

©Copyright 2017

Yishen Wang

Energy Storage Operation with Wind Uncertainty

Yishen Wang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Prof. Daniel S. Kirschen, Chair

Prof. Miguel A. Ortega-Vazquez

Prof. Baosen Zhang

Program Authorized to Offer Degree:
Department of Electrical Engineering

University of Washington

Abstract

Energy Storage Operation with Wind Uncertainty

Yishen Wang

Chair of the Supervisory Committee:
Close Professor Prof. Daniel S. Kirschen
Department of Electrical Engineering

With environmentally oriented energy policies, wind power penetration sharply increases in the recent years. Although operating cost is low, wind power's intrinsic uncertainty and variability have brought another layer of operating challenges for the power system operators and system participants. Since traditional operation is not specially designed for these type of characteristics, it motivates us to revise the current operating frameworks to better embrace a more sustainable generation mix.

Other than wind energy, energy storage systems (ESS) are another key players in this smart grid era. They provide additional operational flexibility to enhance the whole grid operation and planning. Recently, decreasing installation and operation costs for ESS furthers drives the interests to develop advanced operation decision making processes for grid and market applications. Therefore, in this dissertation, our work focuses to understand how to better integrate renewable generation and energy storage with respect to wind uncertainty.

We first improve wind modeling with ensemble approach. The proposed data mining concept provides not only better point forecast but also scenario set and uncertainty region for stochastic optimization. To reduce the computation burden, we compare the scenario reduction techniques and propose a submodular scenario reduction algorithm to further improve the computational efficiency. By optimally selecting representative scenarios, we reduce the stochastic programming solution time without sacrificing solution quality.

In terms of ESS, we first stand from the system operator's perspective and analyze the benefits of ESS operation in the centralized scheduling with uncertain wind. Stochastic programming based operation strategies well balance the trade-off between the system economics and reliability with the assistance from energy storage. Moreover, we also stand from the merchant storage owner's perspective to derive bidding strategies to increase profits through participating the energy market. A look-ahead strategy fully explores the arbitrage opportunities across the time and space.

TABLE OF CONTENTS

	Page
List of Figures	iii
List of Tables	v
Glossary	vii
Chapter 1: Introduction	1
1.1 Operation Changes and Opportunities	1
1.2 Organization	4
Chapter 2: Scenario Generation	6
2.1 Motivation and Literature Review	6
2.2 Time-series Based Approach	12
2.3 Ensemble-based Approach	22
2.4 Summary	36
Chapter 3: Scenario Reduction	38
3.1 Motivation	38
3.2 Comparison of State-of-art Techniques	42
3.3 Scenario Reduction with Submodular Optmization	58
3.4 Summary	66
Chapter 4: Energy Storage Operation in a Centralized ISO Environment	68
4.1 Motivation	68
4.2 Analysis of the Benefits of Energy Storage in UC	70
4.3 Summary	90

Chapter 5: Energy Storage Operation From Merchant Owner Perspective	91
5.1 Motivation	91
5.2 Impact of Local Transmission Congestion on Energy Storage Arbitrage Opportunities	93
5.3 Look-Ahead Bidding Strategy for Energy Storage	102
5.4 Summary	129
Chapter 6: Conclusions and Suggestion for the Future Works	131
6.1 Conclusions	131
6.2 Suggestion for the Future Work	132
Bibliography	133
Appendix A: Nomenclature	156
A.1 Chapter 3 and Chapter 4	156
A.2 Chapter 5	160
Appendix B: Diagrams of IEEE RTS Test Systems	164
B.1 One-Area RTS	164
B.2 Three-Area RTS	165
Appendix C: Author's Bibliography	166
Appendix D: Vita	169

LIST OF FIGURES

Figure Number	Page
2.1 Probabilistic forecasts with Gradient Boosting Machines (GBM)	12
2.2 Flowchart for time series scenario generation	13
2.3 Wind speed power curve example	15
2.4 Wind speed seasonal mean	16
2.5 Wind speed monthly mean	17
2.6 Nataf transformation	19
2.7 Sample ACF and PACF function for Wind Farm 1	20
2.8 Wind power scenarios with ARMA	23
2.9 Flowchart for ensemble scenario generation	25
2.10 Correlation coefficients among wind farms	27
2.11 Correlation histogram among different wind farms	32
2.12 Wind scenarios with ensemble approach	33
2.13 Comparison 1 of different algorithms	34
2.14 Comparison 2 of different algorithms	35
3.1 Schematic representation of a scenario set in the (a) scenario-based with five scenarios, and (b) interval stochastic UC with three scenarios. The gray lines represent scenarios and the blue lines represent the deterministic constraints on ramping [103].	41
3.2 Wind profiles with positive and negative correlation with load	47
3.3 Comparison of the range of uncertainty obtained for different techniques . .	49
3.4 Scenario selection incremental benefit	64
4.1 Comparison of the AOC cost distribution at Day 1	85
4.2 Comparison of the AOC cost distribution at Day 2	86
4.3 Comparison of the AOC cost distribution at Day 3	87
4.4 Comparison of the AOC cost distribution at Day 4	88
4.5 Comparison of the AOC cost distribution at Day 5	89

5.1	Typical patterns of annual storage profits.	99
5.2	Annual profits at various buses under a 50% reduction in line transmission capacity.	100
5.3	Locational marginal price duration curve with a 50% line capacity reduction.	101
5.4	LMP at bus 14 on days 100–102 with a 50% line capacity reduction.	102
5.5	Histogram of daily storage profits at bus 19.	103
5.6	An illustration of the proposed bilevel model and the interfaces between the upper- and lower-level problems.	106
5.7	Time-frame for market-clearing on days D_0 and D_1	106
5.8	Wind and load profiles for the test week	116
5.9	Storage profits with different γ^{D_1} and γ^{CVaR}	119
5.10	Profits for different D+1 forecasting error levels for $\gamma^{D_1} = 0.75$	120
5.11	Storage 4 Offer in Day 4	121
5.12	Storage profits with different look-ahead window lengths	122
5.13	Storage revenue and profits with typical cost numbers	124
5.14	Storage profits with competing generators offers	125
5.15	Storage fSoC level with different system conditions	127
5.16	Storage profits for different ramp limits with normal transmission capacities	128
5.17	Profits under different transmission capacities with normal ramp limits	129
B.1	One-Area RTS Diagram	164
B.2	Three-Area RTS Diagram	165

LIST OF TABLES

Table Number		Page
2.1	Comparison of Algorithms	37
3.1	Day Ahead Cost (In 10^3 \$) of the Scenario-Based schedule with positively correlated wind and load	50
3.2	Day Ahead Cost (In 10^3 \$) of the Scenario-Based Schedule with negatively correlated wind and load	50
3.3	Key statistics for the Day-Ahead Interval UC schedules with positively correlated wind and load	51
3.4	Key statistics for the Day-Ahead Interval UC schedules with negatively correlated wind and load	51
3.5	Computational time (s) for the Scenario-based SUC schedules with positively correlated wind and load	52
3.6	Computational time (s) for the Scenario-based SUC schedules with negatively correlated wind and load	52
3.7	Actual Operating Cost (In 10^3 \$) of the Scenario-based SUC schedules with positively correlated wind and load	54
3.8	Actual Operating Cost (In 10^3 \$) of the Scenario-based SUC schedules with negatively correlated wind and load	55
3.9	Actual Operating Cost (In 10^3 \$) of the Interval UC schedules with positively correlated wind and load	56
3.10	Actual Operating Cost (In 10^3 \$) of the Interval UC schedules with negatively correlated wind and load	57
3.11	Solution time comparison	62
3.12	Number of scenarios selected by both FFS and ccSSR	63
3.13	Selection results in terms of FFS objectives	63
3.14	Selection results in terms of SSR objectives $\lambda = 0.5$	65
3.15	Selection results in terms of SSR objectives $\lambda = 2.5$	65
3.16	Selection results in terms of SSR objectives $\lambda = 10$	65

4.1	Representative day wind generation	75
4.2	Energy storage installation in the system	75
4.3	Day-Ahead Cost (DAC) in \$ 10^6 for Day 1 with 10 % wind generation	76
4.4	Day-Ahead Cost (DAC) in \$ 10^6 for Day 2 with 16 % wind generation	77
4.5	Day-Ahead Cost (DAC) in \$ 10^6 for Day 3 with 26 % wind generation	77
4.6	Day-Ahead Cost (DAC) in \$ 10^6 for Day 4 with 35 % wind generation	78
4.7	Day-Ahead Cost (DAC) in \$ 10^6 for Day 5 with 45 % wind generation	78
4.8	Day-Ahead computing time (seconds) for Day 1 with 10 % wind generation .	79
4.9	Day-Ahead computing time (seconds) for Day 2 with 16 % wind generation .	80
4.10	Day-Ahead computing time (seconds) for Day 3 with 26 % wind generation .	80
4.11	Day-Ahead computing time (seconds) for Day 4 with 35 % wind generation .	80
4.12	Day-Ahead computing time (seconds) for Day 5 with 45 % wind generation .	81
4.13	Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 1 with 10 % wind generation	82
4.14	Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 2 with 16 % wind generation	82
4.15	Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 3 with 26 % wind generation	83
4.16	Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 4 with 35 % wind generation	83
4.17	Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 5 with 45 % wind generation	84
4.18	Expected daily wind spillage (MWh) for Day 5 with 45 % wind generation .	90
5.1	Parameters of the storage devices	116
5.2	Final state of charge (MWh) and profits (\$) without transmission and ramp rate constraints	118
5.3	Total profits (\$) with different system conditions	126

GLOSSARY

AR: Autoregressive model.

ARIMA: Autoregressive integrated moving average model.

ARMA: Autoregressive–moving-average model.

AOC: Actual operating costs.

ANN: Artificial neural networks.

BESS: Battery energy storage systems.

BPA: Bonneville Power Administration.

CAISO: California Independent System Operator.

CAES: Compressed Air Energy Storage.

CCSSR: Cardinality-constrained submodular scenario reduction.

CDF: Cumulative distribution function.

CUSSR: Cardinality-unconstrained submodular scenario reduction.

DA: Day-ahead.

DAC: Day-ahead costs.

DUC: Deterministic unit commitment.

ED: Economic dispatch.

EENS: Expected energy not served.

EPEC: Equilibrium problems with equilibrium constraints.

ERCOT: Electric Reliability Council of Texas.

ESS: Energy storage systems.

EV: Electric vehicles.

FACTS: Flexible AC transmission system.

FFS: Fast forward selection.

FSOC: Final state-of-charge.

GBM: Gradient boosting machines.

GIS: Geographical Information System.

IS: Importance sampling.

ISO: Independent system operator.

IUC: Interval unit commitment.

LL: Lower level.

LMP: Locational marginal prices.

LP: Linear programming.

KNN: K-Nearest-Neighbors.

MA: Moving-average model.

MILP: Mixed-integer linear programming.

MC: Monte-Carlo.

MISO: Midcontinent Independent System Operator.

MPEC: Mathematical programming with equilibrium constraints.

NN: Neural networks.

NREL: National Renewable Energy Laboratory.

NWP: Numerical weather predictions.

OPF: Optimal power flow.

PHES: Pumped hydro energy storage.

PJM: Pennsylvania-New Jersey-Maryland Interconnection.

PMU: Phasor measurement units.

PNNL: Pacific Northwestern National Laboratory.

RF: Random forests.

RT: Real-time.

RTS: Reliability Test System.

SBR: Simultaneous backward reduction.

SP: Stochastic programming.

SSR: Submodular scenario reduction.

SUC: Stochastic unit commitment.

SVM: Support vector machines.

RUC: Robust unit commitment.

UC: Unit commitment.

UL: Upper level.

ACKNOWLEDGMENTS

It has been an amazing journey so far as I started working on the research topics in this dissertation. First and foremost, I would like to express my utmost appreciation to my advisor and mentor Prof. Daniel S. Kirschen for his consistent support and guidance in the past few years. It is always a privilege for me to have such a knowledgeable, inspiring, supportive and patient advisor to walk me through this intriguing power systems research world. More importantly, Prof. Kirschen serves as an exemplary role model for my career. This dissertation would not have been possible without him.

My gratitude goes to my reading committee members, Prof. Miguel A. Ortega-Vazquez and Prof. Baosen Zhang for their insightful feedback and suggestions along my research and teaching assistant periods. It is always fruitful to discuss with these sharp researchers. I would also like to thank Prof. Archis Ghate for serving the GSR committee and providing many constructive suggestions.

I would like to recognize my internship mentors in the past. It has been a great pleasure working with Dr. Audun Botterud and Dr. Zhi Zhou from Argonne National Laboratory, and Dr. Hongbo Sun from Mitsubishi Electric Research Laboratories. For being great mentors and friends, it has been valuable experiences to work with such experts in power systems.

I would like to thank Dr. Kevin Schneider from PNNL for being such an excellent instructor. I learned a lot for being his TA during the distribution system classes. It is always delightful to talk with him as well.

I must acknowledge the friends and collaborators from Renewable Energy Analysis Lab (REALab): Prof. Yury Dvorkin, Dr. Ahlmahz Negash, Ting Qiu, Prof. Hrvoje Pandžić, Marco Au, Ryan Elliot, Zeyu Wang, Dr. Mushfiq Sarker, Yushi Tan, Jesus Elmer Contreras

Ocaña, Bolun Xu, Dr. Ricardo Fernández-Blanco, Daniel Olsen, Ahmad Milyani, Abeer Almainouni, Yao Long, Agustina Gonzalez, Kelly Kozdras, Rebecca Breiding, Dr. Pierre Henneaux, Dr. Tu Nguyen, Dr. Remy Rigo-Mariani, Pan Li, Yuanyuan Shi, Dr. Hao Wang, Yize Chen, Chenghui Tang, Dr. Weiling Zhang, Dr. Yangwu Shen, Dr. Yang Wang, Yao Chang, Jingkun Liu, Yi Wang, Prof. Yunfeng Wen, Dr. Tobias Haring, Leonardo H. Macedo. It is rewarding to work with such a talented group.

I would also like to thank my friends outside the lab: Yuzong Liu, Tong Zhang, Xiang Chen, Zheng Li, Xi Cheng, Yue Yang, Tiangxing Cheng, Jincheng Che, Hao Wu, Ruizhi Sun, Ning Ning, Dian Li, Ce Zhang, Xudong Li, Chuanjia Xing, Jingda Wu, De Meng, Fang Sun, Zeran Wang, Chun Liang, Hewei Li, Di Li, Xiaodong Chen, Billy Tubbs, Yi Luan, Ruiyi Li, Minwen Weng, Zhongming Ye, Steven Xu, Hongdi Fu, Xinwei Fan, Ye Li, Tong Zhang, Zhenyu Tan, Meng Huang, Bo Zhu, Ying Gan, Guoxiang Ma, Pu He, Rui Yao, Bo Li, Zhengshuo Li, Zhigang Li, Jiahui Guo, Junjian Zhao, Cheng Wang, Runze Chen, Ti Xu, Zhiwei Xu, Miao Fan, Yutian Zhou, Lijun He, Zhi Chen, Wei Tao, Tengyue Zhang, Chengyao Zhao, Chen Yang, Zhihao Zhang, Chun Hou, Zhaoying Xu, Keyuan Jin, Peng Chen, Bowen Zhang, Guangyu Feng, Jian Zhang and Chieh-Chi Kao. It is nice to share these beautiful memories with you.

I'm also grateful to all the UWEE faculties and staffs who have helped me in the past few years.

My special thanks to uncle Ken Chang and aunt Li for treating me like a family member and all the help these years. It is warm to have such a nice host family when I first arrived at the Seattle.

Last but definitely not least, I would like to thank my father Huanliang Wang and my mother Hong Shen for all the unconditional love and support. I could not finish this dissertation without them. Let alone the encouragement from them to study electrical engineering and continue a PhD study. This dissertation is dedicated to all of you.

DEDICATION

to My Family

Chapter 1

INTRODUCTION

1.1 Operation Changes and Opportunities

Historically, reliably serving the electricity demand has been the fundamental task for power system operation. With various generation resources, system operators also try to run the system at minimum cost level, and always balance reliability and economics during operation [1]. Thermal power plants contributed the majority of electrical power generation for nearly 100 years [2]. As the pollution caused by fossil fuels became a concern, the classical minimum cost generation scheduling problem is being revisited. Sustainability requirements drive the development of renewable generation technologies, such as wind, solar, geothermal, and biomass. Although these technologies do bring us more environmental benefits, there are also difficult operation and policy challenges. Since wind generation is one of the most widely accepted renewable generation, in this dissertation we will focus on this specific type of renewable generation technology. However, since wind and solar generation share some features like uncertainty modeling, our conclusions are not limited to wind generation and can be generalized to solar as well.

Traditionally, the major uncertainty source for the transmission system has been the contingencies caused by generator or transmission line outages. However, with a sharply increasing penetration of wind power, its intrinsic uncertainty and variability have brought another layer of operating challenges for the independent system operators (ISO) [3]. We need to distinguish the concepts of uncertainty and variability. Uncertainty refers to the unpredictability: since a forecast is never perfect, the actual wind realization can never be exactly known in advance. On the other hand, variability refers to the fluctuations of wind power within a certain time period. Since traditional operation is not specially designed for

these type of questions, it motivates us to revise the current operating frameworks to better embrace a more environmental-friendly generation mix. Numerous work have contributed to a better understanding of how to integrate renewable generation. These efforts can be divided into three directions, improved wind modeling [4, 5, 6, 7], advanced operation decision making processes [3, 8, 9, 10, 11, 12, 13] and the application of energy storage systems (ESS) in the grid [14, 15, 16, 17, 18, 19].

First, researchers have developed advanced wind modeling and forecasting algorithms to better predict wind generation. Many of these prediction algorithms have been successfully deployed in the field. These methods include neural networks (NN), support vector machines (SVM), multiple linear regressions, time series, random forest, gradient boosting tree [4]. These algorithms provide a point-forecast *i.e.* the best single-value possible result of wind realization. Recently, probabilistic forecasting [5] has become more popular due to its ability to provide the full distribution of potential generation. Scenario generation is another widely accepted approach that provides the input needed by stochastic programming methods [6]. Nevertheless, how to best utilize the results from forecasting services still requires a well-designed operation framework. In this dissertation, we will present two schemes to generate wind profiles to assist decision-making processes.

Second, traditional deterministic generation scheduling is not adequate to fully capture the uncertainty on renewable generation, and more advanced formulations have been proposed to improve the economics and reliability of system operation. In deterministic formulations, reserve is co-optimized with unit commitment decisions. Less scheduled reserve lowers the operating cost but may affect reliability. This may also result in higher real-time re-dispatch cost to compensate for wind forecast errors. If the scheduled reserve is larger than the amount that is actually needed, the day-ahead scheduling is too conservative and the cost is unnecessarily high. Ortega-Vazquez [3] proposes to co-optimize the unit schedule and spinning reserve amount assuming that the wind generation follows a certain parametric distribution. Also following the idea of improving the reserve quantification, Zhou *et al.* [9] schedule a dynamic reserve to hedge against wind shortage from probabilistic wind forecast

results and an operating reserve demand curve. For the deterministic formulation, the major wind profile used is still a deterministic wind forecast.

Due to forecast errors, a single trajectory is not enough to describe the potential evolution of wind generation, scenarios and scenarios tree are used for stochastic programming formulations [11]. Wind scenarios are mainly used for a two-stage formulation [10], and a tree structure can be further applied to a multi-stage formulation [12]. Scenario reduction [13] is widely adopted to reduce the problem dimension. However, due to physical computational resource constraints, the number of scenarios used in the formulation is always limited, and some extreme cases could be overlooked.

To prevent neglecting extreme scenarios, robust optimization is used to immunize against worst case scenarios [20]. It ensures system feasibility under the worst case as well as minimizes such worst-case cost. Since the worst case scenario can be hard to identify, Wang *et al.* [21] proposed an interval optimization approach to define extreme scenarios as the ramping requirements among the net load upper bound, central forecast and lower bound. Dvorkin [22] suggested a hybrid approach to combine stochastic formulation and interval formulation together for a better computational and economic performance trade-off. Pandžić [23] modified the original interval approach to use more realistic, data-driven ramping constraints. All proposed formulations try to reduce the system cost and maintain system reliability within an affordable computational time. In this dissertation, we will focus on the stochastic unit commitment approach and describe a scenario reduction technique to limit the problem size.

Power systems must always maintain a real-time balance between the generation and load. Pumped hydro is one option for storing surplus energy and inject it back into the grid later. However, due to the limitations of geography, pumped hydro is not a universal solution for this power balance requirements. Thanks to advances in material science and chemical engineering, battery energy storage systems (BESS) technologies are becoming viable options for system operation [15].

ESS can participate multiple services in the system, such as energy arbitrage [14], reserve market [24], peak shaving, frequency support [16], voltage support, congestion management

as well as contingency support [17]. Generic storage and pumped hydro have been successfully incorporated in stochastic and robust formulations of the unit commitment problem [18], [19]. However, there are still several questions to be answered, such as the cost-benefit analysis of ESS among different UC formulations, the ESS setpoint ranges to ensure maximum system reliability as well as minimum cost, and the ESS profitability under a pool market environment. In this dissertation, we will investigate the applications of ESS from the system operators' and the storage aggregator's perspectives separately to determine the best operation strategies. BESS applications are natural extensions for our work after properly accounting for the degradation costs.

1.2 Organization

The rest of this dissertation is organized as follows.

- Chapter 2 consists of three sections. Section 2.1 summarizes the objectives of constructing renewable generation scenarios and provides a literature review of state-of-the-art approaches. Section 2.2 presents the time-series based ARMA scenario generation method. Section 2.3 describes the ensemble based scenario generation approach.
- Chapter 3 includes three sections. Section 3.1 describes current power system operation formulations with different renewable uncertainty models. The necessity and motivation for scenario reduction is also discussed. Section 3.2 compares different state-of-the-art scenario reduction techniques applied to stochastic unit commitment and interval unit commitment. Section 3.3 describes the proposed submodular scenario reduction technique.
- Chapter 4 includes two sections. Section 4.1 illustrates the motivation for energy storage applications in a centralized ISO environment. Section 4.2 evaluates the effect of energy storage systems in the day-ahead market and real-time adjustments with different UC formulations.

- Chapter 5 includes four sections. Section 5.1 discusses the importance of analyzing energy storage profitability from the storage owner's perspective. Section 5.2 discusses a bilevel formulation to optimize the energy storage bidding and offering price-quantity pair with local network congestion. Section 5.3 extends this work and presents a look-ahead risk-constrained bidding strategy to better allocate stored energy for arbitrage opportunities.
- Chapter 6 concludes the whole dissertation works and suggests some future research directions.

