and F_2 and F_4 to the 69-bus IEEE standard bus feeder [172–174]. The load that is serving entities at a transmission node i are denoted by LSE$_i$. There are five generators connected at the transmission level in nodes 1, 3, 4 and 5. The transmission system data may be found in [97]. To demonstrate how the TSO-DSO coordination schemes can facilitate the integration of DG, we modify the standard IEEE 33- and 69-bus feeders by deploying PV and battery systems at different nodes. We assume that the distributed resources are mostly installed at end-nodes in the distribution level where the voltage drop levels are worst [175]. The modified feeders are depicted in Figs. 5.3, 5.4, respectively. In particular, PV and battery systems are installed in nodes 18, 22, 25 and 33 in the 33-bus feeder and in nodes 2, 3, 27, and 64 in the IEEE 69-bus feeder. The distributed resources data are presented in Table 5.1. Furthermore, we assume that each node's voltage in the distribution system is bounded between 0.95 pu and 1.05 pu and α is equal to 1.

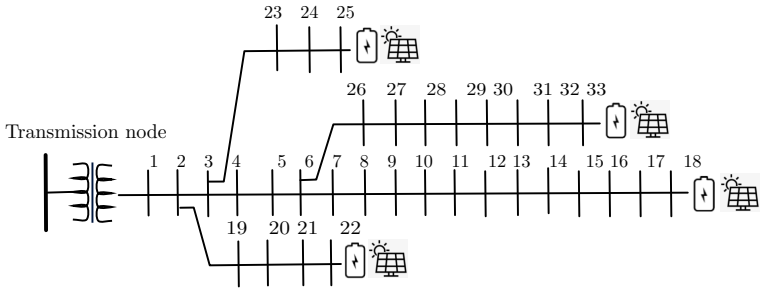

Fig. 5.3 Modified IEEE 33-bus distribution feeder.

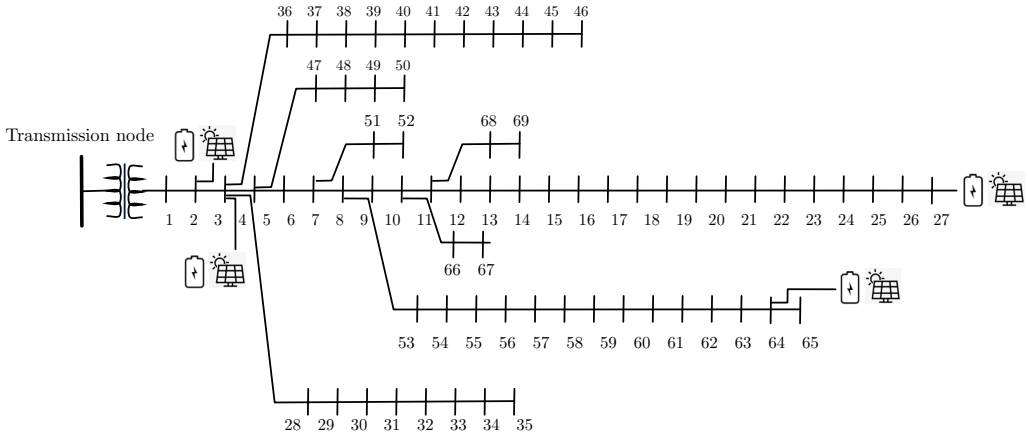

Fig. 5.4 Modified IEEE 69-bus distribution feeder.

Next, we implement both the proposed centralised and the decentralised schemes, and we compare the results with current practice, which refers to when the TSO solves its OPF and determines the LMPs at the substations. Next, the DSOs dispatch distributed DG by optimising cost and considering the LMP at the substation as a fixed parameter. In current practice, there is minimal coordination between TSOs and DSOs. Finally, the three methodologies are compared against a variety of metrics: total cost, hourly LMPs, hourly DG output, hourly generator output at the transmission level, netload, and level of congestion.

5.4.2 Decentralised Coordination Scheme

We apply the scheme proposed in Section 5.2 to the system described above. In order to demonstrate how the decentralised scheme facilitates the integration of distributed energy resources, we compare its optimal operation (method (ii)) against the current practice (method

Feeder	Variable	Value	Unit
All	$P_{PV}^{\min}$	0	MW
All	$P_{PV}^{\max}$	30	MW
All	B_{PV}	2.584	€/MW
F_1, F_3	$P_B^{\text{dis,min}}$	0	MW
F_1, F_3	$P_B^{\text{dis,max}}$	30	MW
F_1, F_3	$P_B^{\text{ch,min}}$	0	MW
F_1, F_3	$P_B^{\text{ch,max}}$	30	MW
F_1, F_3	$B_B^{\text{dis,min}}$	0.380	€/MW
F_2, F_4	$P_B^{\text{dis,min}}$	0	MW
F_2, F_4	$P_B^{\text{dis,max}}$	15	MW
F_2, F_4	$P_B^{\text{ch,min}}$	0	MW
F_2, F_4	$P_B^{\text{ch,max}}$	15	MW
F_2, F_4	$B_B^{\text{dis,min}}$	0.380	€/MW
F_1, F_3	$P_{\text{grid}}^{\min}$	-110	MW
F_2, F_4	$P_{\text{grid}}^{\min}$	-60	MW

Table 5.1 Distributed resources' physical limits and bid information.

(i)), where the current practice as discussed in the introduction section is when the TSO solves its own OPF and determines the LMPs at the substation, and the DSOs dispatch DG by optimising cost and considering the LMP at the substation as a fixed parameter. In other words, there is limited coordination between the operators in the current practice, and each operator tries to meet its own individual objectives. We run both cases for a one day period with hourly intervals. In Fig. 5.5, the TSO operation cost for both cases is depicted. We notice that the proposed decentralised coordination scheme results in a reduced transmission operation cost for all hours of the day. The reason is that distributed energy resources, which are less expensive than generators connected at the transmission level, are used to a greater extent, as seen in Fig. 5.6.

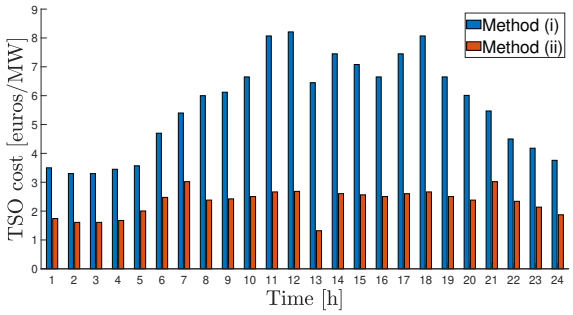

Fig. 5.5 Transmission operation cost for methods (i) current practise and (ii) proposed decentralised TSO-DSO coordination scheme.

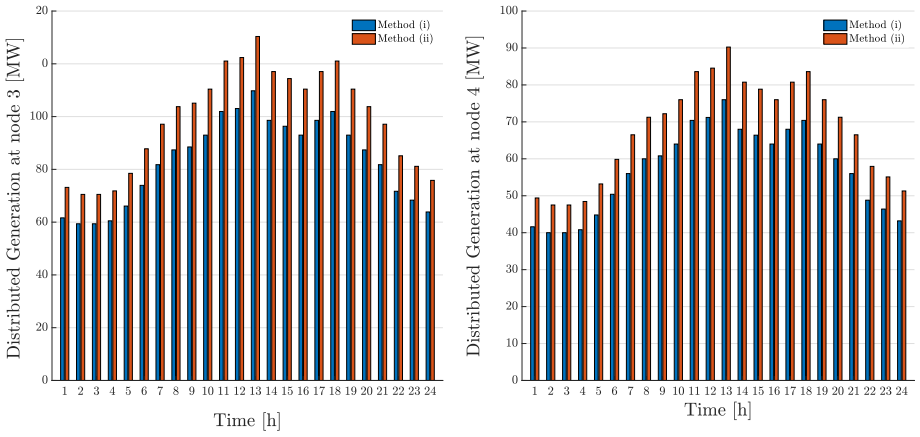

Fig. 5.6 The total amount of distributed generation for methods (i) current practise and (ii) proposed decentralised TSO-DSO coordination scheme at nodes 3 and 4.

Another effect of the increasing use of distributed resources is that they relieve the congestion present in the transmission system, which reduces TSO operational costs. For method (i), the LMPs for each hour at each node may be found in Table 5.2. We notice that for the same hour, each node has a different LMP. This demonstrates, based on the formulation of the augmented DCOPF in (3.1), that some line flows reached their limits. The LMPs of method (ii) are shown in Table 5.3. We notice that the LMP difference between hours has been reduced, reflecting the fact that there is less congestion in the transmission system. In fact, the LMPs are practically the same for all nodes at every hour when the proposed decentralised scheme is implemented. Following the formulation of (3.1) and using the KKT conditions of optimality, the LMP difference is expressed as a function of the

congestion that can be present in the network (see, e.g., [176]), i.e.,:

$$\lambda_k - \lambda_{k'} = \sum_{\ell \in \tilde{\mathscr{L}}} \phi_\ell^{\{k,k'\}} \mu_\ell, \tag{5.6}$$

where μ_ℓ is the dual variable of the power flow limits for line ℓ; $\tilde{\mathscr{L}}$ is the subset of lines that are at their limits, i.e., $\tilde{\mathscr{L}} = \{\ell_i : i = 1, \ldots, L, \mu_{\ell_i} \neq 0\}$; and $\phi_\ell^{\{k,k'\}}$ is the power transfer distribution factor of transaction with node pair $\{k,k'\}$ with respect to line ℓ. We can interpret (5.6) physically by considering an injection at node k and its withdrawal at node k'. We interpret $\phi_\ell^{\{k,k'\}}$ as the fraction of the transaction with node pair $\{k,k'\}$ of 1 MW that flows on line ℓ. As such for every hour the LMP differences are purely a function of the transmission usage costs of the congested lines, thus showing the "level" of congestion.

Hour	Node 1	Node 2	Node 3	Node 4	Node 5
1	12.67	28.15	25.22	17.15	13.46
2	12.62	28.01	25.10	17.08	13.41
3	12.62	28.01	25.10	17.08	13.41
4	12.64	28.08	25.16	17.11	13.44
5	12.76	28.42	25.45	17.30	13.56
6	12.93	28.89	25.87	17.55	13.74
7	13.09	29.36	26.28	17.80	13.92
8	13.21	29.70	26.58	17.99	14.05
9	13.23	29.77	26.64	18.02	14.08
10	13.32	30.04	26.88	18.17	14.18
11	13.51	30.58	27.35	18.46	14.39
12	13.53	30.65	27.41	18.49	14.41
13	13.68	31.05	27.76	18.71	14.57
14	13.44	30.38	27.17	18.35	14.31
15	13.39	30.24	27.05	18.28	14.26
16	13.32	30.04	26.88	18.17	14.18
17	13.44	30.38	27.17	18.35	14.31
18	13.51	30.58	27.35	18.46	14.39
19	13.32	30.04	26.88	18.17	14.18
20	13.21	29.70	26.58	17.99	14.05
21	13.09	29.36	26.28	17.80	13.92
22	12.88	28.75	25.75	17.48	13.69
23	12.81	28.55	25.57	17.37	13.62
24	12.71	28.28	25.34	17.22	13.51

Table 5.2 Locational marginal prices for method (i): current practise for TSO-DSO coordination in €/MW.

Hour	Node 1	Node 2	Node 3	Node 4	Node 5
1	12.27	12.28	12.28	12.27	12.27
2	12.13	12.14	12.14	12.14	12.13
3	12.13	12.14	12.14	12.14	12.13
4	12.20	12.21	12.21	12.21	12.20
5	12.54	12.55	12.55	12.54	12.54
6	13.01	13.02	13.02	13.01	13.01
7	12.55	28.14	25.19	17.06	13.35
8	12.88	12.89	12.89	12.88	12.88
9	12.90	12.91	12.91	12.90	12.90
10	12.98	12.99	12.99	12.99	12.98
11	13.15	13.16	13.16	13.15	13.15
12	13.17	13.18	13.18	13.17	13.17
13	11.93	11.94	11.94	11.94	11.93
14	13.08	13.10	13.10	13.09	13.08
15	13.04	13.06	13.06	13.05	13.04
16	12.98	12.99	12.99	12.99	12.98
17	13.08	13.10	13.10	13.09	13.08
18	13.15	13.16	13.16	13.15	13.15
19	12.98	12.99	12.99	12.99	12.98
20	12.88	12.89	12.89	12.88	12.88
21	12.55	28.14	25.19	17.06	13.35
22	12.87	12.89	12.89	12.88	12.87
23	12.67	12.68	12.68	12.68	12.67
24	12.40	12.41	12.41	12.41	12.40

Table 5.3 Locational marginal prices for method (ii): proposed decentralised TSO-DSO coordination in €/MW.

