January 2020

Uncertainty Quantification And Economic Dispatch Models For The Power Grid

Arun Sukumaran Nair

Uncertainty Quantification and Economic Dispatch Models for the Power Grid

by

Arun Sukumaran Nair
Bachelor of Technology, University of Kerala, 2011
Master of Technology, Tuskegee University, 2016

A Dissertation
Submitted to the Graduate Faculty
of the
University of North Dakota
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

December
2020

This dissertation, submitted by Arun Sukumaran Nair in partial fulfillment of the requirements for the Degree of *Doctor of Philosophy* from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Dr. Prakash Ranganathan, Chairperson

Dr. Hossein Salehfar, Committee Member

Dr. Naima Kaabouch, Committee Member

Dr. Saleh Faruque, Committee Member

Dr. Jeremiah Bartz, Member-At-Large

This dissertation is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Dr. Chris Nelson
Dean of the School of Graduate Studies

Date

PERMISSION

Title	Uncertainty Quantification and Economic Dispatch Models for the Power Grid
Department	School of Electrical Engineering and Computer Science
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the Professor who supervised my dissertation work or, in his absence, by the Chairperson of the department or the Dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Arun Sukumaran Nair

December 7th 2020

ACKNOWLEDGEMENT

First of all, I would like to thank my advisor Dr. Prakash Ranganathan, for his continued guidance, support and the freedom he provided while conducting the doctoral work. I would also like to thank Dr. Naima Kaabouch and Dr. Hossein Salehfar for their guidance and the opportunities provided for collaborative work. I thank my other committee members Dr. Saleh Faruque and Dr. Jeremiah Bartz for their support and guidance for my dissertation. I thank the National Science Foundation (NSF) (award #153756) and School of Electrical Engineering and Computer Science for granting financial support during the course of my Ph.D. I also take this opportunity to acknowledge the financial support from NSF through the INTERN award, North Dakota Established Program to Stimulate Competitive Research (ND EPSCoR) through the Doctoral Dissertation Award and UND for the Summer Doctoral Fellowship Award. I want to thank Dr. Ryan Adams, Department Chair and Dr. Naima Kaabouch, Graduate Programs Director and other faculty members for their continued support during the program. I also thank my research collaborators and colleagues Mitch Campion, Tareq Hossen and Radhakrishnan Chinnathambi for their support. Finally, I would like to thank my parents Mr. Sukumaran Nair, Mrs. Premalatha and my wife Gopika Raveendran, for their continued love and support without which I couldn't have completed my Ph.D. degree program.

DEDICATION

To God and Family

ABSTRACT

The modern power grid is constrained by several challenges, such as increased penetration of Distributed Energy Resources (DER), rising demand for Electric Vehicle (EV) integration, and the need to schedule resources in real-time accurately. To address the above challenges, this dissertation offers solutions through data-driven forecasting models, topology-aware economic dispatch models, and efficient optional power flow calculations for large scale grids. Particularly, in chapter 2, a novel microgrid decomposition scheme is proposed to divide the large scale power grids into smaller microgrids. Here, a two-stage Nearest-Generator Girvan-Newman (NGGN) algorithm, a graphical clustering-based approach, followed by a distributed economic dispatch model, is deployed to yield a 12.64% cost savings. In chapter 3, a deep-learning based scheduling scheme is intended for the EVs in a household community that uses forecasted demand, consumer preferences and Time-of-use (TOU) pricing scheme to reduce electricity costs for the consumers and peak shaving for the utilities. In chapter 4, a hybrid machine learning model using GLM with other methods is designed to forecast wind generation data. Finally, in chapter 5, multiple formulations for Alternating Current Optimal Power Flow (ACOPF) are designed for large scale grids in a high-performance computing environment. The ACOPF formulations, namely, power balance polar, power balance Cartesian, and current balance Cartesian, are tested on bus systems ranging from a 9-bus to 25,000. The current balance Cartesian formulation had an average of 23% faster computational time than two other formulations on a 25,000 bus system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TERMS AND ABBREVIATIONS

A-GN Admittance Girvan-Newman

AMPL A Mathematical Programming Language

DER Distributed Energy Resources

ED Economic Dispatch

EMJ Empirical method of Justus

EML Empirical method of Lysen

EPF Energy pattern factor method

ERCOT Electric Reliability Council of Texas

EV Electric Vehicle

ISO Independent System Operator

L-GN Length Girvan-Newman

LP Linear Programming

MAD Mean Absolute Deviation

MAED Multi-Area Economic Dispatch

MAPE Mean Absolute Percentage Error

MISO Midcontinent Independent System Operator

ML Maximum Likelihood method

MML Modified Maximum Likelihood method

NG-GN Nearest-Generator Girvan-Newman

NWP Numerical Weather Predictions

PEV Plug-in Electric Vehicle

RES Renewable Energy Sources

SCOPF Security Constrained Optimal Power Flow

SD Standard Deviation

SOC State of Charge

SPP Southwest Power Pool

TOU Time-of-use

UC Unit Commitment

UQ Uncertainty Quantification

V2G Vehicle-to-Grid

V2H Vehicle-to-Home

V2V Vehicle-to-Vehicle

CHAPTER 1

Introduction

1.1 Motivation

The electric power industry is experiencing drastic changes in its economic outlook as countries are replacing fossil fuels with Renewable Energy Sources (RES) [1]. The percentage contribution of RES to the power grid needs to increase in order to satisfy the increasing level of energy consumption. While the integration of RES is beneficial to the environment in reducing greenhouse gas emissions, it comes with technological challenges for utilities, such as the variability and reliability of RES [2]. In US, the most widely used RES in 2019 were wind and solar, constituting a total of 33% among the different renewable sources as shown in Figure 1.1 [3]. The energy production from these sources depends on several factors, such as wind speed, wind direction, solar irradiance, temperature, air pressure and humidity, among other variables. Any of these sources may vary instantaneously, influencing the amount of power generation. In the operations and planning of power systems, quantifying the uncertainty associated with RES is vital [4, 5, 6]. Predicting solar power is more accurate when compared to wind power, as there are fewer uncertain variables; the intensity and duration of wind power is both uncertain and variable, while solar power is more variable than wind power but more predictable [7, 8, 9, 10].

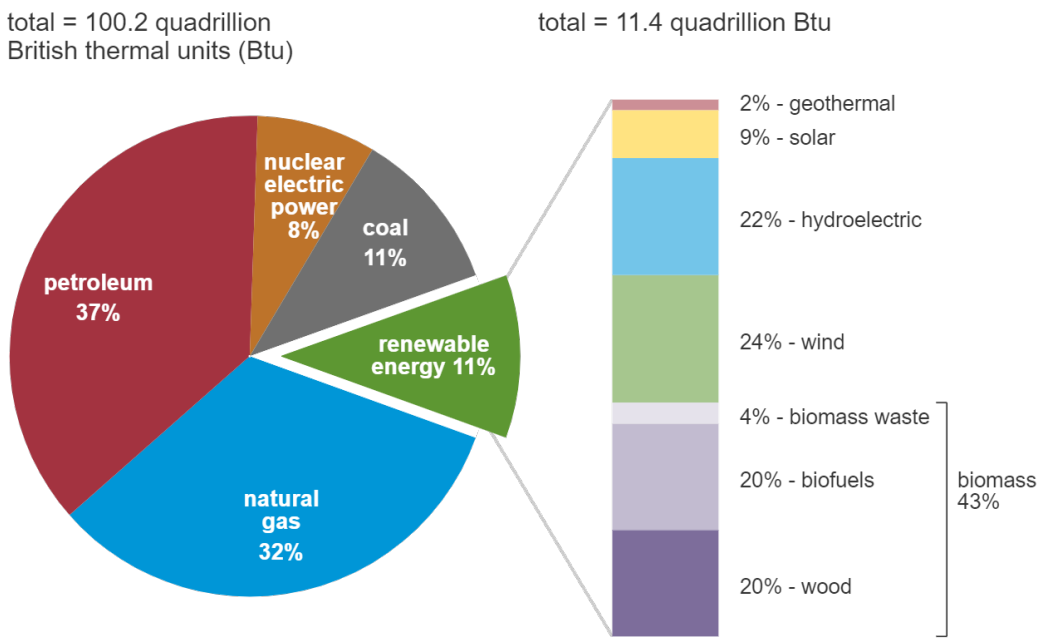


Figure. 1.1 U.S. primary energy consumption by source for 2019 [3].

Uncertainty evaluation and quantification is a topic of great interest in both the industry and the scientific community, as evidenced from the increased amount of literature in the area. The Uncertainty Quantification (UQ) plays an important role in improving the reliability, security, and availability of RES when evaluated within the engineering processes. Earlier concepts of uncertainty focused on probability theory to represent the variation in a variable process. Representation of uncertainty by probability methods may not always suit the scenario, such as when an information needed is unavailable. When using methods such as Economic Dispatch (ED) and Unit Commitment (UC), a more robust methods of representing uncertainty is required.

The uncertainty associated with a process can be categorized as either modeling, parametric, aleatory, epistemic or error. Aleatory uncertainty is the irreducible or stochastic uncertainty which can be better characterized with an increase in knowledge. Epistemic uncertainty is the reducible or systematic uncertainty, gets reduced with an increase in knowledge. Error is considered different from uncertainty; however, some

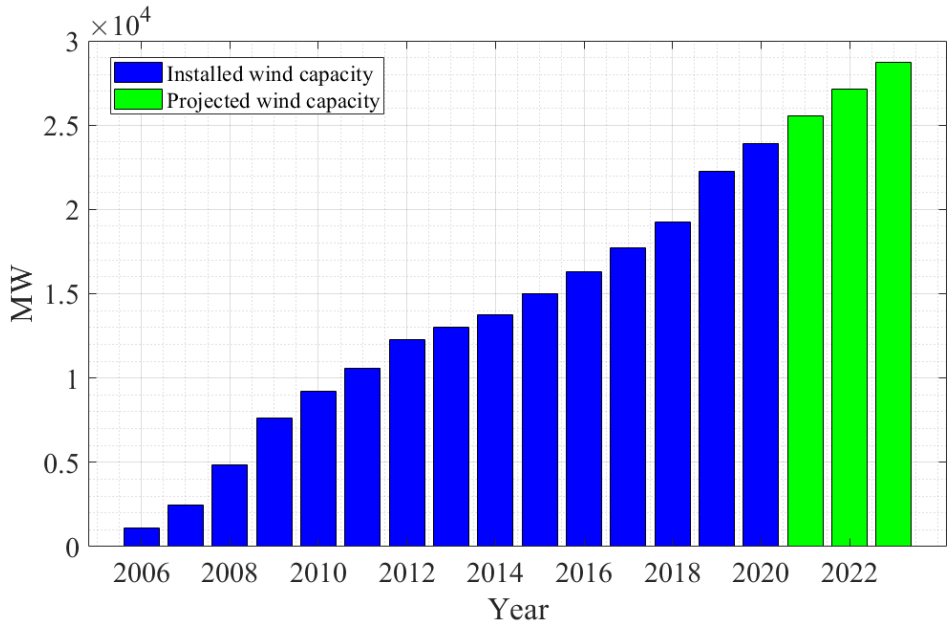


Figure. 1.2 Wind power generation capacity growth in MISO.

researchers refer to it as numerical uncertainty [11]. R. Lima and R. Sampaio [12] feel there is a need to develop an improved statistical measure to represent uncertainty since the utilization of CDF is clumsy, even though it may be the best method to represent variability.

A steady increase in installed wind capacity can be seen in the MISO region, the capacity growth of which is shown in 1.2 [13]. The management of wind power relies heavily on short-term forecasting. For wind power two uncertainty types can be considered: the inherent uncertainty of wind speed due to its variable and intermittent nature (aleatory uncertainty) and the uncertainty associated with the relationship between wind speed and wind power generated due to different parameters influencing the power generation (epistemic uncertainty). The wind power generation values from April 1st through 7th, 2020, in PJM MIDATL region, is shown in Figure 1.4. The wind power generation values does not follow any pattern like the load values (AECO load area) shown in Figure 1.3.

The two important optimization problems in a power system are unit commitment

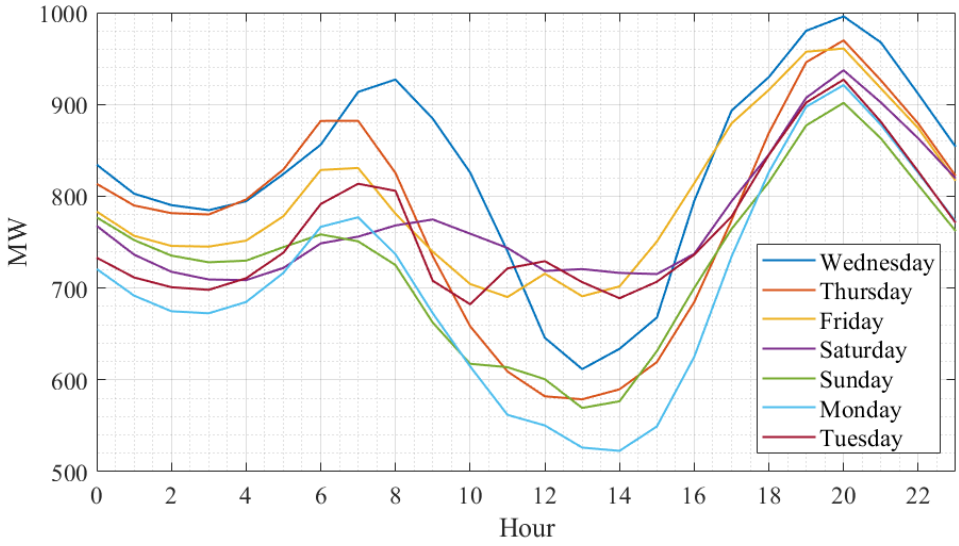


Figure. 1.3 Net load in PJM April 1st - April 7th.

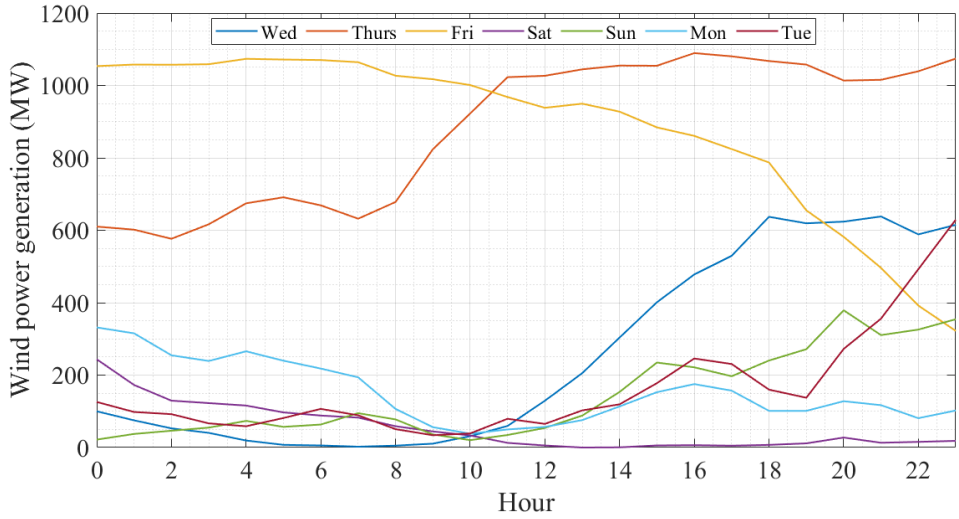


Figure. 1.4 Wind generation in PJM April 1st - April 7th.

(UC) and economic Dispatch (ED), both of which handle resource scheduling and resource allocation operation in power systems. ED is a process of economically allocating generation values to a mix of generating units to satisfy the system load requirements. The allocation process is subject to load, generation, and transmission constraints with the goal to reduce generation costs [14]. The UC involves selecting the generators to meet system demand while reducing the cost of operation and ED determines the generation value for the committed generators. An optimum resource

allocation model requires an accurate value of wind generation in the grid to economically allocate the resources. To increase the percentage of RES in the generation mix of a power grid and to make it a more attractive investment for power companies, requires an accurate forecast model to predict the RES generation.

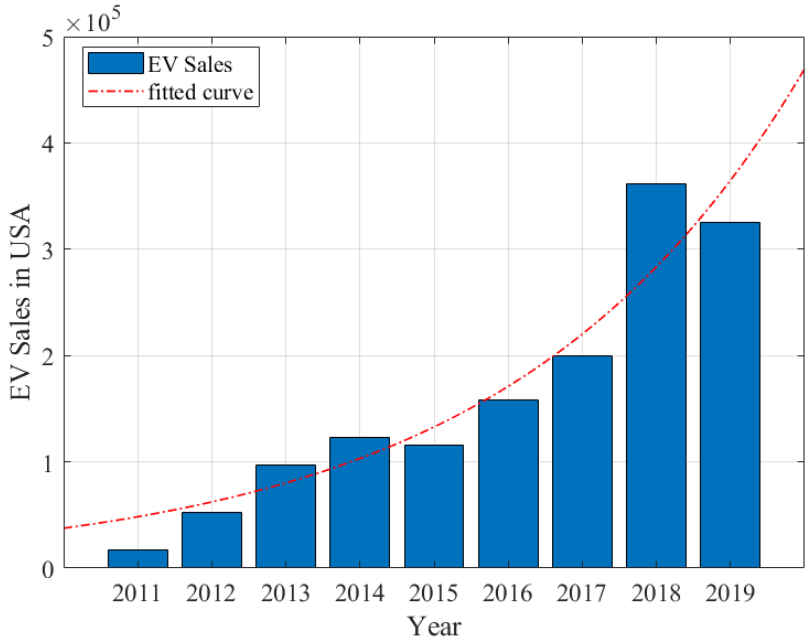


Figure. 1.5 PEV Sales in USA.

A significant and growing demand on the electrical grid is related to providing power to EVs, a field which is growing rapidly. The exponential growth in EV sales within the US from 2011 to 2019 is shown in figure 1.5 [15]. The 2019 sales of EVs alone increased by 329,528 units in the US [15]. California is the leader in US EV adoption with an increase of 61.7% from 2017 to 2018, which accounts for a 7.84% market share. The US has almost 1.2 million EV vehicles on the road as of March 31st, 2019 [16]. Globally EVs are gaining acceptance among consumers, where sales were approximately 2.1 million in 2018, a growth of 64% compared to the sales in 2017. Perhaps the most challenging issue the power industry faces is the increasing demand for electricity during the peak hours, which happens to be at the same time EV owners

connect their cars to the grid after working hours.

Microgrids which can be defined as a low-voltage distribution network with controllable loads, storage devices and DER, can be used to counter the uncertainty from wind generation, increased EV load, system failures and load variation. The formulation of the resource optimization problem can also play a vital role in the efficient operation of a grid management system. The dissertation tries to answer the following research questions:

1. What is an economical way to decompose (or group) a large scale grid into smaller microgrids?

2. How do we optimize electricity consumption by analyzing load profiles (e.g., EV) and consumer preferences?

3. How do we generate accurate forecast models for distributed energy resources (e.g., wind generation) under uncertain conditions?

4. How do we reduce the computation run times of optimal power flow calculations for larger grid systems?

1.2 Thesis Outline

Chapter 1 introduces the concepts and motivation for this thesis. Chapter 2 details the grid clustering and the proposed economic clustering schemes for microgrid decomposition. Chapter 3 establishes the need to schedule the EV load in the system for peak shaving and models the scheduling scheme. In chapter 4, the novel hybrid methods for short-term wind power forecasting is introduced. The chapter 5 discusses the computational and numerical performance evaluation of ACOPF and Security Constrained

Optimal Power Flow (SCOPF) formulations. Chapter 6 describes in detail the conclusions drawn from the thesis and outlines the future work to overcome the limitations and to extend the current research work.

1.3 Contributions of the Dissertation

My key contributions from this work are summarized below:

- Demonstrated the performance of clustered microgrid systems for varying load conditions (Chapter 2,C1-C3).
- Developed economic decomposition of a large-scale system in to smaller microgrids using grid clustering (Chapter 2, B1)
- Presented a novel EV scheduling scheme using linear programming incorporating customer preferences and forecasted demand using clustering technique and DNN (Chapter 3,C5).
- Explored the performance of different forecasting methodologies for short-term wind power generation (Chapter 4,J2).
- Proposed hybrid forecasting methods with superior performance in delivering consistent results for short-term wind power generation (Chapter 4,J2).
- Demonstrated the computational and numerical performance of the three different formulation of ACOPF and SCOPF on large-scale test grids (Chapter 5,J3).
- Proposed the relevant future work regarding the limitations of this thesis and outlined the strategy to improve certain aspects of the research (Chapter 6).

1.4 Publications

Results presented in this work have been published in a book chapter, peer-reviewed journal publications and conference proceedings. The findings from chapter 2 were published in a book chapter (B1) and in conference proceedings (C1, C2, C3). The work presented in chapter 3 has been published in articles C4 and C5. The results from chapter 4 and 5 are in the preparation stage for journal submissions. Collaborations with colleagues during the course of this work resulted in various publications which are listed from C6-C15, which are not included as part of the dissertation.

Book Chapter

B1. M. Campion, **A. S Nair**, E. Nygard, P. Ranganathan, "Decomposition of Microgrids in Large-scale Electric Test Beds for Economic Dispatch Optimization" in Distributed Linear Programming Models in a Smart Grid, Springer, 2017

Journal Publications

J1. **A. S Nair**, P. Ranganathan, T. Hossen, and N. Kaabouch, "Multi-Agent Systems for Resource Allocation and Scheduling in a Smart Grid," in *Technol. Econ. Smart Grids Sustain. Energy*, vol. 3 (15),2018.

Conference Proceedings

C1. **A. S. Nair**, P. Ranganathan and N. Kaabouch, "A Constrained Topological Decomposition Method for the Next-generation Smart Grid," *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, 2017, pp. 1-6.

C2. **A. S. Nair**, P. Ranganathan, H. Salehfar and N. Kaabouch, "Uncertainty Quantification of Wind Penetration and Integration into Smart Grid: A survey," *2017 North American Power Symposium (NAPS)*, Morgantown, WV, 2017, pp. 1-6.

C3. **A. S. Nair**, M. Campion, D. Hollingworth, and P. Ranganathan, "Two-Stage Load Forecasting for Residual Reduction and Economic Dispatch using PJM Datasets," in *2018 IEEE International Conference on Electro/Information Technology*, 2018.

C4. **A. S. Nair**, T. Hossen, M. Campion, and P. Ranganathan, "Optimal Operation of Residential EVs using DNN and Clustering Based Energy Forecast," in *50th North American Power Symposium (NAPS 2018)*, 2018.

C5. T. Hossen, **A. S. Nair**, S. Noghanian, and P. Ranganathan, "Optimal Operation of Smart Home Appliances using Deep Learning," in *50th North American Power Symposium (NAPS 2018)*, 2018.

C6. T. Hossen, **A. S. Nair**, R. K. A. Chinnathambi, and P. Ranganathan, "Residential Load Forecasting Using Deep Neural Networks (DNN)," in *50th North American Power Symposium (NAPS 2018)*, 2018.

C7. L. Peterson, **A. S. Nair**, and P. Ranganathan, "Short-term Forecast for Locational Marginal Pricing (LMP) Data Sets," in *50th North American Power Symposium (NAPS 2018)*, 2018.

C8. K. Schmatz, G. Gilje, **A. S Nair**, and P. Ranganathan, "An Object-oriented Graphical User Interface (GUI) for Optimal Placement of Phasor Measurement Units," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, 2018.

C9. M. Spitzer, C. Eerdmans, **A. S Nair**, and P. Ranganathan, "Evaluation of PMU Placements with SORI and ORC Indices for IEEE Test Feeders," *2018 IEEE Int. Conf. Electro/Information Technol.*, pp. 3–6, 2018.

C10. R. Kedrowski, J. Nelson, **A. S Nair**, and P. Ranganathan, "Short-Term Seasonal Energy Forecast," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, 2018.

C11. W. Kayser, S. J. Buchhop, B. Morrow, A. S. Nair, and P. Ranganathan, "Next-generation Transportation Vulnerabilities," in *ESCS'18 - The 16th Int'l Conf on Embedded Systems, Cyber-physical Systems, and Applications*, 2018.

C12. J. Pagel, M. Campion, **A. S. Nair** and P. Ranganathan, "Clustering Analytics for Streaming Smart Grid Datasets," *2016 Clemson University Power Systems Conference (PSC)*, Clemson, SC, 2016, pp. 1-8.

C13. R. A. Chinnathambi, M. Campion, **A. S. Nair**, and P. Ranganathan, "Investigation of Price-Feature Selection Algorithms for the Day-Ahead Electricity Markets," in *EPEC 18*, 2018.

C14. Z. E. Mrabet, D. F. Selvaraj, **A. S. Nair** and P. Ranganathan, "Detection of the False Data Injection Attack in Home Area Networks using ANN," *2019 IEEE International Conference on Electro Information Technology (EIT)*, Brookings, SD, USA, pp. 176-181,2019.

C15. R. A. Chinnathambi, S. J. Plathottam, T. Hossen, **A. S. Nair**, and P. Ranganathan, "Deep Neural Networks (DNN) for Day-Ahead Electricity Price Markets," in *EPEC 18*, 2018.

CHAPTER 2

Economic Dispatch Models using Decomposition Methods

2.1 Introduction

Sources of RES power generation, such as wind, solar, tidal, and wave power, are heterogenous and have uncertain variables. This uncertainty, coupled with the dynamic nature of distributed sub-system architecture and the need for information synchronization, has made the problem of resource allocation and monitoring a challenge for the next-generation smart grid. The deployment of distributed algorithms across micro-grids has been unfortunately overlooked in the electric grid sector; centralized methods for managing resources and data may not be sufficient to monitor a complex electric grid. We therefore examine a decentralized constrained decomposition using Linear Programming (LP), which optimizes the inter-area transfer across micro-grids, reducing total generation costs for the grid.

2.2 Literature review

Electrical power systems are designed to meet the power demand through various generation sources, making the minimization of generation and operating cost critical. ED is often used to schedule and match the generator outputs to variable demands meeting system and transmission line constraints [17]. Accomplishing an optimal way to

effectively schedule generation resources will result in significant savings in operating cost such as fuel and transportation cost, while enabling appropriate demand response mechanisms.

Significant work has been done on the ED challenge, which obtains optimal solutions to the generator allocation issue; some reported methods include: lambda iteration methods [18], Lagrange Multiplier [19], genetic algorithm [20], hopfield networks [21], particle swarm optimization [22], dynamic programming, simulated annealing [23], firefly algorithm [24] and artificial intelligence approaches [25]. The application of ED on clustered microgrid has not been seen in the literatures reviewed. We propose economic formation of microgrids using clustering techniques for test systems.

2.3 Problem formulation

Formulating objective functions and constraints for electric grid structures, which is a collection of micro grids, is our goal. The standard IEEE test system (14-bus system) is used to model the electric grid, after which an optimization problem, ED, is devised that minimizes the total cost; while satisfying load demand, generation, and line flow constraints. The fuel cost for generating unit i supplying P_{Gi} amount of real power can be represented by a quadratic equation [26] as shown in equation (2.1).

$$F_i(P_{Gi}) = a_i P_{Gi}^2 + b_i P_{Gi} + c_i \tag{2.1}$$

Where a_i, b_i, c_i are the cost-coefficients of generating unit i and P_{Gi} is the real power generation of the unit i. The objective is to minimize the total cost of generation, which

can be represented by the following equation:

$$\text{Min } F = \sum_{i=1}^{n_g} F_i(P_{Gi}) \tag{2.2}$$

Where n_g is the number of generators working in the bus system. The economic dispatch problem is then solved subject to several formulated constraints.

$$P_{Gimin} \leqslant P_{Gi} \leqslant P_{Gimax} \quad i = 1..n_g \tag{2.3}$$

$$\sum_{i=1}^{n_g} P_{Gi} = D \tag{2.4}$$

$$\sum_{i=1}^{m_g} P_{Gi} + \sum_{i=1}^{m} T_k = D_k \tag{2.5}$$

$$T_{mn(min)} \leqslant T_{mn} \leqslant T_{mn(max)} \tag{2.6}$$

In Equation (2.3), the constraint illustrates how the generation from each generator must be within its maximum and minimum limits, while in Equation (2.4) the constraint shows a condition where total generation should meet the total demand in the system.

In Multi-Area Economic Dispatch (MAED) problems, the entire electric grid network is divided into several microgrids or areas, to carry out the ED. The power flow through the inter-connecting transmission lines, or tie lines, connecting the micro grids is an additional constraint in the MAED problem. In each micro grid, the loads must be satisfied by the sum of generation within the microgrid and the power coming to the microgrid from connected areas. The generation load constraint is shown in Equation (2.5), where the first term shows the summation of power generated in the area and the second term indicates the power flow from connected areas. The variables m_g in-

dicate the number of generators in the current area, variable m indicates the number of inter-connected areas to the area under consideration, and T represents the power flow between the areas. An additional constraint is added to the model to restrict the power flow between the areas as shown in Equation (2.6).

The power flow between two areas m and n are subjected to a minimum and a maximum value of $T_{mn(min)}$ and $T_{mn(max)}$ respectively. In the MAED model, the cost of power flow through the tie lines is also taken into consideration. A value of 0.01$ per MW is assumed as the cost and 100 MW is applied as the maximum tie line flow limit [27]. The total cost function to be minimized, therefore will be modified into Equation (2.7) [28]. The variable C_j denotes the cost of tie line power flow which is assumed to be constant for all tie lines. In the model, t represents the number of tie lines, and T_j is the amount of tie line power flow.

$$\text{Min } F = \sum_{i=1}^{n_g} F_i(P_{Gi}) + \sum_{j=1}^{t} C_j T_j \tag{2.7}$$

2.4 ED on PJM Region

Figure 2.1 shows the data flow in the load forecasting – economic dispatch model. The historical load data values were obtained for the PJM system, and different forecasting methods were applied to obtain the day ahead data for the loads, which implies the forecasted load values for the next 24 hours. The ED will be applied to the forecasted load values to obtain the hourly power generation values for the different generators. The data flow for the proposed model is shown in Figure 2.1.

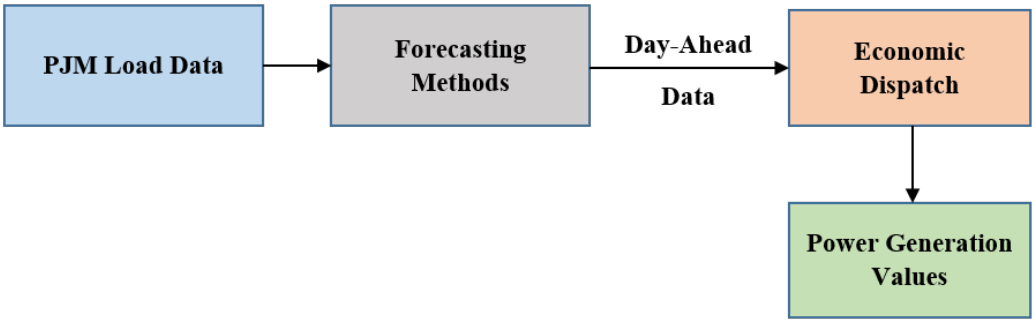


Figure. 2.1 Data flow in the forecasting- economic dispatch model.

2.4.1 PJM Datasets

We utilized the PJM historical load dataset for 2016, which is a publicly available historical dataset with hourly resolution [29]. This dataset consists of load values from 10 load areas in the PJM grid, the region of which is served by the PJM interconnections shown in Figure 2.2 [30]. The territories coordinated by PJM Interconnection include region from New Jersey, North Carolina, Indiana, Illinois, Delaware, Kentucky, Michigan, Maryland, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia.

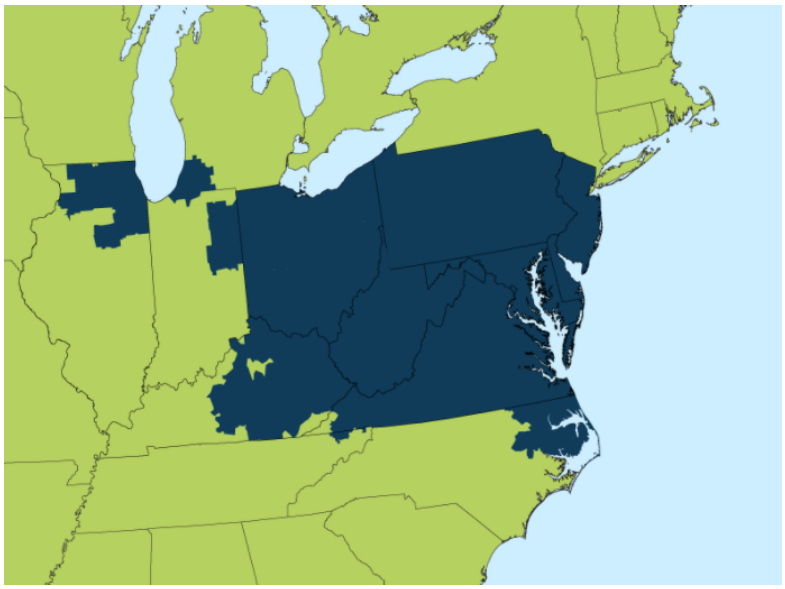


Figure. 2.2 Region served by PJM Interconnection [30].

2.4.2 Forecasting

Forecasting involve the use of different amounts of historical data. The dataset used for this project contains hourly historical demand data ranging from 1/1/2016 to 11/15/2016. The goal of load forecasting is to create a power load forecast for 11/16/2016, to show how the ED of this load could be simulated and examined. ARIMA and exponential smoothing (ES) were applied to the entire past year of power usage data as well as just the preceding 30 days of power demand data. Using ARIMA, ES, and an ensemble ARIMA forecasting method, the hourly loads for 11/16/2016 were forecasted for the PJM data. This ensemble method forecasts the residuals of a first stage fitting model either ARIMA or ES and uses ARIMA to forecast the residuals of the first stage, after which the residuals are added to the forecast from the fitted model. The p,d,q parameters of ARIMA forecasts were determined using the method described in [8]. The accuracy of ARIMA and ES methods were then compared to actual hourly load data for 11/16/2016. The accuracy of the forecasts were quantified according to Mean Absolute Percentage Error (MAPE) and Mean Absolute Deviation (MAD) shown in Equations 2.8 and 2.9 as follows [31]:

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{2.8}$$

$$MAD = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t| \tag{2.9}$$

Day-ahead forecasting using 11-month and 30 day data for a single region are shown in Figures 2.3 and 2.4. The Performance of the methods in day-ahead forecasting using 30 days of data for the total PJM market is presented in Figure 2.5. The forecasting

error from the hybrid methods in compariosn to ARIMA and ES are listed in Table 2.1.