Chapter 2

SCENARIO GENERATION

Increasing penetration of renewable energy adds more uncertainty to the operation of the grid. Since wind energy is the largest source of renewable energy, wind integration needs to be characterized properly. To handle this issue, stochastic optimization is an adequate tool but it requires accurate probabilistic profiles. Monte Carlo Simulation is the traditional way to generate these scenarios, but it is not detailed enough to capture the nature of the source of uncertainty. Moreover, the geographic correlation between wind sites also affects the result of the scenarios. In this section, we propose a time-series approach and a data-mining approach to consider the spatial and temporal correlations to generate wind scenarios.

2.1 Motivation and Literature Review

2.1.1 Motivation

With increasing penetration of uncertain source in power systems, renewable generation, storage devices, electric vehicles and other uncertain components take a more important role in the system operation. Traditional operation is based on a deterministic optimization procedure, which does not explicitly quantify uncertainties [2]. The operators manage the markets and schedules hourly productions according to the results from the unit commitment (UC), economic dispatch (ED) and optimal power flow (OPF). In the classical model, the load is deterministic and given, and the generation dispatch is computed considering all the technology and operation constraints. Unlike thermal units, wind generators are usually not considered as fully dispatchable resources. Instead, we treat it as a negative part of the load using a forecast of wind generation. In this sense, the wind generation is not uncertain and unknown. With this trick, we can then use a solver to handle the complex optimization

problem. The operators can then determine their generation schedule based on the result from the optimization.

However, with stochastic random variables as wind output, this treatment cannot reflect all possible wind realizations, and must be revised. Introducing dynamic reserve moves a step forward for considering uncertainty within a deterministic formulation. However, the amount of reserve is either pre-defined which is generally conservative, or co-optimized which again requires a well-calibrated uncertainty formulation.

Since the short-term load forecast is generally very accurate compared to the renewable forecast [20], we will treat it as a deterministic input and not as a random variable. The results from a wind point forecast is a deterministic wind time series, which indicates the exact wind output at a given time period. However, as an intermittent source, the wind has an intrinsic stochastic nature. It might behave as the forecast suggests, or it might not. To capture its true uncertainty and variability, we should consider all the possibilities, not only the most probable one, but also some of the rare ones. Then, we can establish an upper and lower prediction interval bound for all the scenarios.

Each wind generation scenario represents an operation case. Applying proper algorithms, we can solve the optimization problem and get a cost-effective generation schedule. Therefore, it is important to build these scenarios carefully to ensure their accuracy, otherwise the results will not be optimal. Stochastic optimization should thus be considered for the whole operation processes [25].

Since stochastic optimization is a general optimization term, different models require different features for the input data. Stochastic programming methods (SP) aims at minimizing the expected objective value from a scenario set that contains most possible situations. Generally, SP models the uncertainty with a two-stage or a multi-stage formulation [25]. For a multi-stage formulation, a tree-structured scenario set is desired to capture all the uncertainty dependencies among multiple stages with non-anticipativity constraints. A two-stage formulation is a special case which optimizes “here-and-now” decisions (first-stage decisions) and “wait-and-see” decisions (second stage decisions) with a given scenario set without ex-

plicitly demanding a tree-structure or fan-structure. Similar to the two-stage SP methods, a chance-constrained formulation also requires a large number of scenario samples to model the associated random variables.

Recently, robust optimization [26] and interval optimization [27] have been proposed to solve the stochastic optimization problem. For these two methods, an uncertainty set, rather than a scenario set, is prerequisite. As discussed in [28], the uncertainty set can have different forms, and the most common one is a polyhedral set. This polyhedral uncertainty set has been widely used in unit commitment problems [20], [19]. Since the confidence interval can be efficiently extracted from the scenario set, generating valid probability scenario profiles is therefore essential to the accuracy of solving this optimization problem.

In addition, it is well known that wind is not static and changes differently in different areas. The scale of these changes is affected by the distance. Thus, if we want to describe wind at one location, we should consider the influence of other nearby or faraway regions. The spatial correlations and how they affect our wind scenarios must therefore be analyzed carefully. To be more specific, what we focus on here is not a single wind farm or a single region. Our target is the wind output scenarios in different places considering the spatial correlation with other wind farms.

An adequate representation of the random wind data is thus needed for any SP formulation. It should reflect the underlying probability assumptions and the existing historic data, as well as connected to the purpose of the application, which often requires a compromise between the problem size and the accuracy of the results.

2.1.2 Literature Review

Wind scenario generation techniques can be divided into wind modeling approaches and wind forecast approaches. Both approaches can be further divided into modeling (forecasting) wind power directly, or first modeling (forecasting) wind speed and then converting wind speed to wind power.

For wind modeling approaches, many algorithms have been proposed. Hoyland [29] sug-

gests a moment matching technique which generates synthetic scenarios with respect to historic wind statistics. A nonlinear optimization problem is formulated to minimize the Euclidean distance between the true realizations and expected value of synthetic scenarios. Hoyland [30] also presents a method to construct single- and multi-period scenario tree. As part of the WILMAR project, Lowery [31] further proposes a moment-matching technique to generate a scenario tree with defined scenario mean, variance, skewness and kurtosis for a rolling stochastic unit commitment problem.

Time series models are also widely used for wind scenario generation. Morales *et al.* [32] propose to generate spatially correlated wind scenarios using an Auto Regressive and Moving Average (ARMA) method. Wind speeds are transformed into a Gaussian distributed space with Nataf transformation for better Gaussian-shaped. A univariate ARMA model is trained for each wind farm. A covariance matrix for multiple wind farms training residuals is estimated, and a Cholesky decomposition is then applied to generate spatially correlated errors for sampling by the ARMA models. The ARMA model generates the desired number of scenarios in this Gaussian feature space. Then an inverse Nataf transformation is applied to reverse these pre-processing steps to construct wind speed scenarios. A wind power turbine curve is used later to produce wind power scenarios. Papavasiliou [6], [10] generates wind scenarios with a similar approach.

Billinton *et al.* [33] propose to sequentially sample wind speed data with an ARMA model. Similarly, Karki *et al.* [34] extend this technique by fitting a common wind speed model for all wind farms, then each farm provides individually mean and variance statistics to sample multiple scenarios. Stuart [35] propose to use AR(2) to model a scenario tree for a stochastic unit commitment. An ARIMA model has also been applied. Miranda *et al.* [36] use a vector AR model to capture the spatial correlation of the wind.

Markovian Chain Monte Carlo (MCMC) has also been applied in this field. Papaefthymiou [37] uses a Markov chain model to characterize the temporal dependencies with a higher order transition matrix. Then Monte Carlo sampling is applied to generate multiple scenarios. Some authors [38] improve the method by using a copula to model the stochastic depen-

dency of wind power uncertainty. Zhang [39] also proposes to model the spatial-temporal wind data with a copula. Inspired by this modeling framework, Pinson *et al.* [40] and Papaefthymiou *et al.* [41] generate wind scenarios from a temporal covariance matrix and a spatial-temporal covariance matrix with a probabilistic forecast input. Similarly, Ma *et al.* [42] sample scenarios with an exponential covariance function.

Another common yet simplified approach would be to fit a certain parametric distribution over wind speed or wind power. Then, Monte Carlo sample are generated over this distribution to construct multiple scenarios [43]. However, temporal correlations might not be preserved with such an approach.

As an ensemble version of wind forecast, wind scenarios can be constructed with wind forecasting techniques [44]. Therefore, many wind forecasting techniques can be applied to this context. These forecasting methods are generally categorized into three groups, the physical approach, the statistical approach and the intelligent system approach [4].

For the physical approach, the most state-of-art technique is numerical weather prediction (NWP) [7]. A NWP model uses weather information from a Geographical Information System (GIS). Partial differential equations are used to characterize weather motions. Wind speed data can be extracted from the NWP solutions. By perturbing the initial conditions, an NWP forecast ensemble is generated, which can serve as the scenario set [45, 46]. Although the computation is relatively expensive and complicated, we can directly download NWP results from publicly available NWP datasets. The European wind forecast project ANEMOS [4] and two American wind integration project EWITS [47] and WWSIS [48] used NWP as their forecasting tool. Gaussian Processes have been proposed as input to NWP to produce wind ensemble [49].

Many intelligent algorithms have also been suggested to generate wind scenarios. The most common one is the artificial neural network (ANN) [50]. As ANN training is a non-convex optimization, ANN can produce different results with different initial conditions. ANNs are thus able to produce multiple scenario trajectories. ANNs have been widely used in wind power forecasting as well due to their strong ability to model the nonlinearities

among between the input features and the output [51, 52, 53]. Likewise, particle swarm optimization [54] and fuzzy logic have both been implemented to predict wind. A lack of mathematical proof for these methods makes these algorithms more like a black-box solver. The performance is highly dependent on the engineering design and tuning of the problem.

Statistical techniques thus seem to be the more general. Since time series models have been introduced previously [33, 34, 55, 56], we omitted this discussion here. As wind prediction is essentially a regression problem, numerous successful machine learning algorithms can be applied in this field. Historic wind power, wind speed and other external features such as temperatures, humidity, and date and time can be used as input feature to train these algorithms. These algorithms include, but are not limited to, support vector machine (SVM) [57], random forest [58], gradient boosted tree [59], K-nearest-neighbors (KNN) [60].

With these algorithms, we can generate wind scenarios in two ways. One, we perturb the initial conditions. For example, since we only use a subset of the input features to train the model, we can generate multiple scenarios by different feature randomization. Second, we can combine the results of different forecasting algorithms to obtain scenarios. We will explore these two approaches in section 2.3.

The previous discussion is mainly based on point forecasting. Recently, probabilistic forecasting has attracted a lot of attention both from academia and industry [5, 61]. Rather than producing single or multiple trajectories, the full distribution over every time step is computed. This information gives a better understanding of wind intervals. Quantile regression is one of the most applicable algorithms [44, 62, 63]. In addition, kernel density estimation also performs well in this task [64, 65]. More algorithms and techniques are expected in this field. After obtaining the distribution, the covariance matrix or copula can be used to obtain the scenarios as in [40].

Figure 2.1 shows a probabilistic forecasting prediction results with gradient boosting method. The training and testing data is from Global Energy Forecasting Competition 2014 wind track [61].

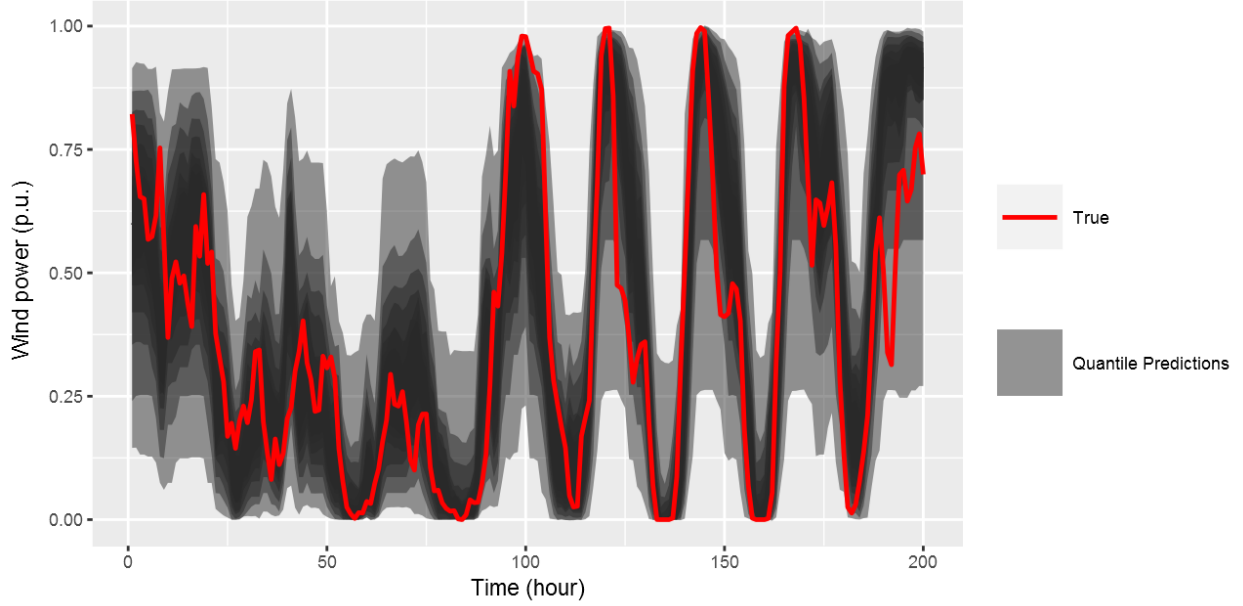


Figure 2.1: Probabilistic forecasts with Gradient Boosting Machines (GBM)

2.2 Time-series Based Approach

In this section, we introduce an efficient time series based scenario generation method. Time series model can capture the temporal and spatial correlation in short-term wind modeling quite well, and therefore it is suitable for this task [4]. Similarly to methods presented in [6, 32], our method captures the stochastic wind dependencies using a covariance matrix. Figure 2.2 provides a detailed flowchart of this approach. Wind data is collected from the NREL Western Wind Dataset [48].

The whole process can be divided into three groups, pre-processing, scenario generation and inverse transformation. In the pre-processing, we first compute the probabilistic speed power curve. Then, wind speed data is normalized and transformed into stationary time series for the ARMA model. Next, in the scenario generation, these series are used to estimate the model parameters, and then to generate time series scenarios repeatedly based on spatial-correlated random errors. Finally, the time series scenarios are inversely transformed into

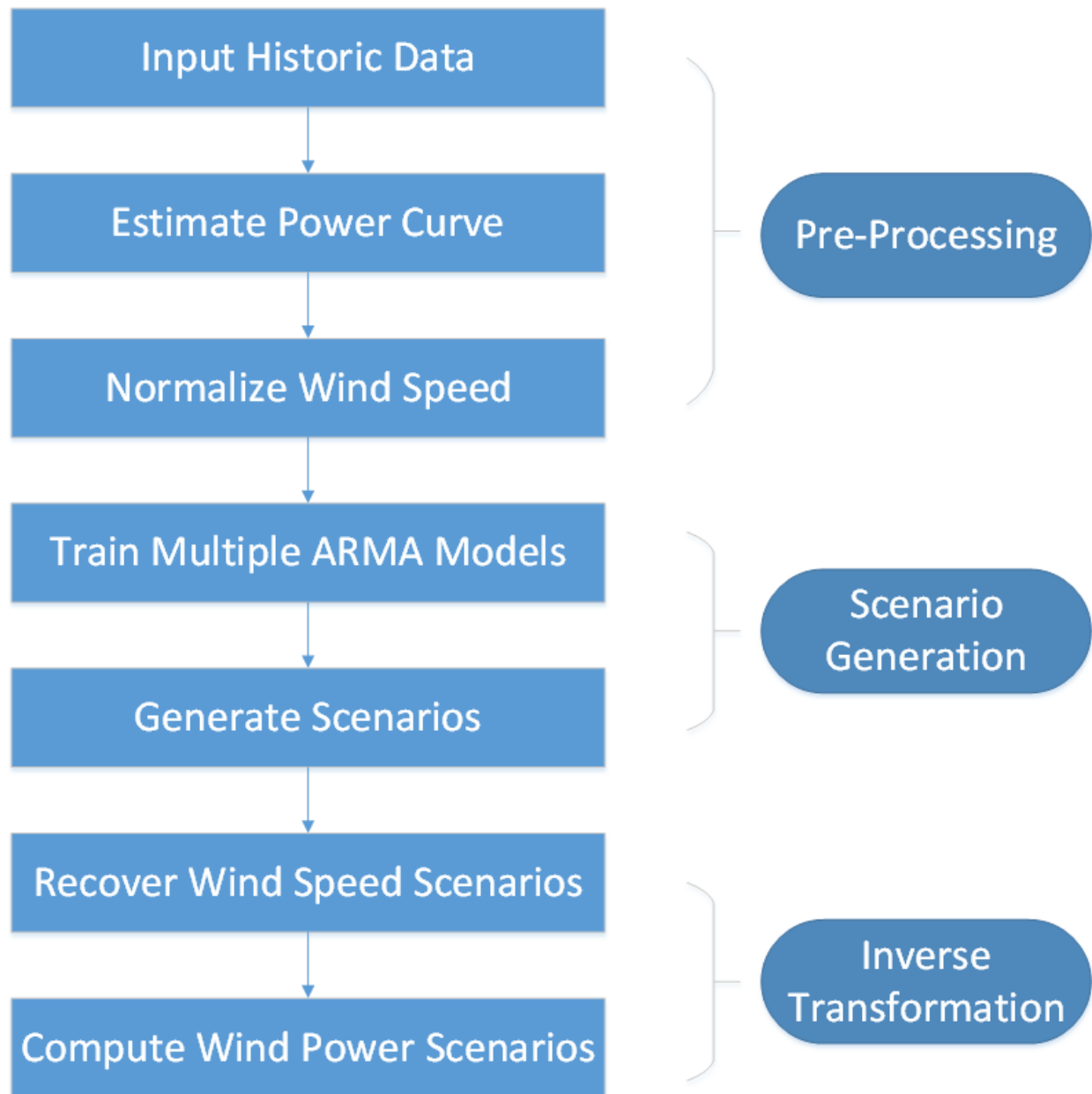


Figure 2.2: Flowchart for time series scenario generation

the wind power data which reverses the pre-processing steps.

2.2.1 Input Historic Data

In order to produce wind power scenarios, we first collect wind speed and wind power data from the NREL Western dataset [48]. This data is a synthetic dataset developed by 3Tier for the NREL renewable integration study. In the dataset, wind power and wind speed in 5-min resolution are provided from 2004 to 2006 for every 30 MW location, *i.e.* a total of 32,043 locations. We group multiple locations together to obtain 19 aggregated wind farms. Then, we average wind speed and power values to generate data with an hourly resolution. Years 2004 and 2005 are used as historic input, and year 2006 is used to update the parameters as we can recast the whole process with a rolling moving window.

2.2.2 Estimate Power Curve

Wind speed data is more suitable than wind power data for time series models [4]. Therefore, we will base our modeling on wind speed data. In order to have a relatively accurate wind power model, we estimate the speed power curve with a least square fit. In addition, we compute the upper and lower bounds to allow flexibility to model more conservative or aggressive wind positions. The bound can be determined with confidence interval. Normally, we choose 99.5 % and 0.5 % to cover 99 % of the data points. A power curve example can be seen in Figure 2.3.

2.2.3 Normalize Wind Speed

Since ARMA model requires a stationary and normally distributed time series, we need to normalize the wind speed data before training the model. A two-step normalization is applied here.

First, we use a mean-variance normalization to de-trend the diurnal component from the data. The daily mean and variance of wind speed are calculated for each season and each

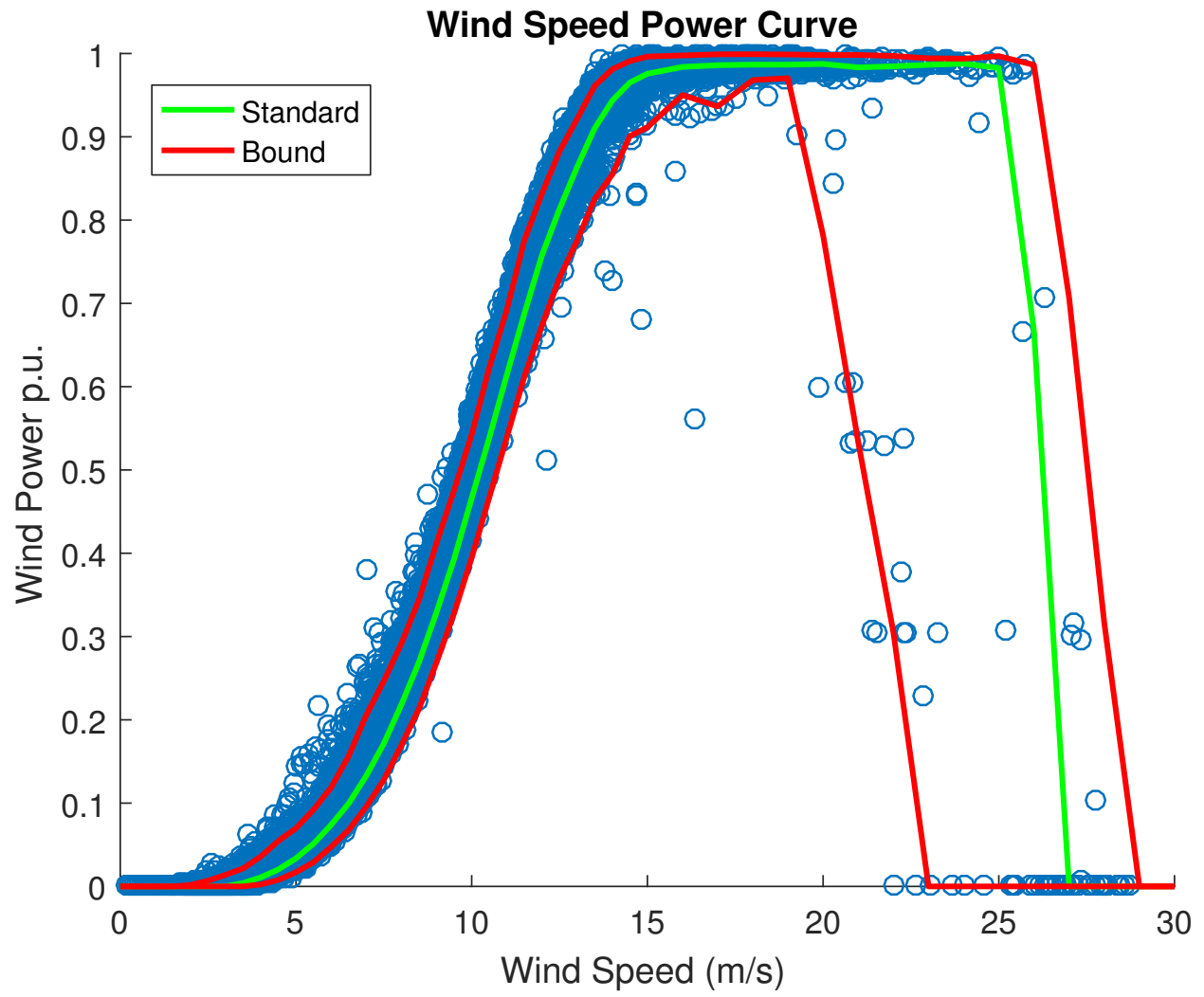


Figure 2.3: Wind speed power curve example

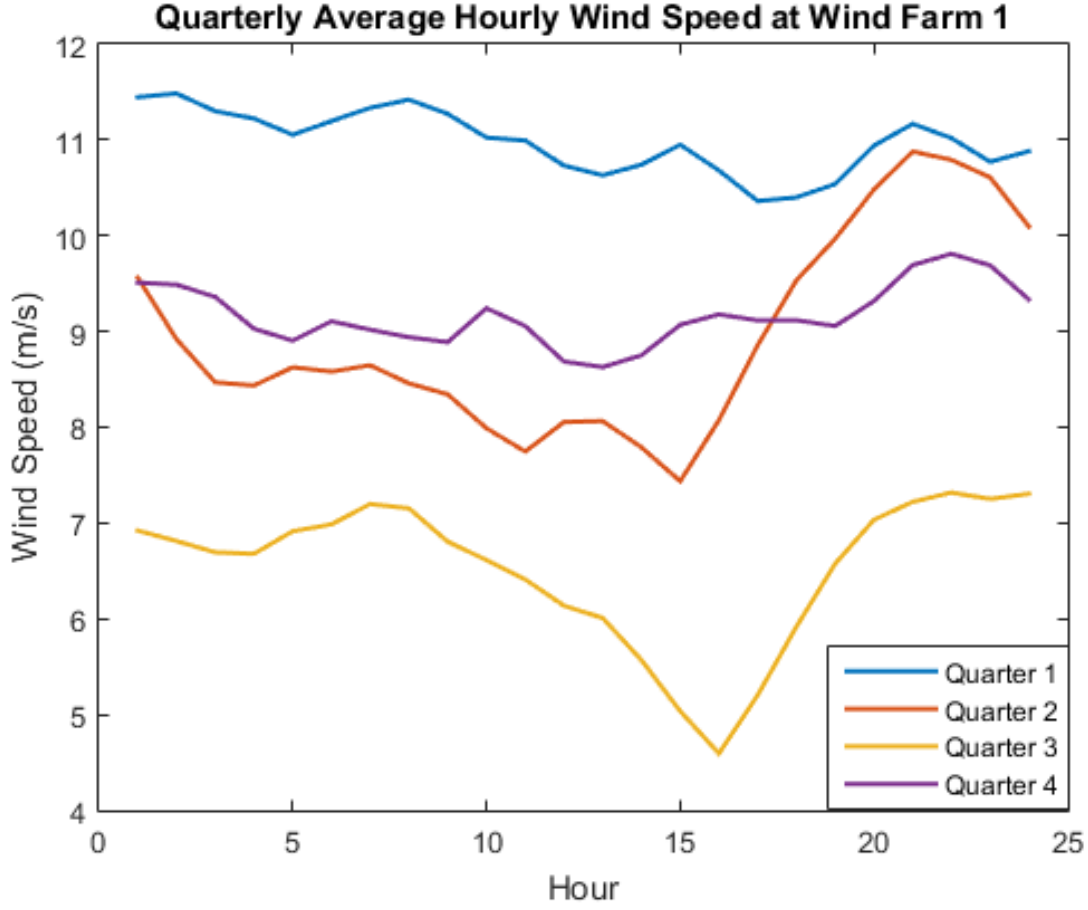


Figure 2.4: Wind speed seasonal mean

hour. In Equation (2.1), w and t denote wind farm w at hour t in season m . ws is the wind speed, and μ , σ are the mean and standard deviation of the wind speed. dws is the de-trended new series.