In Tables 5.4, 5.5, the hourly power output of each transmission generator is shown. We notice that with method (ii), the total power used by generators at the transmission level is reduced compared to method (i). The reason is that the less expensive distributed generators at the distribution level are used to satisfy the load instead. More specifically, we notice that with method (ii), the transmission level generators 2, 3, and 4 have zero output for most hours of the day as they are the most expensive ones.

Hour	P_{G_1}	P_{G_2}	P_{G_3}	P_{G_4}	P_{G_5}
1	110	18.53	19.52	0	110
2	110	15.09	13.36	0	110
3	110	15.09	13.36	0	110
4	110	16.81	16.44	0	110
5	110	25.41	31.84	0	110
6	110	37.45	53.39	0	110
7	110	49.5	74.95	0	110
8	110	58.1	90.35	0	88.4
9	110	59.82	93.43	0	90.88
10	110	60	110	2.45	100.81
11	110	43.78	110	57.07	110
12	94.58	60.36	110.71	60	110
13	62.8	0.03	116.72	42.99	110
14	110	55.25	110	31.2	110
15	110	60	110	16.85	108.26
16	110	60	110	2.45	100.81
17	110	55.25	110	31.2	110
18	110	43.78	110	57.07	110
19	110	60	110	2.45	100.81
20	110	58.1	90.35	0	88.4
21	110	49.5	74.95	0	110
22	110	34.01	47.23	0	110
23	110	28.85	38	0	110
24	110	21.97	25.68	0	110

Table 5.4 The power output in MW of generators at the transmission level for method (i): current practise for TSO-DSO coordination.

Hour	P_{G_1}	P_{G_2}	P_{G_3}	P_{G_4}	P_{G_5}
1	39.14	0	0	0	110
2	30.02	0	0	0	110
3	30.02	0	0	0	110
4	34.58	0	0	0	110
5	57.38	0	0	0	110
6	89.3	0	0	0	110
7	107.99	6.66	6.58	0	110
8	82.98	0	0	0	88.4
9	85.82	0	0	0	90.88
10	91.19	0	0	0	100.81
11	101.05	0.88	0	0	110
12	101.78	1.49	0	0	110
13	9.58	0	0	0	110
14	97.9	0	0	0	110
15	95.22	0	0	0	108.26
16	91.19	0	0	0	100.81
17	97.9	0	0	0	110
18	101.05	0.88	0	0	110
19	91.19	0	0	0	100.81
20	82.98	0	0	0	88.4
21	107.99	6.66	6.58	0	110
22	80.18	0	0	0	110
23	66.5	0	0	0	110
24	48.26	0	0	0	110

Table 5.5 The power output in MW of generators at the transmission level for method (ii): proposed decentralised TSO-DSO coordination.

In Fig. 5.7, we depict the operational cost for each distribution feeder connected to different nodes of the transmission system for methods (i) and (ii). We notice that the proposed coordination scheme results in reduced costs for all DSOs as all resources were utilised in a more efficient way as discussed above.

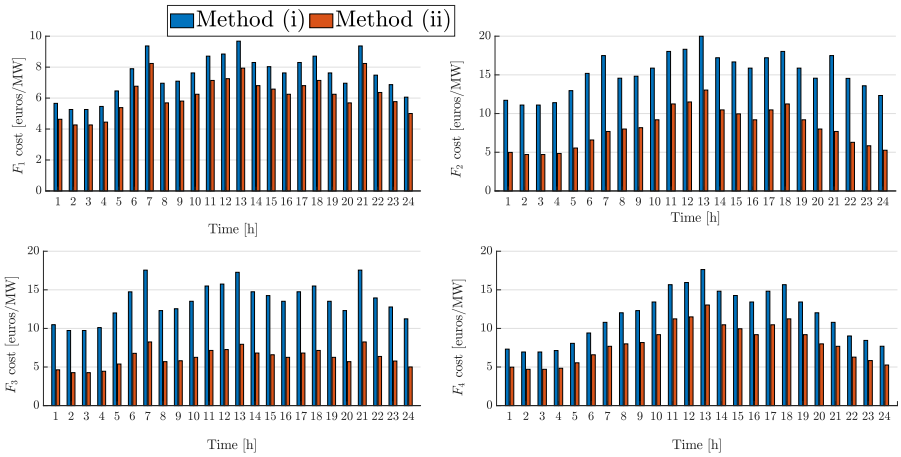

Fig. 5.7 The cost for each feeder for methods (i) and (ii).

We now study the net load at the transmission nodes using both methods. We can see in Fig. 5.8 that the net loads at the transmission system at nodes 2 and 3 decreases, a fact that is also reflected in the OPF in the transmission system and its LMPs. We also notice that there is a sharp fall and rise in the net load, between hours 7 and 8, and 20 and 21, respectively. This is due to the fact that the power flow between nodes 1 and 2 at times 7 and 21 is 75 MW, which is equal to the line's thermal limit. This causes the LMP divergence in these hours, as shown in Table 5.3.

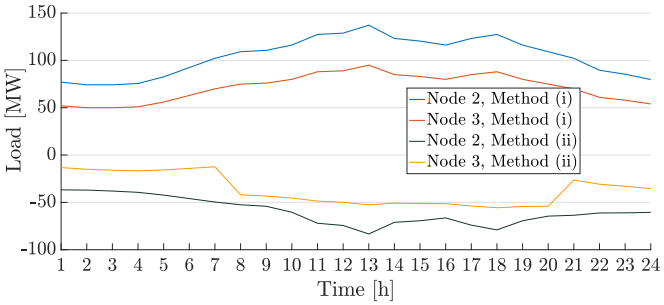

Fig. 5.8 Netload at nodes 2,3 with using methods (i) and (ii).

Last, we depict the hourly operational cost for the TSO and the DSOs in Fig. 5.9 which will be used to compare the two proposed schemes.

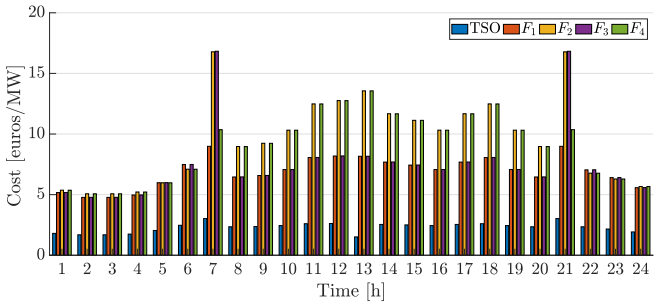

Fig. 5.9 TSO and DSOs operational cost using the proposed decentralised coordination scheme.

We next check the convergence properties of the proposed algorithm. In Figs. 5.10, 5.11, we illustrate the evolution of the hourly objective functions of F_2 and the transmission system for a 24-hour period with respect to the iteration numbers of algorithm. We notice that the algorithm converges after three iterations. To test the sensitivity of the proposed algorithm with respect to the initial point, i.e., the choice of initial load value for the distribution system, we changed the initial point to be full load, 85%, 75% and 65% of the full load. In all cases, the algorithm converges in three iterations. Next, to analyse the sensitivity of the proposed algorithm with respect to the level of distributed resources penetration, we depict in Fig. 5.12 the evolution of F_2 hourly cost for two different levels of penetration with the same initial point (step 3 of the algorithm) concerning the number of iterations. The final cost is different for the two cases since there are hours where the DG price is smaller than the grid price and vice versa.

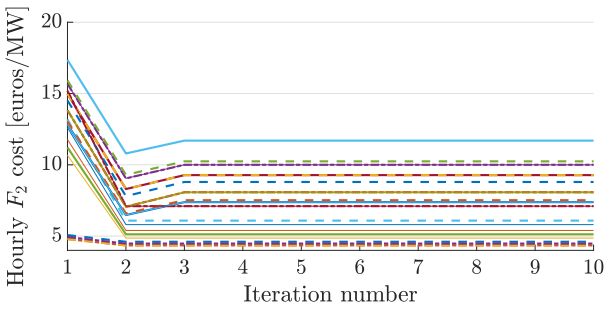

Fig. 5.10 Evolution of the hourly cost for F_2 with respect to the iteration number.

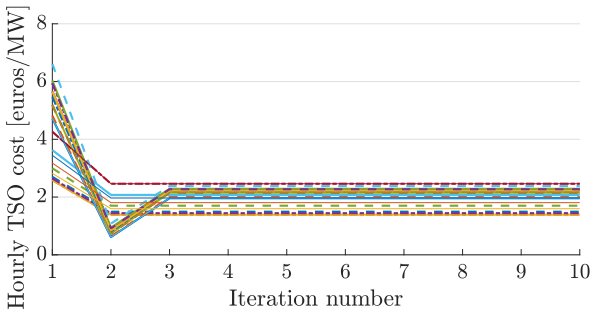

Fig. 5.11 Evolution of the hourly cost for the transmission system with respect to the iteration number.

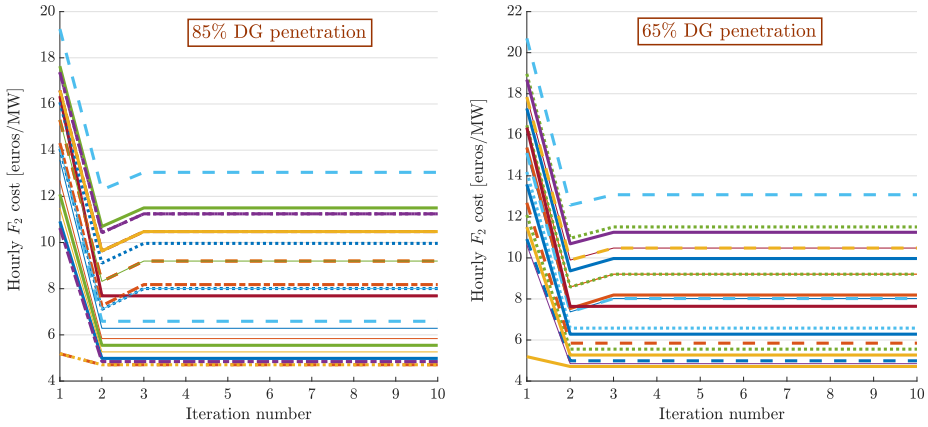

Fig. 5.12 Evolution of hourly cost for F_2 for different penetration levels of distributed generation.

5.4.3 Centralised Coordination Scheme

We apply the proposed scheme developed in Section 5.3 to the system described in Fig. 5.2. In order to demonstrate how the proposed centralised scheme can facilitate the integration of distributed energy resources, we compare method (i), which is the optimal operation with the current practice, with method (iii), which is the proposed centralised scheme. We start the simulation by assigning the same weights to the transmission cost function and the distribution feeders' cost functions as $w_1 = w_2 = 0.5$. The TSO cost, as depicted in Fig. 5.13, is reduced significantly with method (iii), i.e., the centralised scheme, in comparison to the

current practice due to the increase in the integration of the distributed resources at different nodes as shown in Fig. 5.14.

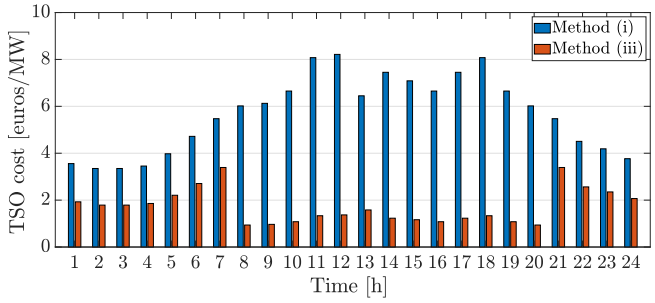

Fig. 5.13 Transmission operation cost for methods (i) current practise and (iii) proposed centralised TSO-DSO coordination scheme.