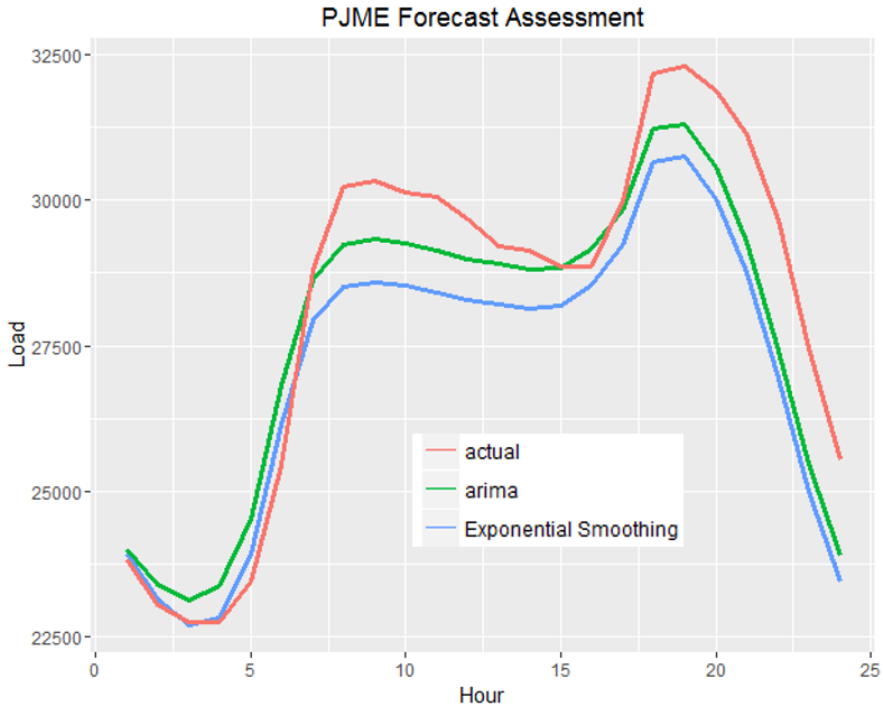


Figure. 2.3 Day-ahead forecast PJM-E using 11-month data

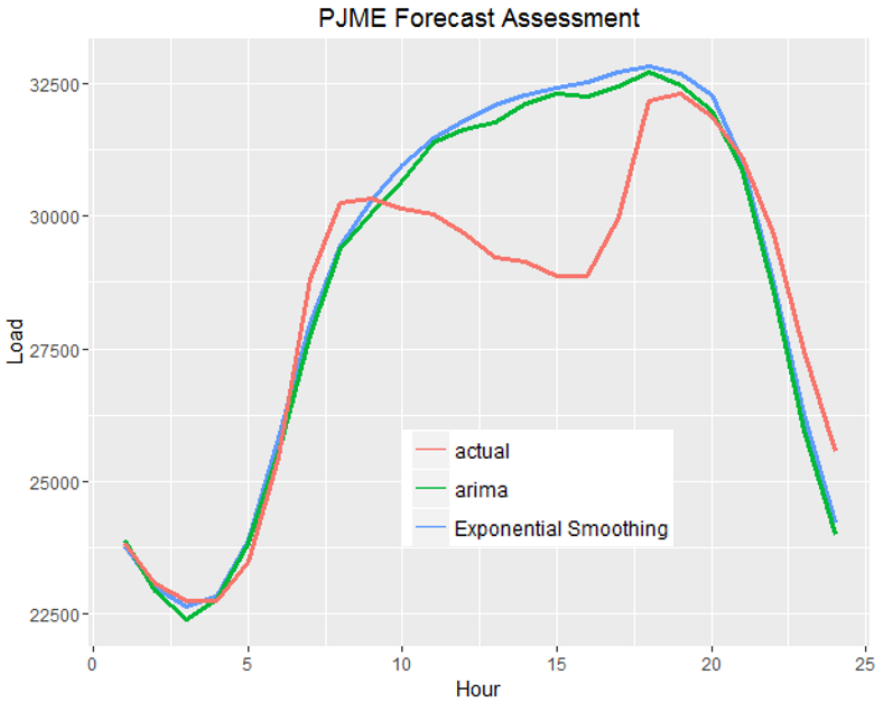


Figure. 2.4 Day-ahead forecast for PJME using 30 days data

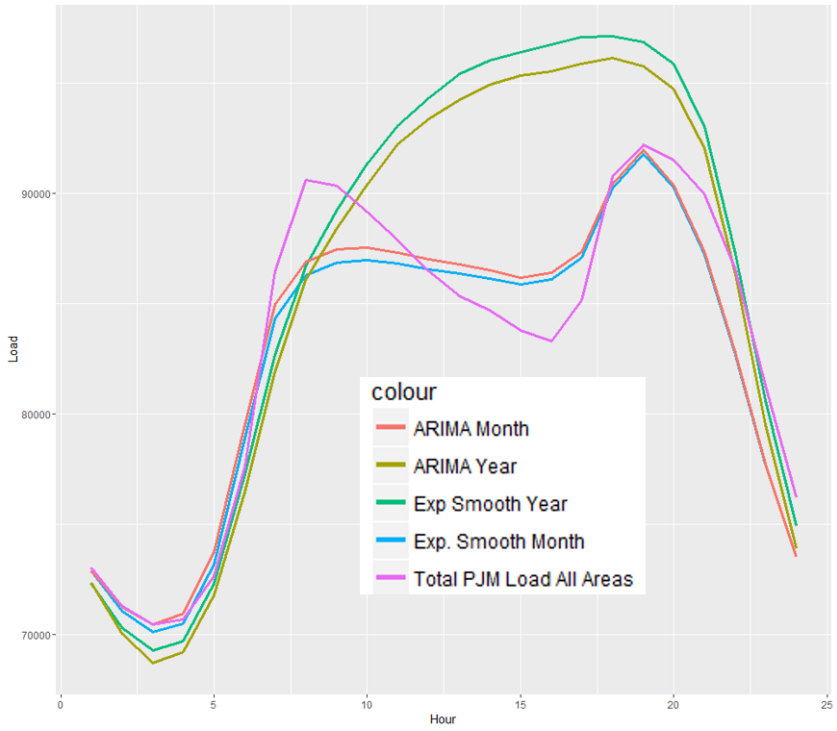


Figure. 2.5 Day-ahead forecast comparison of all the explored methods

Table 2.1 Performance of forecast methodologies for the PJM market

Forecasting Method	***Historical data***	**MAD**	**MAPE**
ARIMA	*1 year*	4287	5.316
Exponential Smoothing	*1 year*	4543	5.0314
ARIMA	*30 days*	1650	1.949
Exponential Smoothing	*30 days*	1687	1.987
ARIMA-ARIMA	*30 days*	884	1.081
Exponential Smoothing-ARIMA	*30 days*	1782	1.544

2.4.3 ED on Forecasted Load

Based on the results of the most accurate load forecasting, a simulated run of ED was carried out for the forecasted hourly load. Figure 2.6 shows the amount of power dispatched by the individual generators to meet the forecasted load demand. For example, at hour 0, GEN 1 to GEN5 should generate 13442 MW, 20754.1MW, 19869.7 MW, 7990.25MW and 10108.9MW respectively. ED for a 24-hour duration based on the

forecasted demand values is shown in Figure 2.6. Combining the forecasting scheme with an ED model allows for the day-ahead scheduling of generators. This combination of techniques provides for operational insight in generator scheduling and planning from a utility perspective. The results indicate that ARIMA performs more accu-

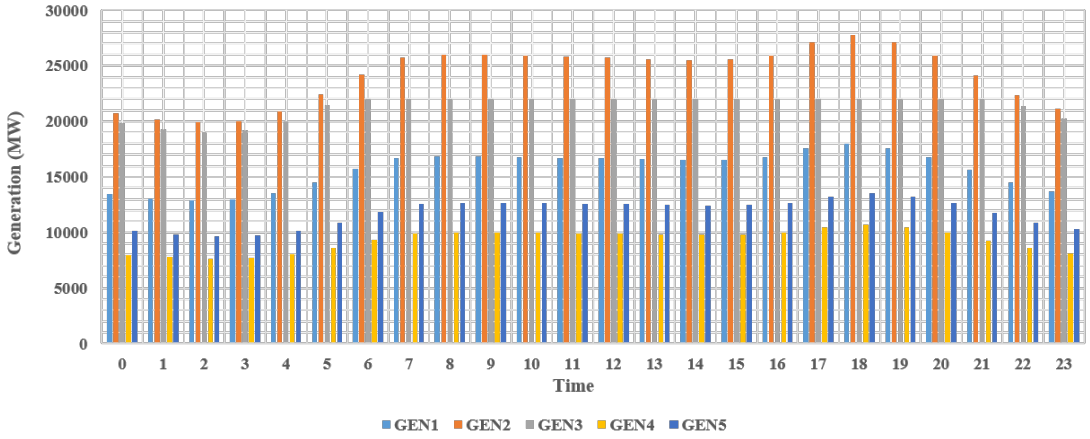


Figure. 2.6 Economic dispatch on PJM data

rately under both training scenarios; in addition, forecasts trained with 30 days prior to forecast day are found to be more accurate. The combination of load forecasting and ED provides a simple framework for day-ahead optimal generation cost and generator scheduling. An ensemble forecasting method, specifically an ensemble ARIMA method in a two-stage combination that forecasts residuals allows for an increase in forecast accuracy. The ensemble ARIMA method produced a forecast error in the range of 1%, which is a considerable increase in accuracy compared to traditional implementations of ARIMA and ES. This increase in prediction accuracy is important for ED and the scheduling of generators.

2.5 System Decomposition and Description

The IEEE 14 bus Test Case, which represents a portion of the American Electric Power System in the Midwestern US as of February, 1962, is used to model the microgrids. The single line diagram of the IEEE 14 bus system model is shown in Figure 2.7 [32], illustrating how a grid can be sectioned into two micro grids. In this division, area 1 or micro grid 1, consists of 3 generators and 4 loads while the area 2 or micro grid 2, consists of 2 generators and 7 loads. The cost coefficients and the maximum and minimum power value of the generators used for the analysis are shown in table 2.2 [26].

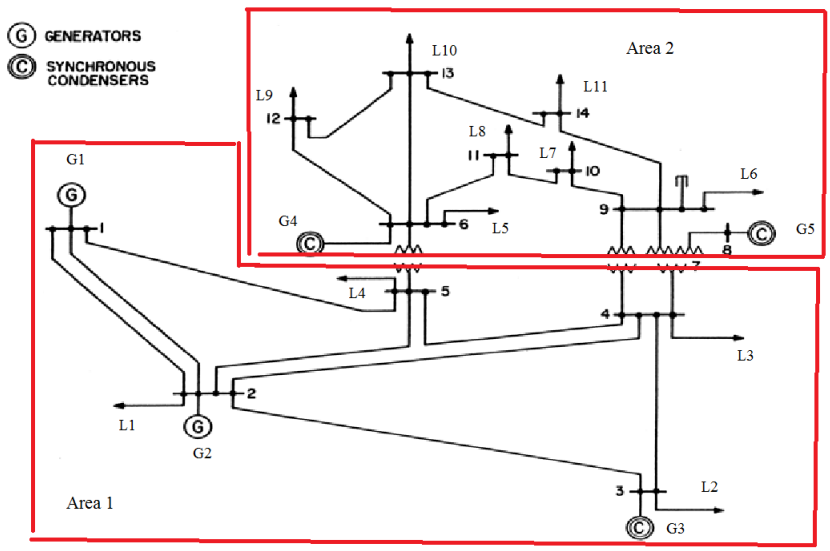


Figure. 2.7 Two grid decomposition model.

The MAED of resource allocation is then performed on the decomposed IEEE 14 bus system, consisting of 5 generators and 11 loads. The developed ED model accommodates two-way power flow between the microgrids. The A Mathematical Programming Language (AMPL) software develops ED model and is solved using the CPLEX (IBM ILOG CPLEX Optimization Studio) solver. The two and three microgrid test

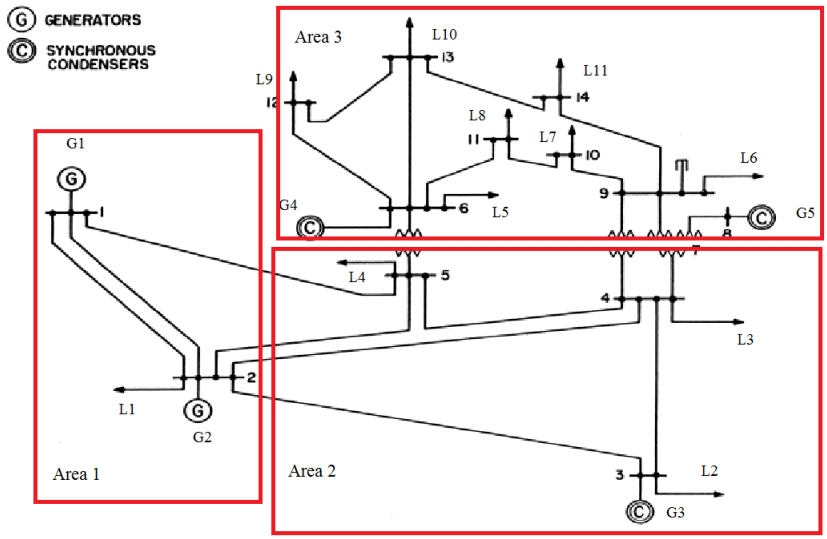


Figure. 2.8 Three grid decomposition model.

Table 2.2 Generator parameters

	a	**b**	**c**	PG_{max}	PG_{min}
G1	0.0301	27.5	750	50	90
G2	0.0195	27.3	1400	30	70
G3	0.0203	30	1050	30	60
G4	0.0507	26.5	450	10	50
G5	0.04	28	600	10	40

system utilized for the study are shown in Figures 2.7 and 2.8.

2.6 Discussion and Results

The model is evaluated on three different test cases.

- *(case 1)* No decomposition
- *(case 2)* Decomposition into two micro grids
- *(case 3)* Decomposition into three micro grids

The variation in the distribution of power generation when there is a load value change is analyzed to study the sensitivity of the generators. The loads in each area is increased in small increments and recorded to observe the generator power distributions until the solution becomes infeasible

2.6.1 No decomposition

The loads L_1 and L_2 are selected randomly and increased in equal amounts, after which the MAED solution is determined using AMPL. The power generation from the 5 generators in the system for the load changes are shown in Figures 2.9 and 2.10.

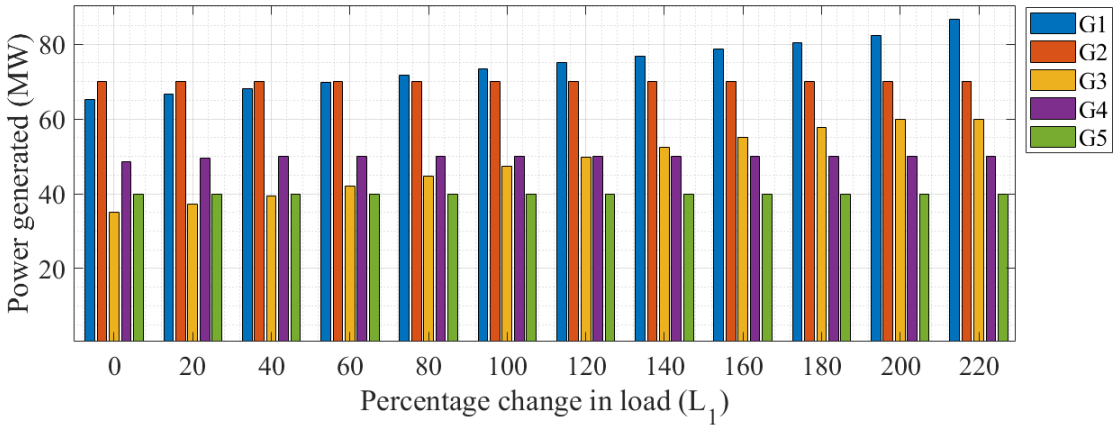


Figure. 2.9 Generation Re-allocation for variation in L_1 with no grid decomposition.

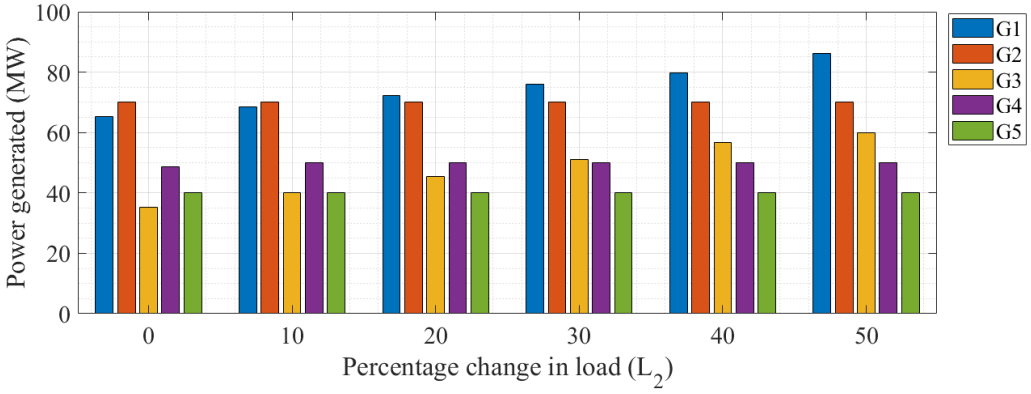


Figure. 2.10 Generation Re-allocation for variation in L_2 with no grid decomposition.

2.6.2 Two Grid Decomposition

Similar to case 1, the loads L_1, and L_2 are increased in equal amounts and the ED solution is determined using AMPL for case 2 with two area decomposition. The power generation from the 5 generators in the system and their respective load changes are shown in Figures 2.11 and 2.12.

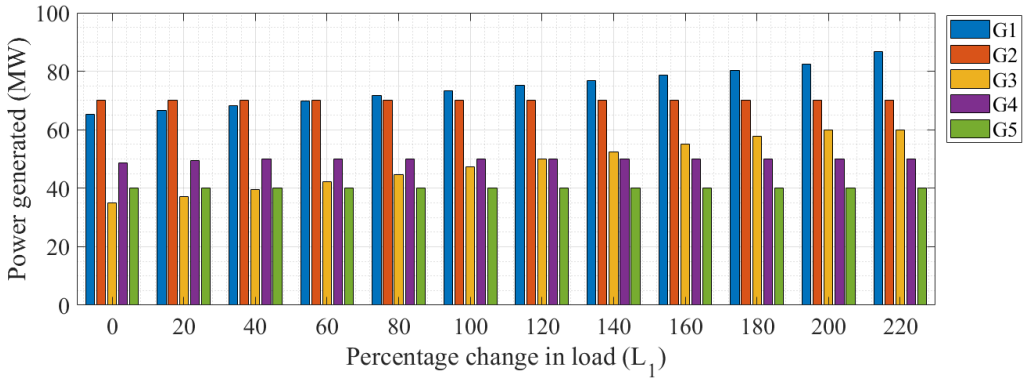


Figure. 2.11 Generation Re-allocation for variation in L_1 in a 2 micro-grid system

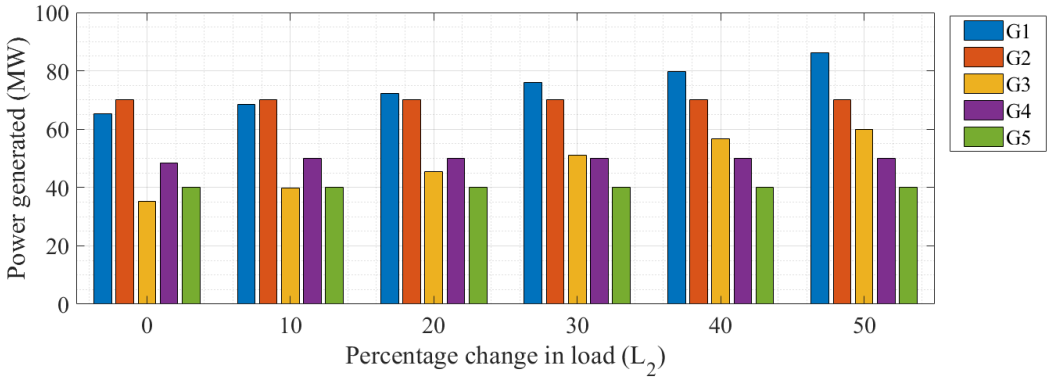


Figure. 2.12 Generation Re-allocation for variation in L_2 in a 2 micro-grid system

2.6.3 Three Grid Decomposition

Similar to case 1, in case 3 the loads $L1$ and L_2 are increased in equal amounts and the ED solution is determined using AMPL with three area decompositions. The power generated by the 5 generators in the system for the load changes are shown in Figures 2.13 and 2.14.

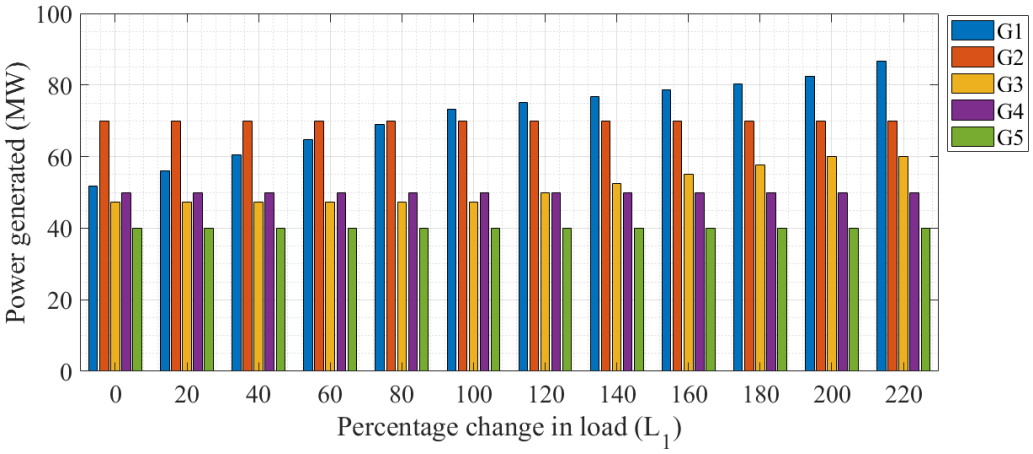


Figure. 2.13 Generation Re-allocation for variation in L_1 in a 3 micro-grid system

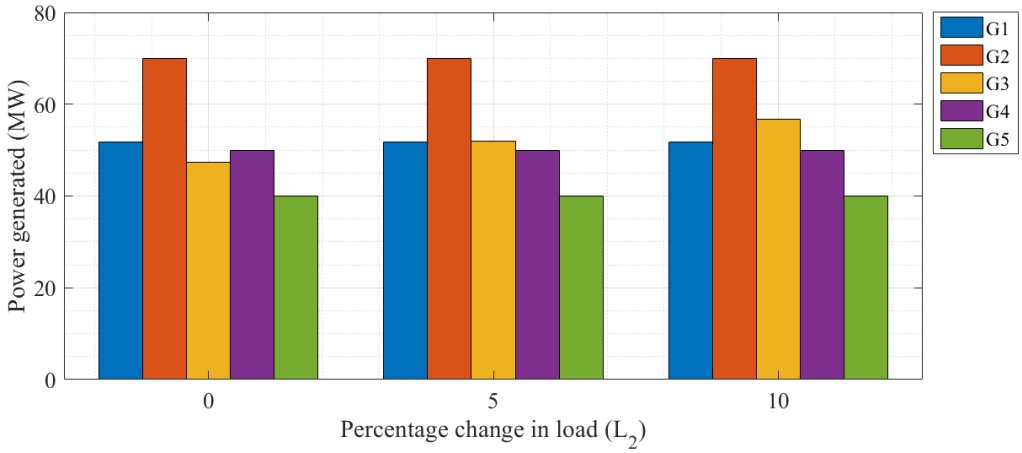


Figure. 2.14 Generation Re-allocation for variation in L_2 in a 3 micro-grid system

Table 2.3 Total variables and constraints in two and three grid decomposition

	Decomposition in to 3 areas			**Decomposition in to 2 areas**		**Without decomposition**
	Area 1	Area 2	Area 3	Area 1	Area 2	
Constraints	5	4	5	7	5	10
Decision variables	3	3	3	4	3	5
Run-time	1.60938			1.59375		1.57812
Total Cost	11859.9			11850.3		11850.3

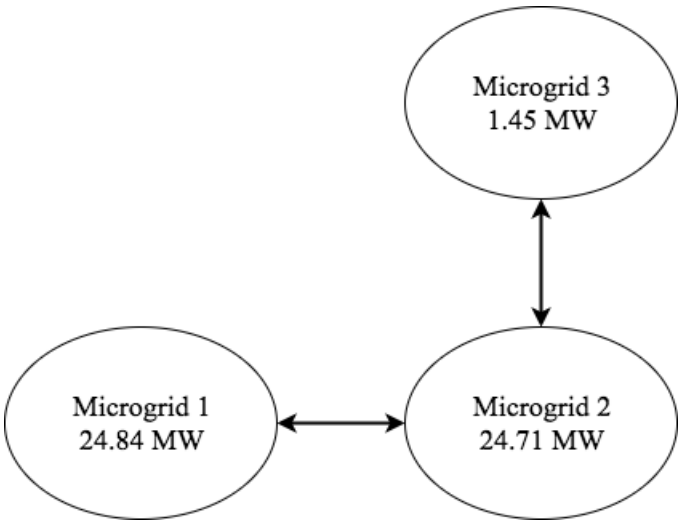


Figure. 2.15 Microgrids with no reserve constraints.

The Figures 2.11 to 2.14 indicate how the grid tries to identify the localized supply of power to the loads within the local micro grid when economic dispatch is carried. In all three cases, generation from G3 increases when the grid is sectioned in to three micro grids to supply power to loads in micro grid 2. Similarly, the generation in G1 reduces in three area decomposition, as other micro grids try to satisfy their demand locally. The number of variables, constraints, run-time and total cost for the grid system are shown in Table 2.3. As an example, allocations of G1 start reducing from 65.24 MW to 51.7 MW, as decomposition is being applied, indicating there is unspent power in the micro grid that can be better utilized in other micro grids where demand needs to be met. Not all generators are optimally re-allocated, which is an important point; the local generations where demand is higher than the supply gets fulfilled first, after which excess supplies are re-routed to other micro grids or areas. Although there is a very small percentage increase (e.g., 0.06%) in total cost of generation, the model successfully optimizes both the inter-transfer and intra-transfer supply among micro grids.

We formulated reserve constraints in the ED model to prevent the loads in the micro grids from utilizing 100% of their generation capacity. These additional constraints will

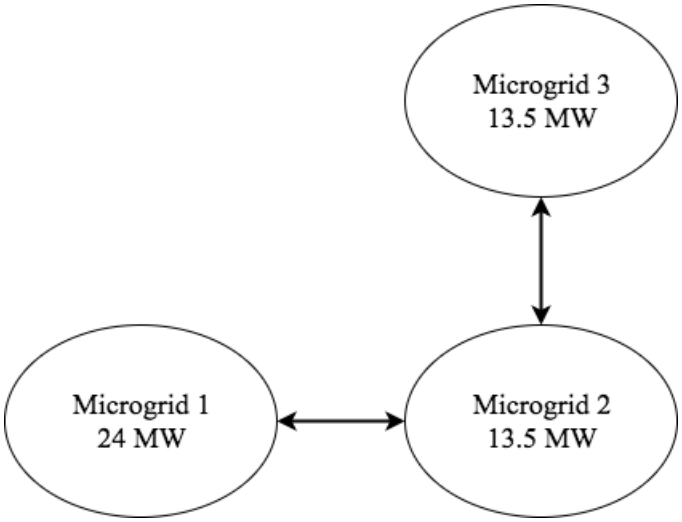


Figure. 2.16 Microgrids with reserve constraints.

make sure the system is not overloaded, and there is at least a certain percentage, 15% in our model, of remaining supply ready to be used in the event of abrupt changes in demand. The reserve capacity in each micro grid for all three area system with and without this additional constraint are shown in Figures 2.15 and 2.16, respectively. In Figure 2.15, area 3 has a very limited supply of power remaining, as there were no reserve constraints, showing how demand is satisfied locally first and it therefore uses its generators heavily. The reserve constraints were added to the model to mandate a minimum requirement of 15% reserve available in all micro grids.

2.7 Graph Clustering for Microgrid Decomposition

In order to develop an economic clustering scheme to decompose a grid system, an ED is applied to the clustered IEEE test systems. We use ED as a method to schedule the generator outputs with respect to its load demands in order to operate the power system most economically. The main objective is to allocate the optimal power generation of different units at the lowest possible cost while meeting all system constraints [19]. Economic-load dispatch is performed in a multi-generator system to schedule the

generators in order to satisfy the loads in the system that are subject to generator and transmission-line limits. In a power system, minimizing the operation cost is important, therefore, we can use ED as an effective way to evaluate the different clustering techniques. The steps followed for the MAED is shown in Figure 2.17. We utilized the IEEE 118 and 300 bus system for the creation of clustered microgrids.

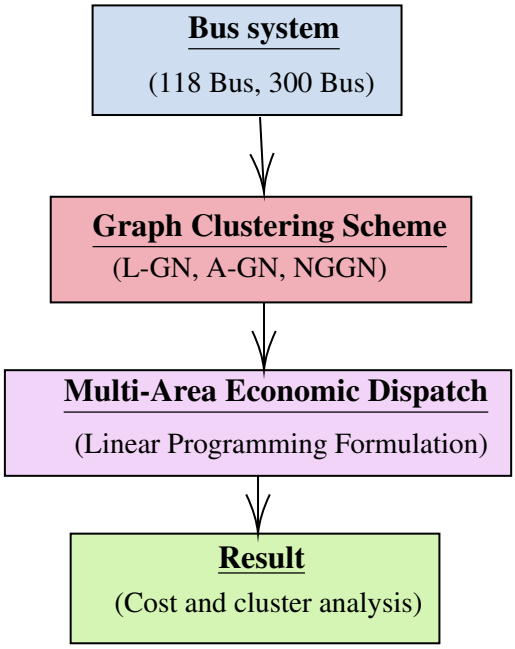


Figure. 2.17 Step followed for multi-area economic dispatch

2.8 Cluster evaluation using Multi-Area Economic Dispatch

The clustering techniques divide the bus system into different zones, or areas and applying ED to a multi-area system is known as MAED. Three clustering techniques, L-GN, A-GN and NG-GN were utilized to group the electric grid structure. The two-stage clustering technique is shown in Figure 2.18. The aim of MAED applications is to minimize the power-generation cost while satisfying the system's load demand subject to the generation and line-flow constraints. The fuel cost for generating unit i to supply a P_{Gi} amount of real power can be represented by a quadratic equation as shown in

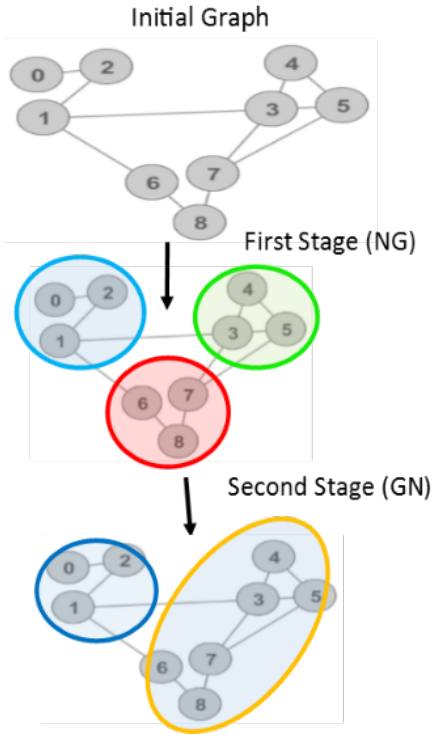


Figure. 2.18 Two-stage process.

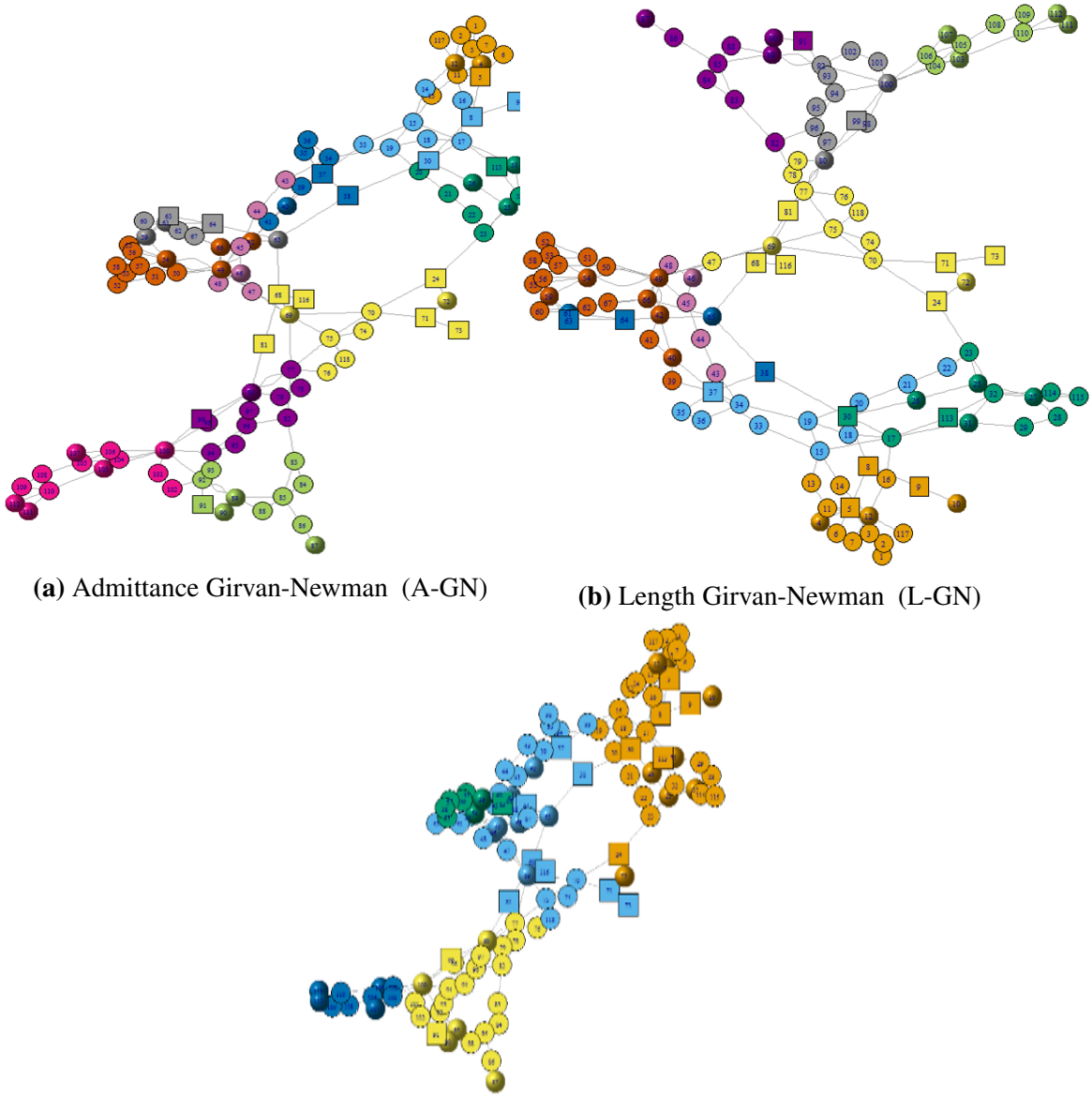
Equation (2.7) [26]. The clustered microgrids for the IEEE 118 and 300 bus test system are shown in Figures 2.19 and 2.20

Tables 2.4 and 2.5 list the generation cost, tie-line flow cost, and total cost for the IEEE 118 and IEEE 300-bus systems for each of the L-GN, A-GN, and NG-GN clustering techniques, respectively. For the 118-bus system, there is a 66.6% reduction in

Table 2.4 ED cost distribution for an IEEE 118-bus system.