$$dws_{w,t,m} = \frac{ws_{w,t,m} - \mu_{i,t,m}^{ws}}{\sigma_{i,t,m}^{ws}} \quad (2.1)$$

The number of seasons can be 12 for monthly or 4 for quarterly. Strong quarterly seasonal variations are shown in Figure 2.4. The distinct seasonal wind speed patterns are very clear.

Figure 2.5 shows monthly wind speed variations. The differences are more subtle though

still with a pattern. Quarterly mean can better represent the general trend in the season, whereas the monthly mean can better capture the detailed variations. In our simulation, the total number of season is chosen as 12, which is monthly de-trended. This selection enables us to remove the monthly trend for a better normalization effect. We can also use other clustering algorithms like K -Means to cluster seasons; however, this clustering does not provide significant improvements over the standard definition of months.

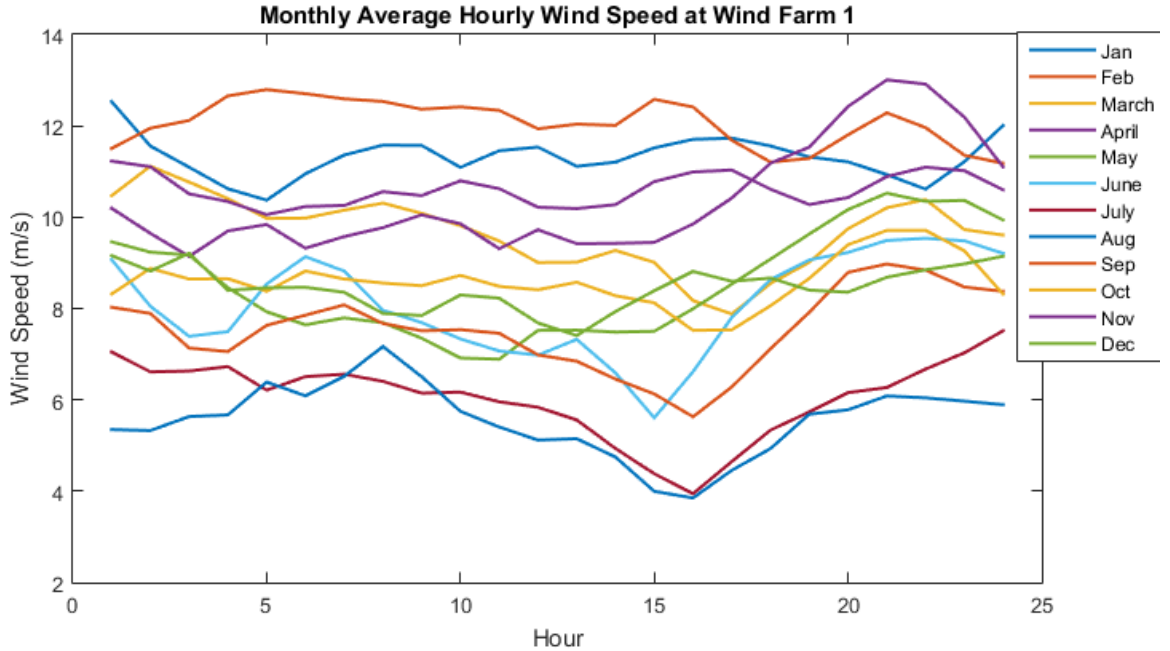


Figure 2.5: Wind speed monthly mean

After seasonal de-trending, we transform the current distribution to the normal distribution because time series models require the series to follow a normal distribution. A Nataf transformation [32], equivalent to normal copula [66], is thus necessary. If we had an infinitely large dataset, we could model the data for each season separately. However, due to the limited size of the data, we no longer distinguish different seasons.

Let $C_w(x)$ represents the empirical cumulative distribution function (CDF) for the de-trended wind data $ts_{w,t}$, $N(x)$ the CDF of the normal distribution, and $M_w(x)$ the Nataf

transformation. $C_w^{-1}(x)$, $N^{-1}(x)$ and $M_w^{-1}(x)$ represents the inverse of these distributions and transformations. Then, we have following equations for this transformation:

$$M_w(x) = N^{-1}(C_w(x)) \quad (2.2)$$

$$M_w^{-1}(x) = C_w^{-1}(N^{-1}(x)) \quad (2.3)$$

$$ts_{w,t} = M(dw s_{w,t}) \quad (2.4)$$

With this transformation, all the series are changed into a more normally distributed series ts . After de-trending and normalization, the time series satisfies the basic Box and Jenkins conditions [67]. We have completed the first pre-processing stage with our data.

2.2.4 Train Multiple ARMA Models

The following multivariate ARMA model is used to represent the time series $ts_{w,t}$ after Nataf transformation. The noise term ϵ has a zero-mean with covariance matrix G .

$$ts_{w,t} = \sum_{j=1}^p \alpha_{w,j} ts_{w,t-j} + \epsilon_{w,t} + \sum_{k=1}^q \beta_{w,k} \epsilon_{w,t-k} \quad (2.5)$$

Since the full multivariate model can be hard to train, we adopt the simplification suggested in [32] to decouple Equation (2.5) into a number of univariate ARMA models as in Equation (2.6).

$$ts_{w,t} = \sum_{j=1}^p \alpha_j^w ts_{w,t-j} + \epsilon_t^w + \sum_{k=1}^q \beta_k^w \epsilon_{t-k}^w \quad (2.6)$$

To retain spatial correlations among wind farms, G needs to be computed and used to generate cross-correlated random errors. For properly estimating from noise residuals, we first need to fit the ARMA models. Motivated by ensemble methods, multiple ARMA model can be tested here. As presented in Figure 2.7, AR (2) could be a good model choice, and

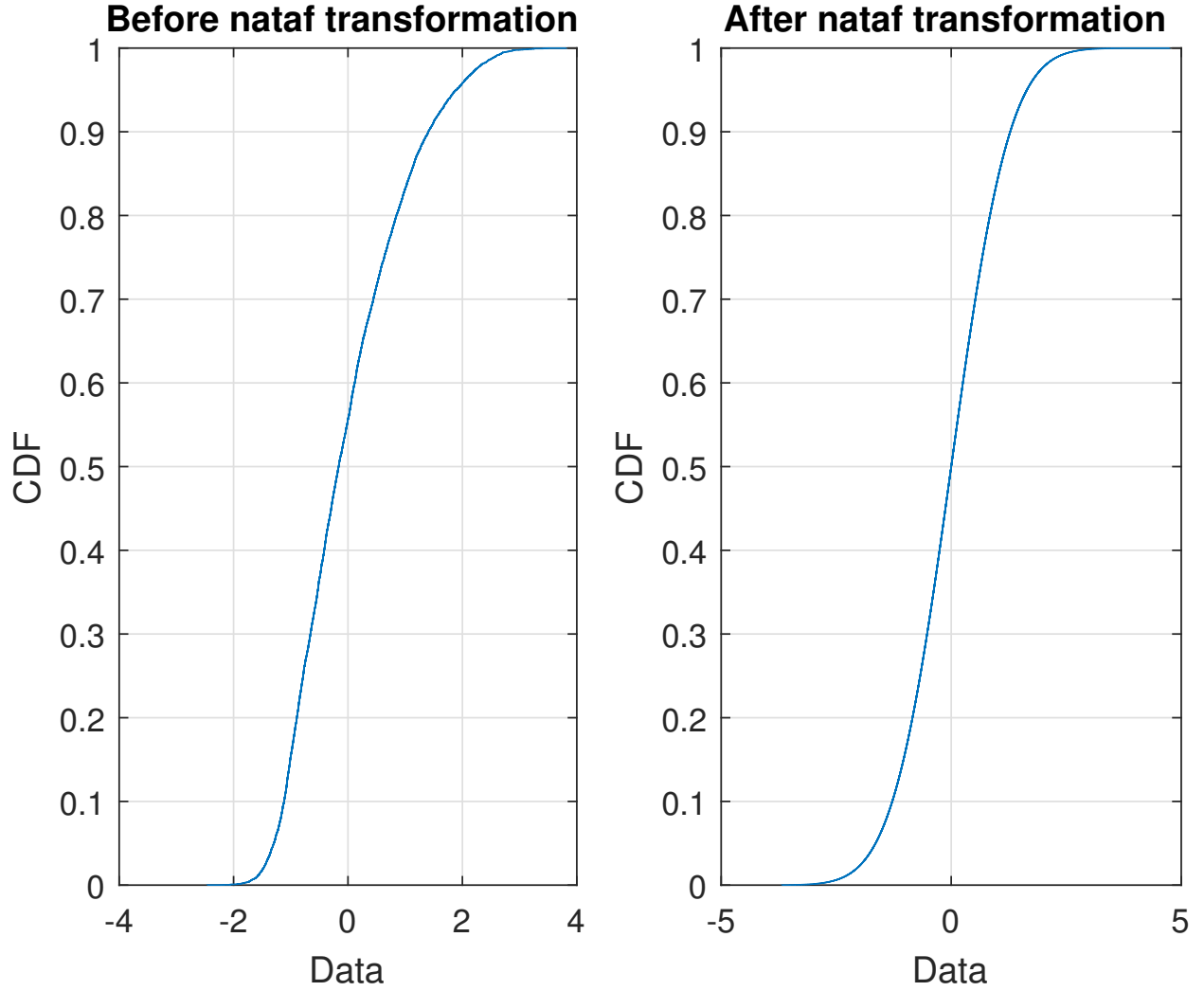


Figure 2.6: Nataf transformation

ARMA (2, 3) can also be explored. Here, we can choose the best two or three models rather than a single one. Akaike information criteria (AIC) and Bayesian information criteria (BIC) are used to select representative models. Readers can refer to [67] for more details.

After fitting ARMA models for each individual wind farms, we can obtain and calculate the residual covariance matrix G . This G matrix should be symmetric or even semidefinite. Then, we compute the orthogonal transformation L by Cholesky decomposition. By varying

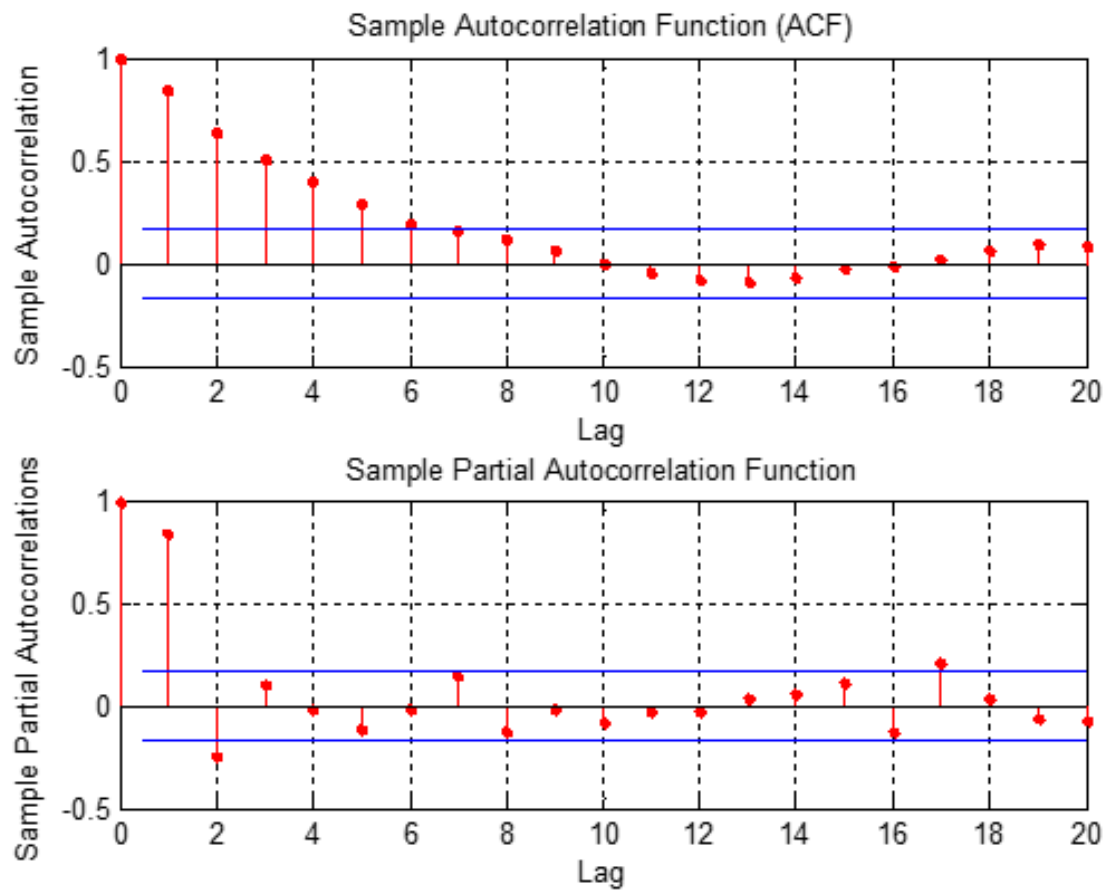


Figure 2.7: Sample ACF and PACF function for Wind Farm 1

ARMA models, we can obtain different G and corresponding L . L is later used to generate cross-correlated random errors.

$$G = LL^T \quad (2.7)$$

2.2.5 Generate Scenarios

We use the following steps to generate scenarios:

Step 1: Generate $N_w \times N_t$ independent standard normally distributed white noise ξ as the error component. N_w is the total number of the wind sites and N_t is the total number of time periods.

Step 2: Apply orthogonal transformation L to produce cross-correlated random errors ζ as in (2.8).

Step 3: Estimate time series with trained ARMA model and cross-correlated random errors ζ as in (2.9).

Step 4: Repeat Step 1 – 3 until desired number of scenarios is reached.

$$\zeta = L\xi \quad (2.8)$$

$$ts_{w,t} = \sum_{j=1}^p \alpha_j^w ts_{w,t-j} + \zeta_t^w + \sum_{k=1}^q \beta_k^w \zeta_{t-k}^w \quad (2.9)$$

2.2.6 Recover Wind Speed Scenarios

By reversing the pre-processing steps, we can recover wind speed scenarios from generated time series scenarios $ts_{w,t}$. Inverse Nataf transformation and mean-variance transformation are applied as in (2.10) and (2.11).

$$dws_{w,t} = M_w^{-1}(ts_{w,t}) \quad (2.10)$$

$$ws_{w,t,m} = dws_{w,t,m}\sigma_{i,t,m}^{ws} + \mu_{i,t,m}^{ws} \quad (2.11)$$

2.2.7 Compute Wind Power Scenarios

Using the fitted wind speed power curve as discussed in Section 2.2.2, we obtain the wind power scenarios as desired. Figure 2.8 provides an example of generated wind scenarios and shows that the model captures the associated uncertainty. The spatial-temporal correlation can be clearly observed at wind farm 1, 10, 11. Wind farm 19 has different wind shapes, and the model can preserve it reasonably well. Due to the way randomness is modeled in the ARMA model, and the fact that three steps are required to transform the normally distributed numbers into wind power, additional errors are introduced for the wind power scenarios. This might lead to spikes being over-estimated. This may cause issues as the related scenarios are quite wide-spread, and sometimes the ARMA model tends to over-estimate large ramps.

2.3 Ensemble-based Approach

Although the wind speed ARMA model provides a well-calibrated wind scenario uncertainty set, they are limited to the very short-term wind series modeling. In addition, the nonlinear relationships among the wind power time series, wind speed time series and other explanatory variables is not been fully explored.

In NWP, a weather prediction ensemble is generated by perturbing the initial starting conditions, which results in a collection of weather scenarios representing different conditions to represent weather uncertainty [7]. Likewise, Breiman [68] proposed his famous Random Forest algorithm to generate a strong predictor by combining several weak predictors through bootstrap sampling and random feature selection.

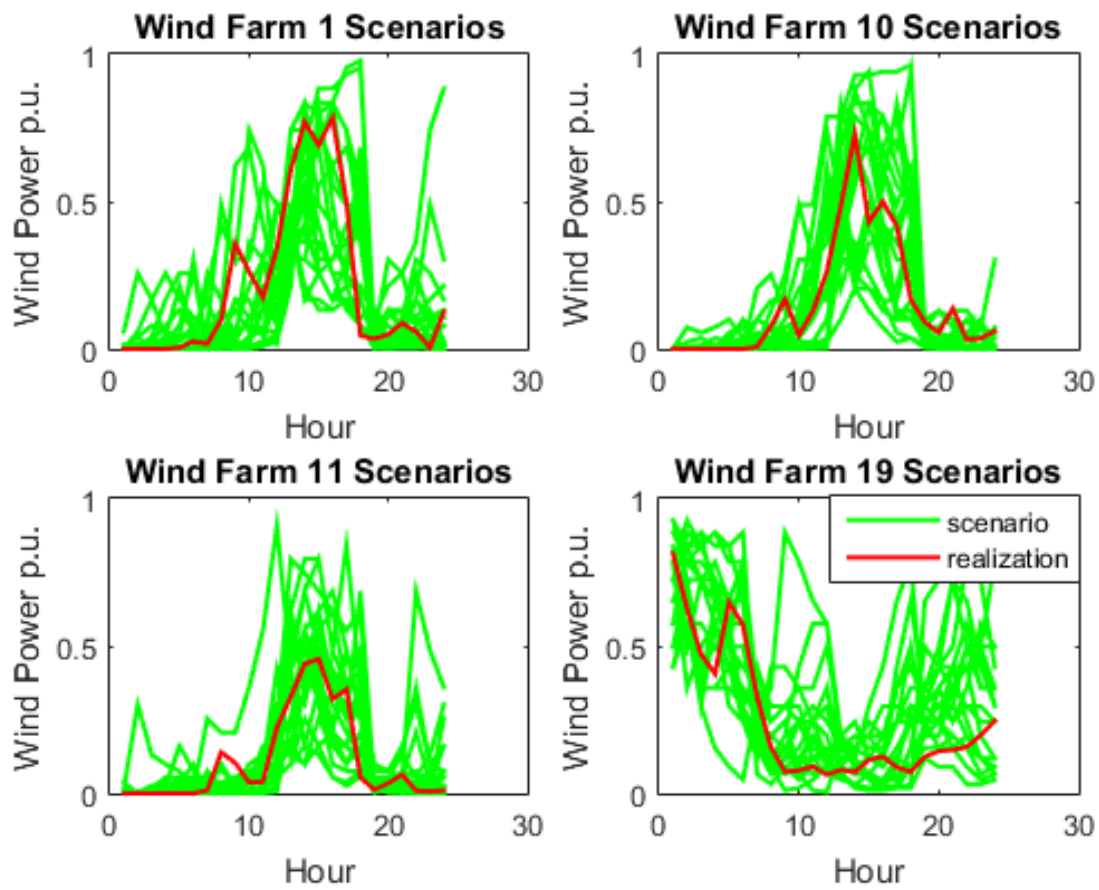


Figure 2.8: Wind power scenarios with ARMA

Motivated by these works, we present in this subsection an ensemble-based approach to directly generate wind power scenarios. ARMA models belong to the wind modeling category, while the proposed ensemble approach belongs to wind forecast approach. Since the forecasting approach is essentially the regression task from the machine learning field, several well-known regression algorithms are applied here to generate scenarios, including support vector machine (SVM), artificial neural network (ANN), bagging and Random Forest (RF). As our ensemble approach can be separated into different modules, future advanced algorithms could be added as a natural extension. The flowchart is presented in Figure 2.9.

2.3.1 *Input Historic Data*

As in Section 2.2.1, the NREL Western Dataset is used to construct the scenario set. In the dataset, the wind power series and the wind speed series are provided to feed the machine learning algorithms as input features.