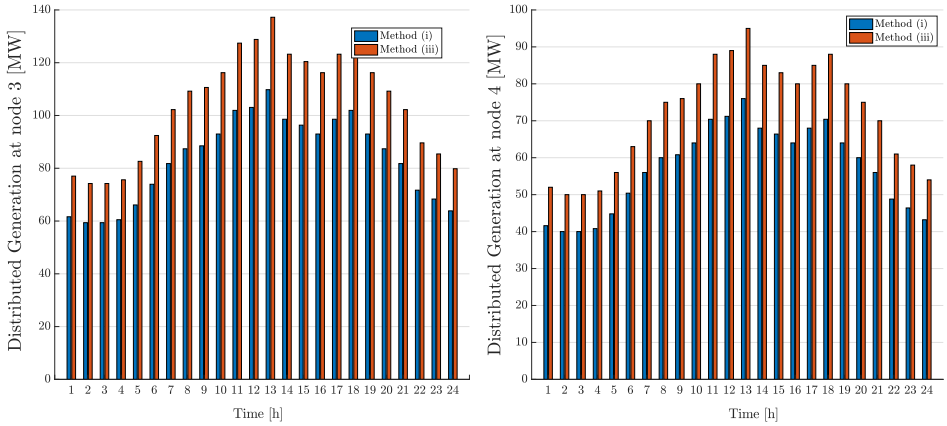

Fig. 5.14 The total amount of distributed generation for methods (i) current practise and (iii) proposed centralised TSO-DSO coordination scheme at nodes 3 and 4.

In Fig. 5.15, the netload at the transmission level using methods (i) and (iii) is depicted. We notice that it is more cost-efficient for the TSO to purchase power from the DG present in the distribution systems. For instance, the negative load at node 2 means that the excess power of the distributed resources is redirected to the transmission system. DGs usually sell at a price equal to the LMP at their PCC. This results in distributed resources' owners gaining revenue by selling power to the TSO, while the TSO also meets its load at a lower cost. In Fig. 5.16, the operational cost for each hour for the TSO and DSOs for the proposed centralised coordination scheme is depicted. Fig. 5.16 shows that the transmission cost for

method (iii) with $w_1 = w_2 = 0.5$ is lower than that of method (ii) as depicted in Fig 5.9. The difference is that more power is being used from the DGs in method (iii) compared to that of method (ii). However, we notice that the cost of feeders in method (iii) is higher than that of method (ii). Again, this is because more power is being used from the DGs in method (iii) compared to method (ii). DSOs and TSOs can use these values to formulate their bids and provide incentives for DG participation, respectively.

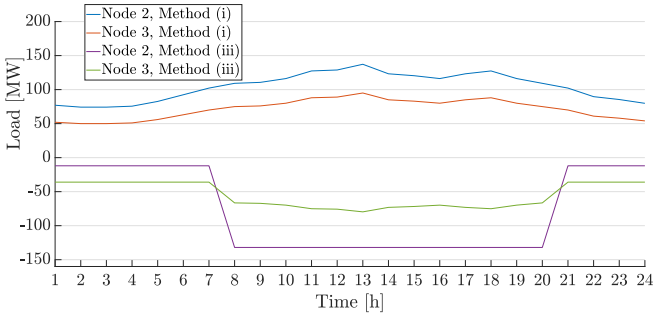

Fig. 5.15 Net load at nodes 2,3 with using methods (i) and (iii).

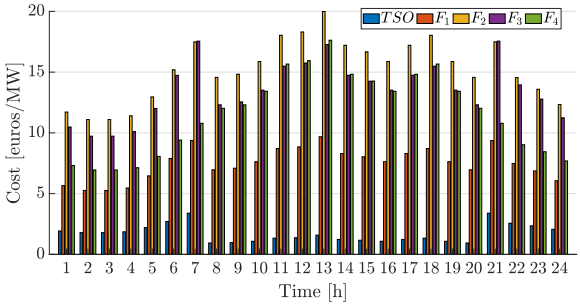

Fig. 5.16 TSO and DSOs operational cost using the proposed centralised coordination scheme.

Hour	P_{G_1}	P_{G_2}	P_{G_3}	P_{G_4}	P_{G_5}
1	52.05	0	0	0	110
2	42.45	0	0	0	110
3	42.45	0	0	0	110
4	47.25	0	0	0	110
5	71.25	0	0	0	110
6	102.64	2.2	0	0	110
7	110	10.87	17.58	0	110
8	0	0	0	0	88.4
9	0	0	0	0	90.88
10	0	0	0	0	100.81
11	10.67	0	0	0	110
12	13.15	0	0	0	110
13	28.05	0	0	0	110
14	3.22	0	0	0	110
15	0	0	0	0	108.26
16	0	0	0	0	100.81
17	3.22	0	0	0	110
18	10.67	0	0	0	110
19	0	0	0	0	100.81
20	0	0	0	0	88.4
21	110	10.87	17.58	0	110
22	95.25	0	0	0	110
23	80.85	0	0	0	110
24	61.65	0	0	0	110

Table 5.6 The power output in MW of generators at the transmission level for method (iii): proposed centralised TSO-DSO coordination.

The hourly power output of transmission generators for method (iii) is presented in Table 5.6. We notice that between hours 8 and 20, the distributed resources located in the distribution systems satisfy the load at the transmission level, whereas at night hours, mostly the TSO is responsible for supplying the load to the customers. This reverse power flow also impacts the LMP as shown in Table 5.7, where we notice a marginal increase in the LMPs for the night hours is achieved. Similar to method (ii), there is congestion at hours 7 and 21 due to the congested line between nodes 1 and 2.

Hour	Node 1	Node 2	Node 3	Node 4	Node 5
1	14.52	14.53	14.53	14.53	14.52
2	14.42	14.43	14.43	14.43	14.42
3	14.42	14.43	14.43	14.43	14.42
4	14.47	14.48	14.48	14.48	14.47
5	14.71	14.72	14.72	14.72	14.71
6	15.03	15.04	15.04	15.03	15.03
7	15.13	27.74	25.35	18.78	15.78
8	11.24	11.24	11.24	11.24	11.24
9	11.27	11.27	11.27	11.27	11.27
10	11.41	11.41	11.41	11.41	11.41
11	14.11	14.11	14.11	14.11	14.11
12	14.13	14.13	14.14	14.13	14.13
13	14.28	14.28	14.29	14.28	14.28
14	14.03	14.03	14.04	14.04	14.03
15	11.52	11.52	11.52	11.52	11.52
16	11.41	11.41	11.41	11.41	11.41
17	14.03	14.03	14.04	14.04	14.03
18	14.11	14.11	14.11	14.11	14.11
19	11.41	11.41	11.41	11.41	11.41
20	11.24	11.24	11.24	11.24	11.24
21	15.13	27.74	25.35	18.78	15.78
22	14.95	14.97	14.97	14.96	14.95
23	14.81	14.82	14.82	14.81	14.81
24	14.62	14.63	14.63	14.62	14.62

Table 5.7 Locational marginal prices for method (iii): proposed centralised TSO-DSO coordination in €/MW.

Next, we analyse the interaction between the TSO and the DSOs. For this, we modify the weights of (5.4) to obtain an approximation of the Pareto front. More specifically, we start with $w_1 = 0$ and $w_2 = 1$, and with increments of 0.05 we reach $w_1 = 1$ and $w_2 = 0$. The Pareto front is depicted in Fig. 5.17. The x axis shows the summation of all hourly costs as daily cost. By moving along the curve, we can minimise DSOs' objective at the expense of the TSO objective or minimise the TSO objective at the expense of DSOs' objective. However, we cannot improve both at once, i.e., there is no mathematical "best" point along the Pareto front.

To provide insights into the potential conflicts between TSOs and DSOs we discuss in greater detail the two extreme cases, i.e., $w_1 = 0$ and $w_2 = 1$ and $w_1 = 1$, and $w_2 = 0$. The

TSO and DSO costs for the first one are 0 €/MW and 500 €/MW, respectively, and for the latter, they are 140 €/MW and 0 €/MW, respectively. In other words, when the objective is only to minimise the TSO cost, all costs are being incurred by the DSOs and vice versa. In both cases, all constraints, e.g., voltage and thermal limits, are met; thus, the power system quality is guaranteed.

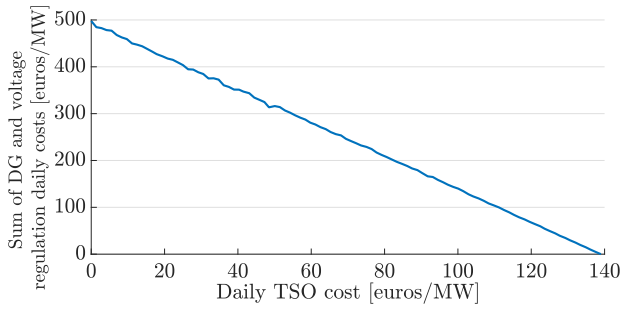

Fig. 5.17 Pareto Front of the sum of all feeders DG and voltage regulation daily cost with respect to the TSO cost.

In Fig. 5.18, we depict the total DSO cost that includes the payments to the TSO given in (3.9), DG cost given in (3.11) and (3.12), and voltage regulation costs given in (3.13). We compare the results for different weights with methods (i) and (ii). We notice that method (ii) results are close to the Pareto front offering a near-optimal solution. The appropriate choice of operation for the Pareto front is a balance of priorities between TSOs and DSOs and the determination of specific incentives, which are part of future work. Another implication of the Pareto front is that any point in the feasible region that is not on the Pareto front is not considered to be a "good" solution, e.g., method (i). Either objective or both can be improved at no penalty to the other. This demonstrates that there are many improvements to be made to the current TSO-DSO coordination practice, i.e., method (i). To determine the priorities of the proposed decentralised scheme, we have to analyse where its solution lies in the Pareto front. More specifically, we notice in Figs. 5.18 and 5.19 that the proposed decentralised scheme provides a balance between the TSO and DSO objective as it lies between the two extreme cases.

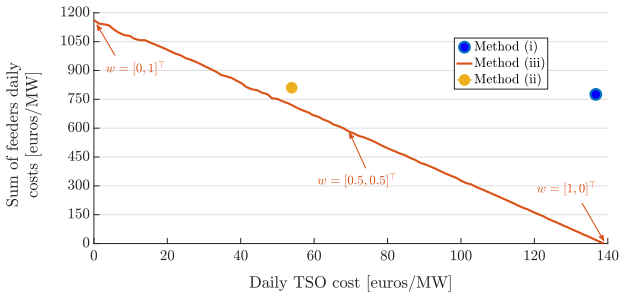

Fig. 5.18 Pareto Front of the sum of all feeders daily cost with respect to the TSO cost.

Next, we depict in Fig. 5.19 the daily cost of individual feeders, which includes the payments to the TSO, the cost of DG and voltage regulation, to investigate how far from the optimal solution each feeder operates for the various schemes. We notice that for method (ii), F_2 operates at the optimum, F_3 at a point that is at the expense of other feeders, and F_1 and F_4 are at points further away from the optimal solutions. However, the summation of these costs corresponds to a near optimal solution as seen in Fig. 5.18.

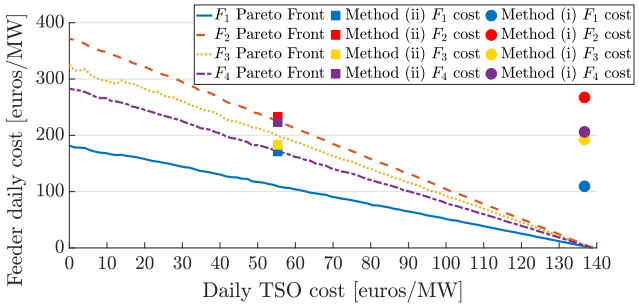

Fig. 5.19 Pareto Front of daily cost for F_i, $i = 1,\ldots,4$ with respect to the TSO cost.

In both schemes, the transmission cost decreases, while for method (iii), the transmission operation cost reduction is higher than that of method (ii). In comparison to the current practice, i.e., method (i), both schemes are more effective in terms of the share contribution of the distributed generators at each transmission node, while the utilisation rate of generation for method (iii) is higher than that of method (ii). Using method (iii), we can see that the output of each generator at the transmission level is lower than that of method (ii) and for method (ii) is lower than that of method (i). Although for method (ii) and method (iii), the congestion level is improved, the LMP for each node at each hour is higher at night hours in

method (iii). This is due to the increased output of transmission generators at night hours. In all case studies, all variables, e.g., voltage levels, transmission line flows, are kept within the limits of acceptable power quality purposes as defined by the constraints of the OPFs. For example, voltage levels of each bus in the distribution system at every time interval are in the range of 0.95 – 1.05 pu. The algorithm running time for the centralised scheme is 12,387 msec, and for the decentralised is 21,800 msec in a Windows machine which is equipped with AMD® FX-9830P RADEON R7 CPU with four Cores at 3.00 GHz and 16 GB of RAM. As expected, the centralised scheme is approximately two times faster; however, both schemes are fast enough for real-time operation purposes.