	L-GN	**A-GN**	**NGGN**
Number of clusters	10	11	5
Generation Cost ($)	9074.86	9137.03	9148.06
Tie-line flow cost ($)	262.871	87.725	36.45
Total Cost ($)	9337.73	9224.76	9184.51

tie-line flow cost for the A-GN clustered system when compared to the L-GN system and there is a significant reduction of 86.13% for the NGGN method compared to the L-GN method. For the total cost, the cost reductions are 1.21% and 1.64%, respectively,

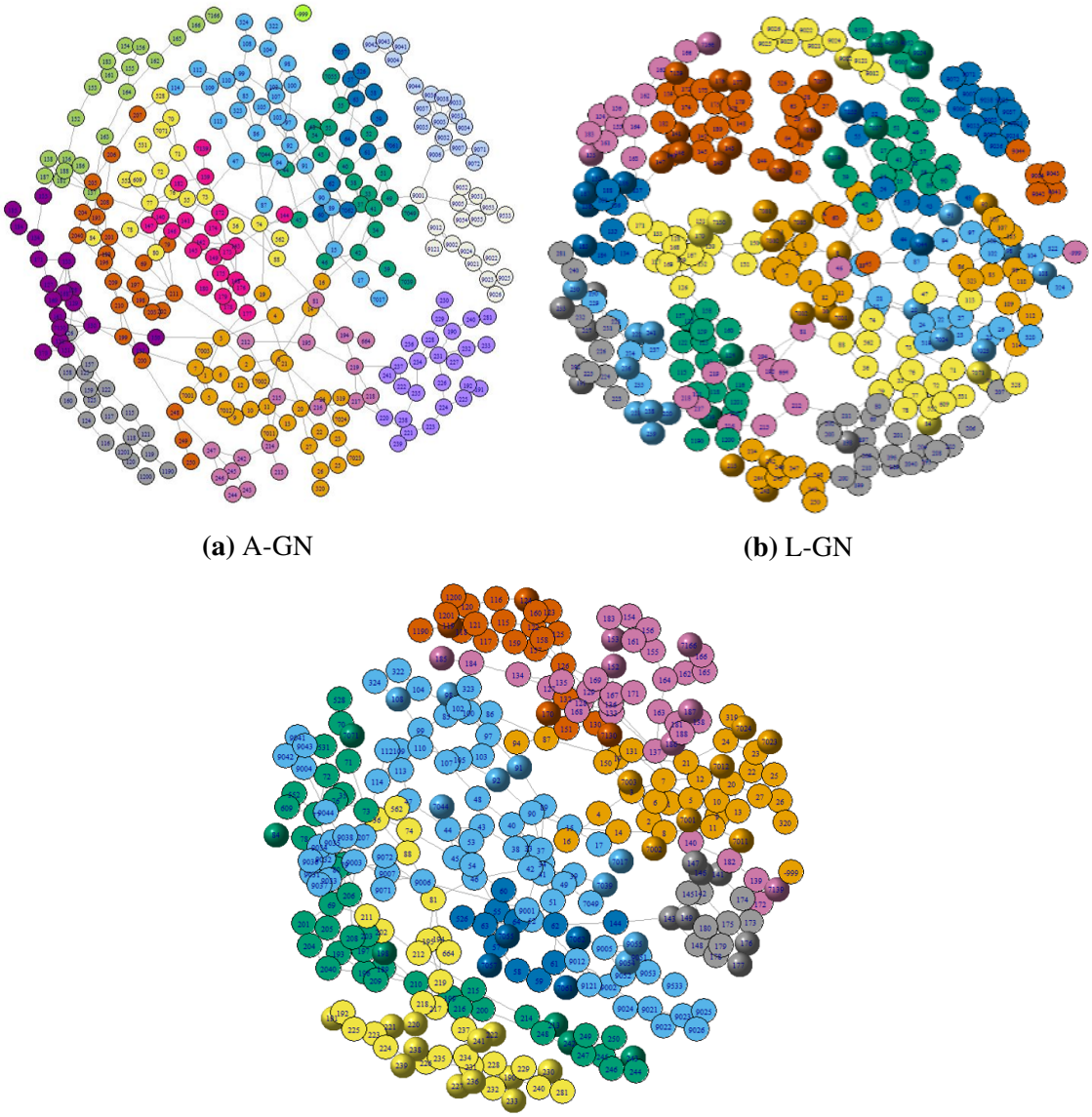


(a) Admittance Girvan-Newman (A-GN)

(b) Length Girvan-Newman (L-GN)

(c) Nearest-Generator Girvan-Newman (NG-GN)

Figure. 2.19 118 Bus system grouped using the different clustering techniques



(a) A-GN

(b) L-GN

(c) NG-GN

Figure. 2.20 300 Bus system grouped using the different clustering techniques

for the A-GN and NGGN method. For the 300-bus system case, the value for the tie-line flow-cost reduction is 0% for the A-GN method as this cluster was identical to the L-GN method and 29.91% for the NGGN method. Similarly, the reduction for the total costs are 0% for the A-GN method and 12.64% for the NGGN method. These results demonstrate a significant reduction in the tie-line flow cost for the NGGN clustering techniques when compared to the L-GN and A-GN technique. The usage of the NGGN method also results in the reduction of total cost for the system. Another parameter pro-

Table 2.5 ED cost distribution for an IEEE 300-bus system.

	L-GN	**A-GN**	**NGGN**
Generation Cost ($)	141704	141704	123824
Number of clusters	14	14	8
Tie-line flow cost ($)	233.089	233.089	163.361
Total Cost ($)	141937	141937	123988

posed to compare grid clusters is the generation to load (G/L) ratio, Equation (2.10), which provides a measure of the self-sufficiency of the micro grid. An important benefit of having microgrids is it's ability to isolate from the main grid in the event of any system wide contingency.

$$\frac{G}{L} = \frac{Generation\ capacity\ in\ the\ microgrid}{Microgrid\ Load} \tag{2.10}$$

A G/L value greater than 100% indicates a self-sufficient grid cluster with excess generation which can be given to other regions. A G/L value less than 100% indicates the generation within the cluster is not sufficient to satisfy its load, therefore requiring resources from neighboring regions to meet the demand. Tables 2.6 and 2.7 list the maximum and minimum values of G/L for the IEEE 118-bus and 300-bus systems,

excluding the zones with no active power generation. The data analysis suggest the NG-GN clusters are more suited due to their self-sufficiency, when compared to the other two cases, since the values are closer to the ideal value of 100.

Table 2.6 Generator/Load ratio of the IEEE 118-bus system clusters.

Generation/Load ratio	**L-GN**	**A-GN**	**NGGN**
Maximum value (%)	198.61	277.08	124.48
Minimum value (%)	49.62	47.22	50.63

Table 2.7 Generator/Load ratio of the IEEE 300-bus system clusters.

Generation/Load ratio	**L-GN**	**A-GN**	**NGGN**
Maximum value (%)	616.61	616.61	411.86
Minimum value (%)	61.9	61.9	86.96

CHAPTER 3

Optimal Operation of Residential EVs using DNN and Clustering based Energy Forecast

3.1 Introduction

The automobile industry is now going through a transitional phase with gas powered vehicles being replaced by Electric Vehicle (EV) and hybrid models. The ability to deliver efficient and clean mode of transportation and government incentives have increased the consumer interest towards EVs [33]. The increased customer demand for EVs have led the change with all the major automakers now manufacturing hybrid and Plug-in Electric Vehicle (PEV) variants of their gas powered vehicle models.

Although EV sales are a small fraction of total US purchases, 1.8% in March of 2019 [16] market share is expected to increase, creating a challenge for utilities who must meet this new energy demand. The flexibility of the grid must, therefore, be increased to accommodate the charging of these vehicles in the form of either Vehicle-to-Grid (V2G), Vehicle-to-Home (V2H), Vehicle-to-Vehicle (V2V) or charging schedule. The V2G system utilizes the EV battery as a storage system to deliver energy back to the grid in times of peak load. The V2V and V2H work similarly by allowing energy exchanges between vehicle to vehicle and vehicle to home in times of need [34].

The increasing percentage of EVs in residential communities necessitates the need

for a scheduling model to charge EVs. The actual load from MISO from July 1st through July 7th, as shown in 3.1, represents an example for a "Duck Curve" pattern. The increased demand for power during peak times for a utility, starting at approximately 5 PM, creates an additional load on an already stressed power grid and increased prices for consumers. High peak loads require innovation on the part of utility companies as they look for methods in which to reduce the peak load, also called 'peak-shaving', such as relying on demand response and TOU pricing schemes.

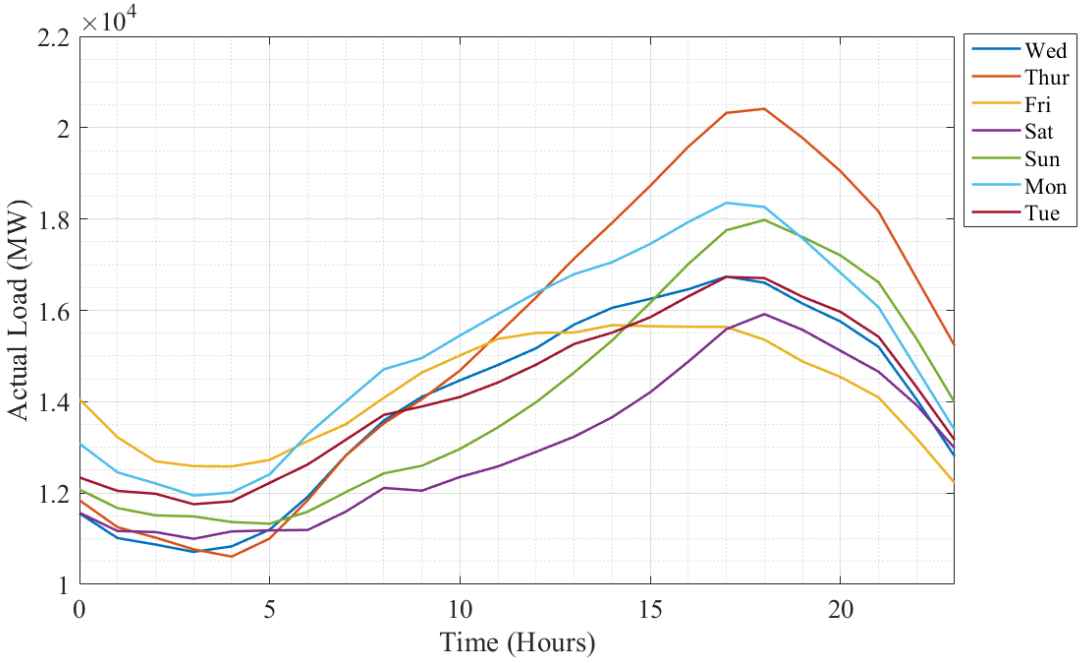


Figure. 3.1 Load from the MISO area from July 1-7, 2020.

The impact on the distribution system from the increased market penetration of EVs is discussed in [35], which also illustrates the effects on electricity generation adequacy,transformer aging and the distribution of power. The authors of [35] also proposes mitigation techniques such as Time-Of-Use (TOU) pricing schemes and smart charging algorithms to mitigate the effects of connected EVs. A coordinated framework to charge an EV fleet was presented by M. Usman et al. [36], proposing a control scheme to maintain the grid capacity while satisfying the needs of the EV fleet.

We therefore propose the creation of a scheduling model for a community with high EV market penetration. The method used will encompass the next-day forecasted EV usage, consumer preference for the amount of charge, and the charger unit to avoid a surge in demand from the community due to price drop or simultaneous charging from a large number of EVs. The proposed EV scheduling model is based on forecasted day-ahead EV demand based on historic values, and will also considers the forecasted household demand. The approach of utilizing household demand, along with demand in the community in the scheduling model, has not been seen in the reviewed literature.

3.2 Dataset

The demand profiles of 200 randomly selected households from the Midwest region of the United States [37] are generated by using a modeling technique proposed by Muratori et al. [38, 39]. This modeling technique produced residential power consumption data with a 10 minute resolution and was validated using metered data, which represented the total electricity usage in Watts and households varying in size and number of residents [40].

The variations in demand for a single household, household 1 in the dataset, for a whole year is shown in the Figure 3.2. There is a surge in demand during the month of June to the end of September, which may be caused by the increased load of air conditioning units used during the summer months. The energy usage of all 200 households for a single day (January 1st), is shown in Figure 3.3. This figure does not show any pattern for single day household energy usage, but at an ISO level, the energy usage for a day usually exhibits the duck curve pattern.

The residential electricity usage from household's L1 and L2 charging ports for the

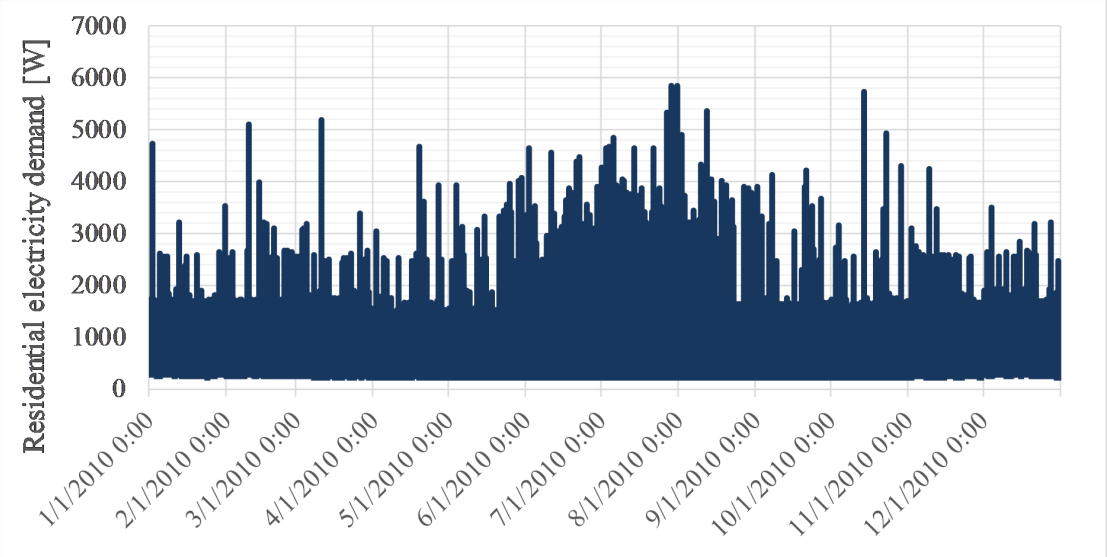


Figure. 3.2 Electricity demand from household 1.

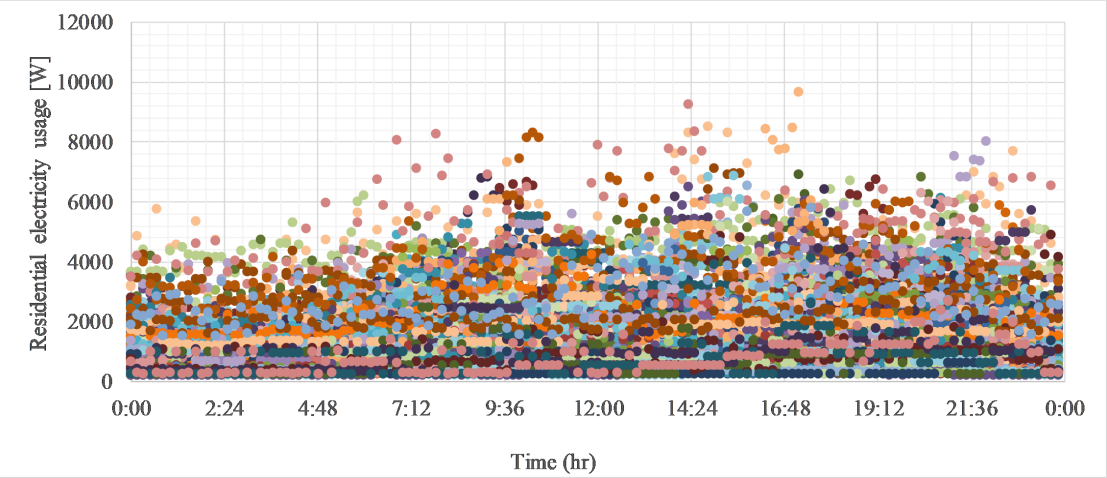


Figure. 3.3 Electricity usage from all the 200 households for a single day (Jan 1st).

year of 2010 is shown in Figures 3.4 and 3.5. The electricity demand from the L1 port will be held 1920 W and the value from L2 port will be held at 6600W.

3.3 DNN based Day-Ahead Energy Forecast

3.3.1 Clustering and k-means cluster analysis

Our model consists of a set of 200 households, with each households owning one to six EVs, for a total of 348 EVs connected. Using the K-means clustering technique,we utilized a 5-cluster system to group the households, 7 to 99 per cluster, based on their en-

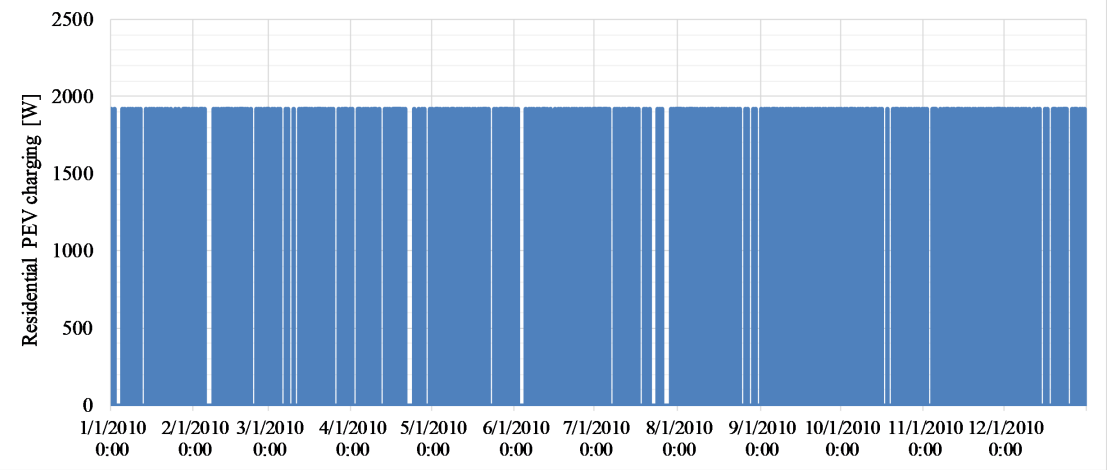


Figure. 3.4 Residential PEV charging from household 1 (L1 port).

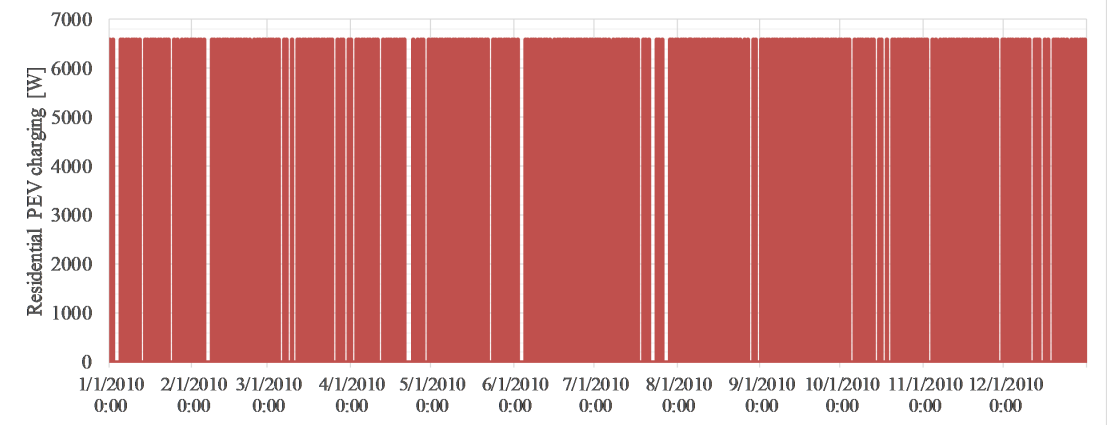


Figure. 3.5 Residential PEV charging from household 1 (L2 port).

ergy consumption pattern. The clusters are listed in Table 3.1. The day-ahead scheduling for EVs are based on the forecasted demand of EVs and household energy consumption. A DNN based forecasting was used to predict the energy consumption of EVs and households based on a year-long dataset with 10-minute resolution.

3.3.2 Designing Deep Neural Networks

In our model, we used Keras on top of TensorFlow to design the DNN. Keras is an open source neural-network library developed using Python, and is capable of running on top of TensorFlow, Theano or R. Now one of the most popular ML libraries, Keras allows the users to utilize GPU to accelerate neural network training. TensorFlow is an

Table 3.1 Household clusters

Cluster	No. of households	No. of EVs
Cluster 1	7	22
Cluster 2	28	111
Cluster 3	9	20
Cluster 4	27	54
Cluster 5	99	141

open source platform for machine learning, and is an interface for the development and implementation of ML algorithms.

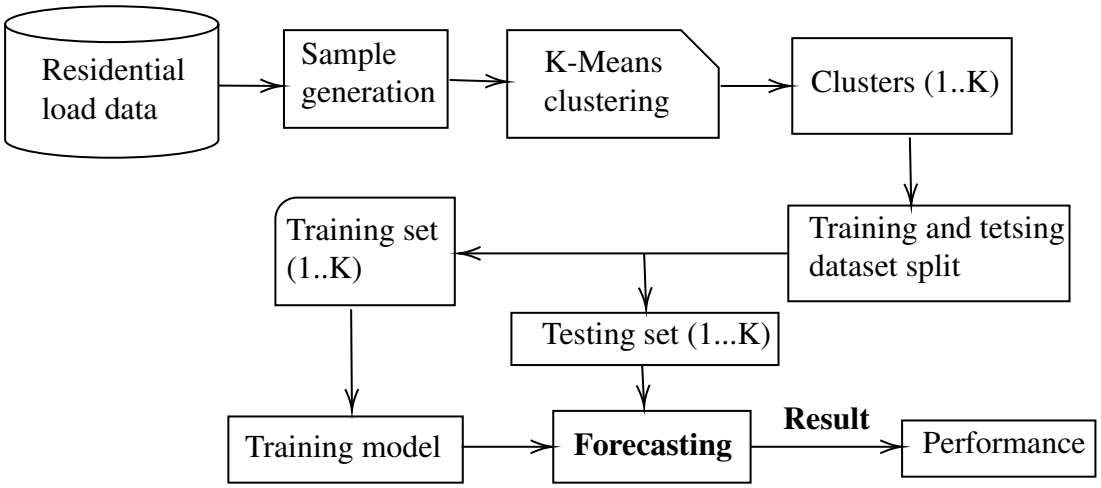


Figure. 3.6 Flows of constructing the DNN based method with K-Means algorithm

3.3.3 Results

The prediction accuracy of the forecasting approach is evaluated using MAPE, which defines the accuracy as a percentage, given in Equation 3.1 [41]. Here, n represents the number of instances in the forecast period, y_t the actual value and $\hat{y}_t$ the forecasted value.

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{3.1}$$

Table 3.2 Deep learning TensorFlow parameters

Parameters	**Values**
Total number of samples	52560
Training samples	52416
Validation samples	1440
epochs	100
Learning rate	0.001
Mini batch size	100
Activation function	Linear
Number of hidden layer	3

The MAPE values for the forecasting results of individual clusters are given in Table 3.3. These values range from 1.0064 for the cluster with the smallest number of households, to 3.2936 for the cluster with the largest number of households, therefore,the forecasting model is able to deliver accurate results.

Table 3.3 Cluster wise residential household MAPE values

Cluster	**MAPE**
Cluster 1	1.0064
Cluster 2	1.1018
Cluster 3	1.3347
Cluster 4	1.3078
Cluster 5	3.2936

3.4 EV Scheduling Model

The future power grid system needs to accommodate the increased energy consumption from EVs without a surge in demand during peak hours. The scheduling model needs to consider user comfort, household energy consumption, EV demand,energy price as well as other factors, while reducing the electricity cost of the total EVs in the system.A cost reduction model will make the use of EVs more cost-effective, therefore customers will adjust power usage according to the scheduling scheme based on TOU price.

Table 3.4 Optimization model parameters

Sl. No	Model Parameters
1	Forecasted EV demand
2	Forecasted household demand
3	Schedule time
4	Day-Ahead energy price
5	Consumer preferred charging time
6	Consumer preferred charger type (L1 or L2)
7	Final State of Charge (SOC)
8	EVs in each household
9	Households in each cluster

3.4.1 System Modeling

The objective of the proposed model is to minimize the total electricity cost for EV charging in a residential location, as shown in Equation (3.2). The different parameters used for the scheduling model are listed in Table 3.4

$$\text{Min } F = \sum_{i=1}^{M} \sum_{j=1}^{N} P_{ij} C_j \tag{3.2}$$

Here, M stands for the number of connected EVs in the system and N represents

the varying energy price time-slots. N represents the energy scheduled for the i^{th} EV in time-period j along with the energy cost. The EV scheduled in the model is subject to several user-defined and power system constraints [42, 43, 44].

1. Demand constraint

   The demand constraint, shown in Equation (3.3), assures the forecasted demand for each individual EVs is met while scheduling for the time-period desired.

$$\sum_{j=1}^{N} P_{ij} = EV_j \tag{3.3}$$

   EV_j represents the forecasted demand for the EVs in the system. The sum of all energy scheduled in N time periods should meet the forecasted demand.

2. Consumer preferred time

   Preferred time duration for charging EVs will vary with the customer. The model allows the residential consumer to specify his preferred time of charging for the EV,therefore the energy scheduling model will only be allocated within the user-specified time period, as represented in Equation (3.4).

$$P_{ij} \geqslant 0 \quad T_{ij} = 1 \; else \; 0 \tag{3.4}$$

   T_{ij} represents the user preference for the i^{th} EV at time j. The program accepts the user preference as a set of ones and zeros for the next 24 hours.

3. Consumer preferred charger

   The model provides an option for customers to identify a preferred charging port.

The energy allocation will be determined based on the user specified charging port, either L1 or L2.

4. EV charger rating

   The charger rating ensures the constraint of the specified charger capacity allocated for each time period, as represented in Equation (3.5). If the rated charger capacity is greater than that allocated by the optimization model, the charger points will not be able to support the increase in power level demand.

$$P_{ij} \leqslant Charger\ Rating \tag{3.5}$$

   The charger rating will be 1.92KW (L1) or 6.6KW (L2) depending on the user preference for each individual EV.

5. State of Charge (SOC)

   The user can also provide the percentage of demand to be met for the next day. If the user does not plan to make longer trips with their EVs, a reduced charging percentages can be allocated to reduce cost.

$$\sum_{j=1}^{N} P_{ij} \geqslant (FSOC_j) * EV_j \tag{3.6}$$

   FSOC represents the user specified percentage of charge for each individual EV.

6. EVs connected to the household

   All of the households evaluated in this dataset have connected EVs, with ownership varying from one to six vehicles. The scheduling of all connected EVs in a

household at the same time can cause a sudden surge in demand and can create power stability issues. The constraint throttles the scheduled EV energy in each hour for a single household to a maximum of 25% of the total demand.

$$\sum_{i=1}^{E_h} \sum_{j=1}^{N} P_{ij} \leqslant K_h H \tag{3.7}$$

For each household in the system, E_h represents the number of connected EVs , H is the corresponding household demand and K_h represents the percentage of household demand allowed in each hour.

7. Households in a cluster

In each cluster, the number of households vary from 7 to 99 and have multiple EVs connected to them. If all of the EVs in a cluster are plugged in at the same time, it can create transmission and generation issues for power companies. This constraint limits the total scheduled energy for all EVs in the cluster for each hour, up to a maximum of 20% of the total household demand in the cluster.

$$\sum_{i=1}^{E_t} \sum_{j=1}^{N} P_{ij} \leqslant K_c \sum_{h=1}^{H_c} H_j \tag{3.8}$$

For each cluster, H_c represents the number of houses, E_c represents the connected EVs for all the households and K_c represents the percentage of total demand allowed in each hour.

The day-ahead energy price from the MISO on December 31st, 2016 was used for this model [45]. MISO is an RTO providing open-access transmission service, and monitors the high-voltage transmission system in the Midwest region of the United

Table 3.5 MISO day-ahead energy price values

Time-period	Energy Price	Time-period	Energy Price
0	0.01322	12	0.01978
1	0.01064	13	0.02021
2	0.00925	14	0.02156
3	0.00821	15	0.02116
4	0.009	16	0.02175
5	0.00883	17	0.02883
6	0.00912	18	0.03114
7	0.01288	19	0.02453
8	0.01514	20	0.02133
9	0.01669	21	0.01933
10	0.01927	22	0.01885
11	0.01976	23	0.01456

States. Figure 3.7, shows the variation in energy price for different hours of the day.

Customers have an option of L1 and L2 charging points. Many consumers prefer to use L2 charging points due to the increased speed of charging. An L1 port is capable of charging a depleted EV battery in 20 hours or more from a standard 110V outlet while an L2 port can do the same in four to six hours from a 220V outlet. While an L2 port can be used to deliver short bursts of power, consumers usually keep their EVs plugged in all night, and may not necessarily need the speed of a L2 charger. The households in the different clusters are assumed to have access to both L1 and L2 charging points, therefore optimization is performed taking consumer preference into account.

The proposed scheduling model also provides an option for households to provide their preferred timing, any period within a 24-hour duration, for charging their individual EVs. Some customers may not consistently use their EVs for long distance travels and therefore do not require fully charged batteries. This model supports the option to

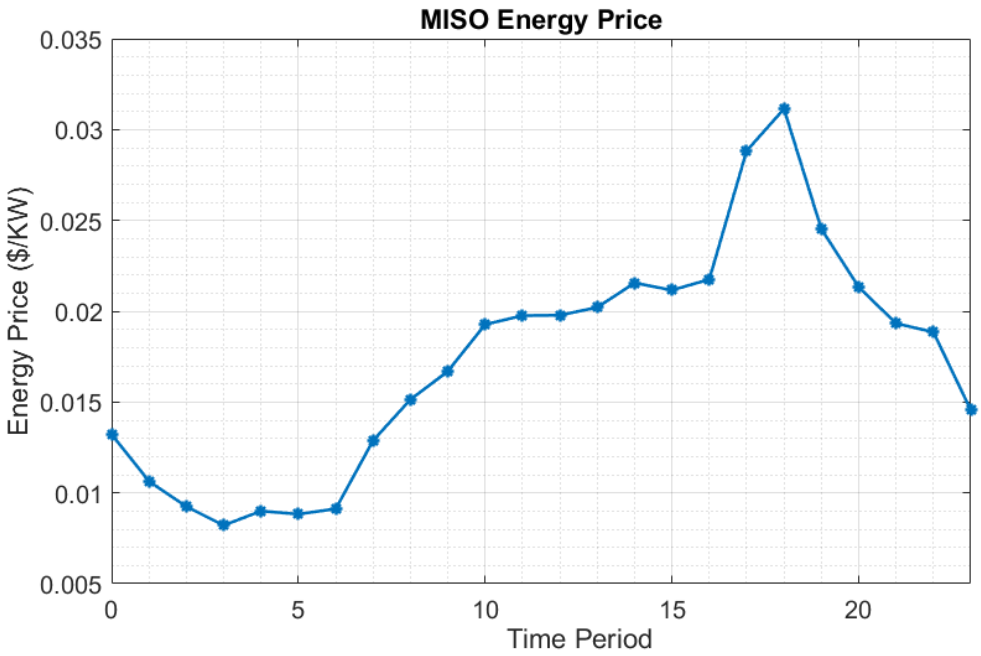


Figure. 3.7 MISO day-ahead energy price

supply a reduced amount of charge for the next-day according to the customer's needs.

3.4.1.1 Optimization

AMPL is used in this research to model the linear programming based scheduling model. AMPL is an algebraic modeling method created at Bell Laboratories by Robert Fourer, David Gay, and Brian Kernighan and is used to solve large scale optimization and scheduling problems. The AMPL coding syntax is similar to the mathematical notation of optimization problems, which helps developers program their models, and supports several open-source and commercial solvers.

3.5 Results and Discussion

The forecasted results were used in the second part to develop a scheduling scheme for EVs in the region to minimize cost and prevent aggregated charging during low demand, high price time slots. The total energy cost of connected EVs in each individual cluster

is represented in Figure 3.10. The energy cost is almost reduced by half for clusters properly scheduled to connect their EVs to the power grid. The individual household energy costs for their EVs are shown in Figure 3.8 and 3.9. All of the households in each cluster show a cost savings from the scheduling process.

The model can perform the EV schedule based on user preference. The Figure 3.10 illustrates the cluster wise cost analysis. The model schedules EVs during off-peak hours when the price is lower to shift the load from peak hours.

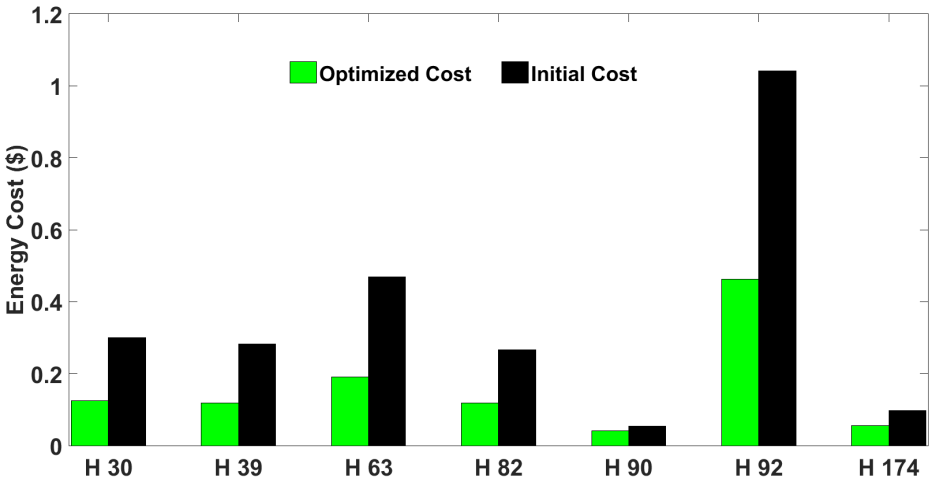


Figure. 3.8 Household cost analysis in cluster 1.

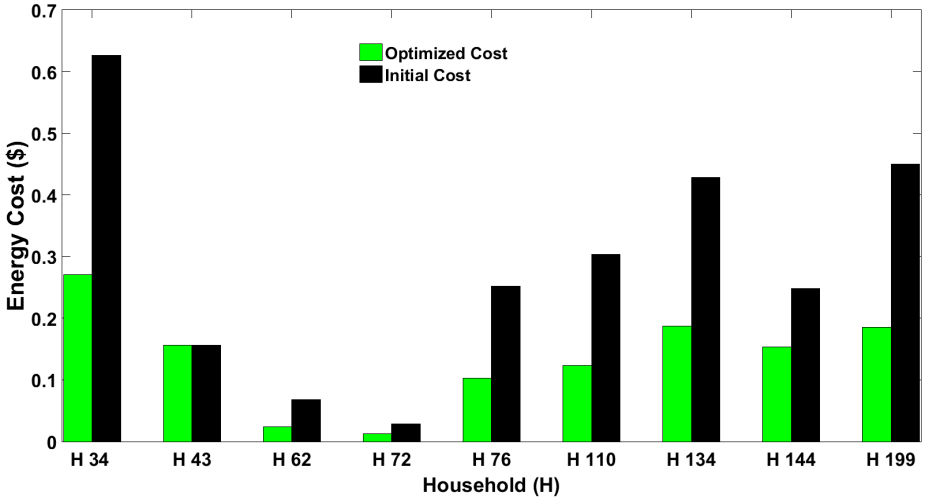


Figure. 3.9 Household cost analysis in cluster 3.