2.3.2 *Design Spatial-Temporal Features*

Wind power series are firstly normalized based on its nameplate capacity. Every wind power is therefore in per unit. Since the target variable is wind power, it is not necessary to over-complicate the processing for wind speed. Therefore, we use mean-variance normalization as introduced in Section 2.2.3. After we normalize these values, we can construct the input feature set for every prediction target or response variable.

A time-lag h^{lead} has to be pre-defined, which indicates how many hours ahead is the prediction algorithm. A smaller h^{lead} leads to a more accurate modeling, whereas a larger h^{lead} results in larger deviation errors. Since the NREL Western Dataset does not include wind forecast information, for the sake of a better calibration effect, we set $h^{lead} = 1$ or $h^{lead} = 6$ depending on the level of uncertainty we want to reach. If forecast data is provided, h^{lead} can be set to a large value (such as 24 or 48) to truly perform a short-term forecast with the proposed ensemble approach. In our simulation, we mainly model wind power behavior

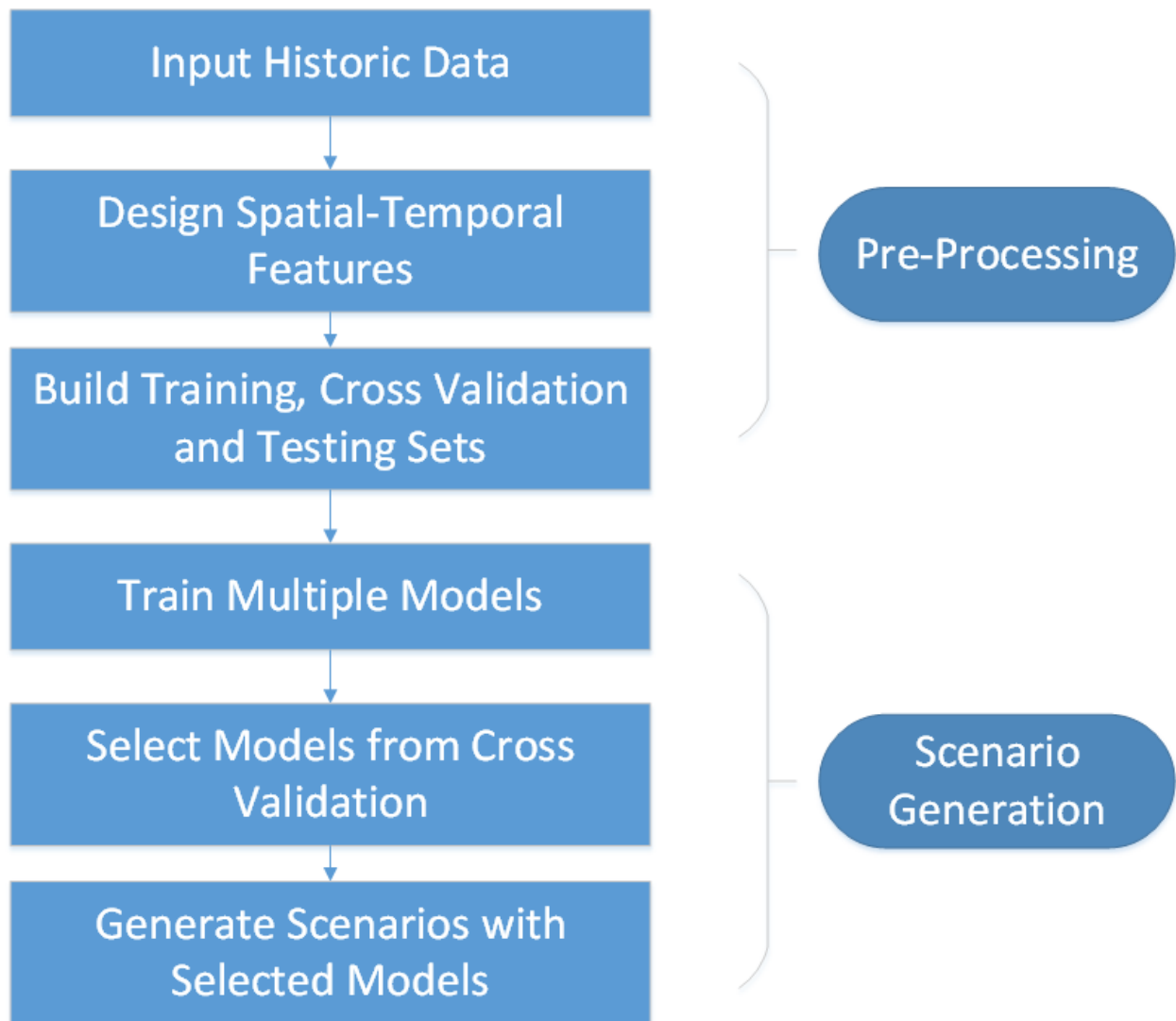


Figure 2.9: Flowchart for ensemble scenario generation

with such an approach. It is a middle ground between pure wind series modeling and pure wind forecast.

After defining the time lag h^{lead} , the temporal features from each site can be determined by choosing N_{hist}^{local} , number of local historical hours, to include the effect of recent information [69]. As wind patterns are relatively volatile, it is not necessary to set N_{hist}^{local} too large. On the other hand, to fully capture wind uncertainty, this value should not be too small either. Therefore, we choose 7 days or 168 hours as a compromise. Experiments show that this number is adequate to represent the wind.

As previously mentioned, it is also crucial to consider the spatial correlation. Therefore, additional spatial features are introduced as well. Ideally, the wind movement along the whole area is related, so we need to take all other wind farm information into consideration. In this simulation, 19 wind farms are included. If all other 18 farms are considered for $N_{hist}^{near} = 168h$, we introduce 3024 additional features. Increasing the size of the feature set increases the training time without necessarily introducing significant modeling improvements.

Figure 2.10 presents graphical illustrations of these correlation coefficients. The more red area indicates stronger correlations whereas the more blue area suggests weaker correlations. This data shows that wind farm 1 is not very correlated with the other wind farms because all correlation coefficients are below 0.7. On the contrary, wind farms 12 – 16 and 19 are quite correlated with all coefficients over 0.85. These observations suggest to reduce the spatial feature range. For example, include only the 4 nearest sites information for spatial correlation with $N^{near} = 4$. This filtering step greatly reduces the size of the feature set and reduces the computational time required for the later training stage.

Both wind speed and wind power at nearby sites can be used for feature construction. As wind power spatial features already provide necessary information, the wind speed spatial features are not included. We set $N_{hist}^{near} = 4$ to further limit the feature size. As mentioned above, it is not always good to have large feature set as feature redundancy reduce computational performances without necessarily improving accuracy. Since feature selection is not the focus here, interested readers can refer to [70] and [71] for more details.

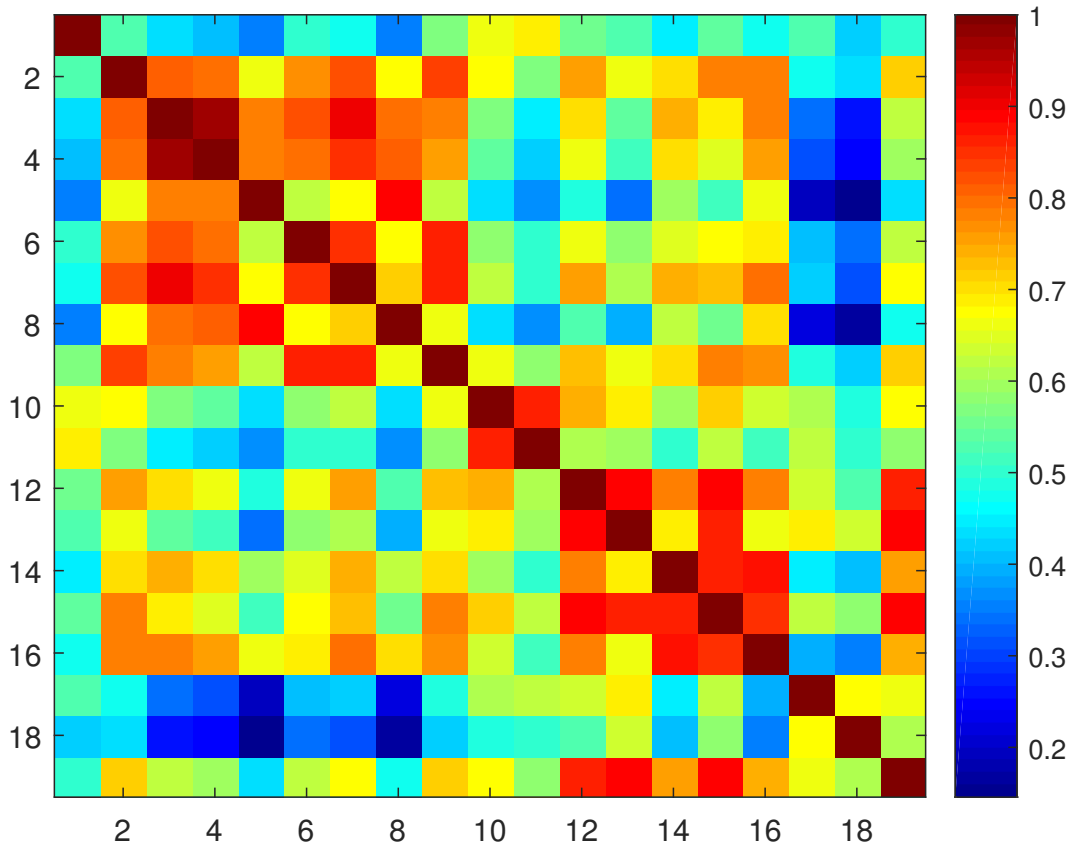


Figure 2.10: Correlation coefficients among wind farms

After the feature construction step, for notation simplicity, we use $x_{w,t}^{input}$ and $y_{w,t}^{target}$ to represent the input features and responsive target. $wp_{w,t}$ and $ws_{w,t}$ are the normalized wind power and wind speed at location w at time t . The goal now is to find a good function $f(x)$ to represent their relationship as in (2.12) – (2.14).

$$\begin{aligned}
 x_{w,t}^{input} = & \quad [wp_{w,t-h^{lead}}, wp_{w,t-h^{lead}-1}, \dots, wp_{w,t-N_{hist}^{lead}-1}, \\
 & \quad ws_{w,t-h^{lead}}, ws_{w,t-h^{lead}-1}, \dots, ws_{w,t-N_{hist}^{lead}-1}, \\
 & \quad wp_{j_1,t-h^{lead}}, wp_{j_1,t-h^{lead}-1}, \dots, wp_{j_1,t-N_{hist}^{near}-1}, \\
 & \quad \dots
 \end{aligned} \tag{2.12}$$

$$\begin{aligned}
 & \quad wp_{j_{N^{near}},t-h^{lead}}, wp_{j_{N^{near}},t-h^{lead}-1}, \dots, wp_{j_{N^{near}},t-N_{hist}^{near}-1}] \\
 y_{w,t}^{target} = & \quad wp_{w,t}
 \end{aligned} \tag{2.13}$$

$$y_{w,t}^{target} = f_w(x_{w,t}^{input}) \tag{2.14}$$

2.3.3 Build Training, Cross Validation and Testing Sets

In order to successfully apply existing machine learning algorithms, we formulate the whole process as a regression problem. It is therefore vital to construct the training, cross validation (CV) and testing sets. The testing set is straightforward to build as it includes all $y_{w,t}^{target}$. During simulation, we use the last 90 days (or 2160 hours) for the full training set. Because wind is highly influenced by recent weather conditions, this number is adequate to model wind uncertainty. This 2160 hour data set is used to construct the actual training and cross validation set.

There are some hyper-parameters that must be tuned for the adopted algorithms, and cross validation is effective to improve modeling generality by avoiding overfitting. The K -fold method is generally adopted for validation [72]. In this method, the whole training set is randomly split into k parts, and every time we train $k-1$ out of k parts and use the remaining part for validation. Therefore, we train models k times, and the hyper-parameters with the

best average validation results will be used for testing purpose. 2-fold is commonly used as well. Due to the strong statistical nature of time series, we follow a special version of 2-fold. The last 3 days of full training data are used for cross validation, and the remaining 2088 hours data are used for actual training. In this way, we ensure that the trained models have the best possible predictive adaptability for recent wind fluctuations.

2.3.4 Training Multiple Models

The key concept for our proposed ensemble approach is not to generate the single best scenario, but to generate multiple scenarios to cover the possible randomness of the wind. It is important to ensure that the whole scenario set represents wind uncertainty. Therefore, three techniques to create reasonable randomness are adopted, bootstrap sampling [73], random feature selection [68] and model ensemble [72].

To obtain a large number of scenarios, using a single training set is not ideal. Bootstrap sampling, *i.e.* sampling with replacement, serves well for providing confidence interval estimation in statistics. By generating multiple bootstrap training set, we are able to explore the randomness of training data without significantly breaking the original statistics.

As mentioned in the previous section, not every feature is necessarily indispensable. Therefore, We randomly sample a feature subset from the full feature set to train a weak predictor. This weak predictor is not as accurate as the one trained with the full feature set, but it does not lose a lot of accuracy either. In addition, by collecting a group of weak predictors, we actually form a strong predictor to better represent the data. We typically use 70 % of the total number of original features.

Forecasting and regression have been discussed for years [7], but there is still no conclusion regarding which algorithm performs consistently the best. Because performance is highly dependent on the data and different algorithms have different capabilities to model the data. By applying the model ensemble, we use several algorithms to the same dataset, and it has a higher chance of capturing the uncertainty when we use the results from all these algorithms together. Here, we explore SVM [74, 75], ANN , Random Forest [68]

and Bagging [76] algorithms for their strong capabilities to model the nonlinearity from the data. With the development from machine learning community, other established algorithms, including LASSO [77], gradient boosting machine [78, 79], additive models [80] and deep neural networks [81], worth investigating in the future work as well. Interested readers can refer to the references for more details.

Each wind farm is trained individually for each model. However, in order to retain the spatial relationship, all wind farms share the same random sampling sequence (seeded from random number generator) for every generated scenario. Different scenarios are seeded with different random number.

In the procedure, N_{sce}^{boost} and N_{sce}^{hyper} denote the number of bootstrap sampling and number of hyper-parameter to be tuned. The following steps are used for training each algorithm for wind farms:

Step 1: Generate random number sequence for sampling.

Step 2: Create a bootstrap training set with current random sequence for every wind farm.

Step 3: Sample random features from the full feature set with current random number sequence.

Step 4: Train each wind farm individually to obtain algorithm parameters.

Step 5: Repeat Step 2 – 4 until N_{sce}^{boost} is reached.

Step 6: Update hyper-parameters to be tuned.

Step 7: Repeat Step 2 – 5 until N_{sce}^{hyper} is reached.

2.3.5 Select Models from Cross Validation

There are multiple hyper-parameters to be determined during training. These include the tree depth for RF and Bagging, ℓ_2 penalty factor for ANN and γ for kernel function in SVM. In addition, there are randomness caused by bootstrap sampling and random feature selection. We select N_{sce} out of $N_{sce}^{boost} \times N_{sce}^{hyper}$ to reach the best performance for each model. Take RF for example, the tree depth is a hyper-parameter to be tuned. If there are 3 tree depth options to select, then $N_{sce}^{hyper} = 3$. The bootstrap sampling is with $N_{sce}^{boost} = 100$; therefore, we have initially $N_{sce}^{boost} \times N_{sce}^{hyper} = 300$ scenarios in total. For 50 desired scenarios $N_{sce} = 50$, we pick the best N_{sce} models out from cross validation.

The cross validation is performed system wide to retain spatial correlation to its maximum extent as each model use the same parameters. The best single scenarios at each farm might not be the most correlated scenarios.

Selection is performed for each algorithm to avoid discriminating. Therefore, we will have 50 scenarios for RF, 50 scenarios for Bagging, 50 scenarios for ANN and 50 scenarios for SVM. So there would be 200 scenarios in total in this specific example.

2.3.6 Generate Scenarios with Selection Models

With N_{sce} selected models, we generate scenarios with the testing set to get the wind power scenarios desired. Some results are presented below.

In Figure 2.11, the correlation histogram indicates that the spatial correlation between wind farms can be well captured by spatially correlated scenarios. Not only highly correlated case can be captured, less correlated or negatively correlated cases can also be well presented. Although the presented histogram does not have a perfect Gaussian shape, the tail length is reasonable.

In Figure 2.12, we can observe that the ensemble approach produces a much more compact scenarios uncertainty set compared to the ARMA approach. The realizations are well captured by the algorithms. The bound obtained with the set can be used for the Interval

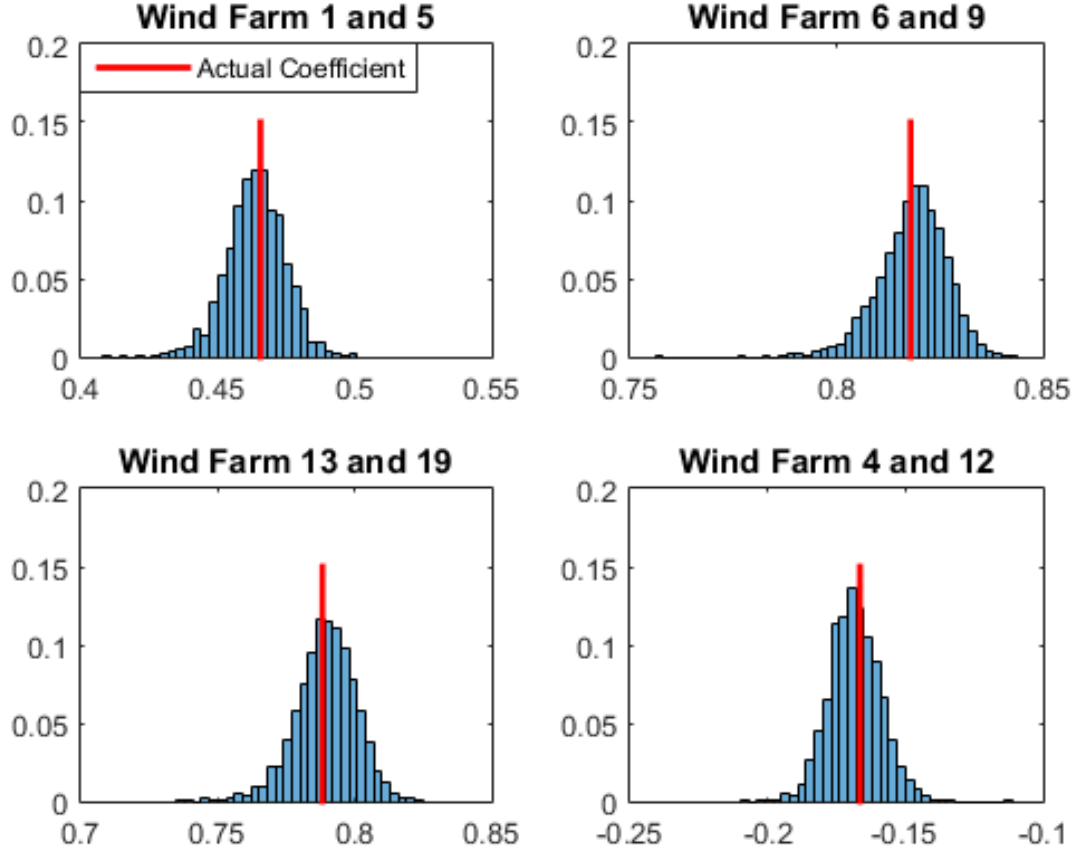


Figure 2.11: Correlation histogram among different wind farms

UC or Robust UC.

Figure 2.13 and Figure 2.14 present two examples. In general, all four algorithms perform well to model the trend of the whole series. Random Forest and Bagging are especially well behaved in capturing the data without introducing a large range of uncertainty. On the contrary, SVM and ANN sometimes may not calibrate well enough to model sharp ramps. There is a small time lag. In addition, they also provide a larger uncertainty range.