5.5 Conclusion

In this chapter, we have presented a novel TSO-DSO coordination framework that increases the efficient use of distributed generation resources. More specifically, we have two coordination schemes: one centralised, another decentralised. The underlying network for both systems is approximated linear, and the OPF formulations result in convex optimisation problems. We have formulated a decentralised TSO-DSO coordination scheme based on an iterative approach where no sensitive information is exchanged that achieves a near-optimal solution. Next, we have analysed the interaction of TSOs and DSOs and how conflicting their objectives are by approximating the Pareto front of a multi-objective OPF problem where the entire system, i.e., transmission and distribution systems, is modelled. Through numerical results, we have demonstrated that both coordination schemes result in (i) reduced operational costs for both TSOs and DSOs, (ii) congestion relief, and (iii) increased use of distributed generation.

In the two proposed schemes, different entities are responsible for making a decision, and thus diverse information is shared between them. In particular, in the centralised scheme, the TSO makes the decisions and has access to all information about the underlying physical distribution systems as well as DG bidding. In the decentralised scheme, both the DSO and TSO share the decision-making-process, and the only information that the TSO sends the

DSO is the LMP at the PCC and the DSO to the TSO its netload. The two proposed methods also differ in the total cost, level of DG integration, voltage levels and level of congestion, as demonstrated in the numerical results' section. These affect the "power quality" of the system. However, all variables, e.g., voltage levels, transmission line flows, are kept within the limits of acceptable power quality purposes as defined by the constraints of the OPFs.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Due to climate concerns, the integration of renewable energy resources into the power system has gained popularity. However, the large-scale penetration of these resources into the conventional power system brings challenges to the system. This dissertation focused on different aspects of this problem by addressing them in a cascade order. Firstly, we address the power system optimal operation by linearising the optimal power flow considering integrating different renewable energy resources into the power distribution system. Solving the first problem, we came up with the new research concern, an accurate probabilistic forecast model for the renewable energy output such as PV. The answers to the previous research questions corroborate the network operators' critical role in transitioning to the carbon-free power system. We addressed the concern in the last research question in the last chapter. The results are elaborated in detail in the following paragraphs.

As we incorporate the DGs into the power system, the distribution problem's operation gets more complicated. The reason is the intermittency of these resources; they inject a considerable amount of uncertainties into the system. However, the active power injection to the distribution power system can help decrease the load demand. In contrast, the price of such RESs makes the operation cost more expensive. To this end, proposing optimal operation management and the output of these RESs can minimise operating costs and mitigate the

voltage deviation. To better make use of renewable energy, microgrids have gained popularity in recent years. Microgrids are distribution systems comprising different renewable energy resources that can satisfy the local load in a stand-alone mode. Simultaneously, they can also connect to the power system, sell the excess energy to the system, and profit.

This dissertation proposed deterministic optimal energy management of a microgrid connected to one 33-bus IEEE standard feeder. Due to the bidirectional load flow and the optimal power flow's nonlinear nature, the optimal power flow problem becomes nonlinear and nonconvex. Therefore, we proposed a linearised optimal power load flow. We solved the problem using Quadratic programming. The results explicitly show that although the PV, WT, and hydroelectric plants are the most expensive units, it is more beneficial to buy energy from them to benefit from using green energy to produce energy. The battery starts to charge when the energy cost is low, while it gets discharged in intervals when the cost is higher. After proposing the deterministic framework for the optimal microgrid management to model the problem in a more realistic environment, we researched the solar output forecast model.

As mentioned above, the inherent uncertainty involved with renewable energy resources such as PV makes integration more difficult. The solar output depends on the weather data, such as solar irradiance and temperature. After exploring relevant studies on different forecast techniques, we found that most models relied on the seasonal trend of solar output rather than the time of the day. Also, we found GPR as one of the widely-used methods in building forecast models. GPR has been used since the 1940s as a tool for prediction, but it is becoming increasingly popular due to its flexibility to be applied to a wide range of time-series data. During the past two decades, energy systems using machine learning (ML) for different purposes such as forecasting are in higher demand. Based on one different energy benchmarking technique, they can be categorised into three different groups. The first ones are based on mathematical and physical constraints that need a lot of simulation and documentation; the second techniques, such as support vector machine (SVM) and artificial neural network method (ANN), require historical statistical data for training. The third method, such as GPR, use historical data and d data fitting approaches to build a robust

model. GPR is a good approach based on meteorological data uncertainties, since it can calculate the parameters' uncertainties. Moreover, the model's nonlinear performance can explicitly be modelled by this method. Compared to other methods, GPR method is more efficient for predicting time-series, with a wide range variation of each hour of a day over one year.

Thus, identifying the highly correlated features, direct solar irradiance, diffused solar irradiance, horizontal solar irradiance, temperature, zenith, and azimuth, we clustered our data in four groups based on day-time. We utilised different techniques, such as Elbow, Gap, and k-means, to determine the optimal number of clusters to cluster the data into four groups. Afterwards, we trained a model for each of the clusters using GPR. The kernel we used is Matérn $5/2$, which has not been used for solar output forecast before; however, it proved that has plenty of capabilities in dealing with uncertainty in other forecast models. We ran each model 30 times and tested the accuracy of the model. We used a mixture of two different test and validation techniques, i.e., 5-fold cross-validation and hold-out validation methods. We proceeded to validate our results based on implementing the model on five various sites. We held out 30 random days as a test set as the representatives of different days in different seasons.

The results clearly showed there are similar data points in different seasons that can be grouped in the same cluster in our model, while in the seasonal model, they happen to be grouped in different seasons. Furthermore, choosing the appropriate number of clusters can significantly improve the results, as RMSE for eight clusters is 2.4%, while this metric value is 8.34% without clustering. For the four clusters, the choice of our study, this value is 4.12%. The largest RMSE and MAE were 4.60 % and 2.06 %, respectively. Testing the model implementation on different sites, we found our methodology to improve the state-of-the-art results significantly. In other words, the clustering technique we used in our model grouped similar points into the same cluster. Therefore, the variance of each cluster reduces and the accuracy increased.

After proposing the forecast model for solar output forecast, we investigated the impact of renewable energy, i.e., PV-Battery systems, on the TSO and DSO interactions. The reason

is that integrating renewable energy into the power system needs collaboration between the DSO and TSO to operate the system reliably. Since the energy that RESs generate are non-dispatchable, they need to be either consumed or sold to the upper-hand power system. In this dissertation, we proposed two different centralised and decentralised linearised coordination schemes.

In the first scheme, as the decentralised scheme, we modelled the problem into an iterative bi-level optimisation problem, while in the upper-level, the TSO operates the system and bids the LMP to the DSO. In other words, the DSO and TSO collaborate to solve the optimal operation problem in this scheme. At the upper level, the TSO operates the system and bids the LMP to the DSO. The DSO optimises its network to evaluate the PV-Batt generations and send back the new load into the TSO. Iteratively, the problem is solved. We demonstrated that the problem is converged to a near-optimal solution. Afterwards, we investigated the interactions between the DSO and TSO through a centralised coordination scheme and a multi-objective optimisation problem. For this, we approximated a Pareto Front to show the existence of any conflicting objectives. These schemes were implemented on a five-bus transmission system with four different distribution feeders, IEEE 33- and 69-bus feeders, connected to its different nodes. The results showed improvement in LMP and congestion management. Also, the decrease in the operation cost and the increase in the share of renewable energy in both transmission and distribution systems proves the proposed framework's effectiveness.

Different entities are responsible for making a decision, and diverse information is shared between them. In particular, in the centralised scheme, the TSO makes the decisions and has access to all information about the underlying physical distribution systems as well as DG bidding. In the decentralised scheme, both the DSO and TSO share the decision making process, and the only information that the TSO sends the DSO is the LMP at the PCC and the DSO to the TSO its netload. The two proposed methods also differ in the total cost, level of DG integration, voltage levels and level of congestion, as demonstrated in the numerical results' section. These affect the "power quality" of the system. However, all variables, e.g.,

voltage levels, transmission line flows, are kept within the limits of acceptable power quality purposes as defined by the constraints of the OPFs.

All projects have their strengths and limitations. This work is also not an exception, and I have faced some challenges during my PhD. The most significant challenge was having to conduct a substantial portion of this research during a global pandemic. As a limitation to the Microgrid optimal operation, we could model the probabilistic model considering the uncertainty of load, WT and utility price. However, time limitations hindered me from pursuing this. One of the limitations that we successfully overcame during the work was the lack of a concrete dataset for the solar output forecast. We spent much time collecting three different datasets from different resources to make a consistent dataset. We made the dataset publicly available to other researchers.

6.2 Future Work

In this part, we are going to have an overview of future work. There are a number of ways we can further improve this research in the future, such as:

6.2.1 More detailed coordination scheme

There are natural extensions of the work we presented in Chapter 5. For instance, a distributed solution of the proposed centralised scheme is necessary, so that system operators do not share sensitive information about their topology and generators bids. Moreover, a more detailed representation of the topology of the distribution system would provide more accurate results as well as the incorporation of uncertainty in renewable-based generation. We will report on these developments in future papers.

6.2.2 Probabilistic model

With new power system technologies and renewable energy generators, traditional power system operators face new challenges. Some of these technologies, such as microgrids,

can locally supply energy for their consumers and change the power system voltage and frequency. At the same time, they insert uncertainty into the system and make the operation more complex and challenging. Therefore, to achieve optimal operation, probabilistic modelling of uncertainty is necessary as much as finding the tools for optimizing generation and voltage regulation costs. In this part, Deep Reinforcement Learning Solution (DRL), Neural Network tools, Monte Carlo (MC) modelling and Fuzzy Logic tools can be helpful for future work.

6.2.3 Electro-thermal model of PV

Using the PV thermal systems as an asset to systems can improve solar systems since we can use them to supply the heating loads in MG problems. The model for one PVT is presented in [42]. The idea of PVT systems is new. Different designs and novel applications for them has been introduced recently in [177–181].In future work, we can use electro thermal PV plant to improve the PV system efficiency as well as using the Electro thermal PV for supplying hot water for residential loads.

6.2.4 Model of pump hydro

Pumped storage hydroelectric plants are designed to serve the peak load at certain times, e.g., peak hours, with hydroelectric energy and then pumping the water back up into the reservoir at other times, e.g., light load periods. Thus, two intervals need to be considered when modelling the operation of pumped storage hydroelectric systems: intervals of generation and intervals of pumping. In any one interval, the plant can be (i) pumping or (ii) generating. The idle case may be presented as either a pump or generate. The model for a Pumped Storage Hydro Cascade is presented in [182]. We try to model pump hydro and utilize the model as a storage device instead of a battery in future work. Then in our future work, we can compare the performance of these storage-based units compared to normal lithium-ion batteries.

6.2.5 Model of hydro plant

In [183] a robust optimal operation scheme is implemented on a real cascade hydroelectric power system to maximizing the head level of the dams as well as minimizing the spillage effects. The focus of the optimization problem is to make the head level of the dams robust against uncertainty. The case study is based on the Seven Fork system located in Kenya, which includes five cascaded hydroelectric power systems. In [184] the authors tried to highlight the importance of low-carbon strategies and technologies by simulating a high-resolution electricity demand model to address the requirements needed for the integration of renewable energy resources into the power system. In [185] an optimization approach is proposed for one photovoltaic-wind hybrid system considering the size, cost, and availability. In [186, 187] the value of hybrid solar-hydro systems are discussed. Basically, in mentioned work, these models were not implemented on a typical microgrid to observe their effect on the performance of microgrid operation. In our future work, we aim to use the hydroelectric and pump hydro storage models to see their effects on the new model's performance.

6.2.6 Hyperparameters Optimization

We can optimise the hyperparameters of the kernel function to improve the forecast accuracy. Based on a very recent study [188], optimising the hyperparameters of a kernel function can have great effects on the performance of a GPR model. Moreover, a thorough analysis of the sky coverage effect in solar power output could be investigated. More specifically, knowledge of the density, thickness, and type of clouds can help to improve the solar output prediction model. The use of sky images [189] can help in this direction.