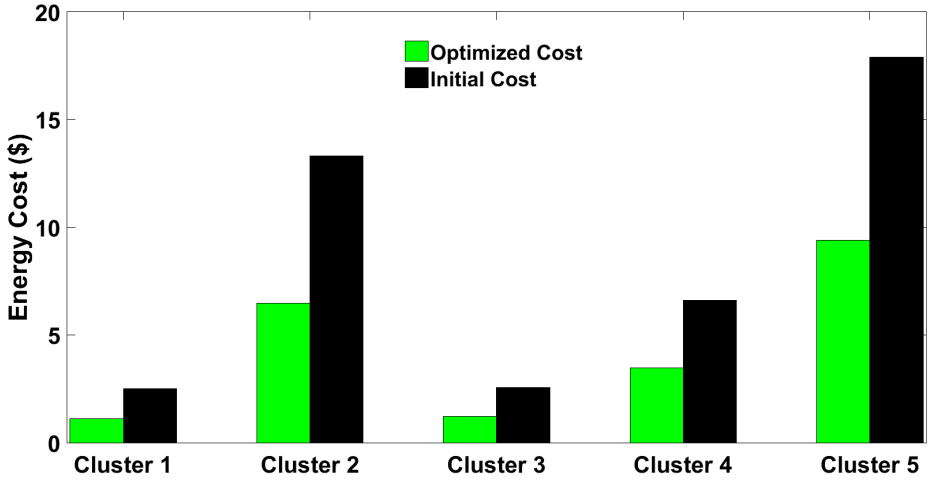


Figure. 3.10 Household cost analysis in the clusters.

3.6 Concluding Remarks

A novel electric vehicle charging schedule is proposed using concepts from DNN, K-means clustering and linear programming. A dataset with 200 households, 348 EVs and 52560 data points were utilized for the study. A K-Means algorithm was employed to segment the data set into subsets in order to construct the household clusters. A DNN based forecasting tool was then applied on the clustered dataset, which improved the forecasting accuracy, and allowed for the generation of the day-ahead demand for residential households and EVs. The linear programming based scheduling model was able to use the forecast demand to generate the charging schedule for all EVs in the cluster. The scheduled EVs reduced the energy cost for consumers along with reduced power surge from households and residential regions.

CHAPTER 4

Quantifying Uncertainty under High DER penetration (Wind)

4.1 Introduction

The installed wind capacity in the US is over 100,000 MW, with an installed capacity of approximately 29,000 MW, the state of Texas produces the most wind power in the country. The state of North Dakota has over 3,500 MW installed wind power [46]. The US wind industry recorded a 117% increase in new wind power installation in the first quarter in 2020 over the same period in 2019 [47]. Forecasting wind power generation is challenging, as there are several uncertain variables during the generation process. Calculation of wind power generation depends on several parameters such as wind speed, wind direction, turbine swept area, air density, power coefficient, turbine height, cut-in speed and cut-out speed. For example in [14], discusses uncertainties such as wind-flow, equipment failure, sensor assemblies and their related inaccuracies and calibration. The authors proposed a probability of exceedance concept to quantify the uncertainty in calculating the total wind power and calculated the net-wind energy production using a normal distribution.

Wind speed can be approximated using historical and Numerical Weather Predictions (NWP) models, but power generated is difficult to predict accurately from wind speed and power curve models. Previous studies have used different statistical mod-

els for wind speed distribution such as Weibull, Rayleigh and β distribution, and have shown the probability distribution of wind speed is adequately represented by Weibull distribution. The Weibull distribution model for a wind model is shown in Equation 4.1 and is plotted in Figure 4.1. Here, v represents the wind speed and a and b are the scale and shape parameters of the Weibull distribution. An extensive review on short-term wind power forecasting is performed in [48]. The authors conclude the superiority of statistical models over physical methods, e.g. NWP models, for short forecasts of 6 hours or less [49, 50]. The NWP model requires hours of computation time and the resultant spatial interpolation can cause imprecision [51].

$$f(v) = \frac{b}{a}(\frac{v}{a})^{b-1} exp[-(\frac{v}{a})^{b}] \tag{4.1}$$

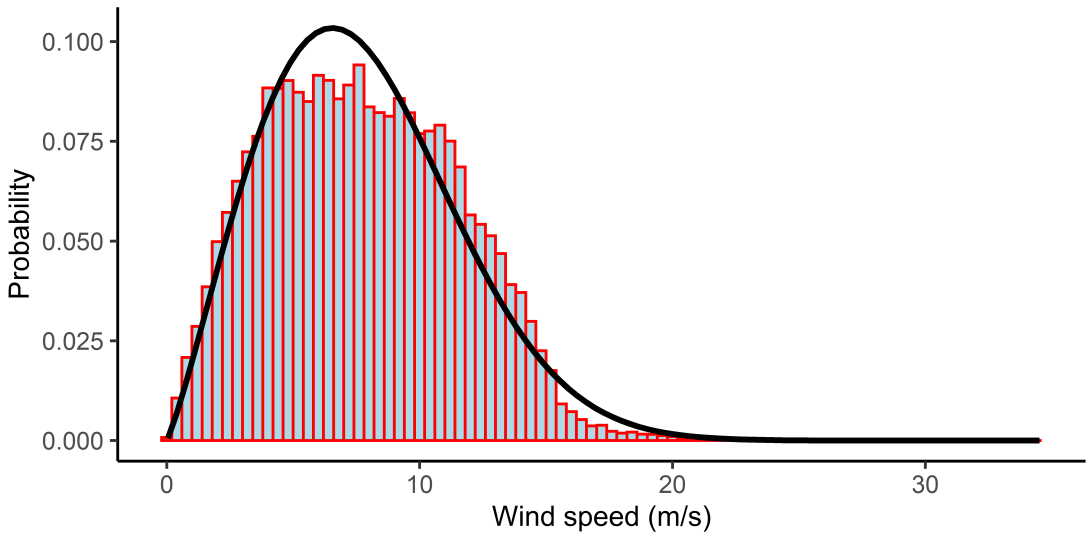


Figure. 4.1 Wind speed probability density distribution plot.

The Weibull distribution of a wind model can be generated using several numerical methods such as the Graphical method, Standard Deviation (SD) method, Maximum Likelihood method (ML), Modified Maximum Likelihood method (MML), Empirical

method of Justus (EMJ), Empirical method of Lysen (EML) and Energy pattern factor method (EPF). Graphical methods use least square regression to fit the data to a curve, which requires sorting the time-series data in to bins. The Standard Deviation method uses the mean and standard deviation to calculate the Weibull parameters. EMJ is a special case of the SD method, where the shape parameter, b, is calculated using Equation 4.2. In EML, shape parameter calculation is similar to EMJ and the scale parameter, a, estimation is shown in Equation 4.3. In the Energy pattern factor method, the scale parameter calculation is similar to the SD method and the shape parameter calculation is related to the average calculated wind speed. The Maximum likelihood method uses numerical iterations to calculate the Weibull parameters; this method uses wind speed data in a frequency distribution format. Although several methods exist to evaluate the parameters for a Weibull distribution, it is difficult to calculate an accurate estimate.

$$b = v_m(0.568 + \frac{0.433^{-\frac{1}{k}}}{a}) \tag{4.2}$$

$$a = (\frac{\sigma}{v_m})^{-1.086} \tag{4.3}$$

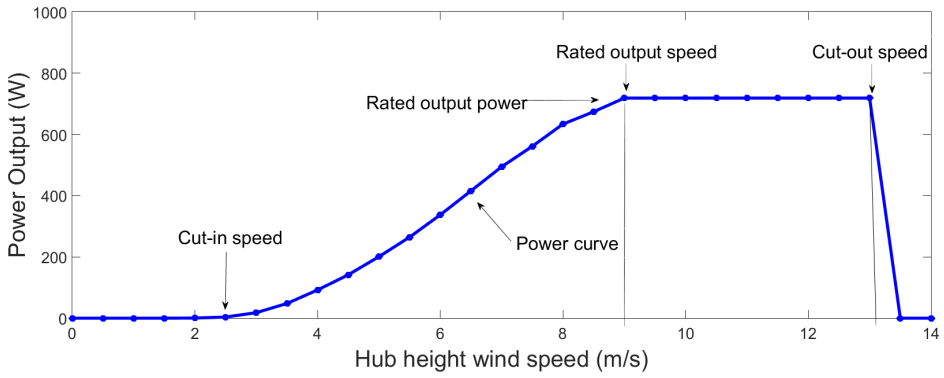


Figure. 4.2 Ideal wind power curve

A general formula for converting wind speed into wind power (P) is given in Equa-

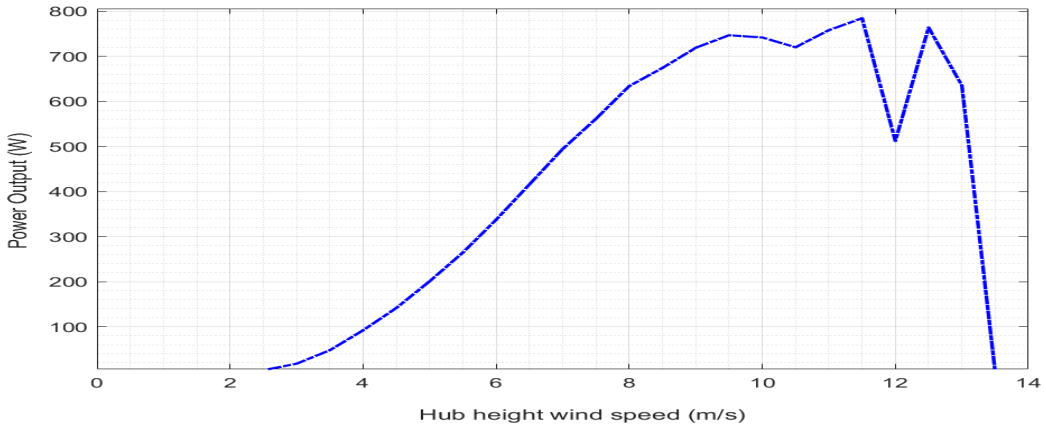


Figure. 4.3 Wind power curve

Table 4.1 Wind Energy assessment uncertainty parameters [52].

Wind Speed Measurement	1. anemometer uncertainty I: calibration uncertainty
	2. anemometer uncertainty II: dynamic over-speeding
	3. anemometer uncertainty III: vertical flow effects
	4. anemometer uncertainty IV: vertical turbulence effects
	5. tower effects
	6. boom and mounting effects
	7. data processing accuracy
Long-term Resource Estimation	8. MCP Correlation Uncertainty
	9. Weibull Parameter Estimation Uncertainty
	10. Changes in the Long-term Average
Wind Resource Variability	11. Inter-Annual Variability Uncertainty
	12. Uncertainty over Turbine Lifetime
Site Assessment	13. Topographic Effects
	14. Wind Shear Model Uncertainty

tion (4.4). Here, ρ represents the air density in kg/m^3, A the area swept by the blades, V the wind speed in m/s at hub height and C_p is the turbine performance coefficient. An example of an ideal wind power curve is given in Figure 4.2. Turbines while in operation will not be able to perform ideally, an example of wind turbine power curve is given in Figure 4.3. The different uncertainties associated with wind power generation are listed in Table 4.1 [52].

$$P = \frac{1}{2} C_p \rho A V^3 \tag{4.4}$$

Independent System Operator (ISO) relies on energy curtailment as the main method to manage the oversupply of RES. Curtailment is the reduction in RES energy generation from its capacity when there is insufficient demand to consume energy production, which results in a loss of resources and an opportunity to produce carbon free energy. Increased solar energy generation in the middle of the day, along with uncertain wind generation creates a challenge for utilities attempting to satisfy the demand which follows the expected duck curve. In 2015, CAISO, curtailed more than 187,000 megawatt-hours (MWh) of wind and solar generation, 308,000 MWh in 2016. Figure 4.4 shows

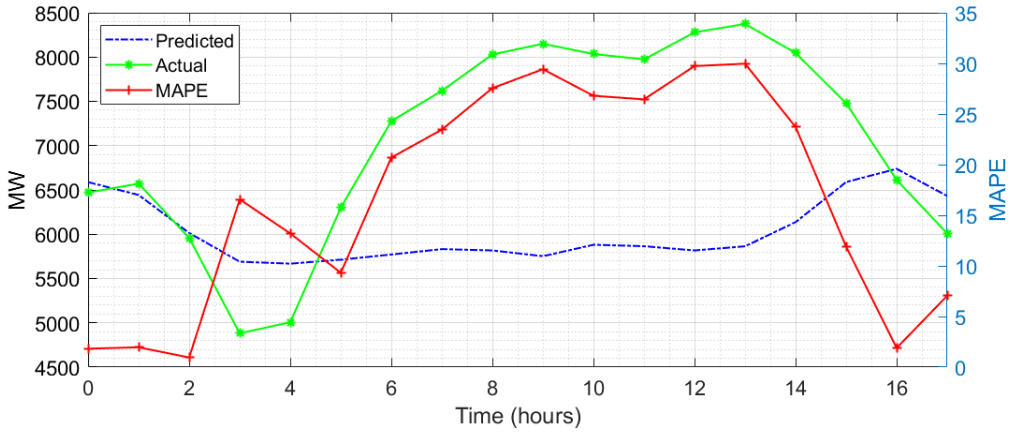


Figure. 4.4 MISO actual and forecasted wind power with their MAPE values

a plot of real-time forecasted wind power generation against actual wind power generation in Midcontinent Independent System Operator (MISO) on May 5th 2020 (from 0 hrs to 17 hrs) and the corresponding MAPE values for each data point. The MAPE values for the forecasting period range from a minimum value of 0.95 to a maximum value of 29.97 in a span of just 11 hours, and had an average MAPE of 16.80.

State-of-the art regression techniques have been applied to wind power forecasting (WPF). In most cases however, the scope is limited to a single time horizon in the dataset or in a portion of the dataset. Independent System Operator (ISO) rely on short-term forecasts of RES to plan daily operations with predictions generated more frequently

than the forecast horizon, such as in the case of short-term forecasts of 6 hours, which are performed after each hour of operation. We study the variation of regressional approaches on moving datasets, and propose the following:

- Performance comparison of different regressional models (ARIMA, SVM, RF, GLM, GAM) on the same dataset.
- Evaluation of the proposed hybrid forecasting methods on wind generation forecasting.
- The capability and performance of these methods for a rolling dataset with a short-term forecast of 6 hours executed at each hour.

4.2 Literature Review

Wind power generation forecasts for a day-head operation using a combination of Gaussian Process and NWP was proposed in [53]. The wind speed information from an NWP model was corrected using a GP and was later converted to wind power using the turbine model. The resulting method had a 17% improvement in MAE compared to ANN. Researchers have used methods such as ARIMA [54, 55, 56, 57], Markov Chain [58], RF [59], SVM [60, 61] and SVM with pattern matching [62] in wind power generation forecasting. Research has been performed on the usage of Neural Networks (NN) in the area of wind power forecasting, such as in [63], where a Recurrent Neural Network (RNN) with a Gated Recurrent Unit (GRU) was used for the short term forecast of wind speed for a 15 minute duration. The method of adaptive neuro-fuzzy inference systems (ANFIS) combines ANN and fuzzy inference systems, which are then used to forecast wind power generation [64]. An NN ensemble approach is proposed in [65], which uses

multiple NN models for the same dataset. A double optimization model with Particle Swarm Optimization (PSO) and the Genetic algorithm is used to adjust the NN parameters in [66]. Some of the other works with NN usage include the application of the radial basis function [67] and integration AI with NWP [68] .The NN based forecasts have yielded acceptable results in wind power forecasting but are susceptible to bias and weight initialization. The issue of over-training may arise if there are numerous parameters to be estimated from training data.

A systematic analysis on wind power generation methods and trends in the past literature is performed in [69], revealing that statistical methods are used 54% of the time, of which, wind generation forecasting had a 43% share. Among the different features utilized for wind generation forecasting, wind speed had a 60% share as an input variable. A comprehensive review on short-term wind power generation forecasting is given in [48], and focuses on methods encompassing Numerical Weather Prediction (NWP) Models and the errors occurring from these approaches.

4.3 Problem Formulation

4.3.1 Formalization

Wind power generation forecasting is modelled as a regression problem with a set of k predictor variables $x(t) = (x_1(t), x_2(t), ..., x_k(t))$ and the predicted variable $y(t)$ with N historical data points $1 \leq t \leq N$ [60]. The objective of this model is to predict the wind energy production $\hat{y}_{t+\theta}$ for a particular time horizon θ, at time t.

4.3.2 Wind Power Generation Data

We utilized wind power generation data from NREL's wind integration dataset, which provides wind generation data with a 5 minute resolution for the years 2009-2012. The data set consists of wind power generation along with 5 independent variables; wind speed (m/s), wind direction (deg), air temperature (K), surface air pressure (Pa), and air density at hub height (kg/m^3). Forecasting models and weather data were used to develop this dataset in a roughly 2km by 2km grid throughout the entire continental United States to generate approximately 126,000 feasible production sites [70].

4.3.3 Performance Metrics

The prediction accuracy of the forecasting approaches are evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The equation for these accuracy metrics are given in Equations (4.5) to (4.7) [41]. Here, n represents the number of instances in the forecast period, y_t the actual value and $\hat{y_t}$ the forecasted value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y_t})^2} \tag{4.5}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y_t}| \tag{4.6}$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y_t}}{y_t}\right| \tag{4.7}$$

4.3.4 Forecast horizon

We have focused on the short-term forecast of wind power generation. The duration of short-term wind power prediction can vary from 30 minutes to six hours depending on the operation and the agency. A duration of six hours is set as the forecast horizon for all of the test cases in this study and is performed with a 5-minute resolution resulting in 72 data points per prediction.

4.3.5 Computing Resource

The forecasting models were programmed using the R language. The computations were performed on a 64-bit Windows operating system with a 2.6 GHz Intel i7 processor,a 4 GB NVIDIA GeForce GTX 1650 graphics card and 16 GB of RAM.

4.4 Correlation with Predictor Variables

The correlation between each of the predictors with the predicted variable are shown in the Figures 4.5 to 4.9.

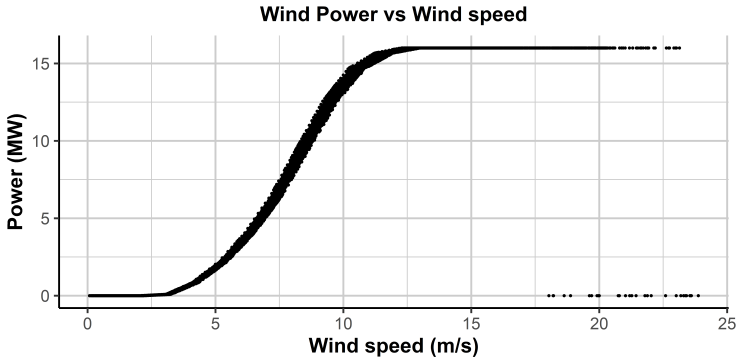


Figure. 4.5 Correlation between wind power and wind speed

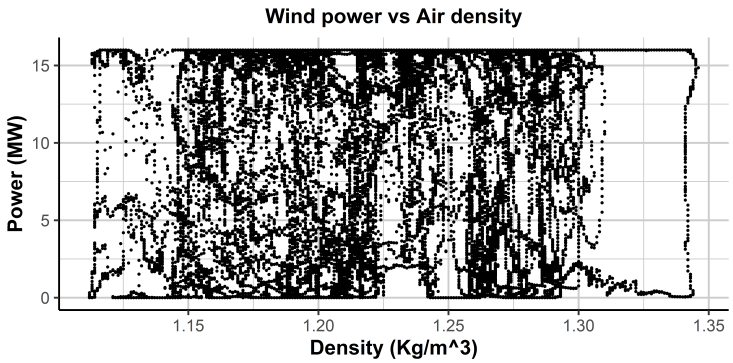


Figure. 4.6 Correlation between wind power and air density

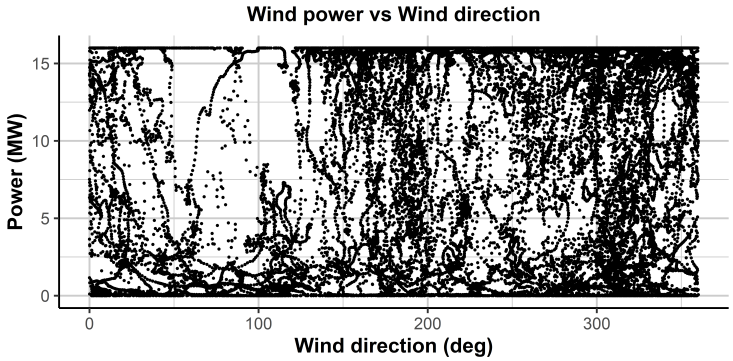


Figure. 4.7 Correlation between wind power and wind direction

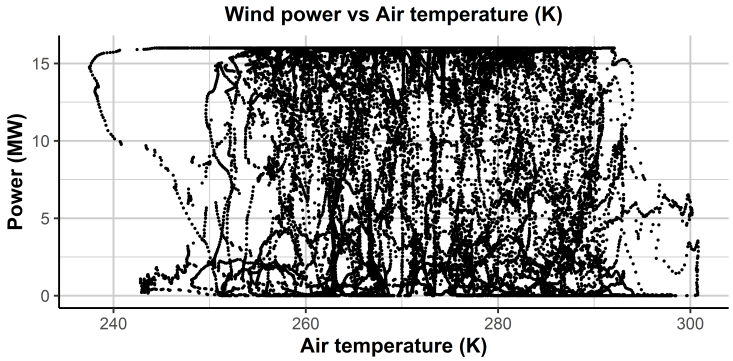


Figure. 4.8 Correlation between wind power and air temperature

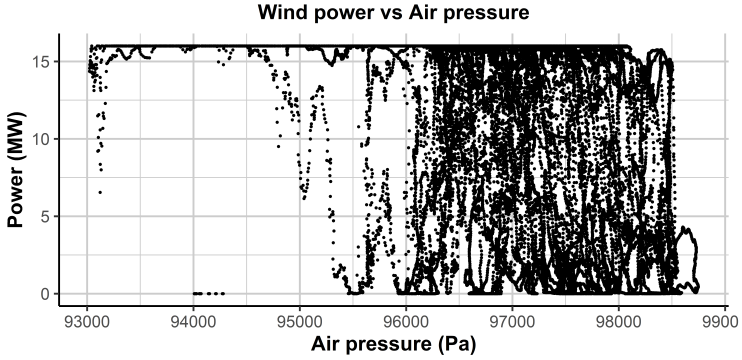


Figure. 4.9 Correlation between wind power and air pressure

4.5 Methodologies

4.5.1 ARIMA

Auto Regressive Integrated Moving Average (ARIMA) is a combination of Autoregressive (AR), Moving Average (MA) and differenced model (I) [56]. In AR models, a linear combination of past values of the variable of interest is used to forecast the future values. The MA model utilizes a regression on past forecast errors, with the differencing used to stationarize the non-stationary data. The full ARIMA model is given in Equation (4.8) [71, 72], where y_{t-i} is the lagged values of y_t, ϕ represents the AR coefficients, ϵ_{t-i} is the MA coefficients and ε_t is the white noise. In the ARIMA(p, q, d) model, the parameters (p, q, d) represent the order of the AR, degree of differencing and order of the MA model.

$$\hat{y}_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{4.8}$$

4.5.2 Generalized Linear Models (GLM)

Generalized Linear Models are a version of normal linear regression, that permits regressed variables with an error distribution other than normal distribution. GLMs include different models such as linear regression, Poisson regression, logistic regression, and Loglinear, among others. There are three components in a GLM model, a linear predictor, a link function and a probability distribution [73]. Here, as shown in Equation (4.9) g is the link function applied to the linear predictor assuming a probability distribution for the data. For a Gaussian distribution with an identity link function, the GLM performs identical to a linear regression model.

$$g(E_Y(y|x)) = \beta_0 + \sum_i \beta_i x_i \tag{4.9}$$

4.5.3 Generalized Additive Models

Generalized Additive Models (GAM) can be considered an extension of GLM, which attempts to incorporate non-linear relationships. GAM provides a structure for the response variable by modeling it based on smooth functions, which is a flexible representation compared to a defined parametric relationship on the independent variables. GAM allows for the adoption of broad range of distributions and the link function measures the effect of the predictors on the predicted variable. A general model for GAM can be written as

$$g(E_Y(y|x)) = \beta_0 + \sum_i f_i(x_i) \tag{4.10}$$

4.5.4 Support Vector Regression

Support Vector Machines (SVM) are supervised machine learning models that can be used for classification and regression, referred to as Support Vector Regression (SVR). Generating a hyperplane for the dataset while maximizing the margin is the core idea behind both SVM and SVR. The intention is to minimize error; however, the algorithms tolerate a small error margin. An SVM can be written as an optimization problem with a minimization objective as in Equation (4.11) subjected to a constraint such as Equation (4.12), where $y = wx + b$ defines the hyperplane, (x_i, y_i) denote the independent and dependent variable in the data, ε stands for the maximum error, ξ represents the training error for each sample n and parameter C is used for regularization [62].

$$min \ \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}|\xi_i| \tag{4.11}$$

subject to

$$|y_i - wx_i| \leq \varepsilon + \xi_i \tag{4.12}$$

4.5.5 Random Forest

Random Forest (RF) regression is a supervised machine learning technique which uses the concept of ensemble learning. This method uses the bagging approach; splitting the data sets into small chunks and feeding it into several learning models or decision trees and aggregating the values to achieve the expected product [74, 75]. The regression technique uses random sampling with the replacement, or bootstraping, of the records and splits the feature set to create the training set for the learning models. The increased diversity of these base models reduces the over fitting and high variance of decision trees resulting in more robust solutions. This class of ensemble models can be represented in the following equation, where the final output $g(x)$ is the aggregate of base models $f_i(x)$.

$$g(x) = f_0(x) + f_1(x) + f_2(x) + ... \tag{4.13}$$

4.5.6 Hybrid Methods

Hybrid forecasting approaches combine the advantages of multiple approaches into one methodology [76]. Combinations of methods performing well are utilized to develop hybrid techniques with greater performance in wind generation forecasting, the procedure of which can be described in these following five steps:

Step 1: (Data preprocessing) The historical data needs to be preprocessed into a

format suitable for forecasting. All explanatory variables are lagged by the forecast horizon and the six explanatory variables are utilized by the model.

Step 2: (Initial wind power generation forecast) The first forecast model creates the wind generation forecast ($\hat{Y}_t$) based on the explanatory variables. We have used Arima and GLM are used for first stage prediction.

Step 3: (Residual calculation) The residual for the forecast from step 2 is calculated (residual (R_t) = actual value (Y) - forecasted value ($\hat{Y}$)).

Step 4: (Residual forecast) The second forecasting model predicts the residual based on the previous residuals and explanatory variables ($\hat{R}_t$).

Step 5: (Final wind power forecast calculation) The final wind power forecast value is calculated based on the forecast from step 2 and residual forecast from step 4 ($\hat{y}_t = \hat{Y}_t + \hat{R}_t$).

Steps 1 through 3 are part of the first stage and step 4 through 5 constitute the second stage of the model. The block diagram for the hybrid forecasting model is shown in Figure 4.10. Detailed representation of the proposed hybrid forecasting framework is given in Figure 4.11.

4.6 Results and Discussion

All forecasting models rely on 29 days of wind generation data with six predictor variables and five-minute resolution, in order to train the forecasting model. The predictor variables include lagged version of wind power, wind direction, wind speed, air temperature, air pressure, and air density. The 30th of every month is chosen as the forecast day, except for February in which the 28th day is chosen. The actual wind generation values on these 12 days are shown in Figure 4.12. We have used 6-hour ahead forecasts, per-

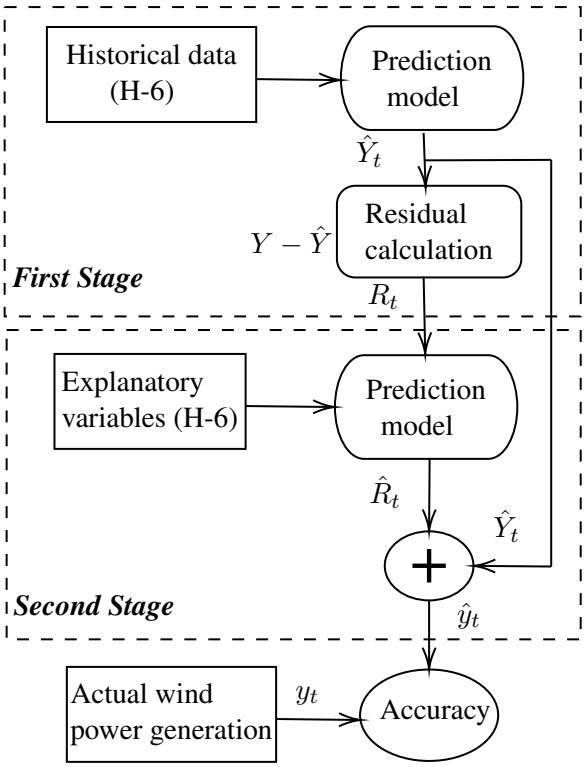


Figure. 4.10 Hybrid forecasting method

formed after each hour in a day. The graphical representation is shown in Figure 4.13. In the reviewed literatures, wind generation forecasting is generally performed starting at either a single point or multiple portions of the same dataset. The performance of the regression methods and the proposed hybrid method's performance is evaluated using this moving window approach. For each of the forecast days, 19 different predictions are executed starting at 12:00AM to 6:00PM.

The ARIMA model utilized the auto.arima function to deliver the p, d, q values. In default mode, auto.arima uses approximations to increase the speed of the model search. To improve model accuracy $approximation$ in the function argument is set to $false$ in order to avoid approximations and $stepwise$ is set to $false$ in order to search a larger model space. To check the accuracy of the auto.arima function in delivering parameters conforming to stationarity, both "Ljung-Box" and "Box-Pierce" tests are conducted. A stationarity test on the prediction for January 30th at 12:00 AM yielded a value of

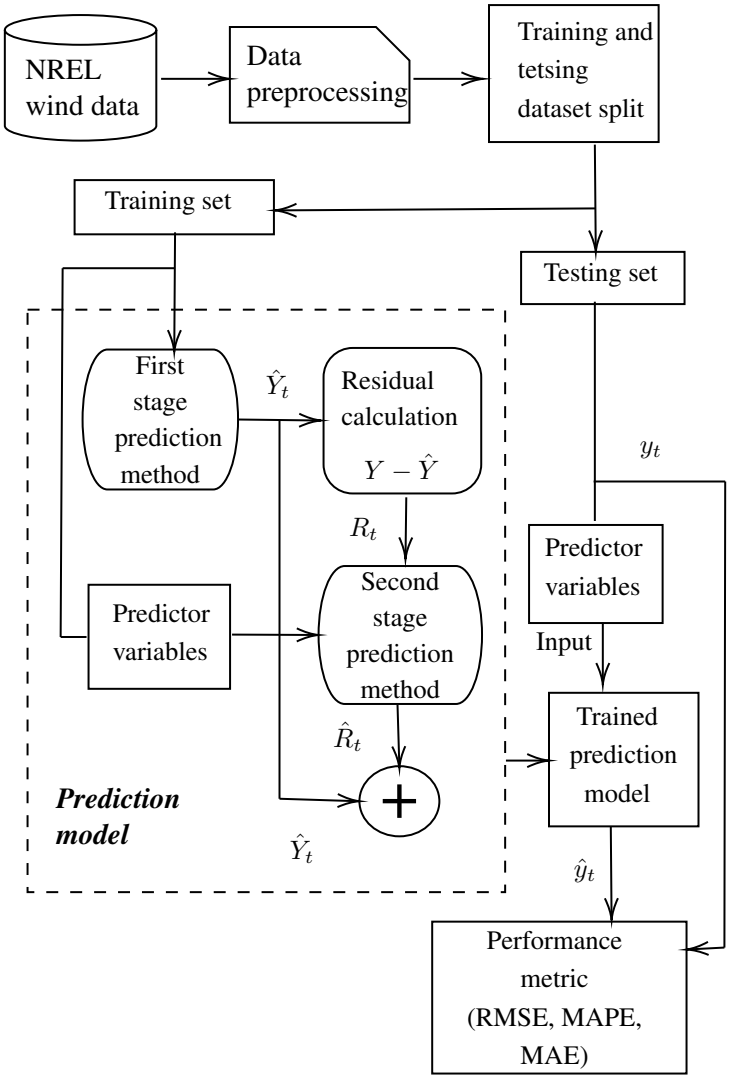


Figure. 4.11 Proposed hybrid forecasting framework

0.8701 for both Ljung-Box and Box-Pierce test. A similar test conducted for August 30th, generated a high value of 0.9585 for an arima model (4,1,1). The larger p-value from the portmanteau tests indicates the randomness of arima residuals, confirming no correlations are left after arima modeling. Another method to test the characteristics of the arima model is to plot the inverse roots [72]. The inverse root characteristics for the ARIMA model (4,1,1) tests the characteristics, as shown in Figure 4.14. The four red dots on the left denote the inverse root of the AR part, while the red dots on the right represent the inverse of the MA. The root values inside the unit circle denote stability, however, if the root values are close to the unit circle may result in numerical instability,

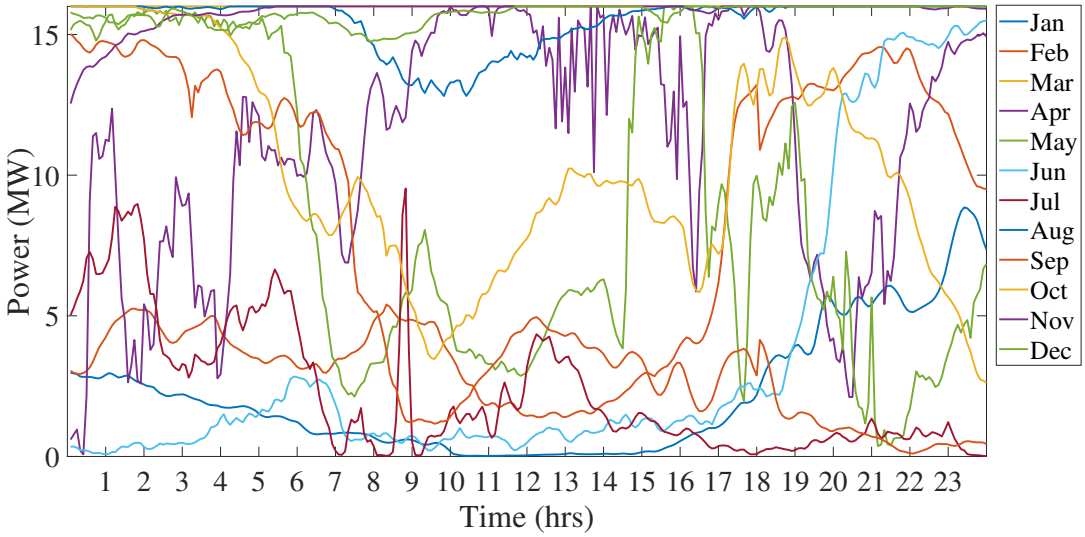


Figure. 4.12 Forecast data points in each month

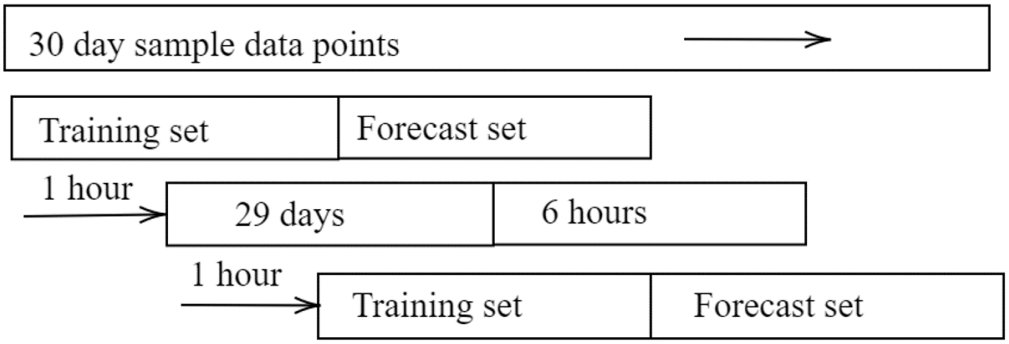


Figure. 4.13 Rolling window approach used for forecasting

which would not be acceptable for forecasting.