In Table 2.1, the Root Mean Square Error (RMSE) for different algorithms and persistence benchmark are presented. It is clear that RF and Bagging perform best, which follows

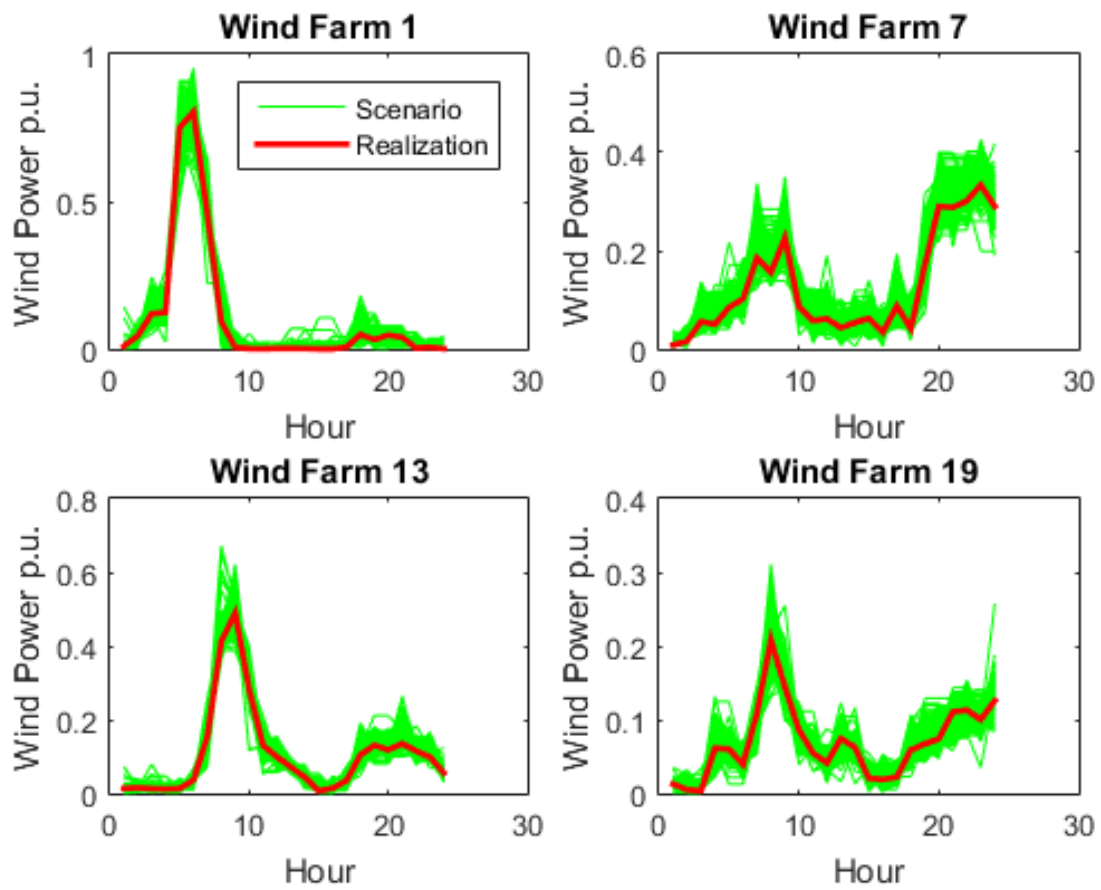


Figure 2.12: Wind scenarios with ensemble approach

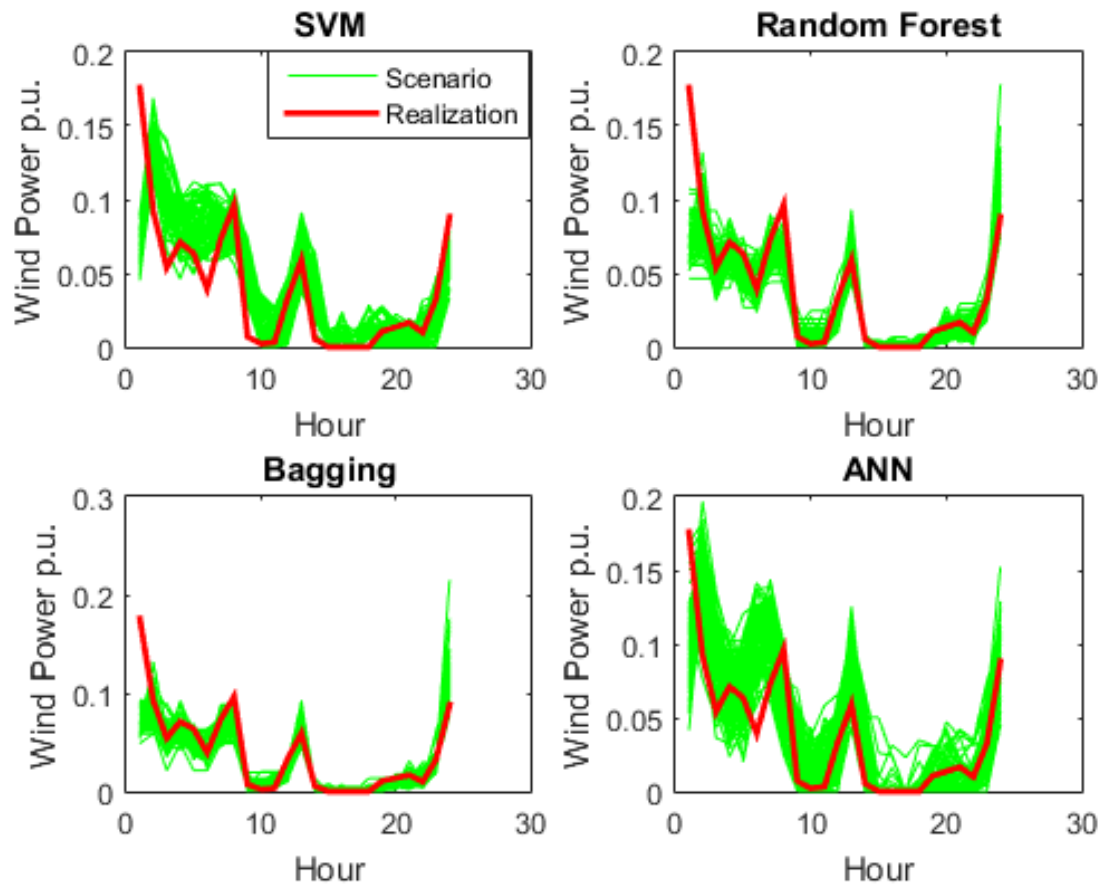


Figure 2.13: Comparison 1 of different algorithms

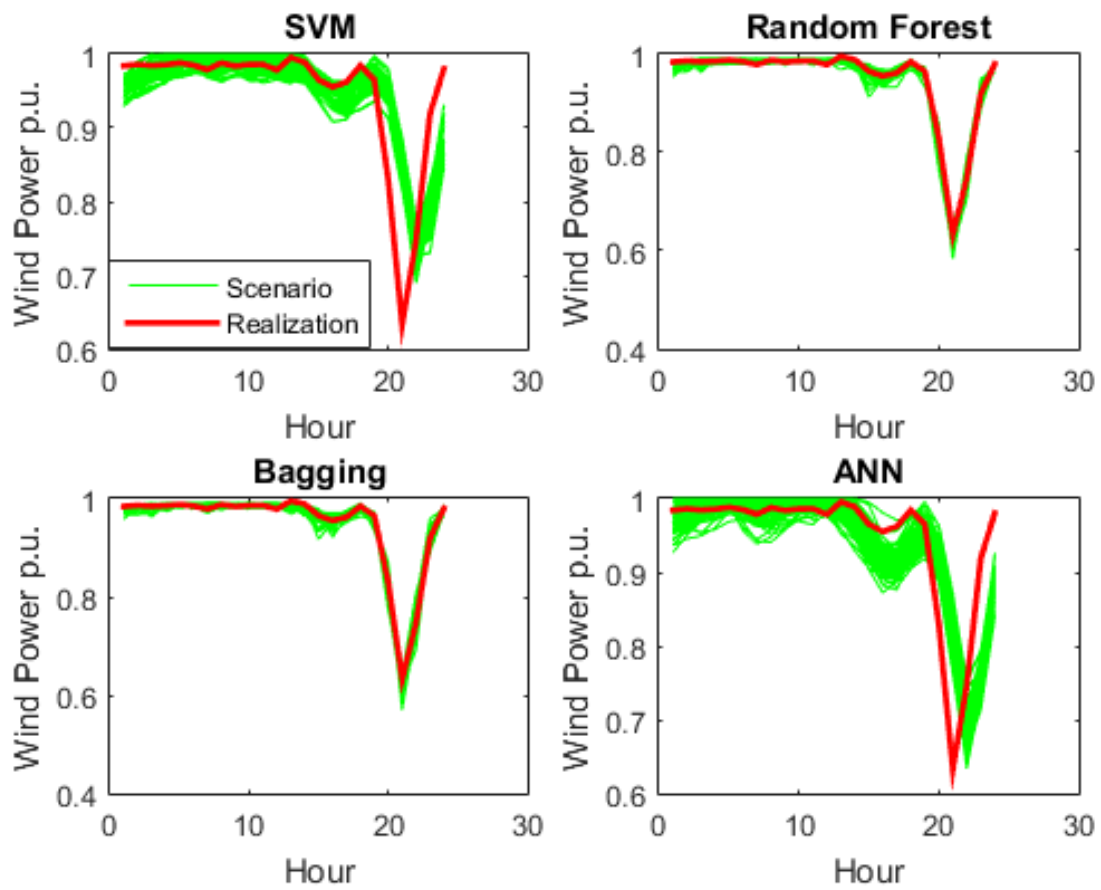


Figure 2.14: Comparison 2 of different algorithms

the pattern of Figure 2.13 and Figure 2.14. The performance of SVM and ANN are not as good as RF and Bagging, but they are still able to beat persistence most of the times. With more explanatory variables like wind forecast, it is expected that SVM and ANN would improve their accuracy.

Noted, there are other metrics to evaluate the quality of scenarios, *e.g.* reliability, skill, sharpness, brier score [40].

2.4 Summary

This section has presented both an ARMA-based time series scenario generation technique and an ensemble-based data mining scenario generation technique. Both approaches capture the spatial temporal correlation between the wind farms to represent a realistic coupling effect. and provide useful uncertainty information to the system operators.

With the fast growing development of machine learning field, the ensemble approach has a better potential to be further improved. As our approach is modular, new advanced algorithms can easily fit in this framework. For example, if SVM, ANN, RF and Bagging are replaced with quantile regression, quantile forecast, or cubic splines, our ensemble approach can produce a probabilistic forecast ensemble. Using copula and covariance matrices, scenarios can be extracted as well.

Table 2.1: Comparison of Algorithms

RMSE	SVM	RF	Bagging	ANN	Persistence
Farm 1	0.032	0.024	0.023	0.027	0.049
Farm 2	0.029	0.007	0.007	0.033	0.037
Farm 3	0.055	0.018	0.019	0.055	0.060
Farm 4	0.089	0.018	0.018	0.089	0.129
Farm 5	0.051	0.013	0.013	0.052	0.057
Farm 6	0.152	0.015	0.014	0.135	0.184
Farm 7	0.025	0.009	0.008	0.029	0.045
Farm 8	0.117	0.026	0.028	0.110	0.145
Farm 9	0.038	0.008	0.008	0.045	0.062
Farm 10	0.033	0.011	0.011	0.033	0.053
Farm 11	0.032	0.010	0.010	0.038	0.050
Farm 12	0.073	0.009	0.010	0.080	0.058
Farm 13	0.121	0.010	0.012	0.109	0.096
Farm 14	0.189	0.079	0.081	0.171	0.108
Farm 15	0.120	0.017	0.016	0.112	0.077
Farm 16	0.183	0.054	0.050	0.169	0.135
Farm 17	0.015	0.004	0.003	0.028	0.021
Farm 18	0.035	0.007	0.006	0.042	0.046
Farm 19	0.083	0.016	0.017	0.073	0.062

Chapter 3

SCENARIO REDUCTION

3.1 *Motivation*

Power system operation is a decision-making processes, and it is important to specify our target and analyze the tools that we have to reach this target. In this dissertation, our focus is on short-term operation planning, and we put our major focus on the unit commitment (UC) applications.

As discussed in Chapter 2, a set of scenarios provides system operators with more insight on the possible wind uncertainty in the day-ahead (DA) and real-time (RT) markets. The question now is how to use this information more effectively. A lot of papers have addressed this problem from various perspectives, including power system, optimization, statistics and computer science.

From the power system perspective, the power system can be protected from wind uncertainty through the provision of additional reserve. Matos [82] determine the system dynamic reserve amount through the probabilistic forecast results. The forecasted median generation minus a lower quantile computes the amount required when a wind generation shortage happens. Botterud *et al.* [83], and Zhou [84] use a similar reserve sizing rule and find that it can produce cost-effective schedules even when compared to a stochastic unit commitment. Zhou [9] further proposes to optimally determine the reserve amount based on the demand curve. Wind uncertainty, load uncertainty and generation outages are considered to construct the demand curve. Then, the optimal amount of reserve can be determined through co-optimization of energy and reserve with this reserve demand curve offer. Similarly, Bruninx [85] proposes to use a lower quantile from the wind forecast error to determine the reserve amount. The same authors further extend their work to probabilistically determine both the

up and down reserve amount with energy and reserve co-optimization [86]. Papavasiliou [10] shows that the “3+5” reserve rule is a good approximation of the results of a stochastic programming scheduling. As suggested in [87], zonal reserve rule can further reduce potential transmission congestion when calling the reserve. MISO implements a performance-based regulating reserve rules as in [88]. Motivated with works in [89], Dvorkin *et al.* [90] proposes a metric to evaluate the system flexibility requirements with wind variability. Same authors also consider to use wind itself to provide reserve products [91]. In addition, system economic and reliability performances are further improved with FACTS devices, PMU and flexible system topology with microgrid connection-mode [92, 93, 94, 95, 96].

A natural way to use scenarios is to apply stochastic programming methods (SP) to do a stochastic optimization for the scheduling. Indeed, SP based stochastic unit commitment (SUC) usually produces a more cost-effective schedule than a deterministic unit commitment (DUC) with reserve rules. Especially when the reserve requirement is excessive, the DUC schedule is much more conservative compared to the schedule produced by a SUC. This topic has been well addressed in the literature [10, 11, 43, 97, 98, 99, 100, 101, 102]. In theory, the more scenarios used in the SP, the better approximation SP can reach to quantify the full distribution of scenarios. However, in reality, due to the limited computational resources, the number of scenarios applied to SUC is restricted. Therefore, a scenario reduction technique is needed to sample scenarios out of the large initial scenario set. Many algorithms have been proposed, including fast forward selection (FFS) [13], simultaneous backward reduction (SBR) [13], K -means clustering [103], importance sampling [6], and other fast forward selection variants as in [104, 105, 106].

Traditional scenario reduction formulates the problem as a transportation problem to minimize the Kantorovic Distance. A continuous probability distribution function is discretized for each scenario. The goal is to preserve the maximum distribution information from the original scenario set in the reduced scenario set. A forward selection and a backward reduction greedy algorithms were first introduced in [107]. Then, modified greedy algorithms were proposed to improve computational efficiency in [13] with fast forward selection and

simultaneous backward reduction. Since they were introduced in the GAMS environment [108], these methods have been widely applied in the literature.

In FFS, the Euclidean distance is commonly used as the metric to quantify the pairwise scenario distance. This metric respects the natural distance between the physical wind energy amounts. However, in power systems, it is not the wind energy uncertainty itself, but the impact caused by this wind uncertainty that really matters to the system operators. Therefore, some FFS variants have been proposed. Morales [104] proposed to compute the Euclidean distance of the system wide market results rather than of the wind energy as the scenario distance for FFS. Feng [109] proposed to first cluster the scenarios and then sample each cluster centroid with FFS. Bruninx [106] proposed to use the absolute difference in the total system costs for FFS. Gomez-Martinez [110] also indicated the benefits with FFS in a UK study.

Papavasiliou [6] proposed a scenario selection algorithm motivated by importance sampling. A deterministic UC is first solved to find the first-stage UC decisions. These decisions are then fixed. For every scenario, a deterministic economic dispatch is run to find their operating cost. These costs actually represent their probability contribution during the scenario sampling.

Essentially, the scenario reduction problem can be viewed as a clustering problem. The centroid of each cluster can be used as the core scenarios to be selected. In [111], Sumali *et al.* propose an information theoretic learning (ITL) mean shift algorithm to cluster and find scenarios. K -means is a most widely accepted clustering technique that can serve this purpose easily [112]. As a K -means centroid is actually a synthetic average of the data points of the cluster, a K -Medoids version [72] is more appropriate in SUC applications. Due to the nonconvex optimization used in the clustering process, the initialization has a big impact on the results. A K -means++ initialization step [113] is used for the K -Medoids version scenario selection.

Although different decomposition algorithms have been applied in this field, such as Benders decomposition [114] and progressive hedging [99, 115], the number of scenarios

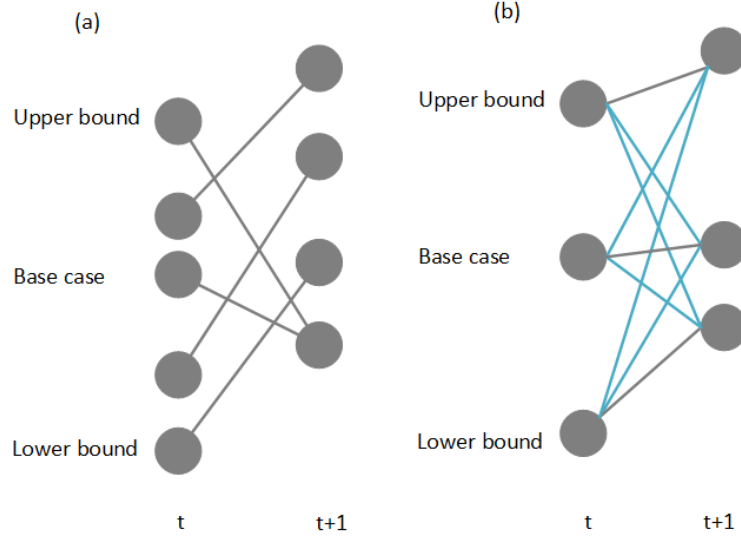


Figure 3.1: Schematic representation of a scenario set in the (a) scenario-based with five scenarios, and (b) interval stochastic UC with three scenarios. The gray lines represent scenarios and the blue lines represent the deterministic constraints on ramping [103].

actually used is still relatively limited compared to the size of the original full set.

Recently, interval optimization [21, 23] has been successfully applied to the UC problem. For these formulations, the uncertainty set is represented by polyhedral uncertainty boundaries or a “box set”. This approach can be thought of as a particular case of scenario reduction.

Figure 3.1 illustrates the difference between the uncertainty models in the scenario-based and in the interval UC. In the interval formulation, the scenario set contains only three scenarios: the central forecast, an upper bound and a lower bound. The bounds for each hour are set in such a way that a given percentage of scenario magnitudes remains within the bounds. Additionally, the inter-hour ramps, depicted by blue lines in Figure 3.1(b), are enforced using additional constraints. The objective function of the interval stochastic UC minimizes the operating cost for the central forecast, while only the feasibility of the solution is guaranteed for the upper and lower bounds. In contrast with the scenario-based

formulation, the interval stochastic UC does not account for the likelihood of the upper and lower bounds. Therefore it produces a schedule that is more expensive but also more robust than the scenario-based UC. In addition, the interval stochastic UC is computationally less demanding than the scenario-based formulation since it considers fewer scenarios. The upper and lower limits for the wind and ramping requirements can be efficiently extracted by finding the confidence interval of the full scenario set. This can be seen as an extreme way to reduce scenario size. Robust optimization [20, 28, 116] can similarly obtain the wind bounds.

Although these techniques have been successfully implemented, some questions remain. Which technique performs best from an operation viewpoint is yet to be ascertained. The proper number of scenarios that needs to be fed to the stochastic UC remains an externally defined parameter.

In this chapter, we first present our comparison of the current scenario reduction techniques for scenario-based SUC and interval UC. The DA costs and the corresponding Monte-Carlo costs are evaluated. The computational performance of the different algorithms are also presented. The interval UC is also included in the evaluation. In the second part of the chapter, a submodular scenario reduction technique is introduced to efficiently select scenarios as well as determine the optimal number of scenarios based on a defined tolerance.

3.2 Comparison of State-of-art Techniques

3.2.1 Scenario-based Stochastic Unit Commitment Formulation

In this subsection, we present the two-stage stochastic scenario-based UC formulation used for simulation. Numerous papers describe computationally efficient or tight and compact formulations for the deterministic unit commitment problem. Here, we adopt a formulation based on the work presented in [117, 101, 118].

3.2.1.1 Objective Function

$$\min \sum_{i,t} csu_{i,t} + \sum_t PROB_s \left(\sum_{i,t} cpg_{i,t,s} + VoLL \sum_{b,t} ens_{b,t,s} + VoWS \sum_{w,t} curt_{w,t,s} \right) \quad (3.1)$$

$$q_{i,j,t} \leq \sum_{tt=T_{i,j}^{MIN}}^{T_{i,j}^{MAX}} y_{i,t-tt}, \quad \forall i, j, t \quad (3.2)$$

$$\sum_j q_{i,j,t} = z_{i,t}, \quad \forall i, t \quad (3.3)$$

$$csu_{i,t} = \sum_j SUC_{i,j} q_{i,j,t}, \quad \forall i, t \quad (3.4)$$

$$cpg_{i,t,s} = NLC_i x_{i,t} + \sum_k MC_{i,k} g_{i,k,t,s}, \quad \forall i, t, s \quad (3.5)$$

As shown in equation (3.1), the SUC model minimizes the sum of the stepwise startup cost and expected sum of the fuel costs and the penalty for load shedding and wind curtailment. In the DUC settings, there is only one scenario and its probability is 1. Equations (3.2) — (3.5) describe the calculation of the startup and fuel costs.

3.2.1.2 System-wide Operating Constraints

Power Balance Constraint:

$$\sum_{i \in r(b)} p_{i,t,s} + \sum_{w \in r(b)} (WF_{w,t,s} - curt_{w,t,s}) - \sum_{b \in f(l)} p_{f,l,t,s} + \sum_{b \in t(l)} p_{f,l,t,s} = D_{b,t} - ens_{b,t,s}, \quad \forall b, t, s \quad (3.6)$$

Power Flow Constraints:

$$p_{f,l,t,s} = \frac{1}{X_l} (\theta_{f(l),t,s} - \theta_{t(l),t,s}), \quad \forall l, t, s \quad (3.7)$$

$$-FL_l^{cap} \leq p_{f,l,t,s} \leq FL_l^{cap}, \quad \forall l, t, s \quad (3.8)$$

Voltage Angle Constraints:

$$-\pi \leq \theta_{b,t,s} \leq \pi, \quad \forall b, t, s \quad (3.9)$$

$$\theta_{b,t,s} = 0, \quad b = slackbus, \quad \forall t, s \quad (3.10)$$

Load Shedding Constraint:

$$ens_{b,t,s} \leq D_{b,t}, \quad \forall b, t, s \quad (3.11)$$

Equations (3.6) – (3.11) represent the constraints related to the system operating constraints, including the power balance constraints, the power flow constraints and the voltage angle constraints.

3.2.1.3 Conventional Generation Constraints

Generator Logical Constraint:

$$x_{i,t} - x_{i,t-1} = z_{i,t} - y_{i,t}, \quad \forall i, t \quad (3.12)$$

Minimum Up and Down Time Constraints:

$$\sum_{tt=t-MUT_i+1}^t z_{i,tt} \leq x_{i,t}, \quad \forall i, t \quad (3.13)$$

$$\sum_{tt=t-MDT_i+1}^t y_{i,tt} \leq 1 - x_{i,t}, \quad \forall i, t \quad (3.14)$$

Generation Block Constraints:

$$p_{i,t,s} = P_i^{MIN} x_{i,t} + \sum_k g_{i,k,t,s}, \quad \forall i, t, s \quad (3.15)$$

$$g_{i,k,t,s} \leq PR_{i,k}, \quad \forall i, k, t, s \quad (3.16)$$

$$P_i^{MIN} x_{i,t} \leq p_{i,t,s} \leq P_i^{MAX} x_{i,t}, \quad \forall i, t, s \quad (3.17)$$

Ramp Rate Constraints:

$$p_{i,t,s} - p_{i,t-1,s} \leq RU_i, \quad \forall t > 1, \quad \forall i, s \quad (3.18)$$

$$p_{i,t-1,s} - p_{i,t,s} \leq RD_i, \quad \forall t > 1, \quad \forall i, s \quad (3.19)$$

$$p_{i,t,s} - PO_i \leq RU_i, \quad \forall t = 1, \quad \forall i, s \quad (3.20)$$

$$PO_i - p_{i,t,s} \leq RD_i, \quad \forall t = 1, \quad \forall i, s \quad (3.21)$$

Equations (3.12) – (3.21) represent all the generation constraints in the UC model, including the minimum and maximum generation limits, the ramping constraints, the UC logic, and the minimum up and down time constraints.