6.2.7 Hosting Capacity

The integration of renewable energy resources into the power system has generated a significant amount of coverage and discussion. As we discussed in this dissertation, the precise estimation of the power system's need to integrate these resources into the system is also

critical. In future work, we can focus on the hosting capacity and model the linearized framework problem.

References

[1] Faranak Golestaneh, Pierre Pinson, and Hoay Beng Gooi. Very short-term nonparametric probabilistic forecasting of renewable energy generation— with application to solar energy. *IEEE Transactions on Power Systems*, 31:3850–3863, 2016.

[2] F. Golestaneh and Hoay Beng Gooi. Batch and sequential forecast models for photovoltaic generation. In *2015 IEEE Power Energy Society General Meeting*, pages 1–5, July 2015.

[3] Mehreen Gul, Yash Kotak, and Tariq Muneer. Review on recent trend of solar photovoltaic technology. *Energy Exploration & Exploitation*, 34(4):485–526, 2016.

[4] D. Apostolopoulou and M. McCulloch. Optimal short-term operation of a cascaded hydro-solar hybrid system: A case study in kenya. *IEEE Transactions on Sustainable Energy*, 10(4):1878–1889, 2019.

[5] Mehreen Gul, Yash Kotak, and Tariq Muneer. Review on recent trend of solar photovoltaic technology. *Energy Exploration & Exploitation*, 34(4):485–526, 2016.

[6] Trevor M. Letcher. 1 - why solar energy? In Trevor M. Letcher and Vasilis M. Fthenakis, editors, *A Comprehensive Guide to Solar Energy Systems*, pages 3 – 16. Academic Press, 2018.

[7] Ricardo J. Bessa, Jethro Dowell, and Pierre Pinson. *Renewable Energy Forecasting*, pages 1–21. American Cancer Society, 2016.

[8] Hao Zhu and Hao Jan Liu. Fast local voltage control under the limited reactive power: optimality and stability analysis. *IEEE trans on power systems*, 31(5):3794 – 3803, Feb 2016.

[9] Baosen Zhang, Albert.Y.S, Alejandro, and D.Dominguez-Garcia. An optimal and distributed method for voltage regualtion in power distribution system. *IEEE Trans. Power Syst*, 30(4):1714–1725, 2015.

[10] Rakibuzzaman Shah, N. Mithulananthan, R.C. Bansal, and V.K. Ramachandaramurthy. A review of key power system stability challenges for large-scale pv integration. *Renewable and Sustainable Energy Reviews*, 41:1423 – 1436, 2015.

[11] D. Apostolopoulou, A. D. Dominguez-Garcia, and P. W. Sauer. An assessment of the impact of uncertainty on automatic generation control systems. *IEEE Transactions on Power Systems*, 31(4):2657–2665, 2016.

[12] R. G. Wandhare and V. Agarwal. Reactive power capacity enhancement of a pv-grid system to increase pv penetration level in smart grid scenario. *IEEE Transactions on Smart Grid*, 5(4):1845–1854, July 2014.

[13] Helena Gerard, Enrique Israel Rivero Puente, and Daan Six. Coordination between transmission and distribution system operators in the electricity sector: A conceptual framework. *Utilities Policy*, 50:40–48, 2018.

[14] Seppo Kärkkäinen et al. Integration of demand-side management, distributed generation, renewable energy sources and energy storages. *Report Task XVII Integration of Demand-Side management, Distributed Generation, Renewable Energy Sources and Energy Storages*, 1:77, 2008.

[15] Zhao Yuan and Mohammad Reza Hesamzadeh. Hierarchical coordination of tso-dso economic dispatch considering large-scale integration of distributed energy resources. *Applied Energy*, 195:600–615, 2017.

[16] Samson Yemane Hadush and Leonardo Meeus. Dso-tso cooperation issues and solutions for distribution grid congestion management. *Energy Policy*, 120:610–621, 2018.

[17] G de Jong, O Franz, P Hermans, and M Lallemand. Tso-dso data management report. *TSO-DSO Project Team, Tech. Rep*, 2016.

[18] Michael Birk, José Pablo Chaves-Ávila, Tomás Gómez, and Richard Tabors. Tso/dso coordination in a context of distributed energy resource penetration. In *Proc. EEIC*, pages 2–3, 2016.

[19] Konstantin Turitsyn, Petr Šulc, Scott Backhaus, and Michael Chertkov. Options for control of reactive power by distributed photovoltaic generators. In *Proceedings of the IEEE*, 2011.

[20] NS Rau and W Yih-Heui. Optimum location of resources in distributed planning. *IEEE Trans Power Syst*, 9:2014–20, 1994.

[21] Dimitra Apostolopoulou, Peter W. Sauer, and Alejandro D. Domínguez-García. Distibuted optimal load frequency control and balancing area coordination. *2015 North American Power Symposium (NAPS)*, 2015.

[22] Shuhui Li, Xingang Fu, Ishan Jaithwa, Michael Fairbank Eduardo Alonso, and Donald C. Wunsch. Control of three-phase grid-connected microgrids using artificial neural networks. *International Joint Conference on Computational Intelligence (IJCCI)*, Aug 2016.

[23] Shuhui Li, Michael Fairbank, Donald C. Wunsch, and Eduardo Alonso. Vector control of a grid-connected rectifier/inverter using an artificial neural network. *IEEE International Joint Conference on Neural Networks*, 2012.

[24] Tine L Vandoorn, Bert Renders, Lieven Degroote, Bart Meersman, and Lieven Vandevelde. Active load control in islanded microgrids based on the grid voltage. *IEEEtrans*, 2(1):139 – 151, 2011.

[25] H.A Mahmoud Pesaran, Huya Phung Dang, and K.Ramachandaramurthy Vigna. A review of the optimal allocation of distributed generation: Objectives, constraints, methods, and algorithms. *Renewable energy and sustainable energy reviews*, 75:293–312, Aug 2017.

[26] Sudipta Chakraborty, Manoja D. Weiss, and M. Godoy Simoes. Distributed intelligent energy management system for a singlephase high-frequency ac microgrid. *IEEE Trans Ind Electronics*, 54(1):97 – 109, 2007.

[27] Chakraborty Sudipta, D. Weiss Manoja, and Simoes M. Godoy. Fuzzy artmap based forecast of renewable generation for a high frequency ac microgrid. *31st Annual Conference of IEEE Industrial Electronics Society*, 2006.

[28] Taher Niknam, Faranak Golestaneh, and Ahmadreza Malekpour. Probabilistic energy and operation management of a microgrid containing wind/photovoltaic/fuel cell generation and energy storage devices based on point estimate method and self-adaptive gravitational search algorithm. *Energy*, 43(1):427–437, 2012.

[29] Taher Niknam, Faranak Golestaneh, and Mehdi Shafiei. probabilistic energy management of a renewable microgrid with hydrogen storage using self-adaptive charge search algorithm. *Energy*, 49(1):252–267, 2013.

[30] Amir-Hamed Mohsenian-Rad and Alberto Leon Garcia. Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE Transactions on Smart Grid*, 1(2):120–133, 2010.

[31] Jamshid Aghaei, Vassilios G. Agelidis, Fatima Raeisi Mansour Charwand, Abdollah Ahmadi, Ali Esmaeel Nezhad, and Alireza Heidari. Optimal robust unit commitment of chp plants in electricity markets using information gap decision theory. *IEEE Transactions On smart grid*, 8(5):2296 – 2304, 2017.

[32] Guodong Liu, Yan Xu, and Kevin Tomsovic. Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization. *IEEE Transactions on Smart Grid*, 7(1):227 – 237, Jan 2016.

[33] Hang Shuai, Jiakun Fang, Xiaomeng Ai, Yufei Tang, Jinyu Wen, and Haibo He. Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming. *IEEE Transactions on Smart Grid*, pages 1 – 1, Jan 2018.

[34] Indragandhi V., Logesh R., , Subramaniyaswamy V, Vijayakumar V, Patrick Siarry, and Lorna Udene. Multi-objective optimization and energy management in renewable based ac/dc microgrid. *Computers and Electrical Engineering*, 70:179–198, Feb 2018.

[35] Wencong Su, Jianhui Wang, and Jaehyung Roh. Stochastic energy scheduling in microgrids with intermittent renewable energy resources. *IEEE Transactions on Smart Grid*, 5(4):1876 – 1883, July 2014.

[36] Alessandra Parisio, Evangelos Rikos, and Luigi Glielmo. A model predictive control approach to microgrid operation optimization. *IEEE Transactions on Control Systems Technology*, 22(5):1813–1827, 2014.

[37] Gang Yang, Junwei Cao, Haochen Hua, and Ziqiang Zhou. Deep learning-based distributed optimal control for wide area energy internet. *IEEE International Conference on Energy Internet (ICEI)*, July 2018.

[38] Z. Zhou, F. Xiong, B. Huang, C. Xu, R. Jiao, B. Liao, Z. Yin, and J. Li. Game-theoretical energy management for energy internet with big data-based renewable power forecasting. *IEEE Access*, 5:5731–5746, 2017.

[39] Bingke Yan, Bo Wang, Lin Zhu, Hesen Liu, Yilu Liu, Xingpei Ji, and Dichen Liu. A novel, stable, and economic power sharing scheme for an autonomous microgrid in the energy internet. *Energies*, 8(11):12741–12764, 2015.

[40] A. Chaouachi, R. M. Kamel, R. Andoulsi, and K. Nagasaka. Multiobjective intelligent energy management for a microgrid. *IEEE Transactions on Industrial Electronics*, 60(4):1688–1699, April 2013.

[41] Aliasghar Baziar and Abdollah Kavousi-Fard. Considering uncertainty in the optimal energy management of renewable micro-grids including storage devices. *Renewable Energy*, 59:158–166, Nov 2013.

[42] Fatemeh Najibi, Taher Niknam, and Abdollah Kavousi-Fard. Optimal stochastic management of renewable MG (micro-grids) considering electro-thermal model of PV (photovoltaic). *Energy*, 97:444–459, Jan. 2016.

[43] Fatemeh Najibi and Taher Niknam. Stochastic scheduling of renewable micro-grids considering photovoltaic source uncertainties. *Energy conversion and management*, 98:484–499, April. 2015.

[44] A Al-Sabounchi, J Gow, M Al-Akaidi, and H. Al-Thani. Optimal sizing and location of a pv system on three-phase unbalanced radial distribution feeder avoiding reverse. *power flow. in: Electrical Power and Energy Conference (EPEC) IEEE*, 2011.

[45] NC Hien, N Mithulananthan, and Bansal RC. location and sizing of distributed generation units for loadabilty enhancement in primary feeder. *IEEE Trans. Syst*, (7):797–806, 2013.

[46] LD Arya, A Koshti, and SC Choube. Distributed generation planning using differential evolution accounting voltage stability consideration. *Int J Electr Power Energy Syst*, 42:196–207, 2012.

[47] Q Kang, M Zhou, J An, and Q Wu. Swarm intelligence approaches to optimal power flow problem with distributed generator failures in power networks. *Automation Sci Eng IEEE Trans*, 10:343–53, 2013.

[48] Z Liu, F Wen, G Ledwich, and X Ji. Optimal sitting and sizing of distributed generators based on a modified primal-dual interior point algorithm in: Electric utility deregulation and restructuring and power technologies (drpt). *Proceedings of the 4th International Conference on. IEEE*, pages 1360–1365, 2011.

[49] W Muneer, K Bhattacharya, and Canizares. Large-scale solar pv investment models, tools, and analysis: the ontario case. *IEEE Trans. Power Syst*, 26:2547–55, 2011.

[50] MF Shaaban, YM Atwa, and EF El-Saadany. Dg allocation for benefit maximization in distribution networks. *IEEE Trans Power Syst*, 28:639–49, 2013.

[51] B Banerjee and SM Islam. Reliability based optimum location of distributed generation. *Int J Electr Power Energy Syst*, 33:1470–8, 2011.

[52] A. Bracale, G. Carpinelli, and P. De Falco. A probabilistic competitive ensemble method for short-term photovoltaic power forecasting. *IEEE Transactions on Sustainable Energy*, 8(2):551–560, April 2017.

[53] M. Rana, I. Koprinska, and V. G. Agelidis. Forecasting solar power generated by grid connected pv systems using ensembles of neural networks. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, July 2015.

[54] Manoja Kumar Behera, Irani Majumder, and Niranjan Nayak. Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. *Engineering Science and Technology, an International Journal*, 21(3):428 – 438, 2018.