RF regression is performed using the $RandomForest$ function in R. This function uses Breiman's RF algorithm for regression and classification [75]. RF is an aggregation method in which the results from multiple decision trees are used, after randomizing the data into chunks. In this example, the number of decision trees, or the $ntree$ argument, is set to the default value of 500, and the number of features sampled at each split $mtry$ is set at 3. The recommended value is $p/3$ for regression where p is the number of predictors. The $importance$ parameter is set to $True$, in order to enable the algorithm to determine variable importance. In the case of SVR, the $ksvm$ function is utilized

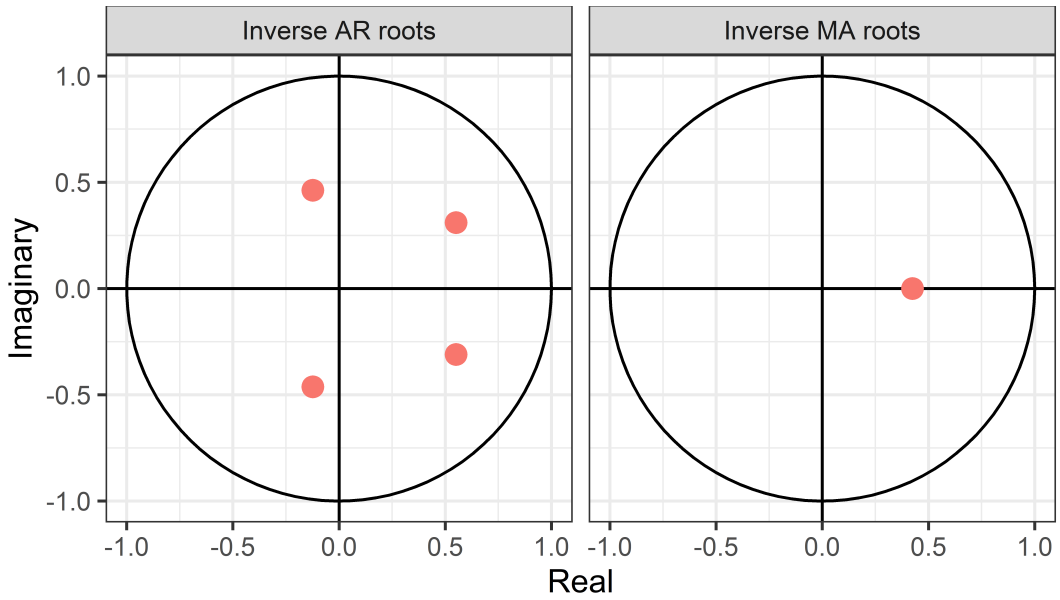


Figure. 4.14 Inverse root characteristics of the ARIMA model (4,1,1)

for the regression problems and can be used for both the classification and regression problems. This function also uses John Platt's SMO algorithm for most of the SVM QP formulation [77]. The default mode in regression $eps - svr$ (epsilon support vector regression) is used to model the problem, which uses the accurate approximation of the $\epsilon-$ insensitive loss function [78]. The radial basis kernel $rbfdot$ is used as the kernel function in all forecast iterations.

The GLM forecasting method utilized the glm function, with family, link and method parameters set to the default value of $gaussian$, $identity$, and $glm.fit$, respectively. In the case of GAM, the gam function from the $mgcv$ package is employed. Gaussian is used for the family parameter and $GCV.Cp$ is used as the smoothing parameter estimation method. This method uses Generalized Cross Validation (GCV) when the scale parameters are unspecified and Un-Biased Risk Estimator (UBRE) when it is specified. These smooth functions were implemented using penalized regression splines and applied basis functions for these splines. All of the hybrid methods used combinations of the above methods.

Table 4.2 Performance of forecasting models on December 30, 12:00 AM for a six hour ahead forecast

	RMSE	MAE	MAPE
Arima	0.299728	0.241168	1.564211
SVM	0.446007	0.38859	2.485144
RF	0.291568	0.234951	1.515214
GLM	0.622082	0.575298	3.711939
GAM	1.304824	1.094148	7.028599
Arima-SVM	0.305738	0.247433	1.60467
Arima-RF	0.299553	0.240562	1.560299
GLM-SVM	0.292583	0.238021	1.541233
GLM-RF	0.524813	0.466921	3.015538

Predictions made using all of the above mentioned forecasting methodologies for December 30th at 12:00AM to create a six-hour ahead forecast utilizing 29 days of training data, for all three error metrics, is shown in Table 4.2. From Table 4.2, we can see that the highest performing methodology is RF followed closely by Arima. The GLM and GAM methodologies had the lowest performance. The hybrid methods of Arima and GLM yielded accuracies higher than their non-hybrid counterparts. The conclusions assumed from the above table and applying the RF to all future wind generation datasets may not deliver the same level of performance. The actual values in each day of the month have a large variance, illustrating how wind power generation does not follow a pattern similar to system load or solar energy generation, as seen in Figure 4.12,. In the case of system load, the shape of the curve generally follows a ”duck curve” pattern, while solar energy generation exhibits a parabolic shape, increasing in the day time and decreasing at night).

Applications of these forecasting techniques on a rolling window approach for each of the 12 months are shown in figures 4.15 to 4.26. In the month of January as shown

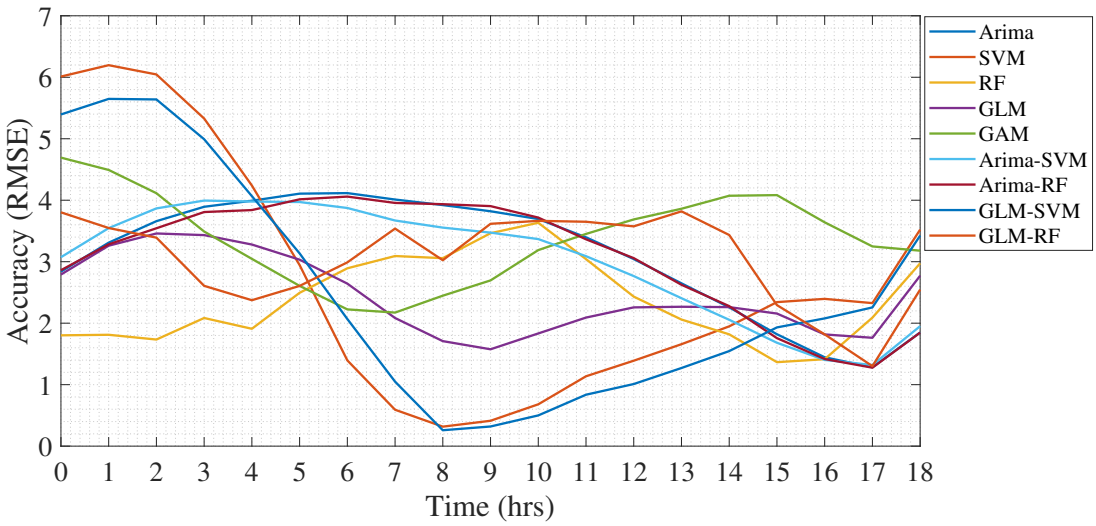


Figure. 4.15 Performance of forecasting models in January

in Figure 4.15, GLM displayed the most consistent results among non-hybrid methods, closely followed by GAM and ARIMA. The hybrid method GLM-SVM shows greater performance when compared to the other three hybrid methods and has better results than GLM, except for the first six iterations. Eventhough, GLM-SVM and GLM generated consistent results, none of the nine methods delivered acceptable accuracy levels for the 30th of January. For the month of February, as illustrated in 4.16, GAM and GLM-RF displayed similar performance and GLM-SVM had greater performance than GAM and GLM-RF, except for the first six iterations. In March, shown in Figure 4.17, GLM-RF had the advantage over other methods in delivering consistent results. All of the other methods had high RMSE values towards the end of the day, but GLM-RF was able to generate fairly accurate results throughout the day.

In the month of April as shown in Figure 4.18, GLM-SVM gave consistent results. Similarly in May (Figure 4.19) GLM-SVM had the most consistent results throughout the day, however, in June GLM and GLM-RF had better performance. For the month of July, GLM-SVM and GLM had the most accurate results. For the month of August and September GLM-SVM and SVM had accurate performance. The month of October

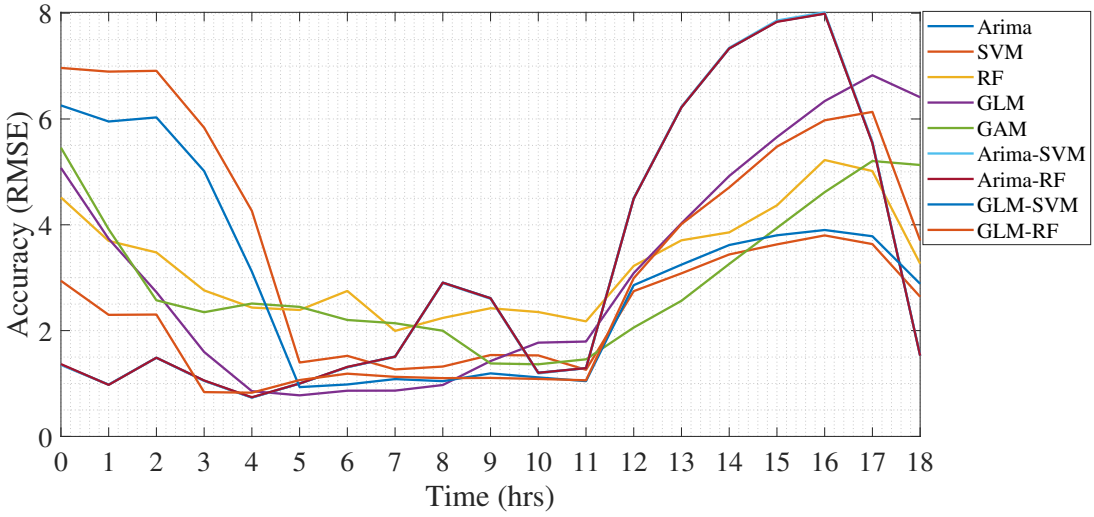


Figure. 4.16 Performance of forecasting models in February

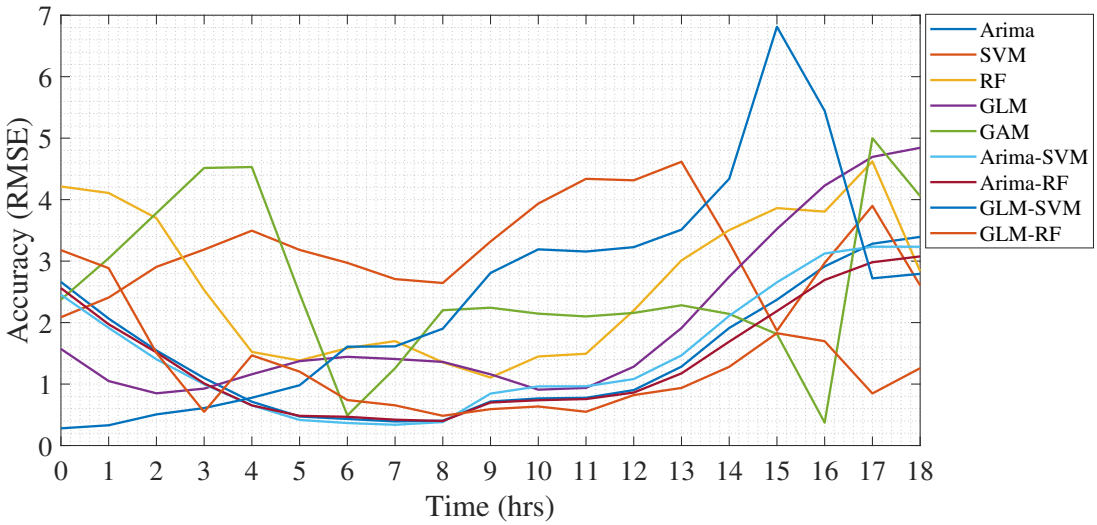


Figure. 4.17 Performance of forecasting models in March

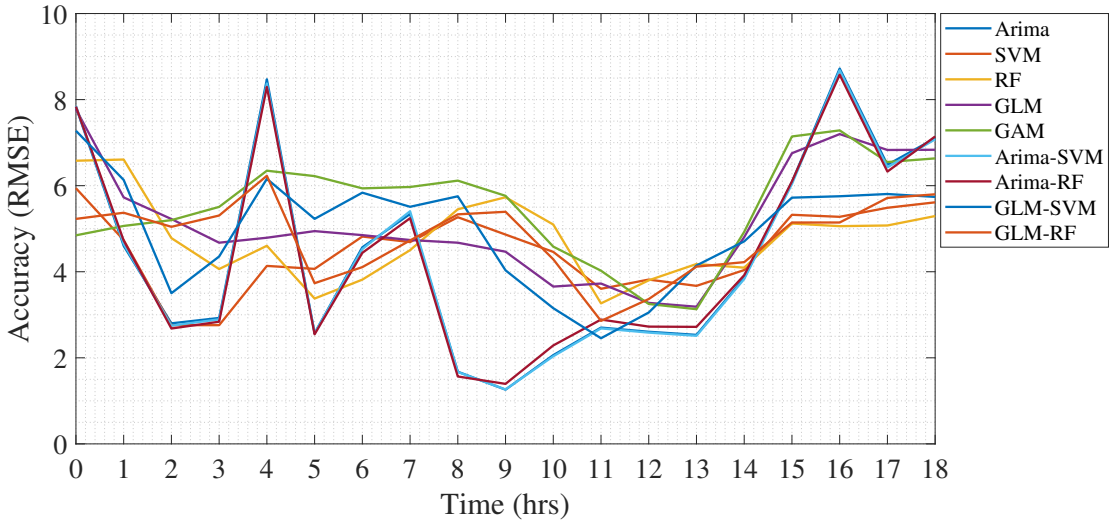


Figure. 4.18 Performance of forecasting models in April

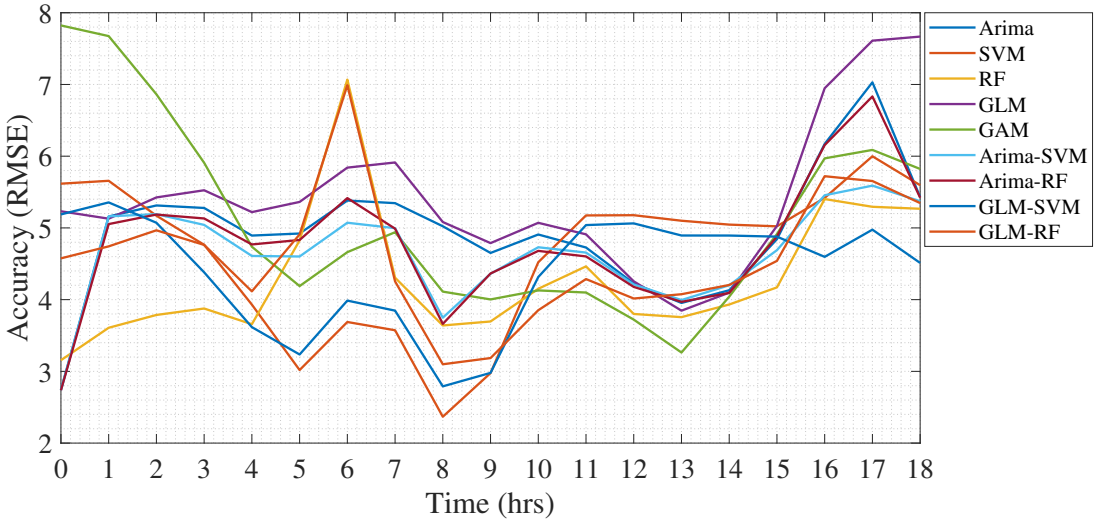


Figure. 4.19 Performance of forecasting models in May

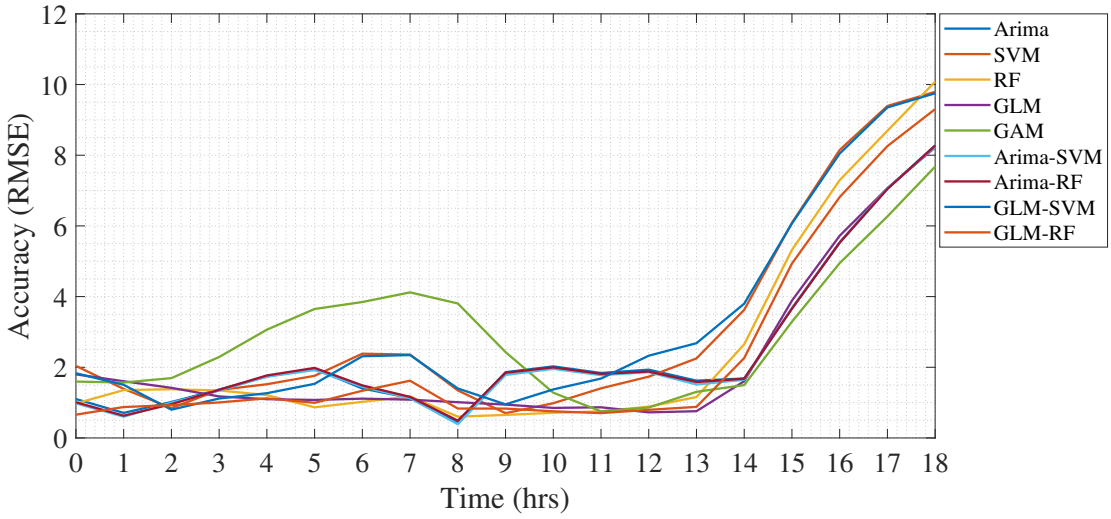


Figure. 4.20 Performance of forecasting models in June

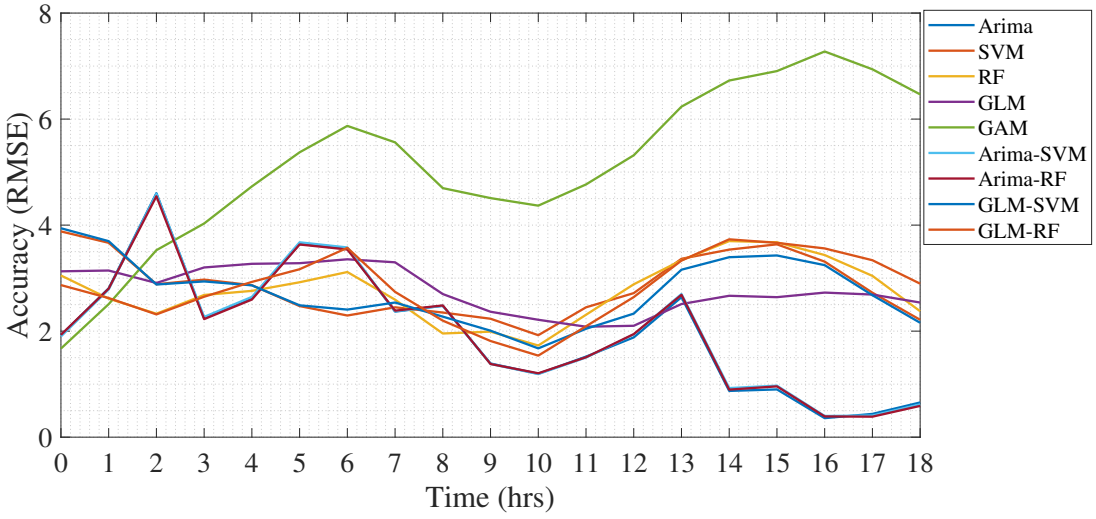


Figure. 4.21 Performance of forecasting models in July

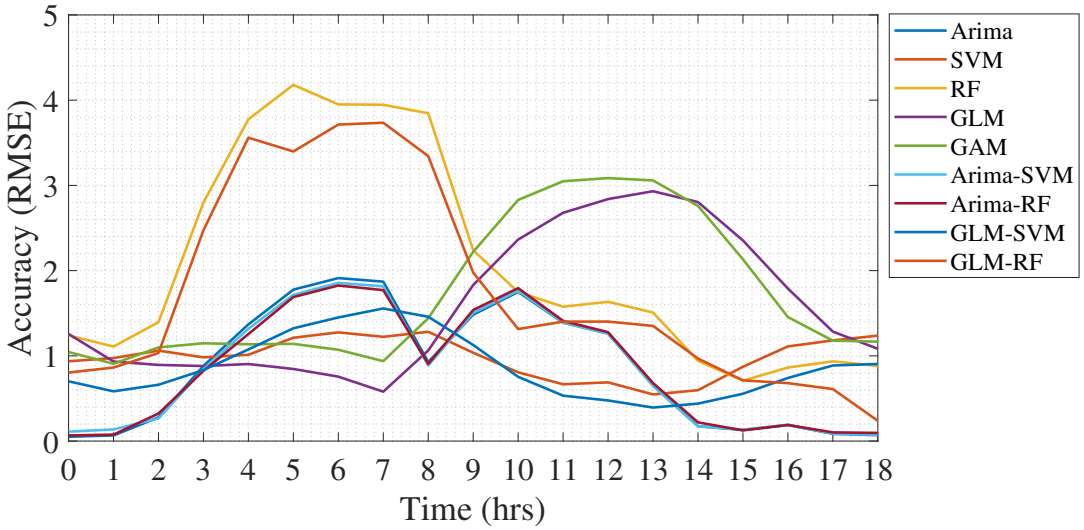


Figure. 4.22 Performance of forecasting models in August

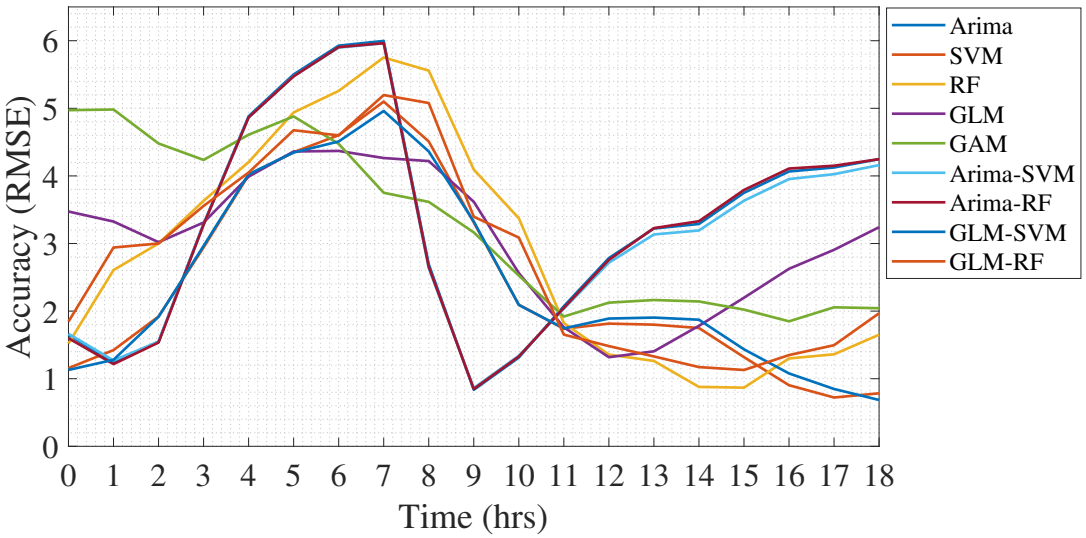


Figure. 4.23 Performance of forecasting models in September

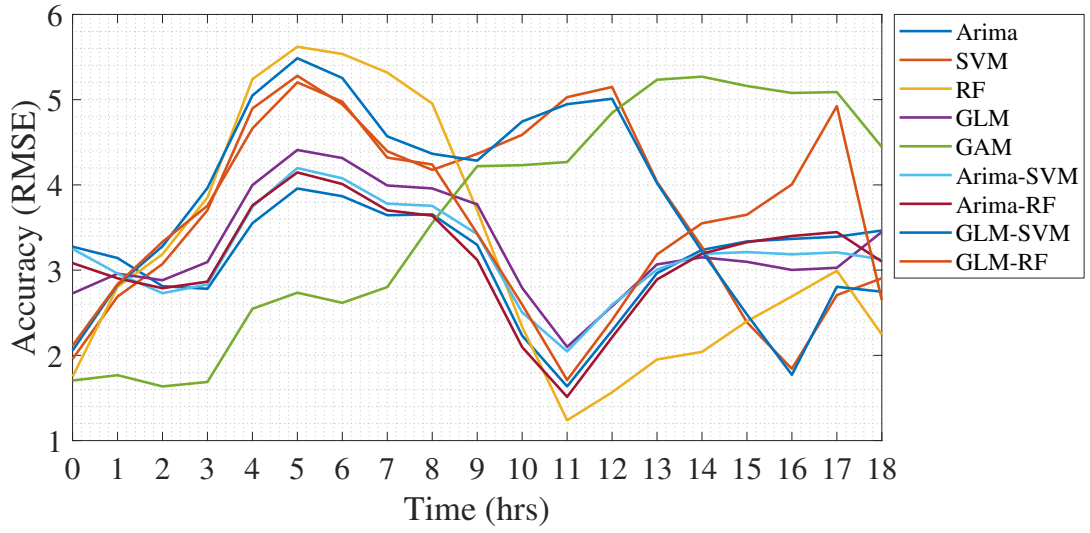


Figure. 4.24 Performance of forecasting models in October

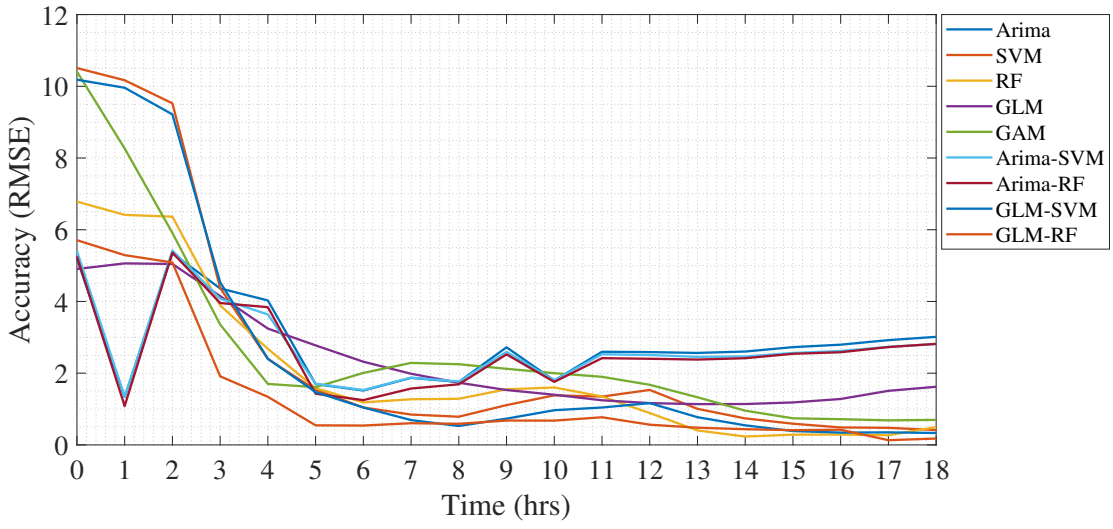


Figure. 4.25 Performance of forecasting models in November

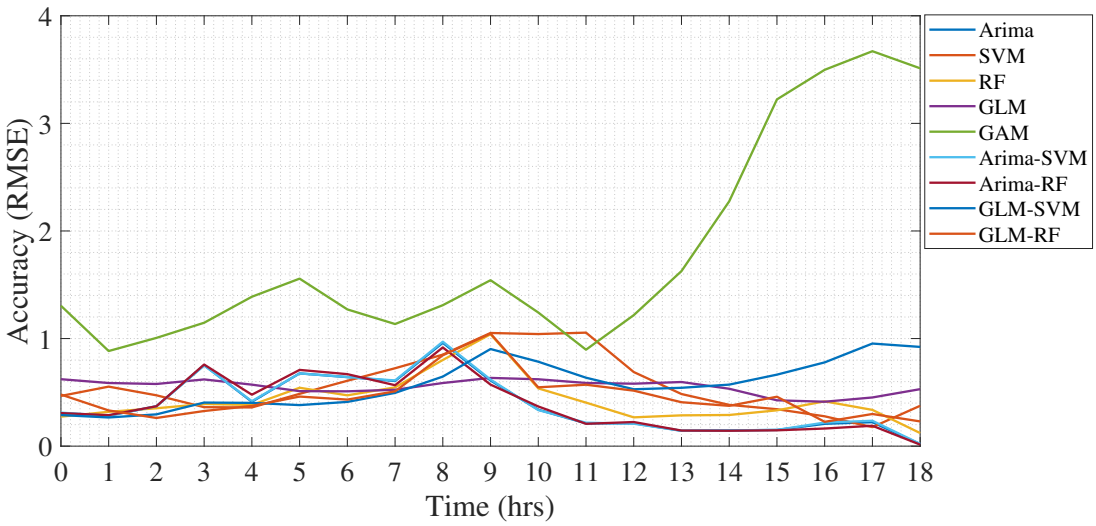


Figure. 4.26 Performance of forecasting models in December

is an anomaly with arima, arima-SVM and arima-RF showing more accurate results than GLM based methods. GLM-RF and GLM were more consistent for the month of November and in December GLM had the most accurate performance.

We have restricted our experiments to short-term forecasts of wind power generation. Short-term forecasts can vary from 30 minutes to 6 hours depending on the functionality. The performance of these methods for different forecast horizons and training data size on December 30th is performed. Predictions of 6, 12 and 24-hours ahead utilizing 28,14 and 7 days of training data are shown in the tables 4.3 to 4.5. Forecasting error was increasing with a longer duration for the prediction. Arima had a RMSE of 0.609 for a 6-hour ahead prediction and increased to 9.453 for a 24-hour prediction. The RF approach had a RMSE of 0.838 for a 6-hour prediction, rising to 9.142 for a 24-hour prediction. The amount of training data required to generate the best prediction also varied depending on the regression method. ARIMA generated the best results for a 6-hour prediction with a 7-day training dataset, while RF had the best results for the same forecast horizon with a 14-day training dataset. A study on the predictor importance calculated using linear regression for varying forecast horizons is

shown in the Tables 4.6 and 4.7. Wind speeds and historic values of wind power were found to be the two most significant covariates in wind generation forecasting.

4.7 Observations and Conclusion

We present a comprehensive performance evaluation of five different forecasting methodologies in addition to four proposed hybrid approaches when predicting wind power generation. The performance of these forecasting techniques is illustrated using the NREL wind generation data with six predictor variables and a resolution of five minutes. Our study was restricted to short-term predictions of wind generation lasting six hours, utilizing 29 days of historic data and employing all five predictor variables. Forecasts were carried out for all 12 months of the year, with 19 one-hour ahead iterations of the six-hour ahead prediction.

Several conclusions can be drawn from this study. ARIMA is able to deliver acceptable results in certain iterations but prone to over-fitting. RF, SVM and GAM behave similarly to ARIMA in delivering acceptable results, but are more robust towards variations. GLM and its hybrid methods outperform others in delivering consistent results and GLM-SVM is able to provide respectable accuracy with more consistency than other methods analyzed. Applying forecasting methods to a small portion of wind generation data is not sufficient to conclude the accuracy of one prediction method over the other; therefore, the methods need to be tested on data from different geographical locations and time duration.

Table 4.3 Performance of forecasting methodologies for variable forecast horizon and training dataset

Method	Training data	Forecast Horizon	RMSE	MAE	MAPE
Arima	28 days	6HA	1.418	0.996	6.995
SVM			4.948	3.311	23.194
RF			0.911	0.653	4.570
GLM			2.230	2.172	14.346
GAM			1.444	1.159	7.696
Arima-SVM			1.402	0.986	6.924
Arima-RF			1.402	0.986	6.924
GLM-SVM			2.947	2.298	15.847
GLM-RF			1.780	1.700	11.313
Arima	14 days	6HA	1.344	0.959	6.720
SVM			2.902	2.244	15.406
RF			0.838	0.594	4.161
GLM			2.041	1.760	12.049
GAM			7.750	6.370	43.214
Arima-SVM			1.347	0.961	6.738
Arima-RF			1.346	0.963	6.749
GLM-SVM			0.900	0.690	4.721
GLM-RF			2.454	1.786	12.440
Arima	7 days	6HA	0.609	0.361	2.583
SVM			0.938	0.820	5.391
RF			1.091	0.744	5.244
GLM			3.451	3.107	21.135
GAM			5.591	4.465	29.799
Arima-SVM			2.303	2.180	14.702
Arima-RF			2.289	2.156	14.553
GLM-SVM			3.360	2.724	18.779
GLM-RF			3.496	2.835	19.522

Table 4.4 Performance of forecasting methodologies for variable forecast horizon and training dataset

Method	Training data	Forecast Horizon	RMSE	MAE	MAPE
Arima	28 days	12HA	8.311	5.441	NaN
SVM			3.998	2.256	NaN
RF			7.653	4.991	NaN
GLM			8.578	6.245	NaN
GAM			13.935	12.240	NaN
Arima-SVM			8.319	5.449	NaN
Arima-RF			8.327	5.458	NaN
GLM-SVM			5.206	3.526	NaN
GLM-RF			8.696	5.898	NaN
Arima	14 days	12HA	8.430	5.507	NaN
SVM			6.279	3.933	NaN
RF			7.735	5.056	NaN
GLM			9.259	7.095	NaN
GAM			9.858	8.097	NaN
Arima-SVM			8.429	5.507	NaN
Arima-RF			8.445	5.522	NaN
GLM-SVM			7.725	4.600	NaN
GLM-RF			9.182	5.817	NaN
Arima	7 days	12HA	8.439	5.525	NaN
SVM			6.275	3.727	NaN
RF			7.744	5.069	NaN
GLM			10.803	8.182	NaN
GAM			17.098	13.769	NaN
Arima-SVM			8.436	5.522	NaN
Arima-RF			8.447	5.533	NaN
GLM-SVM			9.622	6.460	NaN
GLM-RF			10.253	6.539	NaN

Table 4.5 Performance of forecasting methodologies for variable forecast horizon and training dataset

Method	Training data	Forecast Horizon	RMSE	MAE	MAPE
Arima	28 days	24HA	9.453	7.717	NaN
SVM			8.637	6.424	NaN
RF			9.142	7.254	NaN
GLM			11.675	9.887	NaN
GAM			21.224	18.163	NaN
Arima-SVM			9.439	7.704	NaN
Arima-RF			9.454	7.721	NaN
GLM-SVM			10.197	8.154	NaN
GLM-RF			10.328	8.106	NaN
Arima	14 days	24HA	9.566	7.810	NaN
SVM			6.265	4.583	NaN
RF			9.436	7.681	NaN
GLM			14.285	12.795	NaN
GAM			27.036	23.562	NaN
Arima-SVM			9.568	7.811	NaN
Arima-RF			9.571	7.816	NaN
GLM-SVM			9.639	7.926	NaN
GLM-RF			11.779	9.973	NaN
Arima	7 days	24HA	9.484	7.740	NaN
SVM			7.854	6.049	NaN
RF			9.422	7.681	NaN
GLM			15.713	14.345	NaN
GAM			10.512	7.910	NaN
Arima-SVM			9.492	7.746	NaN
Arima-RF			9.495	7.751	NaN
GLM-SVM			12.806	10.812	NaN
GLM-RF			12.341	10.201	NaN

Table 4.6 Predictor importance calculated using linear regression for varying forecast horizon

	Forecast Horizon	speed	power	density	temperature	pressure	direction
Jan	**6H**	0.1887	0.1911	0.0496	0.0820	0.0536	0.0031
	12H	0.0575	0.0636	0.0325	0.0362	0.0587	0.0329
	24H	0.0061	0.0044	0.0723	0.0627	0.0469	0.1027
Feb	**6H**	0.2278	0.2017	0.0196	0.0829	0.0782	0.0224
	12H	0.0737	0.0602	0.0256	0.0075	0.1140	0.0173
	24H	0.0737	0.0602	0.0256	0.0075	0.1140	0.0173
Mar	**6H**	0.1325	0.1053	0.0630	0.0514	0.0204	0.0017
	12H	0.0025	0.0016	0.0290	0.0536	0.0220	0.0086
	24H	0.0101	0.0149	0.0457	0.0610	0.0223	0.0042
Apr	**6H**	0.1852	0.1985	0.0515	0.0785	0.0214	0.0103
	12H	0.0408	0.0539	0.0557	0.1154	0.0218	0.0174
	24H	0.0116	0.0177	0.0905	0.0387	0.0163	0.0006
May	**6H**	0.1331	0.1239	0.0248	0.0385	0.0925	0.0138
	12H	0.0280	0.0209	0.0142	0.0146	0.1209	0.0087
	24H	0.0117	0.0068	0.0408	0.0209	0.0997	0.0033
Jun	**6H**	0.0871	0.0963	0.0720	0.0303	0.0160	0.0160
	12H	0.0175	0.0334	0.0342	0.0395	0.0066	0.0067
	24H	0.0047	0.0064	0.0254	0.0230	0.0055	0.0011
Jul	**6H**	0.1861	0.2125	0.0302	0.0602	0.0123	0.0131
	12H	0.1013	0.1144	0.0169	0.0854	0.0348	0.0139
	24H	0.0577	0.0511	0.0206	0.0306	0.1068	0.0218
Aug	**6H**	0.1962	0.1716	0.0114	0.0318	0.0165	0.0057
	12H	0.0617	0.0525	0.0294	0.0602	0.0318	0.0119
	24H	0.0176	0.0108	0.0313	0.0752	0.0551	0.0350
Sep	**6H**	0.1051	0.0909	0.0406	0.0196	0.1451	0.0058
	12H	0.0169	0.0134	0.0570	0.0769	0.1739	0.0150
	24H	0.0171	0.0158	0.0143	0.0127	0.0653	0.0568
Oct	**6H**	0.1903	0.1643	0.0138	0.0279	0.0288	0.0005
	12H	0.0525	0.0224	0.0305	0.0255	0.0569	0.0001
	24H	0.0270	0.0133	0.0585	0.1664	0.0311	0.0003

Table 4.7 Predictor importance calculated using linear regression for varying forecast horizon

	Forecast Horizon	speed	power	density	temperature	pressure	direction
Nov	**6H**	0.1285	0.1199	0.0241	0.0422	0.0226	0.0008
	12H	0.0170	0.0135	0.0312	0.0454	0.0175	0.0350
	24H	0.0651	0.0352	0.0154	0.0231	0.0449	0.0049
Dec	**6H**	0.2116	0.2221	0.0175	0.0343	0.0479	0.0075
	12H	0.0719	0.0700	0.0198	0.0355	0.0534	0.0053
	24H	0.0100	0.0063	0.0594	0.0183	0.0437	0.0686

CHAPTER 5

Computational and numerical analysis of AC optimal power flow formulations on large-scale power grids

5.1 Introduction

Optimal Power Flow (OPF) is a fundamental tool used in the power industry to ensure security and optimal operation of the grid. OPF is pervasive in the industry with applications spanning across transmission and distribution, real-time and day-ahead operations, and short-term and long-term planning. Grid operators are facing challenges in operating the current grid securely and at the lowest cost possible due to the expanded opportunity and availability of renewable energy resources. These challenges are increased with the introduction of additional monitoring and control introduced by smart grid initiatives as well as increased interaction between transmission and distribution. With these increased complexities, OPF tools will need to be more robust and faster.