3.2.1.4 Wind Constraints

Wind Curtailment Constraint:

$$curt_{w,t,s} \leq WF_{w,t,s}, \quad \forall w, t, s \quad (3.22)$$

3.2.2 Interval Unit Commitment Formulation

In this subsection, we present the Interval UC formulation. Since the upper bound, the lower bound and the central forecast are explicitly modeled in the IUC, its formulation is similar to the scenario-based SUC formulation of Section 3.2.1 with only 3 scenarios. In addition, transition ramping constraints are included to ensure that enough system ramping capability is available for extreme events. Here, we follow the formulations presented in [21, 43, 103].

Objective Function:

$$\min \sum_{i,t} csu_{i,t} + \sum_{i,t} cpg_{i,t,0} + VoWS \sum_{w,t} curt_{w,t,0} \quad (3.23)$$

$$(3.2) - (3.5) \quad (3.24)$$

Equation (3.23) shows that the IUC minimizes the total operating cost for the central forecast (denoted with a 0 scenario index). Since no load shedding is allowed in the IUC formulation, we omitted the penalty cost in the objective function.

System Wide Operating Constraints:

$$(3.6) - (3.10) \quad (3.25)$$

$$ens_{b,t,0} = 0 \quad (3.26)$$

The original load shedding constraint is modified for the IUC setup as in (3.26). Other constraints remain the same as in the SUC formulation.

The wind generation constraints are the same as in (3.22).

Thermal Generation Constraints:

$$(3.12)) - (3.21) \quad (3.27)$$

$$p_{i,t+1,lb} - p_{i,t,0} \leq RU_i, \quad \forall t < T, \quad \forall i \quad (3.28)$$

$$p_{i,t,0} - p_{i,t+1,ub} \leq RD_i, \quad \forall t < T, \quad \forall i \quad (3.29)$$

$$p_{i,t+1,lb} - p_{i,t,ub} \leq RU_i, \quad \forall t < T, \quad \forall i \quad (3.30)$$

$$p_{i,t,lb} - p_{i,t+1,ub} \leq RD_i, \quad \forall t < T, \quad \forall i \quad (3.31)$$

In addition to the technological constraints in (3.27), the IUC adds ramping transition feasibility constraints as in (3.28) – (3.31). Equations (3.28) and (3.29) model the transition between base case scenario 0 to upper bound ub and lower bound lb . Similarly, Equation (3.30) and (3.31) model the transition between the upper bound ub and the lower bound lb cases. Every time a lower wind case moves to a higher one (base case to upper bound or lower bound to upper bound) possibly requires generation ramp down due to the decreased net load as in (3.29) and (3.31). Conversely, when a higher wind case moves to a lower one (base case to lower bound or upper bound to lower bound), the remaining generation must ramp up due to the increased net load as in (3.28) and (3.30). This trend is in the opposite direction as the bound movement because the wind is treated as a negative load.

3.2.3 Simulation Setup

Four scenario reduction techniques using the UC formulations described above have been tested on a modified version of the 24-bus IEEE RTS [119]. The installed capacity of the test system is 3405 MW. The cost curves of the generating units in this system are approximated as three-segment piece-wise linear functions with equally spaced elbow points. Similarly, the start-up cost of each generator is modeled as a piece-wise linear function of the generator's down time. Other details of this system, including the minimum up and down time of generators, ramp rates, the configuration of the transmission network and its power flow limits, can be found in [120]. The wind penetration is assumed to be 20 % in terms of the energy consumed daily system-wide. The value of wind spillage is set to \$ 35 /MWh [121] and the value of lost load is set to \$ 5000 /MWh [122].

The wind and load forecasts are based on BPA data [123]. Wind scenarios are obtained using an ensemble approach as in Section 2.3. Two sets of 1000 wind generation scenarios

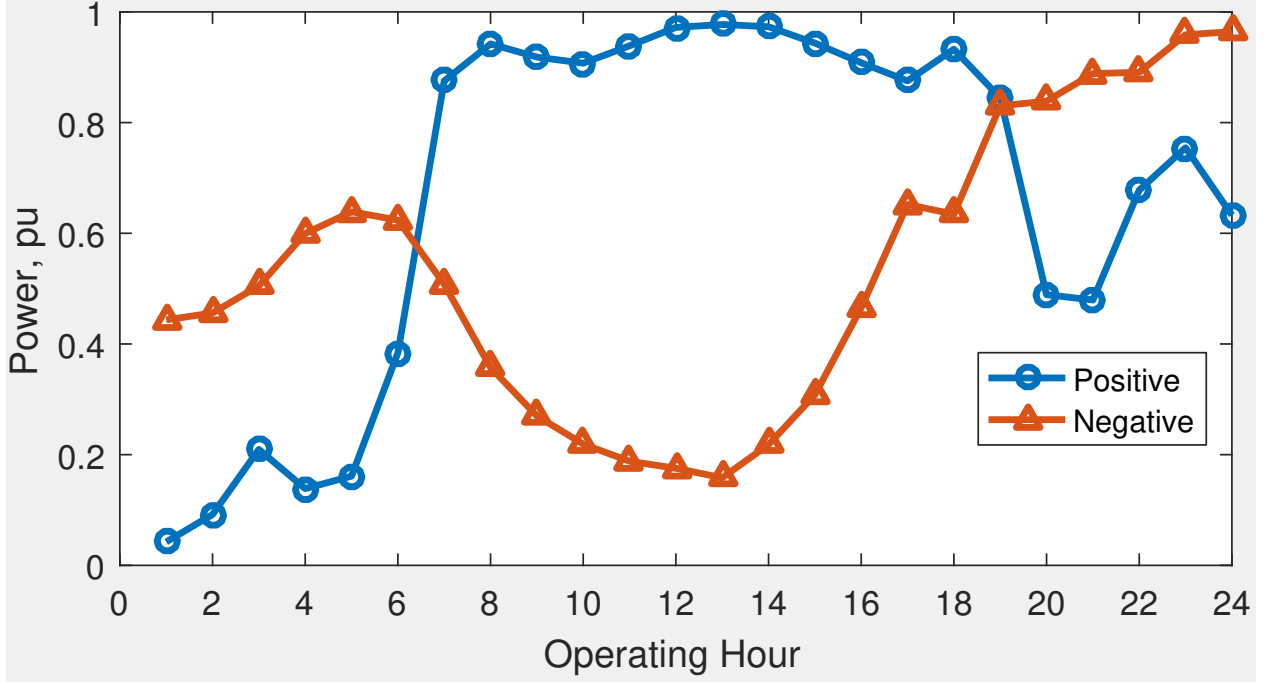


Figure 3.2: Wind profiles with positive and negative correlation with load

were generated with positive and negative correlation between the central wind forecast and the load. The central forecast for each scenario set was calculated as the average of the 1000 scenarios. Figure 3.2 illustrates these central wind forecasts. For positively correlated wind and load forecasts, the peak wind production occurs during daytime hours, and during the night for the negatively correlated wind and load forecasts.

To reflect standard practice in power system operation, the stochastic UC is solved based on the day-ahead load and wind scenarios to produce the optimal day ahead schedule. The value of the objective function for this schedule gives the day-ahead cost (DAC) of operating the power system for the set of scenarios considered. Since these day-ahead scenarios are by definition uncertain, the DAC is a projected or expected cost, rather than an actual cost. A Monte Carlo simulation is required to estimate the actual operating cost (AOC) *i.e.* the cost that would be incurred on the day, when the realization of load and wind uncertainty

is known and the day-ahead schedule must be adjusted to keep the system in balance. The difference between the DAC and the AOC is the cost of corrective dispatch, *i.e.* the cost of the adjustments to the day-ahead schedules that are required in real time to meet the load and wind realizations. Changes in the commitment decisions made in the day ahead schedule are allowed in the Monte-Carlo simulation if constraints on the minimum up- and down- time limits are not violated.

For each Monte-Carlo trial, the day-ahead schedule produced using a given stochastic formulation and scenario reduction technique is dispatched to meet a particular realization of wind and load uncertainty. Realizations of load uncertainty are modeled using a normal distribution, as explained in [8]. Wind power forecast errors are modeled using the skew-Laplace distribution [124]. The cost of this dispatch represents the AOC for a particular realization of uncertainties. This cost includes the start-up cost of the day-ahead and real-time commitments, the cost of dispatch of generators under known realizations of wind and load uncertainties, as well as the wind spillage and load shedding costs. The number of trials required for each schedule is $\max(1000, N_{MC})$ where N_{MC} is the number of trials required to achieve 95 % confidence [125].

3.2.4 Case Study Results

3.2.4.1 Day-Ahead Schedules and Costs

At the day-ahead stage, the scenario-based stochastic UC is solved for 5, 10, 20, and 40 scenarios, obtained using the K -means [113], Simultaneous Backward Reduction (SBR) [13], Fast Forward Selection (FFS) [13] and Importance Sampling (IS) [6] scenario reduction techniques. The interval stochastic UC is solved for the range of uncertainty between bounds that discard 30, 20, 10, and 1 % of the extreme values at each time period. Figure 3.3 shows the range of uncertainty obtained with different numbers of scenarios produced by the four techniques and compares them to the range of uncertainty obtained when the 10 % largest and smallest values are discarded. The range of uncertainty increases with the number of

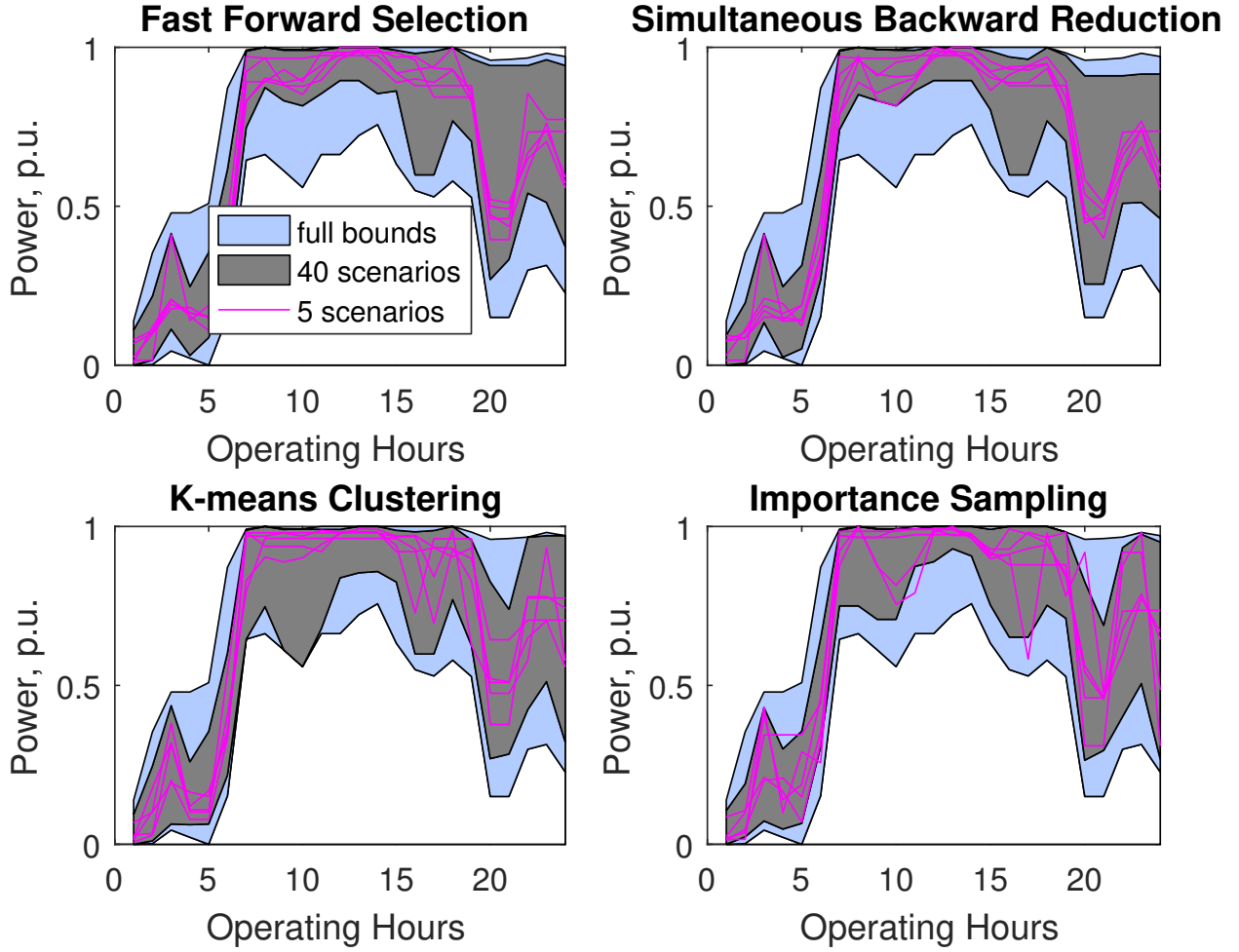


Figure 3.3: Comparison of the range of uncertainty obtained for different techniques

scenarios considered. Note that during some operating hours the ranges of uncertainty of the scenario-based stochastic UC with 40 scenarios and of the interval UC are equal.

Table 3.1 and Table 3.2 show that the Day Ahead Cost (DAC) for the scenario-based UC depends on the number of scenarios considered and increases with the number of scenarios. This increase is explained by the larger range of uncertainty captured by a greater number of scenarios. For a given number of scenarios, the DAC also depends on the scenario reduction techniques but no scenario reduction technique appears to be significantly better than the

others. These variations in DAC remain in the range $[0.37, 0.93]$ % and $[0.44, 0.79]$ % for positively and negatively correlated wind and load forecasts, respectively.

Table 3.1: Day Ahead Cost (In 10^3 \$) of the Scenario-Based schedule with positively correlated wind and load

Scenarios	FFS	SBR	K -means	IS
5	566.7	568.8	567.0	567.7
10	570.6	571.2	570.5*	575.2
20	573.0	572.4*	577.5	577.8
40	576.0*	576.6	578.1	578.2

* - minimum DAC for a given number of scenarios

Table 3.2: Day Ahead Cost (In 10^3 \$) of the Scenario-Based Schedule with negatively correlated wind and load

Scenarios	FFS	SBR	K -means	IS
5	657.1*	657.6	656.7	652.4
10	657.1	659.1	656.9*	660.3
20	658.9	658.5*	659.7	662.1
40	660.6*	662.6	663.5	663.2

* - minimum DAC for a given number of scenarios

Table 3.3 and Table 3.4 show that the DAC of the interval UC increases as the range of uncertainty increases. If the range of uncertainty is relatively small (30 %), the interval UC has a smaller DAC cost than the scenario-based UC with any number of scenarios and any scenario reduction technique. As the range of uncertainty increases, the DAC of the interval formulation also increases. The difference between the DAC of the scenario-based

formulation with 40 scenarios and the interval formulation is in the range $[0.15, 0.53]$ % and $[0.73, 1.1]$ % for positively and negatively correlated wind and load forecasts, respectively.

Table 3.3: Key statistics for the Day-Ahead Interval UC schedules with positively correlated wind and load

	30%	20%	10%	1%
CPU time (s)	61.8	42.42	94.3	49.3
DAC (10^3 \$)	565.2	568.2	572.4	579.3

Table 3.4: Key statistics for the Day-Ahead Interval UC schedules with negatively correlated wind and load

	30%	20%	10%	1%
CPU time (s)	48.7	41.9	37.3	31.4
DAC (10^3 \$)	658.1	658.5	659.3	668.1

Table 3.5 and Table 3.6 present the computation time for the scenario-based stochastic UC. As long as the number of scenarios in the scenario-based approach remains relatively low, there are no significant differences in the scenarios generated by different scenario reduction techniques and, therefore, the computation time is weakly dependent on the scenario reduction technique used. As the number of scenarios increases, scenarios generated by different scenario reduction techniques become more varied and require different amounts of computation time. The FFS scenario selection technique produces scenarios that result in the fastest solution of the scenario-based stochastic UC. On the other hand, the IS scenario reduction technique produces scenarios that require the largest computation time. Table 3.3 and Table 3.4 have already shown that the interval UC usually requires less computing time than the scenario-based stochastic UC and that this computing time does not have a simple

relation to the range of uncertainty considered.

Table 3.5: Computational time (s) for the Scenario-based SUC schedules with positively correlated wind and load

Scenarios	FFS	SBS	K -means	IS
5	53.8	54.101	58.9	59.0
10	168.0	156.6	290.8	271.6
20	864.3	1027.7	1142.5	1942.6
40	2207.9	3516.2	4097	4801

Table 3.6: Computational time (s) for the Scenario-based SUC schedules with negatively correlated wind and load

Scenarios	FFS	SBR	K -means	IS
5	45.5	42.7	52.8	55.8
10	108.0	115.7	101.9	111.9
20	433.2	407.1	384.8	554.0
40	2635	1645	3797	7890

The day-ahead results obtained for the scenario-based and the interval UC indicate that the scenario reduction technique is a factor that significantly affects the computation time of the stochastic UC. On the other hand, if the scenario-based UC is solved with different scenario reduction techniques the DAC variations would be of the same order of magnitude as the DAC difference between the scenario-based and interval formulations.

3.2.4.2 Real-Time Re-dispatch and Actual Operating Cost

Table 3.7 and Table 3.8 show the expected value of the actual operating cost, $E(\text{AOC})$, its standard deviation, $\sigma(\text{AOC})$, the expected cost of corrective dispatch, $E(\Delta)$, the expected start-up cost, $E(\text{SC})$, and its standard deviation, $\sigma(\text{SC})$, for the scenario-based schedules obtained with different scenario reduction techniques and for different numbers of scenarios. The expected value of the start-up cost includes the day-ahead start-up cost and the expected start-up cost of real-time changes to the commitment decisions. The FFS technique results in the least expensive AOC and, therefore, produces the most cost-effective day-ahead schedule. On the other hand, the schedules produced by the FFS technique also result in the largest standard deviation for the AOC. This indicates that the least expensive solution is also less adaptable to worst-case realizations of uncertainty than more expensive schedules obtained with the other scenario-reduction techniques.

The AOC also depends on the number of scenarios considered in the scenario-based stochastic UC. If the number of scenarios is 5, the generation schedule is based on an inaccurate representation of the uncertainty and, therefore, this schedule requires expensive adjustments in real-time. On the other hand, if the number of scenarios is 40, the schedule accommodates scenarios with low probabilities and is therefore unnecessarily robust and expensive. The minimum AOC occurs for 10 or 20 scenarios for all the scenario reduction techniques considered. The standard deviation of the AOC monotonically decreases as the number of scenarios increases. This trend indicates that a larger number of scenario improves the adaptability of the schedule. Although, the expected start-up cost does not vary significantly for different numbers of scenarios or scenario reduction techniques, its standard deviation decreases as the number of scenarios increases. Therefore, a more robust representation of uncertainty causes less cycling of generators.

Table 3.9 and Table 3.10 summarize key statistics of the Monte-Carlo simulations for the interval UC using the same notations as in Table 3.7 and Table 3.8. The AOC increases as the range of uncertainty increases and is very sensitive to the bounds. If the 30 % bounds

Table 3.7: Actual Operating Cost (In 10^3 \$) of the Scenario-based SUC schedules with positively correlated wind and load

scenarios	Parameter	FFS	SBR	K -means	IS
5	E(AOC)	593.0*	593.1	593.1	596.2
	σ (AOC)	17.7	17.7	16.2	16.1
	E(Δ)	26.3	24.3	26.1	28.5
	E(SC)	21.3	21.1	21.3	22.9
	σ (SC)	0.719	0.728	0.715	0.708
10	E(AOC)	590.4*	591.0	590.8	592.2
	σ (AOC)	17.3	17.3	17.3	15.8
	E(Δ)	19.8	19.8	20.3	17
	E(SC)	19.5	21.1	21.3	10.8
	σ (SC)	0.188	0.186	0.182	0.204
20	E(AOC)	590.5*	591.5	590.9	593.1
	σ (AOC)	16.1	16.0	15.7	15.7
	E(Δ)	17.5	19.1	13.4	15.3
	E(SC)	21.4	21.1	23.1	23.1
40	σ (SC)	0.167	0.157	0.193	0.150
	E(AOC)	593.3*	593.5	594.2	596.7
	σ (AOC)	15.7	15.8	15.5	15.7
	E(Δ)	17.3	16.9	16.1	18.5
	E(SC)	22.9	22.9	22.8	22.8
	σ (SC)	0.155	0.191	0.138	0.138

* - minimum AOC for a given number of scenarios

are used, the schedule is not robust enough and results in the largest cost of corrective

Table 3.8: Actual Operating Cost (In 10^3 \$) of the Scenario-based SUC schedules with negatively correlated wind and load

scenarios	Parameter	FFS	SBR	K -means	IS
5	E(AOC)	666.6*	669.0	668.4	667.6
	σ (AOC)	13.9	13.8	13.9	13.0
	E(Δ)	9.5	11.4	11.7	15.2
	E(SC)	12.0	12.2	12.1	12.5
	σ (SC)	0.693	0.754	0.605	0.229
10	E(AOC)	667.4*	669.1	668.7	668.1
	σ (AOC)	13.2	13.0	13.1	12.6
	E(Δ)	10.7	10.0	10.6	7.8
	E(SC)	12.1	12.3	12.2	12.1
	σ (SC)	0.656	0.646	0.584	0.161
20	E(AOC)	665.5*	665.7	665.7	670.6
	σ (AOC)	12.9	12.9	13.0	12.7
	E(Δ)	6.6	6.2	6.0	8.5
	E(SC)	12.3	12.1	12.3	12.2
	σ (SC)	0.596	0.619	0.214	0.129
40	E(AOC)	669.1*	671.9	671.7	671.0
	σ (AOC)	12.6	12.4	12.3	12.4
	E(Δ)	8.5	9.3	8.2	7.8
	E(SC)	12.3	12.1	12.2	12.6
	σ (SC)	0.133	.119	0.136	0.111

* - minimum AOC for a given number of scenarios

actions. On the other hand, the schedule with the 1 % bounds is more expensive than any

Table 3.9: Actual Operating Cost (In 10^3 \$) of the Interval UC schedules with positively correlated wind and load

Parameter	30%	20%	10%	1%
E(AOC)	587.8*	590.1	591.2	596.3
σ (AOC)	18.1	16.8	15.6	1.5
E(Δ)	22.3	21.9	18.8	17.0
E(SC)	12.3	12.5	12.2	12.2
σ (SC)	0.200	0.185	0.207	0.123

* - minimum AOC for a given number of scenarios

scenario-based stochastic UC in Table 3.7 and Table 3.8. In addition, it requires the least costly corrective actions. The AOC of the interval solution is larger than the cost of the scenario-based solution for the positively correlated wind and load profiles and the difference is in the range $[0.13, 0.35]$ % for different numbers of scenarios. On the other hand, for the negatively correlated wind and load profiles, this range is $[-0.48, 0.06]$ %. This indicates that the cost of the interval approach can be lower than the cost of the scenario-based approach (*e.g.* the cost for 5 scenarios) if scenarios do not represent uncertainty accurately and require expensive corrective dispatch.