[55] Fatemeh Najibi and Taher Niknam. Stochastic scheduling of renewable micro-grids considering photovoltaic source uncertainties. *Energy Conversion and Management*, 98:484 – 499, 2015.

[56] C. Wan, J. Lin, Y. Song, Z. Xu, and G. Yang. Probabilistic forecasting of photovoltaic generation: An efficient statistical approach. *IEEE Transactions on Power Systems*, 32(3):2471–2472, May 2017.

[57] M. J. Sanjari and H. B. Gooi. Probabilistic forecast of pv power generation based on higher order markov chain. *IEEE Transactions on Power Systems*, 32(4):2942–2952, July 2017.

[58] N. Liang, G. Huang, P. Saratchandran, and N. Sundararajan. A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Transactions on Neural Networks*, 17(6):1411–1423, Nov 2006.

[59] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, and Alexis Fouilloy. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569 – 582, 2017.

[60] Hugo T.C. Pedro and Carlos F.M. Coimbra. Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy*, 86(7):2017 – 2028, 2012.

[61] J. Shi, W. Lee, Y. Liu, Y. Yang, and P. Wang. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. *IEEE Transactions on Industry Applications*, 48(3):1064–1069, May 2012.

[62] A. Ragnacci, M. Pastorelli, P. Valigi, and E. Ricci. Exploiting dimensionality reduction techniques for photovoltaic power forecasting. In *2012 IEEE International Energy Conference and Exhibition (ENERGYCON)*, pages 867–872, Sep. 2012.

[63] David Knebel. *Simplified energy analysis using the modified bin method*. Atlanta, Ga.; American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 1983.

[64] Patrice Pinel, Cynthia A. Cruickshank, Ian Beausoleil-Morrison, and Adam Wills. A review of available methods for seasonal storage of solar thermal energy in residential applications. *Renewable and Sustainable Energy Reviews*, 15(7):3341 – 3359, 2011.

[65] Mohammad Saad Al-Homoud. Computer-aided building energy analysis techniques. *Building and Environment*, 36(4):421 – 433, 2001.

[66] Carl Edward Rasmussen. Gaussian processes for machine learning. In *Gaussian processes for machine learning*. MIT Press, 2006.

[67] Chao Xu, Zhiyong Feng, and Zhaopeng Meng. Affective experience modeling based on interactive synergetic dependence in big data. *Future Generation Computer Systems*, 54:507 – 517, 2016.

[68] Gábor I. Nagy, Gergő Barta, Sándor Kazi, Gyula Borbély, and Gábor Simon. Gefcom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. *International Journal of Forecasting*, 32(3):1087 – 1093, 2016.

[69] Zhengwei Li, Yanmin Han, and Peng Xu. Methods for benchmarking building energy consumption against its past or intended performance: An overview. *Applied Energy*, 124:325 – 334, 2014.

[70] Yeonsook Heo and Victor M. Zavala. Gaussian process modeling for measurement and verification of building energy savings. *Energy and Buildings*, 53:7 – 18, 2012.

[71] R Garnett, MA Osborne, S Reece, A Rogers, and SJ Roberts. Sequential bayesian prediction in the presence of change points and faults. *Computer Journal*, 53(9):1430–1446, 2010.

[72] Fatemeh Najibi, Dimitra Apostolopoulou, and Eduardo Alonso. Gaussian process regression for probabilistic short-term solar output forecast, 2020.

[73] Sun Bin, Yao Haitao, and Liu Ting. Short-term wind speed forecasting based on gaussian process regression model [j]. *Proceedings of the CSEE*, 32(29):104–109, 2012.

[74] Tao Hong, Pierre Pinson, and Shu Fan. Global energy forecasting competition 2012. *International Journal of Forecasting*, 30(2):357 – 363, 2014.

[75] Xiaoqian Jiang, Bing Dong, Le Xie, and Latanya Sweeney. Adaptive gaussian process for short-term wind speed forecasting. In *ECAI*, pages 661–666, 2010.

[76] N. Chen, Z. Qian, I. T. Nabney, and X. Meng. Wind power forecasts using gaussian processes and numerical weather prediction. *IEEE Transactions on Power Systems*, 29(2):656–665, 2014.

[77] S. Fang and H. Chiang. A high-accuracy wind power forecasting model. *IEEE Transactions on Power Systems*, 32(2):1589–1590, 2017.

[78] Hanmin Sheng, Jian Xiao, Yuhua Cheng, Qiang Ni, and Song Wang. Short-term solar power forecasting based on weighted gaussian process regression. *IEEE Transactions on Industrial Electronics*, 65(1):300–308, 2017.

[79] Astrid Dahl and Edwin V Bonilla. Grouped gaussian processes for solar power prediction. *Machine Learning*, 108(8-9):1287–1306, 2019.

[80] Michael C. Caramanis, Evgeniy Goldis, Pablo A. Ruiz, and Alexandr Rudkevich. Power market reform in the presence of flexible schedulable distributed loads. new bid rules, equilibrium and tractability issues. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1089–1096, 2012.

[81] Fatemeh Najibi, Dimitra Apostolopoulou, and Eduardo Alonso. Enhanced performance gaussian process regression for probabilistic short-term solar output forecast. *International Journal of Electrical Power & Energy Systems*, 130:106916, 2021.

[82] Fatemeh Najibi, Eduardo Alonso, and Dimitra Apostolopoulou. Optimal dispatch of pumped storage hydro cascade under uncertainty. In *2018 UKACC 12th International Conference on Control (CONTROL)*, pages 187–192. IEEE, 2018.

[83] Fatemeh Najibi, Taher Niknam, and Abdollah Kavousi-Fard. Optimal stochastic management of renewable mg (micro-grids) considering electro-thermal model of pv (photovoltaic). *Energy*, 97:444 – 459, 2016.

[84] Fatemeh Najibi and Taher Niknam. Stochastic scheduling of renewable micro-grids considering photovoltaic source uncertainties. *Energy Conversion and Management*, 98:484–499, 2015.

[85] Araz Ashouri, Peter Sels, Guillaume Leclercq, Olivier Devolder, Frederik Geth, and Reinhilde D'hulst. Smart TSO-DSO interaction schemes, market architectures, and ICT solutions for the integration of ancillary services from demand-side management and distributed generation Network and market models, EU Report, Apr. 2017. (691405), 2017.

[86] Arthur Gonçalves Givisiez, Kyriacos Petrou, and Luis F. Ochoa. A review on tso-dso coordination models and solution techniques. *Electric Power Systems Research*, 189:106659, 2020.

[87] Julia Merino, Inés Gómez, Elena Turienzo, Carlos Madina, I Cobelo, A Morch, H Saele, K Verpoorten, ER Puente, S Häninnen, et al. Ancillary service provision by res and dsm connected at distribution level in the future power system. *SmartNet project D*, 1:1, 2016.

[88] L. Kristov, P. De Martini, and J. D. Taft. A tale of two visions: Designing a decentralized transactive electric system. *IEEE Power and Energy Magazine*, 14(3):63–69, 2016.

[89] Nikolaos Savvopoulos, Theodoros Konstantinou, and Nikos Hatziargyriou. Tso-dso coordination in decentralized ancillary services markets. In *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6. IEEE, 2019.

[90] Stephan Dempe, Vyacheslav Kalashnikov, Gerardo A Pérez-Valdés, and Nataliya Kalashnykova. Bilevel programming problems. *Energy Systems. Springer, Berlin*, 2015.

[91] Anthony Papavasiliou. Analysis of distribution locational marginal prices. *IEEE Transactions on Smart Grid*, 9(5):4872–4882, 2017.

[92] Etienne Sorin, Lucien Bobo, and Pierre Pinson. Consensus-based approach to peer-to-peer electricity markets with product differentiation. *IEEE Transactions on Power Systems*, 34(2):994–1004, 2018.

[93] Alexandre Beluco, Paulo Kroeff de Souza, and Arno Krenzinger. A method to evaluate the effect of complementarity in time between hydro and solar energy on the performance of hybrid hydro pv generating plants. *Renewable Energy*, 45(Supplement C):24 – 30, 2012.

[94] Alexandre Beluco, Paulo Kroeff de Souza, and Arno Krenzinger. A dimensionless index evaluating the time complementarity between solar and hydraulic energies. *Renewable Energy*, 33(10):2157 – 2165, 2008.

[95] W.H. Kersting. Radial distribution test feeders. *IEEE Transactions on Power System*, 6(3):975–985, August 1991.

[96] Tim Sternkopf. Feed-in tariff, 2017.

[97] Junjie Sun, Leigh Tesfatsion, Donald Goldfarb, William Hogan, Daniel Kirschen, Chen-Ching Liu, Jim Mccalley, Michael J D Powell, Jim Price, Harold Salazar, Johnny Wong, and Tong Wu. DC Optimal Power Flow Formulation and Solution Using QuadProgJ *. Technical report, 2010.

[98] D.B.West. *Introduction to Graph Theory*. Upper Saddle River: Orentice hall, 2001.

[99] Thomas Morstyn, Alexander Teytelboym, Cameron Hepburn, and Malcolm D. McCulloch. Integrating P2P Energy Trading with Probabilistic Distribution Locational Marginal Pricing. *IEEE Transactions on Smart Grid*, pages 1–1, 2019.

[100] Roger Fletcher. A general quadratic programming algorithm. *IMA Journal of Applied Mathematics*, 7(1):76–91, 1971.

[101] Abinet Tesfaye Eseye, Matti Lehtonen, Toni Tukia, Semen Uimonen, and R John Millar. Machine learning based integrated feature selection approach for improved electricity demand forecasting in decentralized energy systems. *IEEE Access*, 7:91463–91475, 2019.

[102] Salah Bouktif, Ali Fiaz, Ali Ouni, and Mohamed Adel Serhani. Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*, 11(7):1636, 2018.

[103] Shaolong Sun, Shouyang Wang, Guowei Zhang, and Jiali Zheng. A decomposition-clustering-ensemble learning approach for solar radiation forecasting. *Solar Energy*, 163:189–199, 2018.

[104] Tiago Mendes Dantas and Fernando Luiz Cyrino Oliveira. Improving time series forecasting: An approach combining bootstrap aggregation, clusters and exponential smoothing. *International Journal of Forecasting*, 34(4):748–761, 2018.

[105] Mohammed Ali, Ali Alqahtani, Mark W Jones, and Xianghua Xie. Clustering and classification for time series data in visual analytics: A survey. *IEEE Access*, 7:181314–181338, 2019.

[106] Sidra Kanwal, Bilal Khan, and Sahibzada Muhammad Ali. Machine learning based weighted scheduling scheme for active power control of hybrid microgrid. *International Journal of Electrical Power and Energy Systems*, 125:106461, 2021.

[107] S. Ekici, F. Unal, and U. Ozleyen. Comparison of different regression models to estimate fault location on hybrid power systems. *IET Generation, Transmission Distribution*, 13(20):4756–4765, 2019.

[108] Y. Liu, R. J. Patton, and S. Shi. Wind turbine load mitigation using mpc with gaussian wind speed prediction. In *2018 UKACC 12th International Conference on Control (CONTROL)*, pages 32–37, 2018.

[109] J.A. Duffie and W.A. Beckman. *Solar Engineering of Thermal Processes*. Wiley, 2013.

[110] Taghi M Khoshgoftaar, Alireza Fazelpour, Huanjing Wang, and Randall Wald. A survey of stability analysis of feature subset selection techniques. In *2013 IEEE 14th International Conference on Information Reuse & Integration (IRI)*, pages 424–431. IEEE, 2013.

[111] A. A. Shanab, T. M. Khoshgoftaar, and R. Wald. Evaluation of wrapper-based feature selection using hard, moderate, and easy bioinformatics data. In *2014 IEEE International Conference on Bioinformatics and Bioengineering*, pages 149–155, 2014.

[112] M. Rong, D. Gong, and X. Gao. Feature selection and its use in big data: Challenges, methods, and trends. *IEEE Access*, 7:19709–19725, 2019.

[113] Lisa Guan, Jay Yang, and John M Bell. Cross-correlations between weather variables in australia. *Building and Environment*, 42(3):1054–1070, 2007.

[114] P De Jong, AS Sánchez, K Esquerre, Ricardo de Araújo Kalid, and Ednildo Andrade Torres. Solar and wind energy production in relation to the electricity load curve and hydroelectricity in the northeast region of brazil. *Renewable and Sustainable Energy Reviews*, 23:526–535, 2013.