While the performance of power grid applications, such as OPF need to be improved, one must also take into account the rapid growth and change in the computing industry. The computing industry has grown exponentially since OPF was first developed in the 1960's by Carpentier *et al.* [79, 80], and has experienced many eras of computing history, such as mainframe computers, distributed memory clusters with single core nodes, clusters with multicore nodes, and now distributed memory clusters

with heterogeneous nodes with multiple processors and accelerators. The current evolution in the computing industry is on the heavy usage of hardware accelerators, such as Graphical Processing Units (GPUs), which has been spurred on by the gaming industry and machine-learning applications. This new technology has dramatically improved computational performance, but at the same time has imposed new constraints on how mathematical models can be effectively implemented. Under such a changing computing environment, it is prudent to revisit and assess power grid applications and the new ways to adapt them to these newer architectures.

This work attempts at assessing these fundamental building blocks for the AC optimal power flow (AC-OPF) application. We provide an in-depth comparison of three different formulations for AC-OPF – power-balance with polar voltages, power-balance with Cartesian voltages, and current-balance with Cartesian voltages – and compare their structure and characteristics. In addition, we present the numerical and computational performance of these formulations to highlight their differences, point to the most efficient formulation, and provide benchmark comparison metrics on very-large networks. The significant contributions can be summarized as follows

- An in-depth comparison of the AC-OPF model for three formulations:Presenting the different characteristics of power balance polar, power balance Cartesian voltages,and current balance Cartesian voltages.
- Comparison of the structural differences in terms of the number of variables, constraints and non-zeros in Jacobian and Hessian matrices.
- Numerical and computational performance evaluation for the AC-OPF formulations on nine different test cases including large-scale synthetic U.S. networks.

The results from the three model formulations will be compared and validated with

MATPOWER [81], which utilizes the power balance polar formulation for the AC-OPF model.

5.2 AC-OPF literature review

The Optimal Power Flow (OPF) ensures a secure operation of electric power plants for a given transmission network with, typically, the objective of minimizing generation cost subjected to operational and security constraints in the network [82]. It is one of the most important tools used by engineers for power system operation and planning. AC-OPF is an optimization model that considers the full AC power flow equations. It is the most accurate representation of power flow in a network assuming the model parameters are correct. Compared to a DC optimal power flow, the benefits of AC-OPF are increased accuracy, inclusion of reactive power, current, voltage and losses in the network (e.g. transmission losses, active and reactive power load loss)[83]. AC-OPF plays a critical role in the operation of Independent System Operator (ISO) power markets [84]. It is utilized in every important stage of a power system operation and planning such as expansion planning [85], grid management [86], day ahead markets [87], and also for real-time control [88]. AC-OPF is performed yearly for capacity expansion, daily for day-ahead markets, and, in some cases, even for every 5 minutes. An extensive review on the application of AC-OPF in distributed generation planning and operation is provided in [89].

However, AC-OPF remains a computationally complex problem, lacking a rapid and robust solution after 50 years of formulation. The problem is getting more complicated with the introduction of distributed and large scale renewable energy sources [90, 91, 92]. The current approach is to utilize decomposition, approximations and assumptions

for a fast and acceptable solution [93]. Researchers have explored different methods such as linear approximation [94, 95, 96, 97], conic formulation [98, 99], semidefinite programming [100, 101], quadratic convex relaxation [102], decomposition [103, 104] for a faster solution. A detailed study on the effect of inexact convex relaxation in AC-OPF feasibility is given in [105]. The approaches of approximations and assumptions cost the companies in millions of dollars in operational cost, damage to the environment from unnecessary emissions and energy waste. Even a small improvement in dispatch efficiency can save billions of dollars [94]. The authors of [84] have calculated and estimated savings of over twenty billion dollars in the US market with an improvement of 5% in AC-OPF [84].

Much of the AC-OPF literatures utilize the polar power - voltage formulation first introduced by Carpentier in the 1960's. The other two main formulations of OPF are rectangular power-voltage and rectangular current-voltage formulations. Researchers have also explored other formulations of OPF such as current injection and a mix of polar and rectangular coordinates [106, 107]. The hybrid method in [106] used rectangular forms of voltage and current, current mismatch equations for power balance and numerical stability is ensured by PV buses. The equivalent current injection based method also utilized a decoupled optimization for faster processing. M. Jereminov *et al.* [107] proposed a solution methodology for AC-OPF using equivalent circuit formulation. The AC-OPF is represented as a non-linear equivalent circuit in which generator model is represented by conductance and susceptance state variables, and network constraints are handled by the generator admittance state variables. The proposed method solves the convergence problem in the current balance Cartesian formulation.

In [96], authors present a linear approximation method to solve AC-OPF in power

balance polar formulation using the Mixed Integer Linear Programming (MILP) approach. This method of binary expansion discretisation is used to convert the non-linear AC-OPF in to linear problem without losing accuracy. This method also has the advantage of obtaining reactive power and voltage profile at the same time but faced exponential increase in execution time with cplex solver. An approximation of AC-OPF in power balance polar formulation utilizing Langrangian dual is proposed in [90]. A Supervised Deep Learning model with rectified linear unit (ReLU) activation function is modeled to make the generator set-points. The proposed approximation method had a better accuracy and a faster processing time than DC-OPF.

Y. Tang *et al.* in [108] proposed a real-time AC-OPF based on quasi-Newton methods using the current balance Cartesian formulation. The approach utilised the second-order information to provide sub-optimal solutions for real-time applications. A small correction term is used to track the optimal solution assuming a single-phase power flow. The superiority of linear approximation of AC-OPF in current balance Cartesian over traditional quadratic power flow formulation is stated in [109]. It proposes the idea of using AC-OPF in current balance form and its approximations for practical applications for its improved computational performance. A continuation method (homotopy) is used in [110] to covert the DC OPF solution in to AC-OPF by slowly increasing the non-linearities in equality and inequaltiy constraints. The proposed method achieved robust solution with a reasonable computational overhead.

A comparative analysis of different power flow methodologies is performed in [111], but on Power Flow (PF) problem. The authors tested the performance of these formulations on well-conditioned and ill-conditioned networks. A 118, 300 and 730 bus systems were used to test these methods and for well-conditioned networks all three

methodologies showed almost similar performance. The slight performance improvement in power balance polar and current balance Cartesian methods are noted in the work. In the case of ill-conditioned networks rectangular based formulations had better convergence properties. A comparison of three different solvers over different power flow formulations are implemented in [112]. The performance of KNITRO, MATPOWER's MIPS (MATPOWER Interior Point Solver) and FMINCON (Find minimum of constrained nonlinear multivariable function) methods over different bus systems and formulations is studied in this paper. In [113], authors performed a comparative analysis of three different AC-OPF formulation with different generator capability curves, solvers and initial conditions. The evaluations are done on 118 and 2736 bus systems. In the studies, power balance polar and current balance Cartesian performed better in terms of computational time. In case of solvers IPOPTH and KNITRO performed the best and for initial conditions midpoint and flat start as the best choice for AC-OPF.

5.3 Problem formulations

General form for an ACOPF formulation is shown in equations (5.1) to (5.4). It is an optimization model with a minimization objective subjected to a set of equality and inequality constraints. The objective function in AC-OPF can be modelled for the minimization cost, minimization of losses, maintaining constant voltage profile, transmission planning or a combination of objectives.

$$\min_{x} f(x) \tag{5.1}$$

subject to

$$g(x) = 0 \tag{5.2}$$

$$h(x) \leq 0 \tag{5.3}$$

$$x^{min} \leq x \leq x^{max} \tag{5.4}$$

Here, $f(x)$ denotes the objective function for minimizing the generation cost. $g(x)$ represents the nodal power flow balance equations, the inequality constraint $h(x)$ models the branch flow limits and the bounds in equation (5.4) limits the voltage magnitudes, generator power injections and reference bus angles [114, 83].

5.3.1 Power Balance Polar Formulation

Here, AC-OPF does a minimization of the generation cost and the objective function used in the formulation is shown in equation (5.5). The generation cost is assumed to be a second order polynomial function.

$$C = \sum_{k=1}^{ng} \alpha_k P_{G_k}^2 + \beta_k P_{G_k} + \gamma_k \tag{5.5}$$

This formulation employs the polar representation of voltage with voltage magnitude at bus i is V_i and angle θ_i. The equality constraints are shown in equations (5.6) and (5.7) and the inequality constraints in equations (5.8) to (5.10) [111, 96].

$$\begin{aligned} \sum_{A_{br}(f,t)=1} & (G_{ff}(V_f^2) + V_f V_t (G_{ft}\cos(\theta_f - \theta_t) \\ & + B_{ft}\sin(\theta_f - \theta_t)) - \sum_{A_G(f,k)=1} P_{Gk} \\ & + \sum_{A_L(f,j)\neq 0} P_{Dj} = \Delta P_f = 0 \end{aligned} \tag{5.6}$$

$$\sum_{A_{br}(f,t)=1} (-B_{ff}(V_f^2) + V_f V_t (G_{ft}\sin(\theta_f - \theta_t) - B_{ft}\cos(\theta_f - \theta_t)) - \sum_{A_G(f,k)\neq 0} Q_{Gk} + \sum_{A_L(f,j)=1} Q_{Dj} = \Delta Q_f = 0 \tag{5.7}$$

$$V_i^- \leq V_i \leq V_i^+ \tag{5.8}$$

$$P_{Gk}^- \leq P_{Gk} \leq P_{Gk}^+ \tag{5.9}$$

$$Q_{Gk}^- \leq Q_{Gk} \leq Q_{Gk}^+ \tag{5.10}$$

$$0 \leq S_f^2 \leq (S_{ft}^+)^2 \tag{5.11}$$

$$0 \leq S_t^2 \leq (S_{ft}^+)^2 \tag{5.12}$$

The equations (5.8) to (5.10), represents the bounds on the voltage magnitude, active power and reactive power injection at each bus. The bounds on the apparent power flows in the network are shown in equations (5.11) and (5.12). For the model, the reference angle is held constant ($\theta_{ref} = \theta_{ref0}$).

$$S_f = \sqrt{P_{ft}^2 + Q_{ft}^2} \tag{5.13}$$

$$S_t = \sqrt{P_{tf}^2 + Q_{tf}^2} \tag{5.14}$$

$$P_{ft} = G_{ff}(V_f^2) + V_f(G_{ft}V_t\cos(\theta_f - \theta_t) + V_t B_{ft}\sin(\theta_f - \theta_t)) \tag{5.15}$$

$$Q_{ft} = -B_{ff}(V_f^2) + V_f(G_{ft}V_t \sin(\theta_f - \theta_t) - V_t B_{ft} \cos(\theta_f - \theta_t)) \tag{5.16}$$

$$P_{tf} = G_{tt}(V_t^2) + V_t(G_{tf}V_f \cos(\theta_t - \theta_f) + V_f B_{tf} \sin(\theta_t - \theta_f)) \tag{5.17}$$

$$Q_{tf} = -B_{tt}(V_t^2) + V_t(G_{tf}V_f \sin(\theta_t - \theta_f) - V_f B_{tf} \cos(\theta_t - \theta_f)) \tag{5.18}$$

The equations (5.13) to (5.18) model the apparent power flows in the network.

5.3.2 Power Balance Cartesian

In this formulation, the voltage at each bus takes the Cartesian form, the real and imaginary part of the voltage are represented by V_{Ri},V_{Ii} respectively ($\bar{V}_i = V_{Ri} + \sqrt{-1}V_{Ii}$). The objective function to be minimized remains the same as in equation (5.5). The constraints are listed in equations (5.19) to (5.21), (5.24) and (5.25).

$$\sum_{A_{br}(f,t)=1} (G_{ff}(V_{Rf}^2 + V_{If}^2) + V_{Rf}(G_{ft}V_{Rt} - B_{ft}V_{It}) + V_{If}(B_{ft}V_{Rt} + G_{ft}V_{It})) - \sum_{A_G(f,k)=1} P_{Gk} + \sum_{A_L(f,j)\neq 0} P_{Dj} = \Delta P_f = 0 \tag{5.19}$$

$$\sum_{A_{br}(f,t)=1} (-B_{ff}(V_{Rf}^2 + V_{If}^2) + V_{If}(G_{ft}V_{Rt} - B_{ft}V_{It}) - V_{Rf}(B_{ft}V_{Rt} + G_{ft}V_{It})) - \sum_{A_G(f,k)\neq 0} Q_{Gk} + \sum_{A_L(f,j)=1} Q_{Dj} = \Delta Q_f = 0 \tag{5.20}$$

$$(V_i^-)^2 \leq V_i^2 = V_{Ri}^2 + V_{Ii}^2 \leq (V_i^+)^2 \tag{5.21}$$

$$P_{Gk}^- \leq P_{Gk} \leq P_{Gk}^+ \tag{5.22}$$

$$Q_{Gk}^- \leq Q_{Gk} \leq Q_{Gk}^+ \tag{5.23}$$

$$0 \leq S_f^2 \leq (S_{ft}^+)^2 \tag{5.24}$$

$$0 \leq S_t^2 \leq (S_{ft}^+)^2 \tag{5.25}$$

where the maximum flow S^+ is either the normal, short-term, or emergency rating of the line. The apparent power flows S_f and S_t at the from and to ends of the line are given by equations (5.26) to (5.31).

$$S_f = \sqrt{P_{ft}^2 + Q_{ft}^2} \tag{5.26}$$

$$S_t = \sqrt{P_{tf}^2 + Q_{tf}^2} \tag{5.27}$$

$$P_f = G_{ff}(V_{Rf}^2 + V_{If}^2) + V_{Rf}(G_{ft}V_{Rt} - B_{ft}V_{It}) + V_{If}(B_{ft}V_{Rt} + G_{ft}V_{It}) \tag{5.28}$$

$$Q_f = -B_{ff}(V_{Rf}^2 + V_{If}^2) + V_{If}(G_{ft}V_{Rt} - B_{ft}V_{It}) - V_{Rf}(B_{ft}V_{Rt} + G_{ft}V_{It}) \tag{5.29}$$

$$P_t = G_{tt}(V_{Rt}^2 + V_{It}^2) + V_{Rt}(G_{tf}V_{Rf} - B_{tf}V_{If}) + V_{It}(B_{tf}V_{Rf} + G_{tf}V_{If}) \tag{5.30}$$

$$Q_t = -B_{tt}(V_{Rt}^2 + V_{It}^2) + V_{It}(G_{tf}V_{Rf} - B_{tf}V_{If}) - V_{Rt}(B_{tf}V_{Rf} + G_{tf}V_{If}) \tag{5.31}$$

5.3.3 Current Balance Cartesian

This formulation employs a set of equations within current injection equations written in rectangular coordinates [111, 115]. The objective function remains the same as in equation (5.5) and the constraints are listed in equations equations (5.32) to (5.34), (5.37) and (5.38). The equations for apparent power takes the form as shown previously in equations (5.26) to (5.31).

$$\begin{aligned}\sum_{A_{br}(f,t)=1}(G_{ft}V_{Rt} - B_{ft}V_{It}) - (\sum_{A_G(f,k)=1} P_{Gk} \\ - \sum_{A_L(f,j)\neq 0} P_{Dj})V_{Rf}/(V_{Rf}^2 + V_{If}^2) \\ -(\sum_{A_G(f,k)=0} Q_{Gk} - \sum_{A_L(f,j)\neq 1}(Q_{Dj}) \\ V_{If}/(V_{Rf}^2 + V_{If}^2) = 0\end{aligned} \tag{5.32}$$

$$\sum_{A_{br}(f,t)=1}(G_{ft}V_{It}+B_{Rt}V_{It})-(\sum_{A_G(f,k)=1}P_{Gk}$$
$$-\sum_{A_L(f,j)\neq 0}P_{Dj})V_{If}/(V_{Rf}^2+V_{If}^2)$$
$$+(\sum_{A_G(f,k)=0}Q_{Gk}-\sum_{A_L(f,j)\neq 1}Q_{Dj}) \tag{5.33}$$
$$V_{Rf}/(V_{Rf}^2+V_{If}^2)=0$$

$$(V_i^-)^2\leq V_i^2=V_{Ri}^2+V_{Ii}^2\leq(V_i^+)^2 \tag{5.34}$$

$$P_{Gk}^-\leq P_{Gk}\leq P_{Gk}^+ \tag{5.35}$$

$$Q_{Gk}^-\leq Q_{Gk}\leq Q_{Gk}^+ \tag{5.36}$$

$$0\leq S_f^2\leq(S_{ft}^+)^2 \tag{5.37}$$

$$0\leq S_t^2\leq(S_{ft}^+)^2 \tag{5.38}$$

Table 5.1 AC-OPF Formulation Comparison

	Power balance Polar	**Power balance Cartesian**	**Current balance Cartesian**
Variables	$\|v_i\|$, θ_i, P_G, Q_G, P, Q	V_R, V_I, P_G, Q_G, P, Q	V_{Ri}, V_{Ii},P_G, Q_G, I_R, I_I
No. of variables	$2n_b+2n_{br}+2n_g$	$2n_b+2n_{br}+2n_g$	$2n_b+2n_{br}+2n_g$
Network constraints	$2n_b$ Nonlinear	$2n_b$ Quadratic	$2n_b$ locally nonlinear
Network Jacobian	Nonlinear	Linear	locally nonlinear
Voltage magnitude constraints	Variable limit	Non-convex quadratic inequalities	Non-convex quadratic inequalities
No. of equations	(Equality) $2n_b+4n_{br}$ (Inequality) $2n_b+4n_{br}+4n_g$	(Equality) $2n_b+4n_{br}$ (Inequality) $2n_b+2n_{br}+4n_g$	(Equality) $2n_b+2n_{br}$ (Inequality) $2n_b+n_{br}+4n_g$

5.3.4 Model Comparison

The table 5.1 lists the model parameters from the three formulations of AC-OPF [84]. All the three formulation have the same number of variables in the optimization model. In case of power balance polar representation, it has a set of $2n_b$ non-linear equality constraints with sine and,-cosine functions as well as quadratic terms that apply throughout the grid. The formulation has a non-linear network Jacobian and variable limit on the voltage magnitude. The power balance Cartesian formulation models the system with $2n_b$ non-linear equality constraints with quadratic terms. The system has non-convex quadratic inequalities at bus and non-convex inequalities at each set of connected buses. The current balance Cartesian formulation employs locally linear equality and non-convex inequalities at each bus.

5.4 Test Cases

The three formulations of AC-OPF are tested on 9 different bus systems. The case 1 is a 9-bus, 3 generator model based on the data from [116]. The case 2, the 39-bus New England system with generator types of fossil, hydro, nuclear and network interconnections [117]. The IEEE 118 and IEEE 300 bus system models are the 4th and 5th test case. The 500-bus system network is a synthetic model to mimic the 138 and 300 kV transmission network in the northwestern part of South Carolina. The synthetic 2000-bus case is a representation of the 161, 230 and, 500 kV transmission network in the state of Texas [118]. The 3120-bus system is a representation of the Polish system during the morning peak in the summer of 2008. A synthetic representation of a part of the Western Electricity Coordinating Council (WECC) system transmission network with 115, 138, 161, 230, 345, 500, and 765 kV lines is modelled in the 10,000 bus

system case. The synthetic 25,000-bus system case is a representation of the North-East Mid-Atlantic region transmission network in the US [118]. The table 5.2 lists the features of the 9 different test cases used.

Table 5.2 The different bus systems utilized for the analysis

Test Case	No. of Buses	No. of branches	No. of generators
1	9	9	3
2	39	46	10
3	118	186	54
4	300	411	69
5	500	597	90
6	2000	3206	544
7	3120	3693	505
8	10000	12706	2485
9	25000	32230	4834

5.5 Numerical and Computational Performance

The OPF application code developed in this work is written in C language using the numerical computing library PETSc [119, 120]. The optimization problem is solved using the Ipopt library [121]. Ipopt is a widely-used open source software library for solving large-scale non-linear optimization problems. It utilizes primal-dual interior point and line search filter based methods to find the solution. All the performance tests were done in a Macintosh environment with 2.6 GHz Intel Core i7 processor, 8 GB 1600 MHz DDR3 RAM, NVIDIA GeForce GT 650M 1024 MB and Intel HD Graphics 4000 1536 MB graphics card.

The formulation of a problem plays a key role in determining the solution methodology and processing time. This section compares the system model characteristics of the three formulation of AC-OPF. The number of variables in the ACOPF formulation of the different bus systems are shown in Figure 5.1, which illustrates the complexity associated with the ACOPF problem. As an example of this complexity, a 9-bus system only has 24 variables while a 25,000-bus system have 59,668 variables in the optimization model. For a test system, the number of variables remains the same for all three AC-OPF formulations.

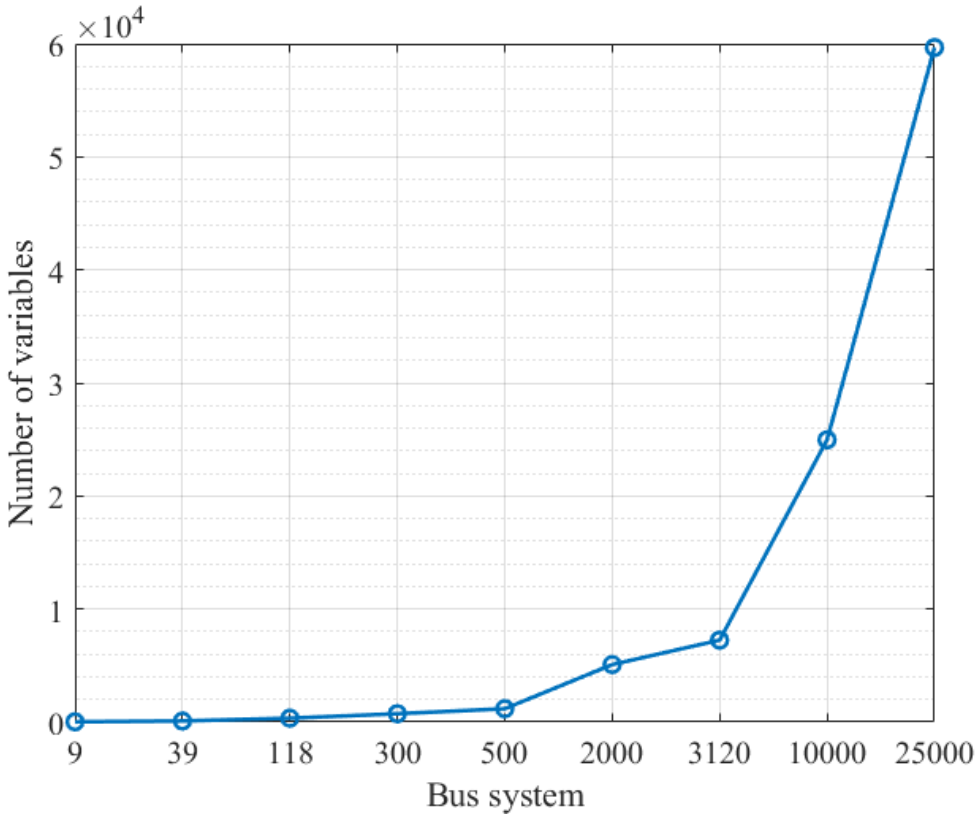


Figure. 5.1 Number of variables in the formulation

The equality and inequality constraints in the Polar and Cartesian formulation are depicted in the figures 5.2 and 5.3. The graphs are plotted in logarithmic scale to show the variation from smaller bus systems. The line flow limits in the case of 118 and 300-bus systems are very high and limits will not be exceeded irrespective of the optimized

value. Therefore the line flow constraints can be excluded in polar formulations as shown in Figure 5.2.

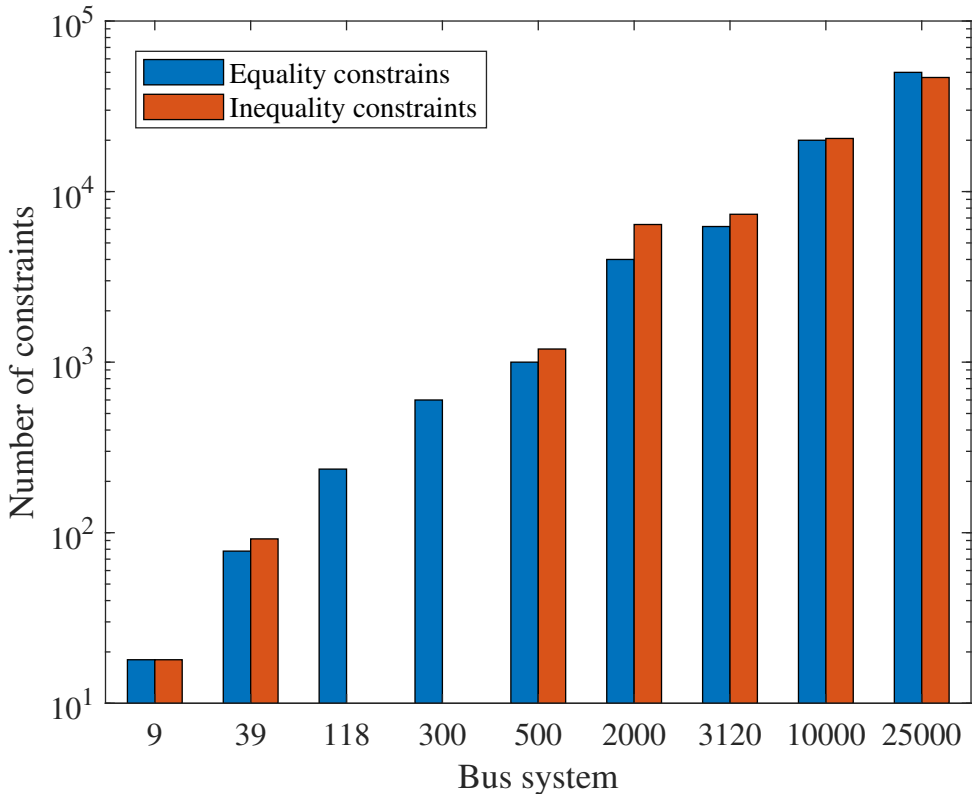


Figure. 5.2 Number of constraints in the power balance Polar formulation.

The number of iterations for the code run to get the solution are shown in figure 5.4. The power balance Cartesian formulation outperforms the other two in smaller bus systems, such as 9,118, and 500-bus, but the current balance Cartesian formulation requires approximately the same number of iterations for larger bus systems, such as the 10,000 or 25,000-bus. Figure 5.6 details the iteration numbers for the three formulations with current balance Cartesian values normalized to 1. The optimization code run-time is shown in figure 5.5, illustrating how the power balance polar form has greater performances in smaller bus system while current balance Cartesian formulation has a lower run time with the largest bus systems analyzed, such as in the case of a 25,000-bus system. The improved performance for the current balance Cartesian form for larger bus

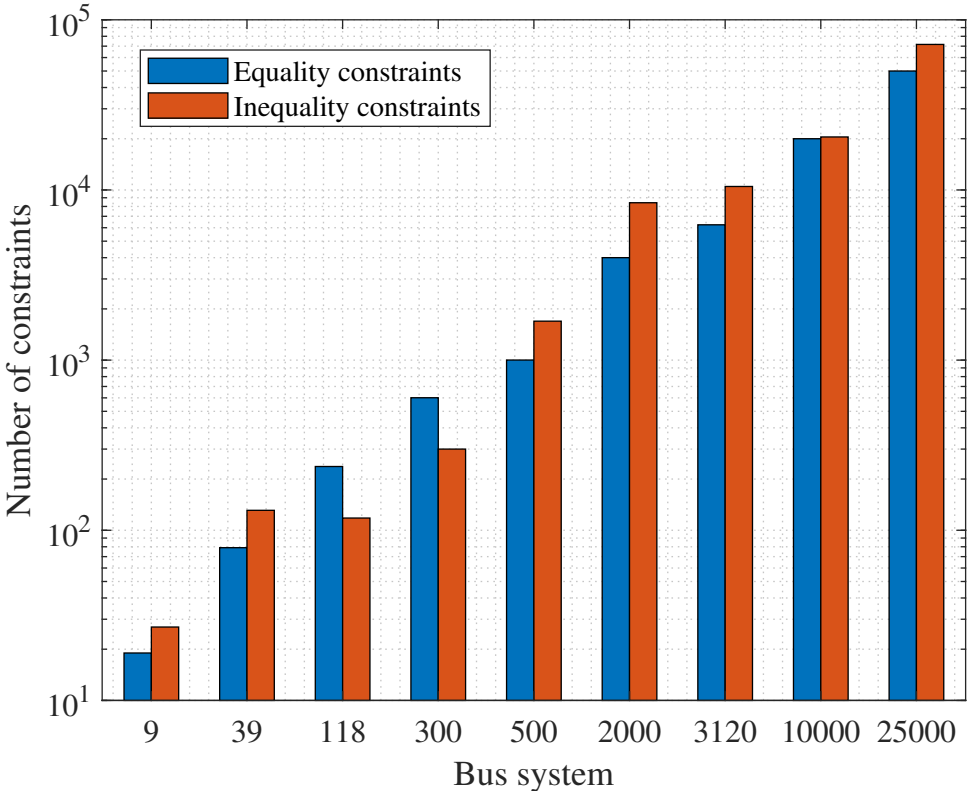


Figure. 5.3 Number of constraints in the power balance Cartesian formulation

system can be clearly seen in Figure 5.6. The performance grows with an increasing number of buses in the network, or with increasing complexity. To understand the modeling framework in a bit more detail, the Jacobian and Hessain values in the formulation are also noted. The values are shown in Table 5.4 and the figure with current balance values normalized to 1 is plotted in Figure 5.8. The acronym Equality (P) stands for the equality constraint Jacobian in power balance Polar formulation, Equality (C) denote equality constraint Jacobian in power balance Cartesian form and similarly for the Inequality terms. The power balance Cartesian and current balance Cartesian show equal or very similar values but power balance polar has lesser values in most cases. The Table 5.4 lists all the values for the the formulations. An extension of this work to SCOPF is also carried out and is shown in Tables 5.6 and 5.7

The results from the code are compared with MATPOWER simulation to check

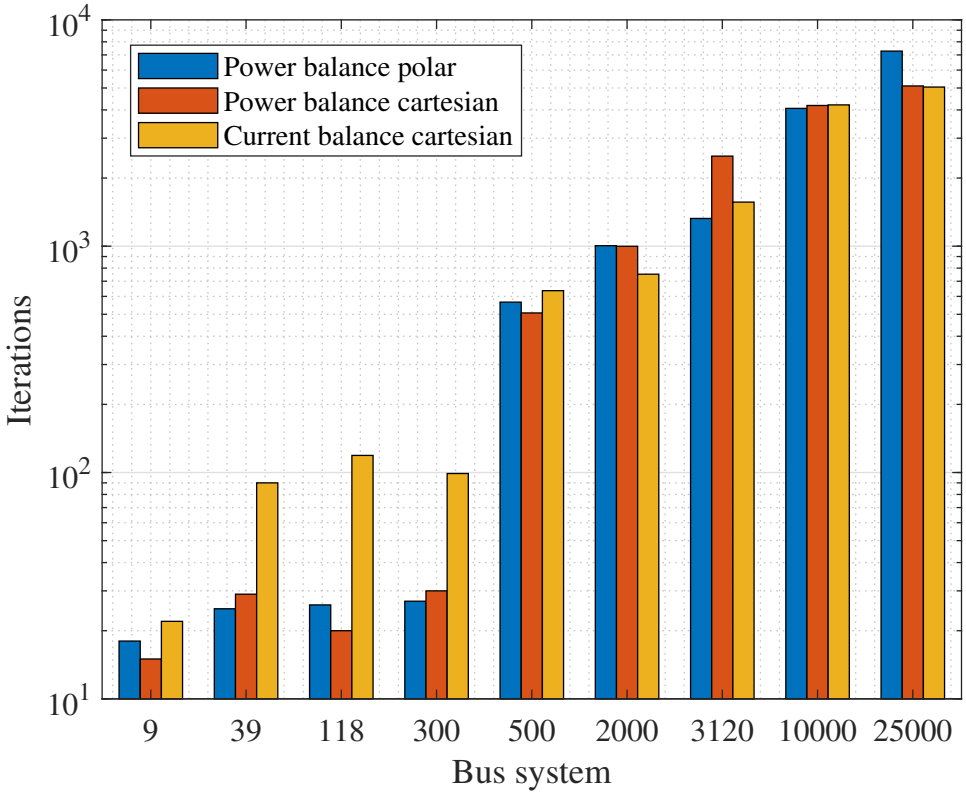


Figure. 5.4 Iterations required to complete the code run

the validity of the results. MATPOWER is an open-source simulations package that provides power flow, OPF, AC-OPF and other tools targeted towards researchers and students [81, 122]. MATPOWER is run in a GNU Octave environment for the power flow simulation with all the cases tested and compared with final objective value [123]. MATPOWER utilizes the MIPS [124] and in polar balance formulation for the analysis. The results are shown in table 5.3, the final objective value from both the executed code and MATPOWER simulations match, up to the first decimal value.