Table 3.10: Actual Operating Cost (In 10^3 \$) of the Interval UC schedules with negatively correlated wind and load

Parameter	30%	20%	10%	1%
E(AOC)	685.3	668.9	665.9*	669.7
σ (AOC)	21.3	20.9	1.30	0.891
E(Δ)	27.2	10.4	6.6	1.6
E(SC)	21.2	21.1	21.2	22.7
σ (SC)	0.825	0.697	0.570	0.113

* - minimum AOC for a given number of scenarios

3.3 Scenario Reduction with Submodular Optimization

Stochastic programming methods have been proven to deal effectively with the uncertainty and variability of renewable generation resources. However, as shown in the previous section, the quality of the solution that they provide (as measured by cost and reliability metrics) depends on the accuracy and the number of scenarios used to model this uncertainty and variability. Scenario reduction techniques are used to manage the computational burden by selecting representative scenarios. The common drawback of existing scenario reduction techniques is that the number of representative scenarios is a user-defined parameter. In this subsection, we propose a new and efficient Submodularity-based Scenario Reduction (SSR) to endogenously optimize the number of scenarios as well as rank these scenarios. This algorithm is compared, both qualitatively and quantitatively, with the state-of-the-art fast forward selection (FFS) algorithm. Numerical results demonstrate that the proposed SSR is related to the FFS but is significantly faster.

3.3.1 Background

Over the last few years, scenario-based Stochastic Programming (SP) has been used to compute generation schedules that take into account the uncertainty on the net load which must be served by conventional generators. Ideally, this optimization should be based on a substantial number of wind generation scenarios. However, to reduce the computational burden, a large initial set of scenarios must be reduced to a smaller subset of representative scenarios. As discussed in the previous subsection, the quality of the SP solution decreases if this reduced set is too small, and the SP solution time increases if it is too large. Among these scenario reduction techniques, FFS has been shown to be the most efficient. However, this technique requires that the reduced number of scenarios be predefined. Furthermore, the scalability of the FFS algorithm to large initial sets of scenarios has not been demonstrated.

Submodular function optimization has recently been shown to provide theoretically-bounded and computationally efficient solutions to feature selection [126, 70, 127, 128], train-

ing data selection [129, 130, 131] and other machine learning applications [132, 133, 134]. Nemhauser [135] showed that if a submodular function is monotone non-decreasing, the problem of maximizing such a function can be approximately solved using a simple greedy algorithm, which is guaranteed to yield a solution within a constant factor of the optimum.

Diminishing return property is one of most important properties for submodular function optimization. It states that the incremental gain from selecting one more member is always non-increasing. Therefore, this gain upper bound for incremental helps to derive efficient algorithms.

Definition: For every $X, Y \subseteq \Omega$ with $X \subseteq Y$ and every $x \in \Omega \setminus Y$, we have that $f(X \cup \{x\}) - f(X) \geq f(Y \cup \{x\}) - f(Y)$.

3.3.2 Method

We first compute the similarity matrix w_{ij} for each scenario pair using a radial basis kernel function (RBF) and an ℓ_2 -norm:

$$w_{ij} = e^{-\frac{d_{ij}}{\lambda}} = e^{-\frac{\|s_i - s_j\|_2}{\lambda}} \quad (3.32)$$

where s_i and s_j are the time series of scenarios i and j . λ is a parameter controlling similarity scaling. Typically we tune λ from empirical values obtained from the pairwise scenario ℓ_2 -norm distribution.

A cardinality-constrained version of the SSR (ccSSR) can be formulated as a facility location problem:

$$\max_R f_1(R) = \max_R \left(\sum_{i \in O} \max_{j \in R} w_{ij} \right) \quad (3.33)$$

$$s.t. : \quad card(R) \leq K \quad (3.34)$$

where O, R are the initial and reduced scenario sets and K is the maximum size of the reduced set. This formulation assumes that each scenario is equiprobable with a uniform

probability distribution p . However, a non-uniform distribution can be easily embedded within the similarity matrix. The probability distribution q of the reduced set R can be computed using the rule proposed in [13].

This type of combinatorial problem is in general NP-hard. Fortunately, our ccSSR maximizes a monotonic non-decreasing submodular function under a cardinality constraint. We can therefore apply an accelerated greedy algorithm [126, 136], as described in Algorithm 1. This accelerated greedy algorithm achieves orders of magnitude performance speedups and scales up well when the initial scenario set is large.

The order of selection represents the relative importance of the contribution of each scenario to the objective function, which can be used as the rank. If K is set to a very large value, we can use this algorithm to rank the scenarios, examine the incremental improvements in the objective function and select the cut-off point on the number of scenarios where the benefit saturate.

Alternatively, we can also formulate the following cardinality-unconstrained SSR (cuSSR) problem:

$$\max_R f_2(R) = \max_R \left[\left(\sum_{i \in O} \max_{j \in R} w_{ij} \right) - \beta \text{card}(R) \right] \quad (3.35)$$

where β controls the weight of the cardinality penalty added to the objective function. A large β tends to select fewer scenarios but reduces the SP solution accuracy, while a small β yields more scenarios but increases the SP computational burden. In our tests, we set $\beta = 1$ because values of the elements of w_{ij} range between 0 and 1. As the objective function is no longer monotone increasing, the bounded optimality condition does not hold. However, the problem in (3.35) can still be solved using Algorithm 1. Testing shows that the solution is still sufficiently fast, and in the ccSSR the parameter β can be used to set the cut-off point for selecting the optimal number of scenarios, which results in same scenario selections as in cuSSR.

Algorithm 1 Submodular scenario reduction algorithm

```

1: Initialize  $R \leftarrow \emptyset$ ,  $f(R) \leftarrow 0$ ,
2: for  $i = 1, 2, \dots, N$  do
3:    $v_i \leftarrow f(i)$ 
4: end for
5: while  $\text{card}(R) \leq K$  do
6:    $j \leftarrow \arg \max_{i \in O \setminus R} v_i$ 
7:    $v_j \leftarrow f(R \cup \{j\}) - f(R)$  (Calculate incremental benefits)
8:   if  $v_j > \max_{i \in O \setminus (R \cup j)} v_i$  then
9:      $R \leftarrow R \cup \{j\}$  {Diminishing return property}
10:     $v_j \leftarrow 0$ 
11:   else if  $v_j \leq 0$  then
12:     Break {Solution is (locally) optimal}
13:   end if
14: end while
15: for  $i = 1, 2, \dots, \text{card}(R)$  do
16:    $q_i \leftarrow p_i + \sum_{j \in O(i)} p_j$ , where  $O(i) \leftarrow \{j \in O \setminus R : i = \arg \max_{i \in R} w_{ij}\}$  {Re-distribute probability}
17: end for
18: return  $R, q$ 

```

3.3.3 Comparison to Fast Forward Selection (FFS) Technique

The FFS technique [13] essentially optimizes $\min_R (\sum_{i \in O} \min_{j \in R} d_{ij})$ with the cardinality constraint (3.34), and solves the above discrete optimization with a forward greedy heuristics. Our proposed formulation is in max max form. The FFS and ccSSR techniques thus solve the same problem but with different pairwise scenario distance scaling. As the results presented in the next section show, these two algorithms indeed select very similar scenarios, but SSR does it much faster than FFS. Furthermore, while FFS can only select a given number scenarios, SSR efficiently ranks these scenarios and cuSSR optimizes the number of scenarios.

3.3.4 Numerical Results

Using neural network as in Section 2.3, we generated an initial set of 10,000 wind scenarios for three wind farms based on the NREL dataset. The upper part of Table 3.11 compares the FFS with the ccSSR in terms of selection speed for various sizes of the initial set of scenarios and different values of the cardinality K of the reduced set. The lower part of this table gives the optimal value of K as calculated by the cuSSR and the corresponding computing time for $\beta = 1$. Except for small scale selections, the ccSSR is consistently faster than the FFS. Moreover, compared to the FFS, the computing time of the ccSSR increases slowly with the initial number of scenarios. The diminishing return property and priority queue motivate the accelerated greedy algorithm [126], which greatly reduces the search space.

Table 3.11: Solution time comparison

	$N = 1000$		$N = 2000$		$N = 5000$		$N = 10000$	
Time (s)	FFS	ccSSR	FFS	ccSSR	FFS	ccSSR	FFS	ccSSR
$K = 20$	0.10	0.15	0.38	0.45	2.34	2.01	9.77	8.02
$K = 100$	0.42	0.23	1.64	0.64	9.96	2.75	40.6	10.7
$K = 1000$	6.46	0.33	20.1	0.94	105	4.20	408	16.0
cuSSR Optimal K^*	19		34		73		136	
Time (s)	0.20		0.68		3.54		15.7	

Table 3.12 shows that, starting from an initial set of 5,000 scenarios, the FFS and ccSSR select mostly the same scenarios. The scenarios obtained by the ccSSR yield similar values of the FFS objective function as those obtained by the FFS. These reductions were carried out for three values of the scaling parameter λ . The overlap between the selections is higher for larger λ because a larger λ results in similar scalings in SSR and in FFS.

Figure 3.4 shows the incremental change in the objective function after each scenario selection. A smaller cardinality penalty β selects more scenarios closer to the saturation

Table 3.12: Number of scenarios selected by both FFS and ccSSR

K	10	20	50	100	500	1000	2500	K^*
$\lambda = 0.5$	8	18	42	90	433	887	2405	240
$\lambda = 2.5$	9	18	49	97	484	961	2449	73
$\lambda = 10$	9	20	50	100	498	986	2472	22

point. By varying the cardinality penalty β , we can therefore achieve a tradeoff between the accuracy and speed of the SP.

The cuSSR cardinality-unconstrained formulation selects about 1.4% of the initial set of scenarios for $\lambda = 10$. These small values of K^* support the observation from [103, 97] that a good enough stochastic unit commitment solution does not necessarily require a large set of scenarios.

In the Table 3.13, we present the selection results in terms of the FFS objective $(\sum_{i \in O} \min_{j \in R} d_{ij})$ (recall that the FFS tries to optimize $\min_R(\sum_{i \in O} \min_{j \in R} d_{ij})$), where d_{ij} is the euclidean distance between two scenarios. The calculation is based on the scenarios selected in the Table 3.12.

Table 3.13: Selection results in terms of FFS objectives

K	10	20	50	100	500	1000	2500
$\lambda = 0.5$	9688.2	9484.6	9237.1	9020.8	8072.4	7066.4	4252.3
$\lambda = 2.5$	9667.6	9482.0	9235.0	9019.4	8071.2	7066.7	4254.0
$\lambda = 10$	9667.6	9482.7	9234.9	9019.4	8071.3	7067.1	4254.8
FFS	9667.5	9482.7	9234.9	9019.3	8071.1	7067.2	4255.1

From Table 3.13, we can observe that our proposed SSR indeed selects scenarios with similar FFS objectives. As explained in section 3.3.3, the FFS and SSR are strongly connected. This table demonstrates this as well. Since our problem is a maximization, it does

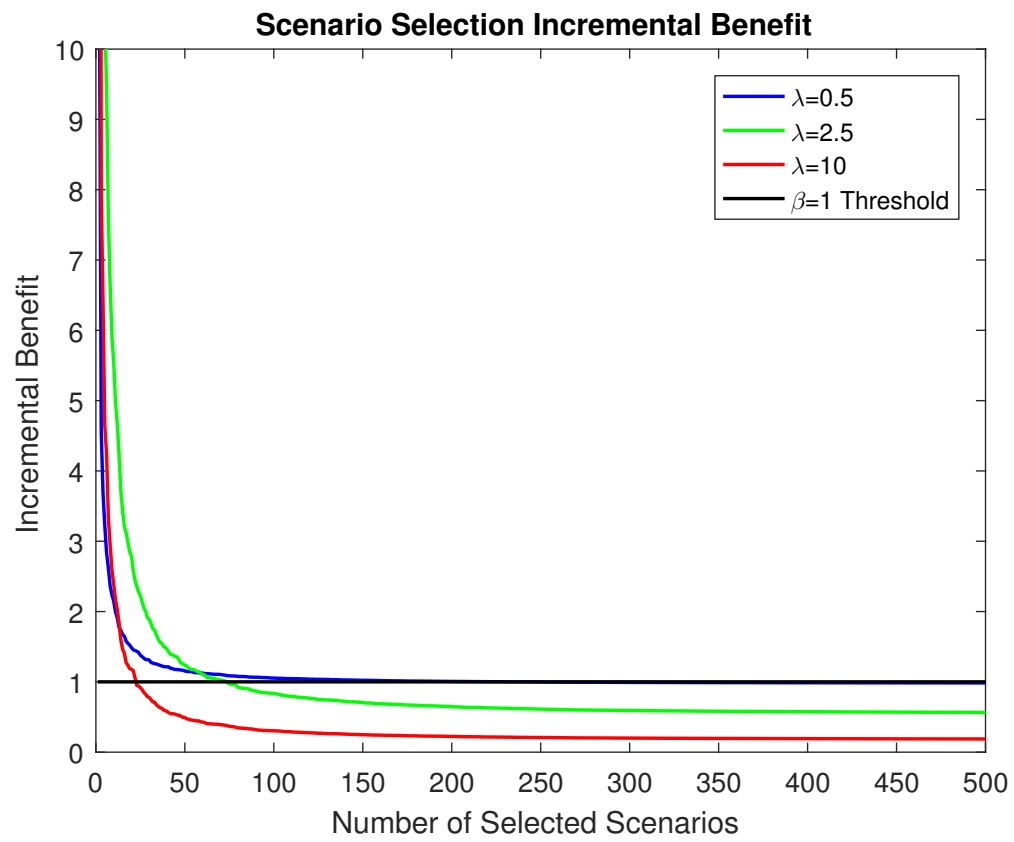


Figure 3.4: Scenario selection incremental benefit

not guarantee that the selection is smaller than what the FFS produces. However, from these empirical results, it is not surprising to see that our results are very close to those of FFS most of the time and even better in some cases. Since FFS is a forward greedy algorithm, there is no guarantee that it produces the global optimal solution.

Table 3.14: Selection results in terms of SSR objectives $\lambda = 0.5$

K	10	20	50	100	500	1000	2500
$\lambda = 0.5$	119.10	135.98	173.94	228.59	628.44	1119.58	2584.32
FFS	118.92	135.89	173.85	228.52	628.26	1119.41	2584.11

Table 3.15: Selection results in terms of SSR objectives $\lambda = 2.5$

K	10	20	50	100	500	1000	2500
$\lambda = 2.5$	2313.77	2349.76	2401.10	2450.43	2698.86	2974.85	3766.32
FFS	2313.68	2349.60	2401.10	2450.43	2698.60	2974.75	3766.11

Table 3.16: Selection results in terms of SSR objectives $\lambda = 10$

K	10	20	50	100	500	1000	2500
$\lambda = 10$	4121.66	4137.04	4157.99	4176.62	4261.40	4352.38	4608.82
FFS	4121.65	4137.04	4157.99	4176.62	4261.40	4352.38	4608.79

In addition, we also present the results in terms of ccSSR objectives $f_1(R) = (\sum_{i \in O} \max_{j \in R} w_{ij})$. As w_{ij} is determined with λ , Table 3.14 – 3.16 present three different cases. In general, FFS and ccSSR result in similar results, and our ccSSR is always no worse than FFS, empirically showing the advantage of our method.

3.3.5 Discussions

The intrinsic uncertainty level of scenarios is not infinite, and the high-dimensional scenario distribution can be captured by a limited number of key scenarios. In this case, similar with the shape in Figure 3.4, the AOC from Monte-Carlo testing with increasing selected scenarios should converge to a saturation level. After the "sweet point", selecting more scenarios does not necessarily bring more economic benefits for modeling the scenario distribution.

Infinite number of scenarios in DA and Monte-Carlo should produce the ideal DA schedule and true AOC results. However, when the scenarios used in Monte-Carlo is not infinite, this saturation level should not be a constant value. As long as the numbers are within the optimality gap from original MILP problem, we should be confident that the sweet point has been found. Since more scenarios bring more computational burden rather than economic savings, this specific scenario size number should be recognized as the optimal scenario number.

Also noted that this optimal scenario number should be case by case. A larger initial dataset is very likely to require a larger reduced set to fully capture the necessary distribution information. Without carefully analyzing the data, it is not correct to simply select a fixed number of scenarios only based on the reference number from the literature.

3.4 Summary

This section compares the scenario-based stochastic UC with different scenario reduction technique and with the interval UC. Results demonstrate that different scenario reduction techniques affect the operating cost and the computation time of the stochastic UC. The fast forward selection technique generates scenarios, which result in the least expensive actual operating cost for the scenario-based stochastic UC formulation. This scenario selection technique is also shown to produce scenarios that cause the lowest computation burden. Although this computation time is relatively small, it is an order of magnitude larger than the computation time of the interval formulation. The results show that the difference

between the cost of the interval and stochastic-based UC can be kept at a minimum if the bounds of the interval method are chosen carefully. Further studies are needed to optimally determine the bounds of the interval formulation and thereby to minimize the cost difference between the interval and scenario-based stochastic UC.

Based on previous results, the proposed SSR optimizes the cardinality of the set of reduced scenarios and outperforms the fast forward selection algorithm in terms of computational speed, while selecting mostly the same scenarios. Because of the submodular property of the objective function and the use of an accelerated greedy algorithm, SSR can be solved much faster than FFS on large dataset. It is thus able to handle the very large initial set of scenarios needed to reflect uncorrelated uncertainties on demand, wind generation, solar generation and prices.

Chapter 4

ENERGY STORAGE OPERATION IN A CENTRALIZED ISO ENVIRONMENT

4.1 Motivation

Energy storage systems (ESS) have attracted significant interests for applications in the electric grid. The International Energy Agency reports that the total installed storage capacity in the US is projected to grow from 22 GW in 2014 to 103 — 152 GW in 2050 [137]. This huge amount of energy and power capacity has the potential to significantly improve grid operation. Due to the capability of storing energy, ESS is a very flexible tool, and can provide multiple services to the grid, including energy arbitrage [14], frequency regulation [138], voltage support [139], emergency support [140, 141], congestion relief [142], demand shifting [143], renewable integration [144], as well as spinning and non-spinning reserve [145]. Moreover, BESS equipped on the electric vehicles (EV) are capable of providing multiple services to the grid as well [146, 147, 148, 149, 150].

In this chapter, we focus on the application of ESS on day-ahead energy markets. In these markets, energy arbitrage is a particularly important service that ESS can provide. When an ISO has full control over the ESS, it can significantly reduce the system operating costs through energy arbitrage. During low-priced periods, energy can be purchased to charge the ESS, and during high-priced time interval, this energy can be discharged back into the grid to lower the operating cost.

A number of papers discuss various formulations for the incorporation of energy storage in the day-ahead market. Pandžić *et al.* [14] formulated a deterministic unit commitment with distributed energy storage, and the DA results are used to make near-optimal storage siting and sizing decisions. O'Dwyer [151] extended the hourly formulation to a sub-hourly

one and showed that this change further reduces system costs.

In addition to this deterministic formulation, stochastic programming scenario-based formulations have also been shown to be effective in dealing with renewable uncertainty. Li [152] evaluated the stochastic unit commitment with energy storage and achieved reduced system constraints violations in an the N-1 analysis. Pozo [18] proposed a stochastic formulation to incorporate ideal storage considering multiple timescales. Das [143] presented a unit commitment and economic dispatch framework to assess storage participation during high wind penetration scenarios.

Based on developments in operations research, robust optimization has gained much popularity in day-ahead scheduling. Jiang [19] formulated a robust unit commitment with pumped hydro units to protect the system against the worst case net injections. Li [153] proposed a storage operation framework based on an improved stochastic optimization framework, which groups scenarios in the same time interval buckets [154]. This achieves a computational and economic trade-off between a two-stage formulation and a multi-stage formulation. In addition, interval formulation were shown to be effective by Wen [16] and Bruninx [145].

Although various formulations have been proposed, no complete comparison has been made between these formulations. With an increasing level of renewable penetration, deterministic models are no longer adequate, and stochastic optimization methods are necessary. The storage planning problem is essential to the future deployment of storage, and an effective base operation model can better reflect its true potential operating advantages and disadvantages.

Storage operating cost is another component during scheduling and operation. For pumped hydro energy storage (PHES) and compress air energy storage (CAES), the operating costs are relatively low [155, 156]. On the contrary, BESS degradation costs can be very expensive. A cycle-based degradation costs are modeled in [157]. At the power electronics level, battery energy management systems are also discussed in [158, 159, 160, 161, 162, 163, 164].

In this chapter, we formulate different unit commitment models with energy storage and compare their performance in terms of day-ahead energy costs, day-ahead computing time, real-time expected operating costs, expected operating cost distribution, operating cost reduction and wind utilization.

4.2 Analysis of the Benefits of Energy Storage in UC

In this subsection, a cost-benefit analysis of energy storage operation is conducted with various UC formations. We compare deterministic UC with reserve rule (DUC), stochastic UC (SUC), interval UC (IUC) and robust UC (RUC) under different penetration levels. As in Chapter 3, both day-ahead (DA) and Monte-Carlo simulation results will be presented to produce a rigorous comparison.

4.2.1 Unit Commitment Formulations with Energy Storage

4.2.1.1 Stochastic Unit Commitment Formulation with Energy Storage

The SUC formulation without energy storage is exactly the same as in Section 3.2.1. With energy storage, the power balance constraints must be modified, and additional storage operating constraints must be added.

Revised power balance constraints to include storage components:

$$\begin{aligned} \sum_{i \in r(b)} p_{i,t,s} + \sum_{w \in r(b)} (WF_{w,t,s} - curt_{w,t,s}) + \sum_{h \in r(b)} (q_{h,t,s}^{dis} - q_{h,t,s}^{chs}) \\ - \sum_{b \in f(l)} pf_{l,t,s} + \sum_{b \in t(l)} pf_{l,t,s} = D_{b,t} - ens_{b,t,s}, \quad \forall b, t, s \end{aligned} \quad (4.1)$$

Storage Operating Constraints:

$$SoC_{h,t,s} = SoC_{h,t-1,s} + q_{h,t,s}^{chs} \eta_h^{chs} - q_{h,t,s}^{dis} / \eta_h^{dis}, \quad \forall t > 1, \forall h, s \quad (4.2)$$

$$SoC_{h,t,s} = SoC_h^{init} + q_{h,t,s}^{chs} \eta_h^{chs} - q_{h,t,s}^{dis} / \eta_h^{dis}, \quad t = 1, \quad \forall h, s \quad (4.3)$$

$$SoC_h^{min} \leq SoC_{h,t,s} \leq SoC_h^{max}, \quad \forall h, t, s \quad (4.4)$$

$$SoC_{h,t,s} = SoC_h^{init}, \quad t = N_T, \quad \forall h, s \quad (4.5)$$

$$0 \leq q_{h,t,s}^{dis} \leq Dis_h^{cap}, \quad \forall h, t, s \quad (4.6)$$

$$0 \leq q_{h,t,s}^{chs} \leq Chs_h^{cap}, \quad \forall h, t, s \quad (4.7)$$

Equations (4.2) – (4.7) implement energy storage state-of-charge (SoC) transitions and impose charging/discharging power limits are imposed. In (4.5), the ESS are required to get back to its initial SoC setpoints. In this formulation, the ESS do not participate in reserve, which would be a future extension.