[115] Fermín Rodríguez, Alice Fleetwood, Ainhoa Galarza, and Luis Fontán. Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control. *Renewable energy*, 126:855–864, 2018.

[116] Gyu Gwang Kim, Jin Ho Choi, So Young Park, Byeong Gwan Bhang, Woo Jun Nam, Hae Lim Cha, NeungSoo Park, and Hyung-Keun Ahn. Prediction model for pv performance with correlation analysis of environmental variables. *IEEE Journal of Photovoltaics*, 9(3):832–841, 2019.

[117] Jiaqi Zhong, Luyao Liu, Qie Sun, and Xinyu Wang. Prediction of photovoltaic power generation based on general regression and back propagation neural network. *Energy Procedia*, 152:1224–1229, 2018.

[118] Avrim L Blum and Pat Langley. Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2):245–271, 1997.

[119] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915.

[120] G. G. Kim, J. H. Choi, S. Y. Park, B. G. Bhang, W. J. Nam, H. L. Cha, N. Park, and H. Ahn. Prediction model for pv performance with correlation analysis of environmental variables. *IEEE Journal of Photovoltaics*, 9(3):832–841, May 2019.

[121] Thomas CR Russell, Rebecca Saive, André Augusto, Stuart G Bowden, and Harry A Atwater. The influence of spectral albedo on bifacial solar cells: A theoretical and experimental study. *IEEE Journal of photovoltaics*, 7(6):1611–1618, 2017.

[122] Y Kotak, MS Gul, T Muneer, and SM Ivanova. Investigating the impact of ground albedo on the performance of pv systems. In *Proceedings of the CIBSE Technical Symposium, London, UK*, pages 16–17, 2015.

[123] Pavel Chrobák, J Skovajsa, and Martin Zálešák. Production of electricity using photovoltaic panels and effects of cloudiness. *WSEAS Transactions on Power Systems*, 12:346–35, 2017.

[124] Luis Hernández, Carlos Baladrón, Javier M. Aguiar, Lorena Calavia, Belén Carro, Antonio Sánchez-Esguevillas, Diane J. Cook, David Chinarro, and Jorge Gómez. A study of the relationship between weather variables and electric power demand inside a smart grid/smart world framework. *Sensors*, 12(9):11571–11591, 2012.

[125] Chaoshun Li, Geng Tang, Xiaoming Xue, Adnan Saeed, and Xin Hu. Short-term wind speed interval prediction based on ensemble gru model. *IEEE Transactions on Sustainable Energy*, 2019.

[126] Rich H. Inman, Hugo T.C. Pedro, and Carlos F.M. Coimbra. Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science*, 39(6):535 – 576, 2013.

[127] Seyyed A. Fatemi, Anthony Kuh, and Matthias Fripp. Parametric methods for probabilistic forecasting of solar irradiance. *Renewable Energy*, 129:666 – 676, 2018.

[128] Tianfeng Chai and R. Draxler. Root mean square error (rmse) or mean absolute error (mae)? *Geosci. Model Dev.*, 7, 01 2014.

[129] Hugo T.C. Pedro, Carlos F.M. Coimbra, Mathieu David, and Philippe Lauret. Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. *Renewable Energy*, 123:191 – 203, 2018.

[130] G. Liu and X. Wu. Time series clustering and evaluation of unknown working conditions of mismatched photovoltaic array systems. In *2018 IEEE 4th International Conference on Control Science and Systems Engineering (ICCSSE)*, pages 164–167, Aug 2018.

[131] Guangyu Liu, Ling Zhu, Xinpeng Wu, and Jiajun Wang. Time series clustering and physical implication for photovoltaic array systems with unknown working conditions. *Solar Energy*, 180:401 – 411, 2019.

[132] R. M. Esteves, T. Hacker, and C. Rong. Competitive k-means, a new accurate and distributed k-means algorithm for large datasets. In *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, volume 1, pages 17–24, Dec 2013.

[133] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Comput. Surv.*, 31(3):264–323, September 1999.

[134] Xiaohui Huang, Yunming Ye, Liyan Xiong, Raymond Y.K. Lau, Nan Jiang, and Shaokai Wang. Time series k-means: A new k-means type smooth subspace clustering for time series data. *Information Sciences*, 367-368:1 – 13, 2016.

[135] Jiawei Han, Micheline Kamber, and Jian Pei. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2011.

[136] M. Aakroum, A. Ahogho, A. Aaqir, and A. A. Ahajjam. Deep learning for inferring the surface solar irradiance from sky imagery. In *2017 International Renewable and Sustainable Energy Conference (IRSEC)*, pages 1–4, Dec 2017.

[137] T Soni Madhulatha. An overview on clustering methods. *arXiv preprint arXiv:1205.1117*, 2012.

[138] Robert Tibshirani, Guenther Walther, and Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.

[139] Yordan P Raykov, Alexis Boukouvalas, Fahd Baig, and Max A Little. What to do when k-means clustering fails: a simple yet principled alternative algorithm. *PloS one*, 11(9):e0162259, 2016.

[140] Anil K. Jain. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8):651 – 666, 2010. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR).

[141] David F Williamson, Robert A Parker, and Juliette S Kendrick. The box plot: a simple visual method to interpret data. *Annals of internal medicine*, 110(11):916–921, 1989.

[142] Venkateswara Reddy Eluri, M Ramesh, Amina Salim Mohd Al-Jabri, and Mare Jane. A comparative study of various clustering techniques on big data sets using apache mahout. In *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)*, pages 1–4. IEEE, 2016.

[143] Xin Jin and Jiawei Han. *K-Medoids Clustering*, pages 564–565. Springer US, Boston, MA, 2010.

[144] Eric Schulz, Maarten Speekenbrink, and Andreas Krause. A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *bioRxiv*, 2017.

[145] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.

[146] Wikimedia Commons. File:k-fold cross validation en.jpg — wikimedia commons, the free media repository, 2018. [Online; accessed 21-August-2019].

[147] S. Yadav and S. Shukla. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, pages 78–83, Feb 2016.

[148] D Cane and M Milelli. Multimodel superensemble technique for quantitative precipitation forecasts in piemonte region. *Natural Hazards & Earth System Sciences*, 10(2), 2010.

[149] W. Tangwongcharoen and W. Titiroongruang. The specific golf swing patterns using k-means clustering with the two-sided confidence interval. In *2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 698–701, June 2017.

[150] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu. Short-term wind speed interval prediction based on ensemble gru model. *IEEE Transactions on Sustainable Energy*, 11(3):1370–1380, July 2020.

[151] Zhixiong Hu, Yijun Xu, Mert Korkali, Xiao Chen, Lamine Mili, and Jaber Valinejad. A bayesian approach for estimating uncertainty in stochastic economic dispatch considering wind power penetration. *IEEE Transactions on Sustainable Energy*, 2020.

[152] Kolbjorn Engeland and Ingelin Steinsland. Probabilistic postprocessing models for flow forecasts for a system of catchments and several lead times. *Water resources research*, 50(1):182–197, 2014.

[153] A. Papavasiliou and I. Mezghani. Coordination schemes for the integration of transmission and distribution system operations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7, 2018.

[154] A. Saint-Pierre and P. Mancarella. Active distribution system management: A dual-horizon scheduling framework for dso/tso interface under uncertainty. *IEEE Transactions on Smart Grid*, 8(5):2186–2197, 2017.

[155] S. Huang, Q. Wu, S. S. Oren, R. Li, and Z.Liu. Distribution locational marginal pricing through quadratic programming for congestion management in distribution networks. *IEEE Transactions on Power Systems*, 30(4):2170–2178, 2015.

[156] Marco Rossi, Gianluigi Migliavacca, Giacomo Viganò, Dario Siface, Carlos Madina, Inés Gomez, Ivana Kockar, and Andrei Morch. Tso-dso coordination to acquire services from distribution grids: Simulations, cost-benefit analysis and regulatory conclusions from the smartnet project. *Electric Power Systems Research*, 189:106700, 2020.

[157] John F Franco, Luis F Ochoa, and Rubén Romero. Ac opf for smart distribution networks: An efficient and robust quadratic approach. *IEEE Transactions on Smart Grid*, 9(5):4613–4623, 2017.

[158] O. D. Montoya, W. Gil-González, and A. Garces. Optimal power flow on dc microgrids: A quadratic convex approximation. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(6):1018–1022, 2019.

[159] Luis Nando Ochoa, Fabrizio Pilo, Andrew Keane, Paul Cuffe, and Giuditta Pisano. Embracing an adaptable, flexible posture: Ensuring that future european distribution networks are ready for more active roles. *IEEE Power and Energy Magazine*, 14(5):16–28, 2016.

[160] Daniel B. Arnold, Michael D. Sankur, Matias Negrete-Pincetic, and Duncan S. Callaway. Model-free optimal coordination of distributed energy resources for provisioning transmission-level services. *IEEE Transactions on Power Systems*, 33(1):817–828, 2018.

[161] Alejandro Vicente-Pastor, Jesus Nieto-Martin, Derek W. Bunn, and Arnaud Laur. Evaluation of flexibility markets for retailer–dso–tso coordination. *IEEE Transactions on Power Systems*, 34(3):2003–2012, 2019.

[162] Rui Yang, Jun Hao, Huaiguang Jiang, and Xin Jin. Machine-learning-driven, site-specific weather forecasting for grid-interactive efficient buildings. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2020.

[163] Thomas Krechel, Francisco Sanchez, Francisco Gonzalez-Longatt, Harold Chamorro, and Jose Luis Rueda. Chapter 11 - transmission system-friendly microgrids: an option to provide ancillary services. In Rajeev Kumar Chauhan and Kalpana Chauhan, editors, *Distributed Energy Resources in Microgrids*, pages 291–321. Academic Press, 2019.

[164] EN Association et al. Open networks project: opening markets for network flexibility. *Energy Networks Association2017*, 2017.

[165] Henrik Carøe Bylling. *Bilevel Optimization with Application in Energy*. PhD thesis, University of Copenhagen, Faculty of Science, Department of Mathematical . . . , 2018.

[166] Heinrich von Stackelberg. *Market Structure and Equilibrium*. Springer, 1934.

[167] Charles Audet, Pierre Hansen, Brigitte Jaumard, and Gilles Savard. Links between linear bilevel and mixed 0–1 programming problems. *Journal of optimization theory and applications*, 93(2):273–300, 1997.

[168] Ankur Sinha, Tharo Soun, and Kalyanmoy Deb. Using karush-kuhn-tucker proximity measure for solving bilevel optimization problems. *Swarm and Evolutionary Computation*, 44:496–510, 2019.

[169] Andreas Britzelmeier, Alberto De Marchi, and Matthias Gerdts. *An Iterative Solution Approach for a Bi-level Optimization Problem for Congestion Avoidance on Road Networks*, pages 23–38. Springer International Publishing, Cham, 2018.

[170] J.Cohon. Academic Press, New York, 1978.

[171] Songqing Shan and G. Gary Wang. An Efficient Pareto Set Identification Approach for Multiobjective Optimization on Black-Box Functions. *Journal of Mechanical Design*, 127(5):866–874, 11 2004.

[172] Bharat Vardani. Optimum location of svc in an ieee 33 bus radial distribution system using power sensitivity index. In *2019 International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, pages 1–5, 2019.

[173] J. S. Savier and Debapriya Das. Impact of network reconfiguration on loss allocation of radial distribution systems. *IEEE Transactions on Power Delivery*, 22(4):2473–2480, 2007.

[174] M.E. Baran and F.F. Wu. Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Transactions on Power Delivery*, 4(2):1401–1407, 1989.

[175] A.A. Sadiq, S.S. Adamu, and M. Buhari. Optimal distributed generation planning in distribution networks: A comparison of transmission network models with facts. *Engineering Science and Technology, an International Journal*, 22(1):33–46, 2019.

[176] Dimitra Apostolopoulou, George Gross, and Teoman Güler. Optimized ftr portfolio construction based on the identification of congested network elements. *IEEE Transactions on Power Systems*, 28(4):4968–4978, 2013.

[177] Canan Kandilli. Performance analysis of a novel concentrating photovoltaic combined system. *Energy Conversion and Management*, 67:186–196, March 2013.

[178] Majed BenAmmar, Maher Chaabene, and Zied Chtourou. Artificial neural network based control for pv/t panel to track optimum thermal and electrical power. *Energy Conversion and Management*, 65:372–380, Jan 2013.