5.6 Observations and Conclusion

This work presented the computational and performance evaluation of three different AC-OPF formulations: power balance polar, power balance Cartesian and current balance Cartesian. The formulations were tested with a wide variety of bus systems, rang-

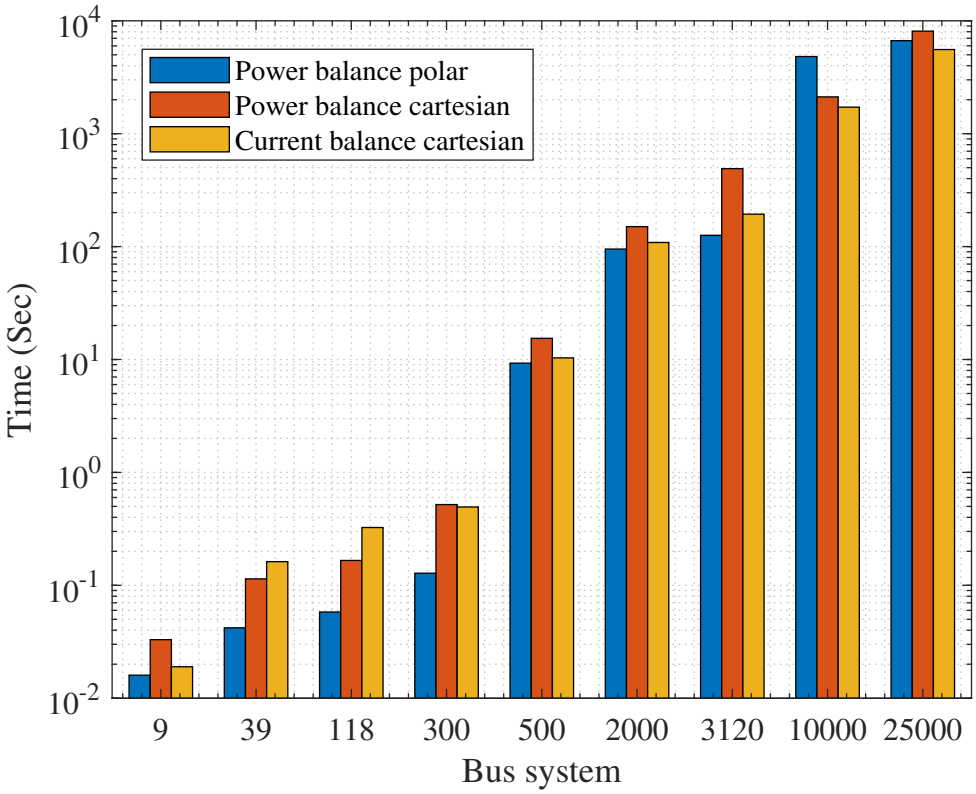


Figure. 5.5 Run time comparison

ing from smaller systems, such as a 9-bus,to larger systems,such as a 25,000-bus. A steady increase in problem complexity as in the number of variables and constraints, with increasing bus number, can be identified from the model. The three formulations converged to the same final solution even though with varying number of iterations and run-time.

Power balance polar form showed the best computational time for smaller bus systems while the current balance Cartesian form showed improvement in computational time with increasing problem complexity, outperforming the other two formulation for the 25,000 bus system case. The results show a similar pattern for power balance polar, having the least iteration number for smaller bus systems, while current balance Cartesian performed the least iteration number for the 25,000-bus system.

The validity of the results is tested in matpower for all cases. The values of non-

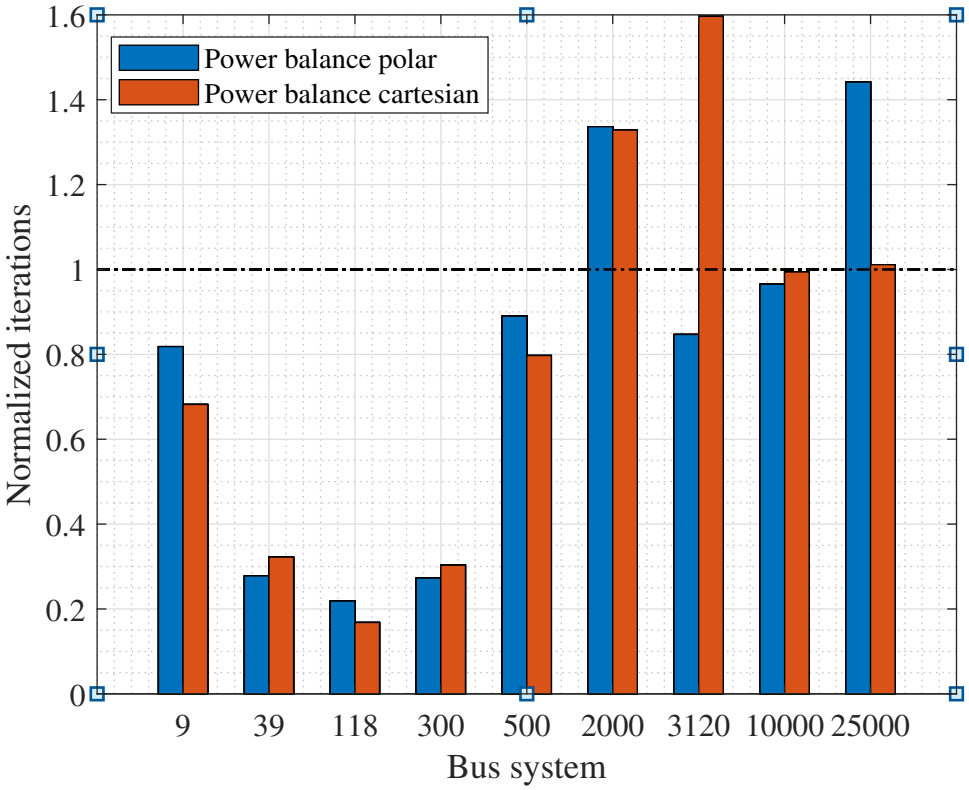


Figure. 5.6 No. of Iterations with current balance iterations normalized to 1

zeros in equality constraint Jacobian, inequality constraint Jacoboian and LaGrangian Hessian in the problem formulation are noted to evaluate the formulations. The current balance Cartesian formulation had the largest number of non-zeros in the equality and inequality constraint Jacobian matrix closely followed by the power balance Cartesian formulation. In the case of non-zeros in Hessian matrix, all three formulations had the same number of non-zero values.

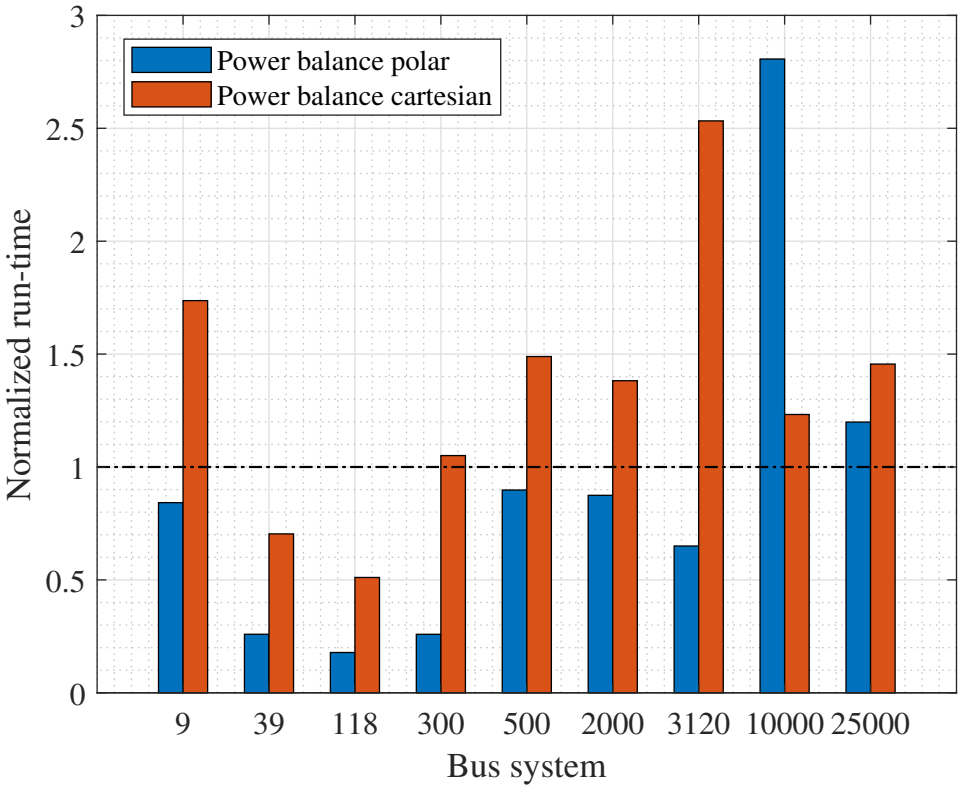


Figure. 5.7 Run-time with current balance run-time normalized to 1

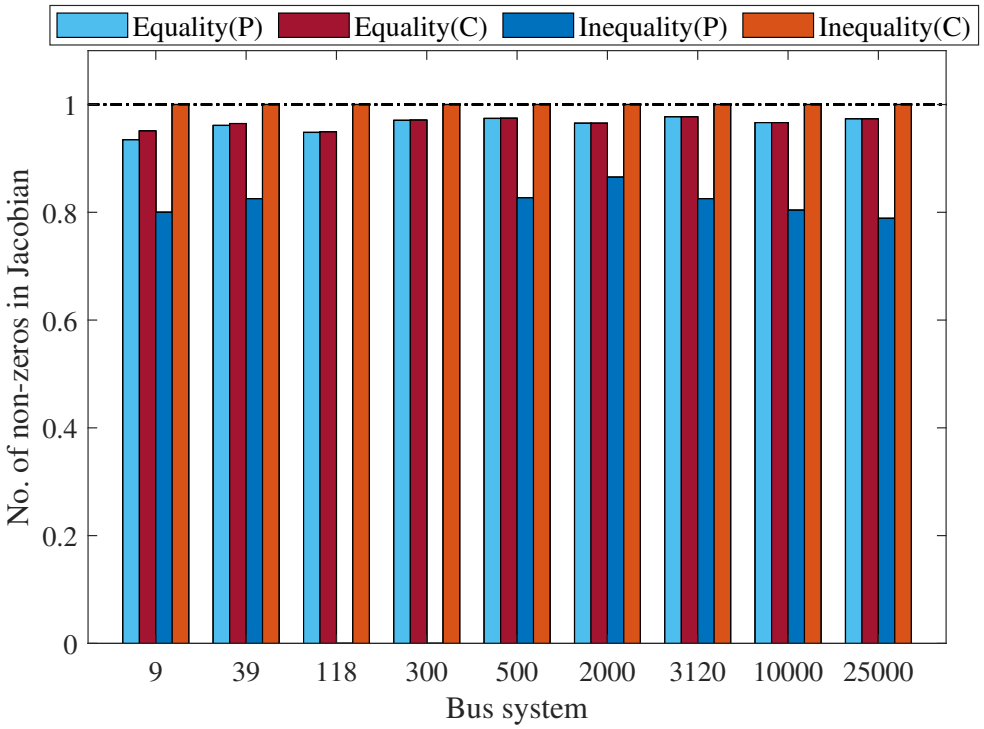


Figure. 5.8 Number of non-zeros in Jacobian with Current balance Cartesian values normalized to 1.

Table 5.3 Objective value comparison with MATPOWER.

Bus System	**MATPOWER**	**Code**
9	5296.69	5297.406
39	41864.18	41864.177
118	129660.7	129660.684
300	719725.11	719725.098
500	72578.3	72578.295
2000	$1.2288 * 10^6$	$1.2288 * 10^6$
3120	$2.1427 * 10^6$	$2.1427 * 10^6$
10000	$2.4858 * 10^6$	$2.4858 * 10^6$
25000	$6.0178 * 10^6$	$6.0178 * 10^6$

Table 5.4 Number of non-zeros in Jacobian and Hessian

Test case	No. of non-zeros	**Power balance polar**	**Power balance Cartesian**	**Current balance Cartesian**
9 Bus System	Equality constraint Jacobian	114	116	122
	Inequality constraint Jacobian	72	90	90
	Lagrangian Hessian	96	96	96
39 Bus System	Equality constraint Jacobian	544	546	566
	Inequality constraint Jacobian	368	446	446
	Lagrangian Hessian	415	415	415
118 Bus System	Equality constraint Jacobian	2012	2014	2122
	Inequality constraint Jacobian	0	236	236
	Lagrangian Hessian	2408	2408	2408
300 Bus System	Equality constraint Jacobian	4610	4612	4750
	Inequality constraint Jacobian	0	600	600
	Lagrangian Hessian	3591	3591	3591
500 Bus System	Equality constraint Jacobian	6852	6854	7034
	Inequality constraint Jacobian	4776	5776	5776
	Lagrangian Hessian	4826	4826	4826
2000 Bus System	Equality constraint Jacobian	30424	30426	31524
	Inequality constraint Jacobian	25648	29648	29648
	Lagrangian Hessian	24552	24552	24552
3120 Bus System	Equality constraint Jacobian	42962	42964	43974
	Inequality constraint Jacobian	29448	35688	35688
	Lagrangian Hessian	35923	35923	35923
10000 Bus System	Equality constraint Jacobian	142706	142708	147678
	Inequality constraint Jacobian	81952	101952	101952
	Lagrangian Hessian	119755	119755	119755
25000 Bus System	Equality constraint Jacobian	350548	350550	360218
	Inequality constraint Jacobian	186640	236640	236640
	Lagrangian Hessian	269342	269342	269342

Table 5.5 AC-OPF Results

Test case	Parameters	**Power balance polar**	**Power balance Cartesian**	**Current balance Cartesian**
9- Bus system	Objective value	5297.406	5297.406	5297.406
	Iterations	18	15	22
	Time (secs)	0.016	0.033	0.019
39-Bus system	Objective value	41864.177	41864.177	41864.177
	Iterations	25	29	90
	Time (secs)	0.042	0.114	0.162
118-Bus system	Objective value	129660.68	129660.68	129660.68
	Iterations	26	20	119
	Time (Secs)	0.058	0.166	0.325
300-Bus system	Objective value	719725.1	719725.1	719725.1
	Iterations	27	30	99
	Time (Secs)	0.128	0.519	0.494
500 Bus system	Objective value	72578.295	72578.295	72578.295
	Iterations	566	507	636
	Time (Secs)	9.284	15.396	10.338
2000 Bus system	Objective value	$1.23 * 10^6$	$1.23 * 10^6$	$1.23 * 10^6$
	Iterations	1005	999	752
	Time (Secs)	95.121	150.327	108.77
3120 Bus System	Objective value	$2.1427 * 10^6$	$2.1427 * 10^6$	$2.1427 * 10^6$
	Iterations	1326	2500	1566
	Time (secs)	125.93	490.638	193.714
10,000 Bus System	Objective value	$2.4858 * 10^6$	$2.4858 * 10^6$	$2.4858 * 10^6$
	Iterations	4063	4185	4210
	Time (secs)	4824.988	2118.738	1719.023
25,000 Bus System	Objective value	$6.017 * 10^6$	$6.017 * 10^6$	$6.017 * 10^6$
	Iterations	7276	5105	5048
	Time (secs)	6672.571	8102.079	5565.438

Table 5.6 SCOPF Results

	Parameters	**Power Balance Polar**	**Power balance cartesian**	**Current balance cartesian**
9- Bus system	Objective value	1299.52	1299.52	1299.52
	Iterations	41	65	95
	Time (secs)	0.232	1.345	0.685
	Variables	224	234	234
	Equality constraints	180	190	190
	Inequality constraints	192	282	282
	Equality constraints Jacobian	1046	1106	1160
	Inequality constraints Jacobian	700	900	900
	Lagrangian Hessian	867	927	927
39-Bus system	Objective value	37054.7	37054.7	37054.7
	Iterations	94	101	209
	Time (secs)	1.547	4.082	3.927
	Variables	970	980	980
	Equality constraints	780	790	790
	Inequality constraints	1004	1394	1394
	Equality constraints Jacobian	5376	5436	5636
	Inequality constraints Jacobian	3816	4616	4616
	Lagrangian Hessian	4090	4150	4150
118-Bus system	Objective value	4646.04	4646.04	4646.04
	Iterations	20	35	50
	Time (Secs)	0.972	3.633	1.943
	Variables	3428	3438	3438
	Equality constraints	2360	2370	2370
	Inequality constraints	485	1665	1665
	Equality constraints Jacobian	19962	20122	21200
	Inequality constraints Jacobian	970	3330	3330
	Lagrangian Hessian	23845	24065	24065

Table 5.7 SCOPF Results (Continued)

	Parameters	**Power Balance Polar**	**Power balance cartesian**	**Current balance cartesian**
300-Bus system	Objective value	46433.9	46433.9	46433.9
	Iterations	22	25	32
	Time (Secs)	1.358	4.763	2.212
	Variables	7370	7380	7380
	Equality constraints	6000	6010	6010
	Inequality constraints	621	3621	3621
	Equality constraints Jacobian	46052	46112	47492
	Inequality constraints Jacobian	1242	7242	7242
	Lagrangian Hessian	35850	35910	35910
500-Bus system	Objective value	26576.2	26576.2	26576.2
	Iterations	982	320	363
	Time (Secs)	245.896	126.307	94.086
	Variables	11100	11110	11110
	Equality constraints	10000	10010	10010
	Inequality constraints	12442	17442	17442
	Equality constraints Jacobian	67782	67842	68952
	Inequality constraints Jacobian	48720	58760	58760
	Lagrangian Hessian	44400	44460	44460

CHAPTER 6

Conclusions and Future Directions

6.1 Conclusions

The electric grid system is undergoing a transitional phase with an increased market penetration of renewable energy sources, increased sales of electric vehicles, the introduction of a time-of-use pricing scheme for consumers, and advancement in high-performance computing. There has also been significant increase in the installation of RES in the US, the largest percentage of which belongs to the wind energy industry setting records in the US with a record high of 56.16% in Electric Reliability Council of Texas (ERCOT) and 71.3% in Southwest Power Pool (SPP). A similar trend can be seen in EVs with a sales growth of 79% in 2018. Operational methods such as economic dispatch and optimal power flow perform key roles in the stability and economic operation of changing grid demands. We investigated four key areas of the grid operation: microgrid decomposition, EV scheduling, wind forecasts and ACOPF formulation.

6.1.1 Microgrid decomposition

Microgrid decomposition, or determining community structures within an electric grid is important for the optimal management of the transmission system. Different clustering techniques were employed to generate microgrids from the larger electric grid

system and a linear programming based ED model was developed to find the most economic microgrid systems. The proposed novel approach of the NGGN clustering scheme generated the best results in MAED for a microgrid system.

6.1.2 EV Scheduling

The number of EVs in the system load of a utility is increasing, with demand growing as the EVs become more affordable. Most consumers who own EVs are likely to begin charging at a time that coincides with peak load in the electric grid, which the utilities strive to mitigate. This mitigation ultimately helps with the integration of RES and the reduction of generation cost, and necessitates the need for an EV scheduling scheme to reduce the cost for consumers and peak load in the system. A novel linear programming based EV scheduling scheme model is proposed, which takes in to account consumer preferences, household demand, forecasted EV demand, and community load. A representative dataset of a modern residential community with 200 households and 348 EVs were utilized to model the scheduling scheme. The proposed strategy will reduce the energy costs for consumers as an incentive for buy in, along with a reduction in peak load from the residential community.

6.1.3 Wind power generation forecasting

To increase the percentage of RES in the generation mix of utility power generation, forecasting plays a critical role. We have investigated the performance of five different forecasting methodologies along with four proposed hybrid approaches in handling wind power generation using, the NREL wind integration dataset with six predictors variables and a data resolution of five minutes. A short-term prediction of wind generation within 6 hours was performed, utilizing 29 days of historic data and employing

all predictor variables. Forecasts were deployed in a rolling window approach for all 12 months of the year, with 19 one-hour ahead iterations of the 6-hour ahead predictions. Instead of relying on a small portion of the dataset, we utilized a rolling window approach to investigate the most accurate forecasting approacheses. The GLM based hybrid approaches, GLM-SVM and GLM-RF, were able to deliver consistent results for the varying datasets.

6.1.4 ACOPF formulation

ACOPF is one of the key operation for the stability and security of the grid. It is the most accurate representation of power flows in a network, assuming model parameters are correct. Operators rely on assumptions and approximations to generate the best possible accuracy for ACOPF, even a minor improvement in ACOPF accuracy or run-time is beneficial in the integration of RES with millions of dollars in potential cost saving. We investigated the computational and numerical performance of ACOPF formulations on large-scale networks, investigating the three formulations of ACOPF: power balance polar, power balance Cartesian and current balance Cartesian. The formulations were tested on networks ranging from a 9-bus to a 25,000 bus system to study variations from small-scale networks to large-scale networks. All three formulations arrived at the same solution, and the results were validated using Matpower. The power balance polar formulation performed well with smaller networks, but the current balance cartesian displayed improvement with increasing network complexity. The current balance cartesian formulation had a 16.59% improvement in computational time when compared to power balance formulation, and a 31.31% improvement when compared to power balance cartesian formulation on a 25,000 bus system.

6.2 Future Work

Multiple microgrids working in tandem is the ideal state for the future electric grid. The network is expected to have high EV and RES market penetration. The proposed clustering scheme for the economic decomposition of electric grid systems was evaluated on IEEE 118 and 300-bus systems, which need to be expanded to large scale test systems. One of the key benefit of using microgrids is the ability to isolate the system in times of electric grid disturbances. The performance of these clustered microgrids needs to be evaluated during electrical disturbances, such as tie-line failures and generation loss.

The proposed EV scheduling utilized data for a high-end community with all households owning connected EVs. The dataset does not specify the EV variant of the owner, which can influence the scheduling scheme. The idea of vehicle-to-grid (V2G) and vehicle-to-home (V2H) can be incorporated in the scheduling scheme for the benefit of both consumers and utilities. The power flow from V2G during peak load can help the utilities in peak-shaving. Most of the households in the dataset are having multiple EVs, so the idea of vehicle-to-vehicle (V2V) can also be investigated. The effects of these schemes on battery life and cost to customers during maintenance need to be explored to convince users to buy in to these technologies.

The wind power forecasts were focused on the short-term analysis of wind power generation with 6 hour ahead predictions. Short-term forecasts can vary from 30 minutes to 6 hour depending on the functionality. For future work, the performance of these methods for different forecast horizons needs to be evaluated. The prediction methods made use of all available covariates but the concept of feature selection can be used to further improve the forecast accuracy. The idea of hyperparameter tuning for RF and

SVM can also be employed to improve accuracy. An emerging topic in wind generation forecast is the application of NN based methods, with necessary testing of the different approaches in their ability to handle varying wind generation data.

The AC OPF task increases the accuracy of current-balance formulation over the other two models, power balance polar and power balance cartesian, on large scale test grids. The formulations were tested on bus systems upto a 25,000-bus case, but even larger network models such as the 75,000-bus system can be explored for validation. The 75,000-bus system was not evaluated in our study due to excessive computational time on commodity hardware. The performance of these models needs to be evaluated on a high-performance framework, and the ability of these models for parallel computing also needs to be explored.

REFERENCES

[1] J. Dillon and M. O. Malley, "Impact of Uncertainty on Wind Power Curtailment Estimation," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, pp. 3011–3016.

[2] W. Warsono, D. J. King, C. S. Özveren, and D. A. Bradley, "Economic Load Dispatch Optimization of Renewable Energy in Power System Using Genetic Algorithm," in *2007 IEEE Lausanne Power Tech*, no. August, 2007, pp. 1–6.

[3] "U.s. energy information administration - eia - independent statistics and analysis," Jun 2020. [Online]. Available: https://www.eia.gov/energyexplained/renewable-sources/

[4] North American Electric Reliability Corporation, "Accommodating High Levels of Variable Generation," Tech. Rep. April, 2009.

[5] S. Surender Reddy, P. R. Bijwe, and A. R. Abhyankar, "Real-Time Economic Dispatch Considering Renewable Power Generation Variability and Uncertainty over Scheduling Period," *IEEE Systems Journal*, vol. 9, no. 4, pp. 1440–1451, 2015.

[6] M. Birk and R. Tabors, D., "The Impact of Distributed Energy Resources on the Bulk Power System : A Deeper Dive," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, pp. 3034–3043.

[7] K. E. Nygard, S. B. Ghosn, M. M. Chowdhury, D. Loegering, R. McCulloch, and P. Ranganathan, "Optimization models for energy reallocation in a smart grid," *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 186–190, 2011.

[8] M. Zidar, P. S. Georgilakis, N. D. Hatziargyriou, T. Capuder, and D. Škrlec, "Review of energy storage allocation in power distribution networks: applications, methods and future research," *IET Generation, Transmission and Distribution*, vol. 10, no. 3, 2016.

[9] B. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1248–1259, 1990.

[10] P. Ranganathan and K. Nygard, "An Optimal Resource Assignment Problem in Smart Grid," in *The Second International Conference on Future Computational Technologies and Applications*, 2010, pp. 28–34.

[11] H. Agarwal, J. E. Renaud, E. L. Preston, and D. Padmanabhan, "Uncertainty quantification using evidence theory in multidisciplinary design optimization," *Reliability Engineering and System Safety*, vol. 85, no. 1-3, pp. 281–294, 2004.

[12] R. Lima and R. Sampaio, "What is uncertainty quantification?" *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 40, no. 3, pp. 1–8, 2018.

[13] Midcontinent Independent System Operator, "Dispatchable Intermittent Resources (DIR) Wind Forecasting Workshop," MISO, Tech. Rep., 2019.

[14] A. Lira, P. Rosas, A. Araujo, and N. Castro, "Uncertainties in the estimate of wind energy production," in *Proceedings of the Energy Economics Iberian Conference*, 2016.

[15] INSIDEEVs, "Monthly Plug-In EV Sales Scorecard: Historical Charts," 2020. [Online]. Available: https://insideevs.com/photo/4028700/monthly-plug-in-ev-sales-scorecard-historical-charts/

[16] EEI, "Electric Vehicle Sales : Facts and Figures," Tech. Rep., 2019.

[17] K. B. Sahay, A. Sonkar, and A. Kumar, "Economic Load Dispatch Using Genetic Algorithm Optimization Technique," in *2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)*. Asian Institute of Technology, 2018, pp. 1–5.

[18] J. P. Zhan, Q. H. Wu, C. X. Guo, and X. X. Zhou, "Fast λ-iteration method for economic dispatch with prohibited operating zones," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 990–991, 2014.

[19] T. Ding and Z. Bie, "Parallel Augmented Lagrangian Relaxation for Dynamic Economic Dispatch Using Diagonal Quadratic Approximation Method," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1115–1126, 2017.

[20] C. Chiang, "Genetic-based algorithm for power economic load dispatch," *IET Generation, Transmission and Distribution*, vol. 1, no. 2, pp. 261–269, 2007.

[21] V. B. Kasangaki, H. M. Sendaula, and S. K. Biswas, "Stochastic Hopfield artificial neural network for unit commitment and economic power dispatch," *Electric Power Systems Research*, vol. 42, no. 3, pp. 215–223, 1997.

[22] Z.-l. Gaing, "Particle swarm optimization to solving the economic dispatch considering the generator constraints," *IEEE Transactions on power systems*, vol. 18, no. 3, pp. 1187–1195, 2003.

[23] K. P. Wong and C. C. Fung, "Simulated annealing based economic dispatch algorithm," *IEE Proceedings C: Generation Transmission and Distribution*, vol. 140, no. 6, pp. 509–515, 1993.

[24] Y. Yang, B. Wei, H. Liu, Y. Zhang, J. Zhao, and E. Manla, "Chaos Firefly Algorithm with Self-Adaptation Mutation Mechanism for Solving Large-Scale Economic Dispatch with Valve-Point Effects and Multiple Fuel Options," *IEEE Access*, vol. 6, pp. 45 907–45 922, 2018.

[25] J. Kumar and G. Sheble, "Clamped State Solution of Artificial Neural Network for Real-Time Economic Dispatch," *IEEE Trans. Power Systems*, vol. 10, no. 2, pp. 925–931, 1995.

[26] M. Y. Hassan, M. N. Suharto, M. P. Abdullah, M. S. Majid, and F. Hussin, "Application of particle swarm optimization for solving optimal generation plant location problem," *International Journal of Electrical and Electronic Systems Research*, vol. 5, no. 1, pp. 47–56, 2012.

[27] A. V. Sudhakar, K. Chandram, and A. J. Laxmi, "Multi Area Economic Dispatch with Tie Line Loss using Secant Method and Tie Line Matrix," *International Journal of Applied Power Engineering (IJAPE)*, vol. 2, no. 3, pp. 115–124, 2013.

[28] D. Streiffert, "Multi-area Economic Dispatch with Tie Line Constraints," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 1946–1951, 1995.

[29] PJM, "PJM Energy Market," 2016. [Online]. Available: https://www.pjm.com/markets-and-operations/energy.aspx

[30] ——, "Map of PJM Territory served," 2020. [Online]. Available: https://www.pjm.com/about-pjm/who-we-are/territory-served.aspx

[31] P. Chaiwuttisak, "Forecasting export value in the automobile industry," in *2018 5th International Conference on Business and Industrial Research (ICBIR)*, 2018, pp. 95–99.

[32] "UW Power System Test Case Archive," 2017. [Online]. Available: http://www.ee.washington.edu/research/pstca/.

[33] K. Rajashekara, "Present status and future trends in electric vehicle propulsion technologies," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 1, no. 1, pp. 3–10, 2013.

[34] C. Liu, K. T. Chau, D. Wu, and S. Gao, "Opportunities and Challenges of Vehicle-to-Home, Vehicle-to-Vehicle, and Vehicle-to-Grid Technologies," *Proceedings of the IEEE*, vol. 101, no. 11, pp. 2409–2427, 2013.

[35] A. Dubey and S. Santoso, "Electric Vehicle Charging on Residential Distribution Systems: Impacts and Mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[36] M. Usman, L. Knapen, A. U. H. Yasar, Y. Vanrompay, T. Bellemans, D. Janssens, and G. Wets, "A coordinated Framework for Optimized Charging of EV Fleet in Smart Grid," *Procedia Computer Science*, vol. 94, no. Fnc, pp. 332–339, 2016.

[37] M. Muratori, "Impact of uncoordinated plug-in electric vehicle charging on residential power demand," *Nature Energy*, vol. 3, pp. 193–201, 2018.

[38] M. Muratori, M. C. Roberts, R. Sioshansi, V. Marano, and G. Rizzoni, "A highly resolved modeling technique to simulate residential power demand," *Applied Energy*, vol. 107, pp. 465–473, 2013.

[39] M. Muratori, V. Marano, R. Sioshansi, and G. Rizzoni, "Energy consumption of residential HVAC systems: A simple physically-based model," in *IEEE Power and Energy Society General Meeting*, 2012.

[40] M. Muratori, M. J. Moran, E. Serra, and G. Rizzoni, "Highly-resolved modeling of personal transportation energy consumption in the United States," *Energy*, vol. 58, pp. 168–177, 2013.

[41] D. B. De Alencar, C. De Mattos Affonso, R. C. L. De Oliveira, J. L. M. Rodríguez, J. C. Leite, and J. C. R. Filho, "Different Models for Forecasting Wind Power Generation: Case Study," *Energies*, vol. 10, no. 12, 2017.

[42] N. Korolko and Z. Sahinoglu, "Robust optimization of EV charging schedules in unregulated electricity markets," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 149–157, 2017.

[43] M. J. Risbeck, C. T. Maravelias, J. B. Rawlings, and R. D. Turney, "A mixed-integer linear programming model for real-time cost optimization of building heating, ventilation, and air conditioning equipment," *Energy and Buildings*, vol. 142, pp. 220–235, 2017.

[44] Maigha and M. L. Crow, "Multi-objective electric vehicle scheduling considering customer and system objectives," in *IEEE Manchester PowerTech*, 2017, pp. 1–6.

[45] "MISO Day-Ahead price data." [Online]. Available: http://www.energyonline.com/Data/GenericData.aspx?DataId=9&MISO___Day-Ahead_Energy_Price

[46] "U.s. installed and potential wind power capacity and generation." [Online]. Available: https://windexchange.energy.gov/maps-data/321

[47] "Wind industry market reports." [Online]. Available: https://www.awea.org/resources/publications-and-reports/market-reports

[48] M. Giebel, G., Brownsword R., Kariniotakis, G Denhard and C. Draxl, "The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview," DTU, Tech. Rep., 2011.

[49] X. Zhu and M. G. Genton, "Short-Term Wind Speed Forecasting for Power System Operations," *International Statistical Review*, vol. 80, no. 1, pp. 2–23, 2012.

[50] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview," *Technical Report, ANEMOS.plus*, pp. 1–109, 2003.

[51] J. Dowell and P. Pinson, "Very-Short-Term Probabilistic Wind Power Forecasts by Sparse Vector Autoregression," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 763–770, 2016.

[52] M. Lackner, A. Rogers, and J. Manwell, "Uncertainty analysis in wind resource assessment and wind energy production estimation," in *45th AIAA Aerospace Sciences Meeting and Exhibit*, 2007, pp. 1–16. [Online]. Available: http://arc.aiaa.org/doi/pdf/10.2514/6.2007-1222

[53] 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*, vol. 29, no. 2, pp. 656–665, 2014.

[54] E. Grigonytė and E. Butkevičiūtė, "Short-term wind speed forecasting using ARIMA model," *Energetika*, vol. 62, no. 1-2, pp. 45–55, 2016.

[55] I. Colak, S. Sagiroglu, M. Yesilbudak, E. Kabalci, and H. Ibrahim Bulbul, "Multi-time series and-time scale modeling for wind speed and wind power forecasting part I: Statistical methods, very short-term and short-term applications," in *2015 International Conference on Renewable Energy Research and Applications, ICRERA 2015*, vol. 5. IEEE, 2015, pp. 209–214.

[56] A. K. Nayak, K. C. Sharma, R. Bhakar, and J. Mathur, "ARIMA based statistical approach to predict wind power ramps," in *IEEE Power and Energy Society General Meeting*. IEEE, 2015, pp. 1–5.

[57] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 667–676, 2010.

[58] L. Yang, M. He, J. Zhang, and V. Vittal, "Support-vector-machine-enhanced markov model for short-term wind power forecast," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 791–799, 2015.

[59] S. Scher and J. Molinder, "Machine learning-based prediction of icing-related wind power production loss," *IEEE Access*, vol. 7, pp. 129 421–129 429, 2019.

[60] O. Kramer and F. Gieseke, "Short-term wind energy forecasting using support vector regression," *Advances in Intelligent and Soft Computing*, vol. 87, pp. 271–280, 2011.

[61] R. Yang and G. Hug, "Potential and efficient computation of corrective power flow control in cost vs. Risk Trade-Off," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2033–2043, 2014.

[62] Z. Zhu, D. Zhou, and Z. Fan, "Short term forecast of wind power generation based on SVM with pattern matching," in *2016 IEEE International Energy Conference, ENERGYCON 2016*, no. 1. IEEE, 2016, pp. 1–6.

[63] Y.-D. Syu, J.-C. Wang, C.-Y. Chou, M.-J. Lin, W.-C. Liang, L.-C. Wu, and J.-A. Jiang, "Ultra-Short-Term Wind Speed Forecasting for Wind Power Based on Gated Recurrent Unit," in *2020 8th International Electrical Engineering Congress (iEECON)*, 2020, pp. 1–4.

[64] D. Zheng, A. T. Eseye, J. Zhang, and H. Li, "Short-term wind power forecasting using a double-stage hierarchical ANFIS approach for energy management in microgrids," *Protection and Control of Modern Power Systems*, vol. 2, no. 1, p. 13, 2017.

[65] S. Li, P. Wang, and L. Goel, "Wind Power Forecasting Using Neural Network Ensembles with Feature Selection," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1447–1456, 2015.

[66] D. T. Viet, V. V. Phuong, and M. Q. Duong, "Models for Short-Term Wind Power Forecasting Based on Improved Artificial Neural Network Using Particle Swarm Optimization and Genetic Algorithms," *Energies*, vol. 13, no. 11, pp. 1–22, 2020.

[67] G. W. Chang, H. J. Lu, Y. Y. Chen, and Y. R. Chang, "Forecasting Wind Power Generation by A New Type of Radial Basis Function-based Neural Network," in *2017 IEEE Power and Energy Society General Meeting*, 2017.

[68] B. Kosovic, S. E. Haupt, D. Adriaansen, S. Alessandrini, G. Wiener, L. D. Monache, Y. Liu, S. Linden, T. Jensen, W. Cheng, M. Politovich, and P. Prestopnik, "A Comprehensive Wind Power Forecasting System Integrating Artificial Intelligence and Numerical Weather Prediction," *Energies*, vol. 13, no. 6, pp. 1–16, 2020.

[69] S. A. Vargas, G. R. T. Esteves, P. M. Maçaira, B. Q. Bastos, F. L. Cyrino Oliveira, and R. C. Souza, "Wind power generation: A review and a research agenda," *Journal of Cleaner Production*, vol. 218, pp. 850–870, 2019.

[70] C. Draxl, A. Clifton, B.-m. Hodge, and J. Mccaa, "The Wind Integration National Dataset (WIND) Toolkit," *Applied Energy*, vol. 151, pp. 355–366, 2015.

[71] Q. Chen and K. A. Folly, "Wind Power Forecasting," *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 414–419, 2018.

[72] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.

[73] J. Fox, "Generalized Linear Models," in *Applied regression and generalized linear models*. SAGE Publications, 2015.

[74] U. Kruger, L. Lamont, and L. El Chaar, "Seasonal Analysis and Prediction of Wind Energy Using Random Forests and ARX Model Structures," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1994–2002, 2015.

[75] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[76] R. Angamuthu Chinnathambi, A. Mukherjee, M. Campion, H. Salehfar, T. Hansen, J. Lin, and P. Ranganathan, "A Multi-Stage Price Forecasting Model for Day-Ahead Electricity Markets," *Forecasting*, vol. 1, no. 1, pp. 26–46, 2018.

[77] J. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization," in *Advances in Kernel Methods — Support Vector Learning*. MIT Press, 1999.

[78] D. Basak, S. Pal, and D. C. Patranabis, "Support vector regression," *Neural Information Processing – Letters and Reviews*, vol. 11, no. 10, pp. 203–224, 2007.

[79] J. Carpentier, "Optimal power flows," *International Journal of Electrical Power and Energy Systems*, vol. 1, no. 1, pp. 3–15, 1979.

[80] W. F. Tinney, "Optimal Power Flow Solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, 1968.

[81] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[82] S. Frank and S. Rebennack, "A Primer on Optimal Power Flow : Theory , Formulation , and Practical Examples," Tech. Rep. October, 2012.

[83] S. Chatzivasileiadis, "Optimization in Modern Power Systems," Technical University of Denmark (DTU), Tech. Rep. September, 2018.

[84] M. B. Cain, R. P. O'neill, and A. Castillo, "History of Optimal Power Flow and Formulations," Tech. Rep., 2012.

[85] P. N. Vovos and J. W. Bialek, "Optimal power flow as a generation expansion and network reinforcement planning tool," in *2006 IEEE Power Engineering Society General Meeting, PES*, 2006.

[86] T. Soares, R. J. Bessa, P. Pinson, and H. Morais, "Active distribution grid management based on robust AC optimal power flow," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6229–6241, 2018.

[87] R. Moreno, J. Obando, and G. Gonzalez, "An integrated OPF dispatching model with wind power and demand response for day-ahead markets," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 4, pp. 2794–2802, 2019.

[88] Y. Liu, Y. Sun, D. Infield, Y. Zhao, S. Han, and J. Yan, "A Hybrid Forecasting Method for Wind Power Ramp Based on Orthogonal Test and Support Vector Machine (OT-SVM)," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 451–457, 2017.

[89] L. F. Ochoa and G. P. Harrison, "Using AC optimal power flow for DG planning and optimisation," in *IEEE PES General Meeting, PES 2010*. IEEE, 2010, pp. 1–7.

[90] F. Fioretto, T. W. K. Mak, and P. Van Hentenryck, "Predicting AC Optimal Power Flows: Combining Deep Learning and Lagrangian Dual Methods," *arXiv:1909.10461*, no. 1, 2019.

[91] A. S. Nair, T. Hossen, M. Campion, and P. Ranganathan, "Optimal Operation of Residential EVs using DNN and Clustering based Energy Forecast," in *2018 North American Power Symposium, NAPS 2018*. IEEE, 2018, pp. 1–6.

[92] A. S. Nair, T. Hossen, M. Campion, D. F. Selvaraj, N. Goveas, N. Kaabouch, and P. Ranganathan, "Multi-Agent Systems for Resource Allocation and Scheduling in a Smart Grid," *Technology and Economics of Smart Grids and Sustainable Energy*, vol. 3, no. 1, 2018.

[93] A. Castillo, "Essays on the ACOPF problem: formulations, approximations, and applications in the electricity markets," Ph.D. dissertation, 2016.

[94] A. Castillo, P. Lipka, J. P. Watson, S. S. Oren, and R. P. O'Neill, "A Successive Linear Programming Approach to Solving the IV-ACOPF," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2752–2763, 2016.

[95] D. Shchetinin, T. T. De Rubira, and G. Hug, "Conservative linear line flow constraints for AC optimal power flow," in *2017 IEEE Manchester PowerTech, Powertech 2017*. IEEE, 2017, pp. 1–6.

[96] T. Akbari and M. T. Bina, "Linear approximated formulation of AC optimal power flow using binary discretisation," *IET Generation, Transmission and Distribution*, vol. 10, no. 5, pp. 1117–1123, 2016.

[97] D. Shchetinin, T. T. De Rubira, and G. Hug, "On the Construction of Linear Approximations of Line Flow Constraints for AC Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1182–1192, 2019.

[98] M. Baradar and M. R. Hesamzadeh, "AC power flow representation in conic format," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 546–547, 2015.

[99] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification-part i," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2564, 2013.

[100] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex Relaxations of Chance Constrained AC Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2829–2841, 2018.

[101] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, 2015.

[102] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "The QC Relaxation: A Theoretical and Computational Study on Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3008–3018, 2016.

[103] S. Mhanna, G. Verbic, and A. C. Chapman, "Adaptive ADMM for distributed ac optimal power flow," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2025–2035, 2019.

[104] R. R. Shoults and D. T. Sun, "Optimal power flow based on P–Q decomposition," *IEEE Trans Power Appar Syst*, vol. PAS-101, no. 1, pp. 745–751, 1982.

[105] A. Venzke, S. Chatzivasileiadis, and D. K. Molzahn, "Inexact Convex Relaxations for AC Optimal Power Flow: Towards AC Feasibility," *arXiv:1902.04815*, no. 6154, pp. 1–11, 2019.

[106] W. M. Lin, C. H. Huang, and T. S. Zhan, "A hybrid current-power optimal power flow technique," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 177–185, 2008.

[107] M. Jereminov, A. Pandey, and L. Pileggi, "Equivalent circuit formulation for solving ac optimal power flow," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2354–2365, 2019.

[108] Y. Tang, K. Dvijotham, and S. Low, "Real-Time Optimal Power Flow," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2963–2973, 2017.

[109] R. P. O'Neill, A. Castillo, and M. B. Cain, "The IV formulation and linear approximations of the AC optimal power flow problem (OPF Paper 2)," *FERC Staff Technical Paper*, no. December, pp. 1–18, 2012.

[110] S. Cvijic, P. Feldmann, and M. Hie, "Applications of homotopy for solving AC power flow and AC optimal power flow," in *IEEE Power and Energy Society General Meeting*, 2012.

[111] V. M. Da Costa and A. L. S. Rosa, "A comparative analysis of different power flow methodologies," in *2008 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America*. IEEE, 2008, pp. 1–7.

[112] B. Sereeter, C. Vuik, C. Witteveen, and P. Palensky, "Optimal power flow formulations and their impacts on the performance of solution methods," in *IEEE Power and Energy Society General Meeting*, 2019.

[113] B. Park, L. Tang, M. C. Ferris, and C. L. Demarco, "Examination of Three Different ACOPF Formulations with Generator Capability Curves," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2913–2923, 2017.

[114] J. Kardos, D. Kourounis, O. Schenk, and R. Zimmerman, "Complete results for a numerical evaluation of interior point solvers for large-scale

optimal power flow problems," Tech. Rep., 2018. [Online]. Available: http://arxiv.org/abs/1807.03964

[115] L. R. Araujo, D. R. R. Penido, S. Carneiro, and J. L. R. Pereira, "A three-phase optimal power-flow algorithm to mitigate voltage unbalance," *IEEE Transactions on Power Delivery*, vol. 28, no. 4, pp. 2394–2402, 2013.

[116] J. H. Chow, *Time-Scale Modeling of Dynamic Networks with Applications to Power Systems*. Springer-Verlag, 1982.

[117] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 2, pp. 573–584, 1979.

[118] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid Structural Characteristics as Validation Criteria for Synthetic Networks," *IEEE Transactions on Power Systems*, 2017.

[119] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. D. Gropp, D. Karpeyev, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, R. T. Mills, T. Munson, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, H. Zhang, and H. Zhang, "PETSc users manual," Argonne National Laboratory, Tech. Rep. ANL-95/11 - Revision 3.13, 2020. [Online]. Available: https://www.mcs.anl.gov/petsc

[120] ——, "PETSc Web page," https://www.mcs.anl.gov/petsc, 2019. [Online]. Available: https://www.mcs.anl.gov/petsc

[121] A. Wachter and L. T. B. On, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.

[122] C. E. Murillo-Sánchez, R. D. Zimmerman, C. Lindsay Anderson, and R. J. Thomas, "Secure planning and operations of systems with stochastic sources, energy storage, and active demand," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2220–2229, 2013.

[123] J. Eaton, D. Batema, S. Hauberg, and R. Wehbring, "GNU Octave version 5.1.0 manual: a high-level interactive language for numerical computations," Tech. Rep., 2019.

[124] H. Wang, C. E. Murillo-Sánchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185–1193, 2007.

Appendices

Appendix A

ED based on grid clustering

1. **118 Bus system A-GN Clustering** (MOD File)

```
set GEN;              #Generator
set Loads;        #Loads
set GEN_Area1;  #Gens in Area 1
set GEN_Area2;  #Gens in Area 2
set GEN_Area3;  #Gens in Area 3
set GEN_Area4;  #Gens in Area 4
set GEN_Area5;  #Gens in Area 5
set GEN_Area6;  #Gens in Area 6
set GEN_Area7;  #Gens in Area 7
set GEN_Area8;  #Gens in Area 8
set GEN_Area9;  #Gens in Area 9
set GEN_Area10; #Gens in Area 10
set GEN_Area11; #Gens in Area 11
set Load_Area1; #Loads in Area 1
set Load_Area2; #Loads in Area 2
set Load_Area3; #Loads in Area 3
set Load_Area4; #Loads in Area 4
set Load_Area5; #Loads in Area 5
set Load_Area6; #Loads in Area 6
set Load_Area7; #Loads in Area 7
set Load_Area8; #Loads in Area 8
set Load_Area9; #Loads in Area 9
set Load_Area10;#Loads in Area 10
set Load_Area11;#Loads in Area 11
set TieLines;         #Tie lines b/w areas
set TieLinesA1t;      #Tie lines to the area 1
set TieLinesA2t;      #Tie lines to the area 2
```

```
set TieLinesA3t;    #Tie lines to the area 3
set TieLinesA4t;    #Tie lines to the area 4
set TieLinesA5t;    #Tie lines to the area 5
set TieLinesA6t;    #Tie lines to the area 6
set TieLinesA7t;    #Tie lines to the area 7
set TieLinesA8t;    #Tie lines to the area 8
set TieLinesA9t;    #Tie lines to the area 9
set TieLinesA10t;   #Tie lines to the area 10
set TieLinesA11t;   #Tie lines to the area 11
set TieLinesA1f;    #Tie lines from the area 1
set TieLinesA2f;    #Tie lines from the area 2
set TieLinesA3f;    #Tie lines from the area 3
set TieLinesA4f;    #Tie lines from the area 4
set TieLinesA5f;    #Tie lines from the area 5
set TieLinesA6f;    #Tie lines from the area 6
set TieLinesA7f;    #Tie lines from the area 7
set TieLinesA8f;    #Tie lines from the area 8
set TieLinesA9f;    #Tie lines from the area 9
set TieLinesA10f;   #Tie lines from the area 10
set TieLinesA11f;   #Tie lines from the area 11

param GenMax {GEN}; #Generator capacity
param GenMin {GEN}; #Min generation value

param TieLineFlowMax {TieLines};    #Generator capacity
param TieLineFlowMin {TieLines};    #Min generation value

param LoadValue{Loads}; #Value of all the loads in the
   system

param Load >=0; #Total Load in the system
param a {GEN};  #'a' parameter for the cost function
param b {GEN};  #'b' parameter for the cost function
param c {GEN};  #'c  parameter for the cost function

var Generation {GEN} >= 0 ;         # Power from GEN to
   the LOAD
var TieLineFlow {TieLines};
```

```
minimize Gen_Cost: sum { g in GEN}  (a[g]*Generation[g]*
   Generation[g]+ b[g]*Generation[g]+ c[g]) + sum {t in
   TieLines } 0.01*TieLineFlow[t];

subject to GenValue  {g in GEN} :                #
   Generation limit inequality constraint
GenMin[g] <=  Generation[g]  <= GenMax[g];

subject to Area1 :
sum {g in GEN_Area1 } Generation[g]  - sum {t in
   TieLinesA1f } TieLineFlow[t] + sum {t in TieLinesA1t }
   TieLineFlow[t]= sum {l in Load_Area1} LoadValue[l];

subject to Area2 :
sum {g in GEN_Area2 } Generation[g] - sum {t in
   TieLinesA2f } TieLineFlow[t] + sum {t in TieLinesA2t }
   TieLineFlow[t] = sum {l in Load_Area2} LoadValue[l];

subject to Area3 :
sum {g in GEN_Area3 } Generation[g] - sum {t in
   TieLinesA3f } TieLineFlow[t] + sum {t in TieLinesA3t }
   TieLineFlow[t]= sum {l in Load_Area3} LoadValue[l];

subject to Area4 :
sum {g in GEN_Area4 } Generation[g]  - sum {t in
   TieLinesA4f } TieLineFlow[t] + sum {t in TieLinesA4t }
   TieLineFlow[t]= sum {l in Load_Area4} LoadValue[l];

subject to Area5 :
sum {g in GEN_Area5 } Generation[g] - sum {t in
   TieLinesA5f } TieLineFlow[t] + sum {t in TieLinesA5t }
   TieLineFlow[t] = sum {l in Load_Area5} LoadValue[l];

subject to Area6 :
sum {g in GEN_Area6 } Generation[g] - sum {t in
   TieLinesA6f } TieLineFlow[t] + sum {t in TieLinesA6t }
   TieLineFlow[t]= sum {l in Load_Area6} LoadValue[l];

subject to Area7 :
```

```
sum {g in GEN_Area7 } Generation[g] - sum {t in
   TieLinesA7f } TieLineFlow[t] + sum {t in TieLinesA7t }
   TieLineFlow[t]= sum {l in Load_Area7} LoadValue[l];

subject to Area8 :
sum {g in GEN_Area8 } Generation[g] - sum {t in
   TieLinesA8f } TieLineFlow[t] + sum {t in TieLinesA8t }
   TieLineFlow[t]= sum {l in Load_Area8} LoadValue[l];

subject to Area9 :
sum {g in GEN_Area9 } Generation[g] - sum {t in
   TieLinesA9f } TieLineFlow[t] + sum {t in TieLinesA9t }
   TieLineFlow[t]= sum {l in Load_Area9} LoadValue[l];

subject to Area10 :
sum {g in GEN_Area10 } Generation[g] - sum {t in
   TieLinesA10f } TieLineFlow[t] + sum {t in TieLinesA6t }
    TieLineFlow[t]= sum {l in Load_Area10} LoadValue[l];

subject to Area11 :
sum {g in GEN_Area11 } Generation[g] - sum {t in
   TieLinesA11f } TieLineFlow[t] + sum {t in TieLinesA11t
   } TieLineFlow[t]= sum {l in Load_Area11} LoadValue[l];

subject to GenerationBalance:
sum {g in GEN} Generation[g] = sum {l in Loads} LoadValue
   [l];

subject to TieLineFlowConstraint {t in TieLines }:
TieLineFlowMin[t] <= TieLineFlow[t] <= TieLineFlowMax[t];
```

Appendix B

EV Scheduling

1. **EV Scheduling** (MOD File)

```
set Hours   ordered;    #Time slots (24 hours)
set EVs ordered;        #Electric vehicles
set Household   ordered;        #House holds
set Cluster ordered;    #Set of cluster of houses
#set ChargerType;
set EVHousehold {Household} within EVs;
          #EV associated with each consumer
set ClusterHouse {Cluster}      within Household;
          #Consumers in each cluster

param Demand {EVs};                     #Forecasted
   demand of the EV
param DemandHousehold {Household};
                  #Forecasted demand of the household
param SOC_Final {EVs};          #Consumer provided State
   of Charge (SOC)
param Charger {EVs};

param Energy_price {Hours};     #Day-Ahead energy price

param Consumer {EVs, Hours};
                  #Consumer preferred timing for charging

var Schedule { EVs, Hours} >= 0;
                  #Schedule for each EV

# Minimize the cost of charging for all the EVs
```

```
minimize Cost :
sum {j in EVs, k in Hours}  Schedule [j, k]*Energy_price[
   k];

#Consumer preferred timing
subject to Consumer_Preference {j in EVs, k in Hours}:
Schedule[j,k] <= (if Consumer[j,k] = 1  then Demand[j]
   else 0);

#Meeting the EV demand
subject to Equipment_Demand {j in EVs}:

sum {k in Hours} Schedule [j,k] >= SOC_Final[j]*Demand[j
   ];

#EV rating constraint
subject to EV_Rating_Constraint {j in EVs,k in Hours}:
Schedule[j,k] <= ( if  Demand [j] = 0 then 0 else ( if
   Charger[j]=1 then 1.92 else if Charger[j]= 2 then 6.6
   else 8.52));

#Restricting the connected EVs from a single consumer in
   each hour
subject to Household_Capacity {h in Household, k in Hours
   }:
sum {j in EVHousehold[h]} Schedule [j,k] <= 0.25*(
   DemandHousehold[h]) ;

#Restricting the connected EVs from a consumer cluster
subject to Cluster_Capacity {c in Cluster, k in Hours}:
sum {h in ClusterHouse[c], j in EVHousehold[h]} Schedule[
   j,k] <= 0.2*(sum {h in ClusterHouse[c]} DemandHousehold
   [h]);
```

Appendix C

Wind Power Generation Forecasting

```
library(forecast)
library(dplyr)
library(xts)
library(ggplot2)
library(XLConnect)
library(tseries)
library(seasonal)
library(fitdistrplus)
library(lubridate)

# dev.off() # To clear all the previous graphs
#setwd("E:/UND/OneDrive/UND/Research/Data/MN/2010")
setwd("C:/Users/EE User/OneDrive/UND/Research/Data/MN/
    2010")
dataset <- read.csv("MNWind2010.csv")
head(dataset)
dataset$DateTime <- as.POSIXct(dataset$DateTime, format=
    "%m/%d/%y %H:%M")
head(dataset)
tail(dataset)
#Arima Hybrid methods
library(randomForest)
library(kernlab)
library(mgcv)
counter =0
Date_total = data.frame()
Result_total = data.frame()
for(n in 0:18){
  # n <- 11
```

```
counter = counter + 1
# #Training dataset
Date1 <- as.POSIXct("01/30/10 0:0", format="%m/%d/%y %H
   :%M")
Date2 <- as.POSIXct("02/27/10 23:55", format="%m/%d/%y
   %H:%M")
Date3 <- as.POSIXct("02/28/10 0:0", format="%m/%d/%y %H
   :%M")
Date4 <- as.POSIXct("02/28/10 5:55", format="%m/%d/%y %
   H:%M")
Date1 <- Date1+n*hours(1)
Date2 <- Date2+n*hours(1)
Date3 <- Date3+n*hours(1)
Date4 <- Date4+n*hours(1)
pastset <- subset(dataset, dataset$DateTime >= Date1 and
   dataset$DateTime <= Date2)
print(head(pastset))
print(tail(pastset))
nrow(pastset)
forecastset <- subset(dataset, dataset$DateTime >=
   Date3 and dataset$DateTime <= Date4)
print(head(forecastset))
print(tail(forecastset))
nrow(forecastset)
DateTime <- pastset$DateTime
Yts <- as.xts(pastset$power, order.by = DateTime,
   frequency = 288)
YtsPower.1 <- as.xts(pastset$power.1, order.by =
   DateTime, frequency = 288)
YtsDirection.1 <- as.xts(pastset$direction.1, order.by
   = DateTime, frequency = 288)
YtsSpeed.1 <- as.xts(pastset$speed.1, order.by =
   DateTime, frequency = 288)
YtsTemperature.1 <- as.xts(pastset$temperature.1, order
   .by = DateTime, frequency = 288)
YtsPressure.1 <- as.xts(pastset$pressure.1, order.by =
   DateTime, frequency = 288)
YtsDensity.1 <- as.xts(pastset$density.1, order.by =
   DateTime, frequency = 288)
```

```
Xreg <- cbind(YtsPower.1, YtsDirection.1, YtsSpeed.1,
    YtsTemperature.1, YtsPressure.1, YtsDensity.1)
DateTime <- forecastset$DateTime
YtsPower <- as.xts(forecastset$power, order.by =
    DateTime, frequency = 288)
YtsPower.1 <- as.xts(forecastset$power.1, order.by =
    DateTime, frequency = 288)
YtsDirection.1 <- as.xts(forecastset$direction.1, order
    .by = DateTime, frequency = 288)
YtsSpeed.1 <- as.xts(forecastset$speed.1, order.by =
    DateTime, frequency = 288)
YtsTemperature.1 <- as.xts(forecastset$temperature.1,
    order.by = DateTime, frequency = 288)
YtsPressure.1 <- as.xts(forecastset$pressure.1, order.
    by = DateTime, frequency = 288)
YtsDensity.1 <- as.xts(forecastset$density.1, order.by
    = DateTime, frequency = 288)
Zreg <- cbind(YtsPower.1, YtsDirection.1, YtsSpeed.1,
    YtsTemperature.1, YtsPressure.1, YtsDensity.1)
# fit <- auto.arima(Yts, xreg = Xreg,seasonal = FALSE,
    stepwise = FALSE, approximation = FALSE)
fit <- auto.arima(Yts, xreg = Xreg)
InitializedModel<- Arima(Yts,order=arimaorder(fit),xreg
    =Xreg)
Prediction <- forecast(InitializedModel,h=72 ,xreg=Zreg
    )
accuracy <- accuracy(Prediction$mean, forecastset$power
    )
print(accuracy)
mape=mean(abs(forecastset$power - Prediction$mean)/
    forecastset$power)*100
print(mape)
RMSE0 <- accuracy[,2]
MAPE0 <- accuracy[,5]
res <- InitializedModel$residuals
print(arimaorder(fit))
# checkresiduals(res)
# autoplot(InitializedModel)
# print(Box.test(res, type = "Ljung-Box"))
```

```
residuals.ts <- as.numeric(res)
residual.data <- cbind(pastset, residuals.ts)
attach(forecastset)

#Hybrid Arima-SVM
svm.model <- ksvm(residuals.ts ~ power.1+direction.1+
    speed.1+temperature.1+pressure.1+density.1, data =
    residual.data, kernel = "rbfdot")
svm.pred  <- predict(svm.model, newdata=data.frame(
    power.1,direction.1,speed.1,temperature.1,pressure.1,
    density.1, data=forecastset))
wind.forecast2 <- Prediction$mean + svm.pred
accuracy2 <- accuracy(ts(wind.forecast2), forecastset$
    power)
print(accuracy2)
mape2=mean(abs(wind.forecast2 - forecastset$power)/
    forecastset$power)*100
print(mape2)
RMSE2 <- accuracy2[,2]
MAPE2 <- accuracy2[,5]

#Hybrid Arima-RF
rf.model <- randomForest(residuals.ts ~ power.1+
    direction.1+speed.1+temperature.1+pressure.1+density
    .1, data = residual.data, importance = TRUE, ntree
    =500,mtry = 2)
rf.pred  <- predict(rf.model, newdata=data.frame(power
    .1,direction.1,speed.1,temperature.1,pressure.1,
    density.1, data=forecastset))
wind.forecast3 <- Prediction$mean + rf.pred
accuracy3 <- accuracy(ts(wind.forecast3), forecastset$
    power)
print(accuracy3)
mape3=mean(abs(wind.forecast3 - forecastset$power)/
    forecastset$power)*100
print(mape3)
RMSE3 <- accuracy3[,2]
MAPE3 <- accuracy3[,5]
```

```
  # Hybrid Arima-GAM
  gam.model <- gam(residuals.ts ~ s(power.1)+s(direction
      .1)+s(speed.1)+s(temperature.1)+s(pressure.1)+s(
      density.1),data = residual.data)
  gam.pred  <- predict(gam.model, newdata=data.frame(
      power.1,direction.1,speed.1,temperature.1,pressure.1,
      density.1, data=forecastset))
  wind.forecast4 <- Prediction$mean + gam.pred
  accuracy4 <- accuracy(ts(wind.forecast4), forecastset$
      power)
  print(accuracy4)
  mape4=mean(abs(wind.forecast4 - forecastset$power)/
      forecastset$power)*100
  print(mape4)
  RMSE4 <- accuracy4[,2]
  MAPE4 <- accuracy4[,5]

  # Hybrid Arima-GLM
  glm.model <- glm(residuals.ts ~ power.1+direction.1+
      speed.1+temperature.1+pressure.1+density.1,data =
      residual.data)
  glm.pred  <- predict(glm.model, newdata=data.frame(
      power.1,direction.1,speed.1,temperature.1,pressure.1,
      density.1, data=forecastset))
  wind.forecast5 <- Prediction$mean + glm.pred
  accuracy5 <- accuracy(ts(wind.forecast5), forecastset$
      power)
  print(accuracy5)
  mape5=mean(abs(wind.forecast5 - forecastset$power)/
      forecastset$power)*100
  print(mape5)
  RMSE5 <- accuracy5[,2]
  MAPE5 <- accuracy5[,5]
  detach(forecastset)
  ResultTable <- data.frame(counter, Date3, RMSE0, MAPE0,
     RMSE2, MAPE2,RMSE3, MAPE3,RMSE4, MAPE4,RMSE5, MAPE5)
  Result_total <- rbind(Result_total,ResultTable)
}
Result_total
```

```
writeWorksheetToFile("C:/Users/EE User/OneDrive/UND/
    Research/Data/MN/2010/Result/
    ArimaHybridPredictionsResult.xlsx", Result_total, sheet
    = "Feb")

# GLM Forecasting

counter =0
Date_total = data.frame()
Result_total = data.frame()
for(n in 0:18){
  # n <- 0
  counter = counter + 1
  # #Training dataset
  Date1 <- as.POSIXct("01/30/10 0:0", format="%m/%d/%y %H
     :%M")
  Date2 <- as.POSIXct("02/27/10 23:55", format="%m/%d/%y
     %H:%M")
  Date3 <- as.POSIXct("02/28/10 0:0", format="%m/%d/%y %H
     :%M")
  Date4 <- as.POSIXct("02/28/10 5:55", format="%m/%d/%y %
     H:%M")
  Date1 <- Date1+n*hours(1)
  Date2 <- Date2+n*hours(1)
  Date3 <- Date3+n*hours(1)
  Date4 <- Date4+n*hours(1)
  pastset <- subset(dataset, dataset$DateTime >= Date1 and
     dataset$DateTime <= Date2)
  print(head(pastset))
  print(tail(pastset))
  nrow(pastset)
  forecastset <- subset(dataset, dataset$DateTime >=
     Date3 and dataset$DateTime <= Date4)
  print(head(forecastset))
  print(tail(forecastset))
  nrow(forecastset)
  DateTime <- pastset$DateTime
  attach(forecastset)
```

```
glm.model <- glm(power ~ power.1+direction.1+speed.1+
    temperature.1+pressure.1+density.1,data = pastset)
# print(summary(glm.model))
glm.pred  <- predict(glm.model, newdata=data.frame(
    power.1,direction.1,speed.1,temperature.1,pressure.1,
    density.1, data=forecastset))

accuracy <- accuracy(glm.pred, forecastset$power)
print(accuracy)
mape=mean(abs(glm.pred - forecastset$power)/forecastset
    $power)*100
print(mape)
RMSE0 <- accuracy[,2]
MAPE0 <- accuracy[,5]
   res <- glm.model$residuals
   residuals.ts <- as.numeric(res)
   residual.data <- cbind(pastset, residuals.ts)

   #Hybrid GLM-SVM
   svm.model <- ksvm(residuals.ts ~ power.1+direction.1+
      speed.1+temperature.1+pressure.1+density.1, data =
      residual.data, kernel = "rbfdot")
   svm.pred  <- predict(svm.model, newdata=data.frame(
      power.1,direction.1,speed.1,temperature.1,pressure
      .1,density.1, data=forecastset))
   wind.forecast2 <- glm.pred + svm.pred
   accuracy2 <- accuracy(ts(wind.forecast2), forecastset
      $power)
   print(accuracy2)
   mape2=mean(abs(wind.forecast2 - forecastset$power)/
      forecastset$power)*100
   print(mape2)
   RMSE2 <- accuracy2[,2]
   MAPE2 <- accuracy2[,5]

   #Hybrid GLM-RF
   rf.model <- randomForest(residuals.ts ~ power.1+
      direction.1+speed.1+temperature.1+pressure.1+
      density.1, data = residual.data, importance = TRUE,
```

```
        ntree=500,mtry = 2)
    rf.pred <- predict(rf.model, newdata=data.frame(
        power.1,direction.1,speed.1,temperature.1,pressure
        .1,density.1, data=forecastset))
    wind.forecast3 <- glm.pred + rf.pred
    accuracy3 <- accuracy(ts(wind.forecast3), forecastset
        $power)
    mape3=mean(abs(wind.forecast3 - forecastset$power)/
        forecastset$power)*100
    RMSE3 <- accuracy3[,2]
    MAPE3 <- accuracy3[,5]
    detach(forecastset)
    ResultTable <- data.frame(counter, Date3, RMSE0,
        MAPE0,RMSE2, MAPE2, RMSE3, MAPE3)
    Result_total <- rbind(Result_total,ResultTable)
}
Result_total
writeWorksheetToFile("C:/Users/EE User/OneDrive/UND/
    Research/Data/MN/2010/Result/GLMHybridResult.xlsx",
    Result_total, sheet = "Feb")

# GLM Hybrid method Forecasting

counter =0
Date_total = data.frame()
Result_total = data.frame()
for(n in 0:18){
  # n <- 0
  counter = counter + 1
  # #Training dataset
  Date1 <- as.POSIXct("01/30/10 0:0", format="%m/%d/%y %H
      :%M")
  Date2 <- as.POSIXct("02/27/10 23:55", format="%m/%d/%y
      %H:%M")
  # Forecast day January 3st 1-6 hours (Winter)
  Date3 <- as.POSIXct("02/28/10 0:0", format="%m/%d/%y %H
      :%M")
  Date4 <- as.POSIXct("02/28/10 5:55", format="%m/%d/%y %
      H:%M")
```

```
Date1 <- Date1+n*hours(1)
Date2 <- Date2+n*hours(1)
Date3 <- Date3+n*hours(1)
Date4 <- Date4+n*hours(1)
pastset <- subset(dataset, dataset$DateTime >= Date1 and
    dataset$DateTime <= Date2)
forecastset <- subset(dataset, dataset$DateTime >=
    Date3 and dataset$DateTime <= Date4)
DateTime <- pastset$DateTime
attach(forecastset)
glm.model <- glm(power ~ power.1+direction.1+speed.1+
    temperature.1+pressure.1+density.1,data = pastset)
# print(summary(glm.model))
glm.pred <- predict(glm.model, newdata=data.frame(
    power.1,direction.1,speed.1,temperature.1,pressure.1,
    density.1, data=forecastset))
accuracy <- accuracy(glm.pred, forecastset$power)
print(accuracy)
mape=mean(abs(glm.pred - forecastset$power)/forecastset
    $power)*100
print(mape)
RMSE0 <- accuracy[,2]
MAPE0 <- accuracy[,5]
res <- glm.model$residuals
residuals.ts <- as.numeric(res)
residual.data <- cbind(pastset, residuals.ts)

#Hybrid GLM-SVM
svm.model <- ksvm(residuals.ts ~ power.1+direction.1+
    speed.1+temperature.1+pressure.1+density.1, data =
    residual.data, kernel = "rbfdot")
svm.pred <- predict(svm.model, newdata=data.frame(
    power.1,direction.1,speed.1,temperature.1,pressure.1,
    density.1, data=forecastset))
wind.forecast2 <- glm.pred + svm.pred
accuracy2 <- accuracy(ts(wind.forecast2), forecastset$
    power)
print(accuracy2)
```

```
  mape2=mean(abs(wind.forecast2 - forecastset$power)/
     forecastset$power)*100
  print(mape2)
  RMSE2 <- accuracy2[,2]
  MAPE2 <- accuracy2[,5]

  #Hybrid GLM-RF
  rf.model <- randomForest(residuals.ts ~ power.1+
     direction.1+speed.1+temperature.1+pressure.1+density
     .1, data = residual.data, importance = TRUE, ntree
     =500,mtry = 2)
  rf.pred <- predict(rf.model, newdata=data.frame(power
     .1,direction.1,speed.1,temperature.1,pressure.1,
     density.1, data=forecastset))
  wind.forecast3 <- glm.pred + rf.pred
  accuracy3 <- accuracy(ts(wind.forecast3), forecastset$
     power)
  print(accuracy3)
  mape3=mean(abs(wind.forecast3 - forecastset$power)/
     forecastset$power)*100
  print(mape3)
  RMSE3 <- accuracy3[,2]
  MAPE3 <- accuracy3[,5]
  detach(forecastset)
  ResultTable <- data.frame(counter, Date3, RMSE0, MAPE0,
     RMSE2, MAPE2, RMSE3, MAPE3)
  Result_total <- rbind(Result_total,ResultTable)
}
Result_total
writeWorksheetToFile("C:/Users/EE_User/OneDrive/UND/
   Research/Data/MN/2010/Result/GLMHybridResult.xlsx",
   Result_total, sheet = "Feb")
```