[179] Pei Gang, Fu Huide, Ji Jie, Chow Tin-taib, and Zhang Tao. Annual analysis of heat pipe pv/t systems for domestic hot water and electricity production. *Energy Conversion and Management*, 56:8–21, April 2012.

[180] Francesco Calise, Massimo Dentice d'Accadia, and LauraVanoli. Design and dynamic simulation of a novel solar trigeneration system based on hybrid photovoltaic/thermal collectors (pvt). *Energy Conversion and Management*, 60:214–225, August 2012.

[181] Jiafei Zhao, Yongchen Song, Wei-Haur Lam, Weiguo Liu, Yu Liu, Yi Zhang, and DaYong Wang. Solar radiation transfer and performance analysis of an optimum photovoltaic/thermal system. *Energy Conversion and Management*, 52(2), Feb 2011.

[182] Fatemeh Najibi, Eduardo Alonso, and Dimitra Apostolopoulou. Optimal dispatch of pumped storage hydro cascade under uncertainty. *UKACC International Conference on Control*, pages 187– 192, Sep 2018.

[183] Dimitra Apostolopoulou, Zacharie De Grève, and Malcolm McCulloch. Robust optimization for hydroelectric system operation under uncertainty. *IEEE trans power system*, 33(3):3337–3348, 2018.

[184] Ian Richardson, Murray Thomson, David Infield, and Conor Clifford. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, 42(10):1878–1887, October 2010.

[185] Mohammad B. Shadmand and Robert S. Balog. Multi-objective optimization and design of photovoltaic-wind hybrid system for community smart dc microgrid. *IEEE trans smart grid*, 5(5):2635–2643, 2014.

[186] Karampelas Panagiotis and Ekonomou Lambros. *Intelligent Solutions for Electricity Transmission and Distribution Networks*. Springer-Verlag Berlin Heidelberg, 2016.

[187] Dimitra Apostolopoulou and Malcolm McCulloch. Cascade hydroelectric power system model and its application to an optimal dispatch design. *10th Bulk Power Systems Dynamics and Control Symposium*, 2017.

[188] Matthias Feurer and Frank Hutter. Hyperparameter optimization. In *Automated Machine Learning*, pages 3–33. Springer, Cham, 2019.

[189] S. Tiwari, R. Sabzehgar, and M. Rasouli. Short term solar irradiance forecast based on image processing and cloud motion detection. In *2019 IEEE Texas Power and Energy Conference (TPEC)*, pages 1–6, 2019.

Appendix

Nomenclature

A	reduced branch-to-node incidence matrix
B_d	diagonal branch susceptance matrix
B_{B_i}	cost of the battery system at node i
B_{PV_i}	cost of PV generation at node i
$c_i(t)$	cost of generator i at time t
$E_{0,i}$	initial value of the energy stored at battery system at node i
$E_{\min,i}$	minimum energy that can be stored at battery system at node i
$E_{\max,i}$	maximum energy that can be stored at battery system at node i
$f(t)$	vector of line power flows at time t
f^M	vector of maximum real power flows
σ_f	kernel parameter, signal standard deviation
σ_l	kernel parameter, characteristic length scale
σ_ε	standard deviation of the errors
f^m	vector of minimum real power flows
$f_1(\cdot)$	TSO objective function

$f_2(\cdot)$ all DSOs objective functions

$g_1(\cdot)$ TSO inequality constraints

$g_2(\cdot)$ DSO inequality constraints

$h_1(\cdot)$ TSO equality constraints

$h_2(\cdot)$ DSO equality constraints

$\mathscr{I}$ set of I generators

$\mathscr{J}$ set of J loads

$\mathscr{J}_k$ set of loads connected to bus k

$\mathscr{K}$ set of K nodes

$\mathscr{L}$ set of L lines

M graph incidence matrix

$\mathscr{N}_{PV}^{d}$ set of PVs connected to distribution system d

$\mathscr{N}_{B}^{d}$ set of battery systems connected to distribution system d

$p_i^d(t)$ net real power at node i at time t in distribution system d

$P_{B_i}^{\text{dis}}(t)$ discharging power of the battery system at node i at time t

$P_{B,i}^{\text{dis,min}}$ discharging power of the battery system at node i lower limit

$P_{B,i}^{\text{dis,max}}$ discharging power of the battery system at node i upper limit

$P_{B_i}^{\text{ch}}(t)$ charging power of the battery system at node i at time t

$P_{B,i}^{\text{ch,min}}$ charging power of the battery system at node i lower limit

$P_{B,i}^{\text{ch,max}}$ charging power of the battery system at node i upper limit

$P_{L_j}(t)$ load j at time t

$P_{\text{load}_i}(t)$ real load at node i at time t

$P^d_{\text{grid}}(t)$ amount of power purchased from the transmission system at time t for distribution system d

$P^{d,\min}_{\text{grid}}$ minimum amount of amount of power purchased from the transmission system for distribution system d

$P_{G_i}(t)$ the power injection of generator i at time t

P^m_G vector of lower generation limits

P^M_G vector of upper generation limits

$P_{PV_i}(t)$ power output of PV at node i at time t

$P^{\min}_{PV,i}$ power output of PV at node i lower limit

$P^{\max}_{PV,i}$ power output of PV at node i upper limit

$q^d_i(t)$ net reactive power at node i at time t in distribution system d

$Q_{\text{load}_i}(t)$ reactive load at node i at time t

R positive definite matrix representing the network

$\mathscr{T}$ time period of interest

$V_i(t)$ voltage level at node i at time t

$V^{\min}_i$ voltage level at node i lower limit

$V^{\max}_i$ voltage level at node i upper limit

V_{ref} voltage reference value

X positive definite matrix representing the network

α the voltage regulation cost

Δt simulations time interval

$\eta_{\text{ch},i}$ charging efficiency of battery system at node i

$\eta_{\text{dis},i}$ discharging efficiency of battery system at node i

$\theta_k(t)$ angle at node k at time t

$\lambda_k(t)$ locational marginal price at node k at time t

$(\cdot)_t$ value of variable at time t

CI Confidence Interval

c_k cluster k centroid

$C(i)$ data point cluster number

$d(x_i, c_k)$ distance between x_i and c_k

$f(x^{(t)})$ mapping function of features to solar output

f vector of mapping functions

$h(x^{(t)})$ basis function

H vector of basis functions

K number of clusters

$K(\cdot,\cdot)$ covariance matrix

M number of features

MAE Mean Absolute Error

MSE Mean Square Error

N_k number of points assigned to cluster k

q number of selected features

RMSE Root Mean Square Error

T period of T hours

$T_\star$ number of hourly intervals the solar output is predicted

$x^{(t)}$ vector of features at time t

x_i vector with n elements

X_i vector of feature i for T time intervals

X matrix of features for entire period

$\mathscr{X}$ set of N elements

$y_\star^{(t)}$ forecasted value for solar generation at time t

$\tilde{y}^{(t)}$ actual value at time t

$y^{(t)}$ solar output at time t

y vector of solar output

β basis function coefficient

$\hat{\beta}$ estimated basis function coefficient

$\varepsilon^{(t)}$ noise variable at time t

ε vector of noise variables

$\bar{\varepsilon}$ mean value of the errors

θ hyperparameter of the kernel function

$\hat{\theta}$	estimated hyperparameter of the kernel function
θ_f	kernel parameter
θ_l	kernel parameter
μ	mean value
$\rho(\cdot,\cdot)$	correlation coefficient
σ	standard deviation
$\hat{\sigma}$	estimated standard deviation

Nomenclature

A	reduced branch-to-node incidence matrix
B_d	diagonal branch susceptance matrix
B_{B_i}	cost of the battery system at node i
B_{PV_i}	cost of PV generation at node i
$c_i(t)$	cost of generator i at time t
$E_{0,i}$	initial value of the energy stored at battery system at node i
$E_{\min,i}$	minimum energy that can be stored at battery system at node i
$E_{\max,i}$	maximum energy that can be stored at battery system at node i
$f(t)$	vector of line power flows at time t
f^M	vector of maximum real power flows
σ_f	kernel parameter, signal standard deviation
σ_l	kernel parameter, characteristic length scale
σ_ε	standard deviation of the errors
f^m	vector of minimum real power flows
$f_1(\cdot)$	TSO objective function

$f_2(\cdot)$ all DSOs objective functions

$g_1(\cdot)$ TSO inequality constraints

$g_2(\cdot)$ DSO inequality constraints

$h_1(\cdot)$ TSO equality constraints

$h_2(\cdot)$ DSO equality constraints

$\mathscr{I}$ set of I generators

$\mathscr{J}$ set of J loads

$\mathscr{J}_k$ set of loads connected to bus k

$\mathscr{K}$ set of K nodes

$\mathscr{L}$ set of L lines

M graph incidence matrix

$\mathscr{N}_{PV}^{d}$ set of PVs connected to distribution system d

$\mathscr{N}_{B}^{d}$ set of battery systems connected to distribution system d

$p_i^d(t)$ net real power at node i at time t in distribution system d

$P_{B_i}^{\text{dis}}(t)$ discharging power of the battery system at node i at time t

$P_{B,i}^{\text{dis,min}}$ discharging power of the battery system at node i lower limit

$P_{B,i}^{\text{dis,max}}$ discharging power of the battery system at node i upper limit

$P_{B_i}^{\text{ch}}(t)$ charging power of the battery system at node i at time t

$P_{B,i}^{\text{ch,min}}$ charging power of the battery system at node i lower limit

$P_{B,i}^{\text{ch,max}}$ charging power of the battery system at node i upper limit

$P_{L_j}(t)$ load j at time t

$P_{\text{load}_i}(t)$ real load at node i at time t

$P^d_{\text{grid}}(t)$ amount of power purchased from the transmission system at time t for distribution system d

$P^{d,\min}_{\text{grid}}$ minimum amount of amount of power purchased from the transmission system for distribution system d

$P_{G_i}(t)$ the power injection of generator i at time t

P^m_G vector of lower generation limits

P^M_G vector of upper generation limits

$P_{PV_i}(t)$ power output of PV at node i at time t

$P^{\min}_{PV,i}$ power output of PV at node i lower limit

$P^{\max}_{PV,i}$ power output of PV at node i upper limit

$q^d_i(t)$ net reactive power at node i at time t in distribution system d

$Q_{\text{load}_i}(t)$ reactive load at node i at time t

R positive definite matrix representing the network

$\mathscr{T}$ time period of interest

$V_i(t)$ voltage level at node i at time t

$V^{\min}_i$ voltage level at node i lower limit

$V^{\max}_i$ voltage level at node i upper limit

V_{ref} voltage reference value

X positive definite matrix representing the network

α the voltage regulation cost

Δt simulations time interval

$\eta_{\text{ch},i}$ charging efficiency of battery system at node i

$\eta_{\text{dis},i}$ discharging efficiency of battery system at node i

$\theta_k(t)$ angle at node k at time t

$\lambda_k(t)$ locational marginal price at node k at time t

$(\cdot)_t$ value of variable at time t

CI Confidence Interval

c_k cluster k centroid

$C(i)$ data point cluster number

$d(x_i, c_k)$ distance between x_i and c_k

$f(x^{(t)})$ mapping function of features to solar output

f vector of mapping functions

$h(x^{(t)})$ basis function

H vector of basis functions

K number of clusters

$K(\cdot,\cdot)$ covariance matrix

M number of features

MAE Mean Absolute Error

MSE Mean Square Error

N_k number of points assigned to cluster k

q number of selected features

RMSE Root Mean Square Error

T period of T hours

$T_\star$ number of hourly intervals the solar output is predicted

$x^{(t)}$ vector of features at time t

x_i vector with n elements

X_i vector of feature i for T time intervals

X matrix of features for entire period

$\mathscr{X}$ set of N elements

$y_\star^{(t)}$ forecasted value for solar generation at time t

$\tilde{y}^{(t)}$ actual value at time t

$y^{(t)}$ solar output at time t

y vector of solar output

β basis function coefficient

$\hat{\beta}$ estimated basis function coefficient

$\varepsilon^{(t)}$ noise variable at time t

ε vector of noise variables

$\bar{\varepsilon}$ mean value of the errors

θ hyperparameter of the kernel function

$\hat{\theta}$ estimated hyperparameter of the kernel function

θ_f kernel parameter

θ_l kernel parameter

μ mean value

$\rho(\cdot,\cdot)$ correlation coefficient

σ standard deviation

$\hat{\sigma}$ estimated standard deviation