4.2.1.2 Deterministic Unit Commitment with Reserve

The DUC formulation is similar to the formulation presented in Section 3.2.1. Most constraints are indeed the same. Since only one scenario is used in this formulation, the scenario subscript has been removed.

Objective Function:

$$\min \sum_{i,t} (csu_{i,t} + cpg_{i,t}) + VoLL \sum_{b,t} ens_{b,t} + VoWS \sum_{w,t} curt_{w,t} \quad (4.8)$$

The DUC minimizes the operating cost and violation penalties, including startup cost, production cost, load shedding penalty and renewable spillage penalty.

System-wide Constraints:

$$(3.6)) - (3.11) \quad (4.9)$$

$$\sum_i r_{i,t,s} \geq \alpha^{load} \sum_b D_{b,t} + \alpha^{wind} \sum_w (WF_{w,t} - curt_{w,t}), \quad \forall t \quad (4.10)$$

The reserve constraint (4.10) is added to ensure that the system has enough reserve to handle the potential wind uncertainty. The “3+5” rule is applied as the reserve requirement [10].

Generation Constraints:

$$(3.12) - (3.16), \quad (3.19), \quad (3.21) \quad (4.11)$$

$$P_i^{MIN} x_{i,t} \leq p_{i,t}, \quad \forall i, t \quad (4.12)$$

$$p_{i,t} + r_{i,t} \leq P_i^{MAX} x_{i,t}, \quad \forall i, t \quad (4.13)$$

$$p_{i,t} - p_{i,t-1} + r_{i,t} \leq RU_i, \quad \forall t > 1, \forall i \quad (4.14)$$

$$p_{i,t} - PO_i + r_{i,t} \leq RU_i, \quad \forall t = 1, \forall i \quad (4.15)$$

$$r_{i,t} \leq MSR_i \Delta T, \quad \forall i, t \quad (4.16)$$

Wind Constraint and storage operation constraints are the same as (3.22) and (4.2) – (4.7).

4.2.1.3 Interval Unit Commitment

Without energy storage, the formulation is exactly as the same as the one presented in Section 3.2.2. With energy storage, in addition to the revised power balance (4.1), storage operation constraints (4.2) – (4.7), additional ramp transitional constraints must be added for the energy storage. These constraints ensure that every storage operating target is reachable during the bound transitions.

$$SoC_{h,t,0} - SoC_{h,t+1,lb} \leq Dis_h^{cap}, \quad \forall t < T, \forall h \quad (4.17)$$

$$SoC_{h,t+1,ub} - SoC_{h,t,0} \leq Chs_h^{cap}, \quad \forall t < T, \forall h \quad (4.18)$$

$$SoC_{h,t,ub} - SoC_{h,t+1,lb} \leq Dis_h^{cap}, \quad \forall t < T, \forall h \quad (4.19)$$

$$SoC_{h,t,lb} - SoC_{h,t+1,ub} \leq Chs_h^{cap}, \quad \forall t < T, \forall h \quad (4.20)$$

4.2.1.4 Robust Unit Commitment

In the RUC formulation, the objective is to minimize the operating cost under the worst case scenario. The problem is in a min max min form. The middle and lower levels are transformed into a single level equivalent using strong duality or KKT conditions [20]. Benders decomposition with outer approximation [19] and Column & Constraint-Generation (CCG)

[116] are widely used to solve this problem. Here, we present the basic formulation, and a detailed derivation for the dual forms and solution techniques can be found in the [116].

Objective Function:

$$\min \sum_{i,t} csu_{i,t} + \max_{w,t} \min_{p_{i,t}} \left(\sum_{i,t} cpg_{i,t} + VoWS \sum_{w,t} curt_{w,t} \right) \quad (4.21)$$

The generation constraints related to binary variables are included in the upper level. The middle and lower levels solve a worst case scenario dispatch problem with given first-stage schedules. The other dispatch related constraints are the same as in the deterministic case without reserve. The wind uncertainty constraints are defined in (4.22) – (4.25). In equations (4.22) – (4.23), uncertain wind should stay within the lower and upper bounds from the forecast. In order to control the conservative level, Γ^w and Γ^t are defined in (4.24) and (4.25) as spatial and temporal uncertainty budgets to limit wind variations. Through these constraints, we separate the original RUC problem into a master problem and a subproblem. Master problem is optimized to determine the optimal generation commitment, and the subproblem to find the worst case scenario candidates. These candidates generate feasibility and optimality cuts back to the master problem at each iteration until it converges.

$$w_{w,t} \leq WF_{w,t}^{UB}, \quad \forall w, t \quad (4.22)$$

$$w_{w,t} \geq WF_{w,t}^{LB}, \quad \forall w, t \quad (4.23)$$

$$\sum_w \frac{|w_{w,t} - WF_{w,t,0}|}{WF_{w,t,0}} \leq \Gamma^w, \quad \forall t \quad (4.24)$$

$$\sum_t \frac{|w_{w,t} - WF_{w,t,0}|}{WF_{w,t,0}} \leq \Gamma^t, \quad \forall w \quad (4.25)$$

4.2.2 Simulation Setup

The four UC formulations (DUC, SUC, IUC and RUC) described above have been tested on a modified version of the 73-bus three-area IEEE RTS system [119]. This system includes a total of 96 conventional generators with 10,215 MW of installed capacity. The generator

configurations and cost curves are as in Section 3.2.3 and reference [120]. There are also 19 wind farms with 6,900 MW of nameplate capacity. Detailed location and capacity information can be found in [14]. The value of wind spillage is set at \$ 0/MWh and the value of lost load is set at \$ 7500/MWh.

To present the value of energy storage in operation, near-optimal ESS siting and sizing decisions are based on [14] and listed in Table 4.2. These storage configurations are determined through a greedy selection algorithm where the increased number of storage affects the power and energy rating decisions. Therefore, ESS at the same bus have different ratings when the number of storage changes. The case without energy storage serves as a benchmark to evaluate the cost savings from ESS.

Five typical days were selected to represent different wind penetration levels with 10 %, 16 %, 26 %, 35 % and 45 % as shown in Table 4.1. The wind scenarios were obtained using the ensemble approach discussed in Section 2.3. A total of 2000 scenarios were generated for each typical day, 1000 scenarios for DA scheduling, and the remaining 1000 for Monte-Carlo testing purposes.

Adopting the same operating structure as in Section 3.2.3, the day-ahead procedure determines the optimal unit schedules. For the deterministic unit commitment, the “3+5” rule is applied for the reserve as in [10]. For the stochastic unit commitment, 10 scenarios are used to minimize the expected costs [103, 165]. For the interval unit commitment, a 90 % confidence interval from the full scenario set was adopted. The robust unit commitment used the same bounds as the IUC with a budget of uncertainty of 4 for both time and space. The cost calculated with the SUC is the expected DA operating cost, while the cost calculated with the RUC is the worst case operating cost. The day-ahead costs are thus projected DA costs. Monte-Carlo simulations are used to assess the actual operating costs and the schedule performances based on these various DA schedules.

Table 4.1: Representative day wind generation

Day	1	2	3	4	5
Wind Penetration Level	10 %	16 %	26 %	35 %	45 %

Table 4.2: Energy storage installation in the system

Storage	Total Power Rating (MW)	Total Energy Rating (MWh)	Itemized Power Ratings (MW)	Itemized Energy Ratings (MWh)	Location
1	84	457	84	457	Bus-325
2	231	1500	171	1146	Bus-121
			60	354	Bus-325
3	262	1719	150	1014	Bus-121
			60	403	Bus-202
			52	302	Bus-325
4	290	1915	85	597	Bus-116
			104	698	Bus-121
			49	318	Bus-202
			52	302	Bus-325

4.2.3 Case Study Results

4.2.3.1 Day-Ahead Operation

Table 4.3 – Table 4.7 present the day-ahead total operating cost for 5 representative days for the different UC formulations and system storage installations. As would be expected, when the system is equipped with more storage, the operating cost decreases for all formulations

and most cases. However, different UC formulations produce different reductions. The Interval UC with $[0.31, 1.70]$ % and the Stochastic UC with $[0.21, 1.60]$ % achieve the best daily cost savings. On average they obtain around 0.50 % daily savings most of the time. For higher wind generations, higher cost savings are expected from the storage as shown in Day 3 and Day 5. In Day 4, although the wind generation is high, the high energy prices at the end of the optimization period result in a larger cost to recover the storage stage-of-charge and a smaller cost saving. In general, these four formulations are all effective in achieving spatial-temporal energy arbitrage using storage. IUC and SUC perform especially well.

Comparing these four formulations, the scenario-based stochastic formulation consistently provides the lowest operating cost. The robust formulation schedules against the worst case scenario and its cost is thus higher. However this DAC is just a projection. The interval formulation performance is between those of the deterministic and stochastic formulations.

Table 4.3: Day-Ahead Cost (DAC) in \$ 10^6 for Day 1 with 10 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	2.2868	2.2813	2.2788*	2.3035
1	2.2837	2.2764	2.2750*	2.2995
2	2.2780	2.2708	2.2690*	2.2946
3	2.2784	2.2698	2.2673*	2.2939
4	2.2765	2.2695	2.2679*	2.2923
Maximum Cost Savings	0.45 %	0.52 %	0.50 %	0.49 %

* - minimum DAC for a given number of storages

Table 4.4: Day-Ahead Cost (DAC) in \$ 10^6 for Day 2 with 16 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	2.5957	2.5862	2.5864*	2.6248
1	2.5925	2.5829	2.5831*	2.6210
2	2.5819	2.5732	2.5735*	2.6138
3	2.5812	2.5724	2.5718*	2.6118
4	2.5807	2.5713	2.5693*	2.6108
Maximum Cost Savings	0.58 %	0.58 %	0.66 %	0.53 %

* - minimum DAC for a given number of storage

Table 4.5: Day-Ahead Cost (DAC) in \$ 10^6 for Day 3 with 26 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3613	1.3605	1.3585*	1.3955
1	1.3601	1.3587	1.3571*	1.3946
2	1.3491	1.3464	1.3460*	1.3916
3	1.3488	1.3460	1.3458*	1.3909
4	1.3472	1.3445	1.3443*	1.3971
Maximum Cost Savings	1.04 %	1.18 %	1.04 %	0.33 %

* - minimum DAC for a given number of storage

Table 4.6: Day-Ahead Cost (DAC) in \$ 10^6 for Day 4 with 35 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3148	1.3160	1.3120*	1.3807
1	1.3138	1.3141	1.3105*	1.3803
2	1.3131	1.3119	1.3092*	1.3797
3	1.3129	1.3124	1.3093*	1.3799
4	1.3134	1.3120	1.3092*	1.3798
Maximum Cost Savings	0.14 %	0.31 %	0.21 %	0.07 %

* - minimum DAC for a given number of storage

Table 4.7: Day-Ahead Cost (DAC) in \$ 10^6 for Day 5 with 45 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3355	1.3346	1.3308*	1.3881
1	1.3319	1.3292	1.3254*	1.3843
2	1.3189	1.3166	1.3136*	1.3715
3	1.3166	1.3141	1.3114*	1.3695
4	1.3145	1.3120	1.3095*	1.3673
Maximum Cost Savings	1.57 %	1.70 %	1.60 %	1.50 %

* - minimum DAC for a given number of storage

Table 4.8 – Table 4.12 give the computing time for five representative days and different system configurations. Since the stochastic formulation uses a set of uncertainty scenarios, it requires the longest computing time among the four UC formulations. The UC formulation with deterministic reserve tends to be the fastest because only the central forecast is used in the decision-making processes. The IUC and RUC perform similarly well and in a consistent manner. Another interesting finding is that the energy storage can significantly reduce the computing time especially for the IUC and SUC formulations. The reason is probably that the energy storage provides more flexibility to the system and helps the MILP solver find an integer solution faster. Although the SUC computing time is relatively high, using a computer with more cores and memory would reduce the computing time. In addition, the development of optimization theory and decomposition algorithms also contributes to solution time reductions for stochastic formulations. In the future, the computational gap between these formulations is expected to be smaller. However, under current computational resources, stochastic programming still takes longer than other formulations.

Table 4.8: Day-Ahead computing time (seconds) for Day 1 with 10 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	67	1112	9018	181
1	68	73	1578	88
2	37	110	1004	66
3	21	61	1845	67
4	23	65	485	62

4.2.3.2 Real-Time Re-dispatch and Actual Operating Cost

Table 4.13 – Table 4.17 present the expected actual operating cost under different system configurations. Similarly to the conclusions drawn in the previous section, the installation

Table 4.9: Day-Ahead computing time (seconds) for Day 2 with 16 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	342	53	1796	156
1	52	145	792	182
2	30	157	552	719
3	18	290	310	139
4	23	211	891	112

Table 4.10: Day-Ahead computing time (seconds) for Day 3 with 26 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	72005	64843	72011	518
1	127	422	582	143
2	36	70	489	100
3	2154	118	395	130
4	33	105	332	157

Table 4.11: Day-Ahead computing time (seconds) for Day 4 with 35 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	709	475	1084	715
1	16	40	683	731
2	10	63	1056	731
3	13	42	338	162
4	10	129	514	201

Table 4.12: Day-Ahead computing time (seconds) for Day 5 with 45 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1327	8377	10099	273
1	151	184	752	802
2	57	43	766	54
3	30	129	338	223
4	156	95	804	50

of energy storage significantly reduces the operating cost in all cases. The cost savings is similar at the DA stage, which shows the consistency of the results. High cost savings are obtained for high wind generation (Day 3 and Day 5), moderate cost savings at low wind generation (Day 1 and Day 2), small cost savings at moderate wind generation (Day 3). Since the cost savings not only depends on the wind generation level, but also on the load profile and the locational marginal prices (LMP), the cost savings are satisfactory because millions of dollars could be saved through this arbitrage annually, on top of potential revenue gained through participation in the ancillary services markets.

The IUC and DUC formulations optimize the central forecast operating cost at the DA stage, and these storage siting and sizing decisions are selected based on the optimized average daily operating cost from the central forecast. Therefore, this comparison is slightly biased in favor of the IUC and DUC results. Even so, the SUC performs best among all cases, and always produces the lowest operating costs. This demonstrates the importance of properly constructed uncertainty sets and the value of stochastic programming in UC applications. The performance of the RUC falls between those of the IUC and DUC. Recently proposed advances in RUC formulations [166] might further reduced the AOC for this formulation.

Table 4.13: Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 1 with 10 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	2.2864	2.2811	2.2759*	2.2792
1	2.2832	2.2762	2.2717*	2.2749
2	2.2775	2.2705	2.2670*	2.2706
3	2.2777	2.2696	2.2659*	2.2695
4	2.2760	2.2693	2.2651*	2.2683
Maximum Cost Savings	0.45 %	0.52 %	0.48 %	0.48 %

* - minimum AOC for a given number of storage

Table 4.14: Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 2 with 16 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	2.5952	2.5864	2.5841*	2.5886
1	2.5917	2.5829	2.5807*	2.5839
2	2.5812	2.5731	2.5717*	2.5753
3	2.5800	2.5720	2.5697*	2.5732
4	2.5792	2.5708	2.5680*	2.5728
Maximum Cost Savings	0.61 %	0.60 %	0.62 %	0.61 %

* - minimum AOC for a given number of storage

Table 4.15: Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 3 with 26 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3616	1.3612	1.3569*	1.3604
1	1.3601	1.3592	1.3562*	1.3595
2	1.3484	1.3474	1.3455*	1.3566
3	1.3481	1.3462	1.3451*	1.3551
4	1.3473	1.3451	1.3437*	1.3610
Maximum Cost Savings	1.05 %	1.19 %	0.97 %	0.39 %

* - minimum AOC for a given number of storage

Table 4.16: Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 4 with 35 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3160	1.3170	1.3107*	1.3188
1	1.3146	1.3151	1.3099*	1.3183
2	1.3133	1.3125	1.3086*	1.3158
3	1.3134	1.3130	1.3086*	1.3152
4	1.3131	1.3125	1.3086*	1.3150
Maximum Cost Savings	0.22 %	0.35 %	0.16 %	0.29 %

* - minimum AOC for a given number of storage

Table 4.17: Expected Actual Operating Cost (AOC) in \$ 10^6 for Day 5 with 45 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	1.3359	1.3348	1.3294*	1.3363
1	1.3321	1.3295	1.3247*	1.3320
2	1.3190	1.3170	1.3131*	1.3194
3	1.3166	1.3144	1.3109*	1.3173
4	1.3145	1.3122	1.3089*	1.3150
Maximum Cost Savings	1.60 %	1.69 %	1.54 %	1.60 %

* - minimum AOC for a given number of storage

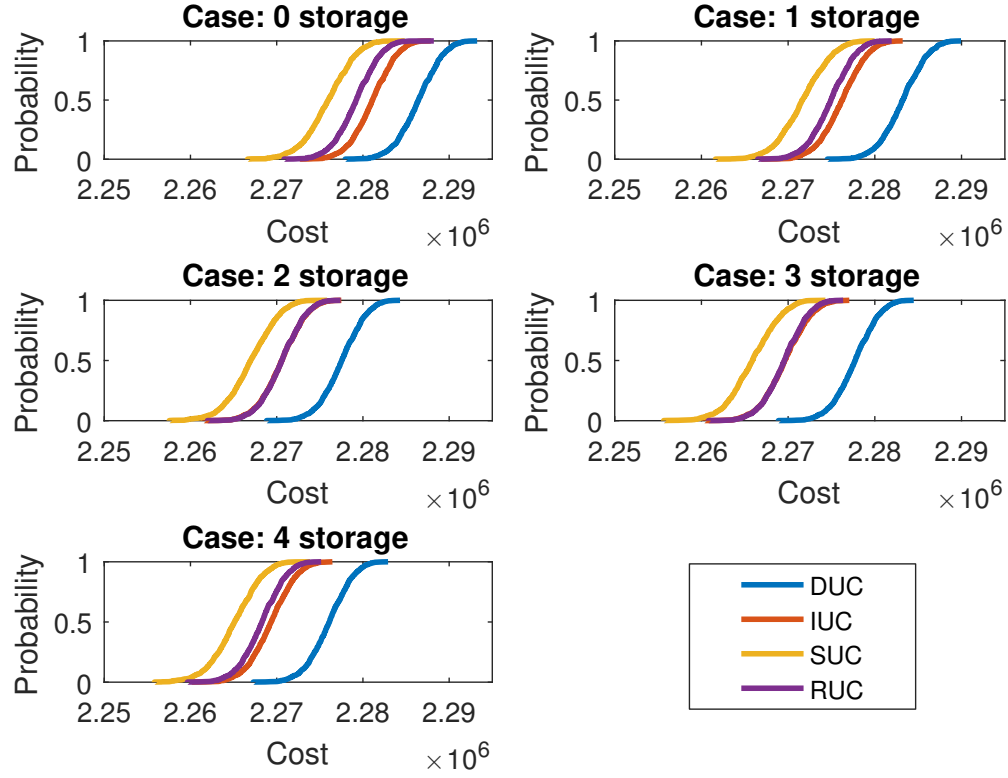


Figure 4.1: Comparison of the AOC cost distribution at Day 1

Figure 4.1 – Figure 4.5 present the AOC distribution under different system configurations. The trend of cost reduction is clear, but the variance does not change much with more storage. In addition, the SUC causes the largest span of cost, whereas the RUC results in the least span. This confirms the findings of reference [23].

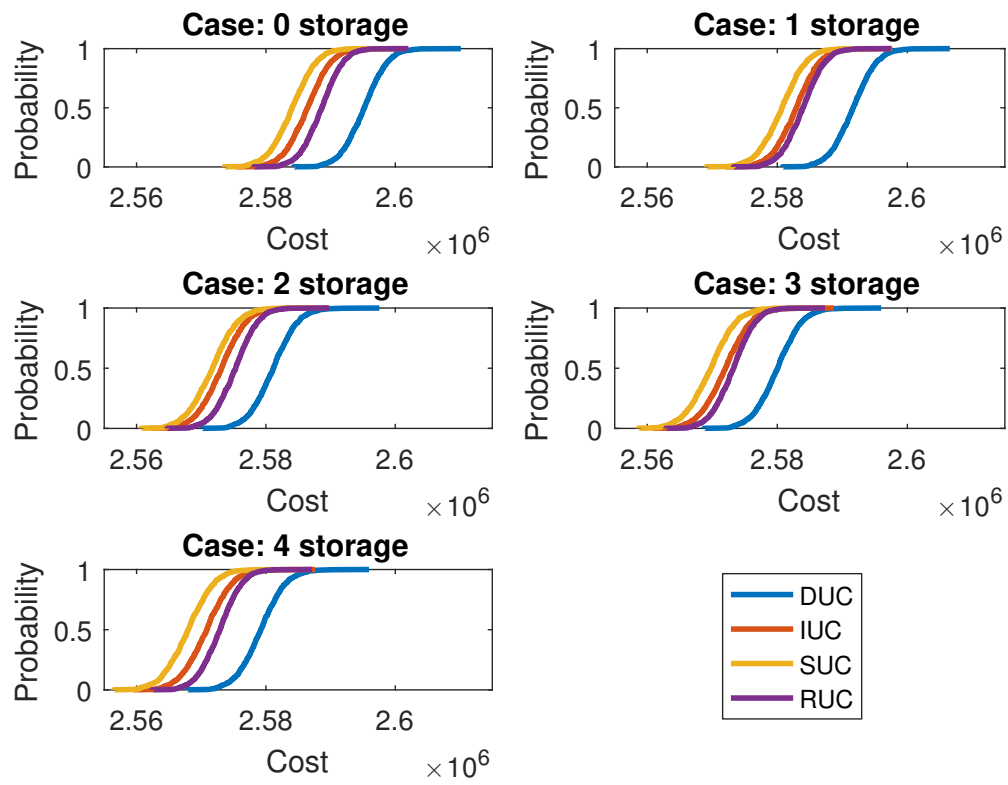


Figure 4.2: Comparison of the AOC cost distribution at Day 2

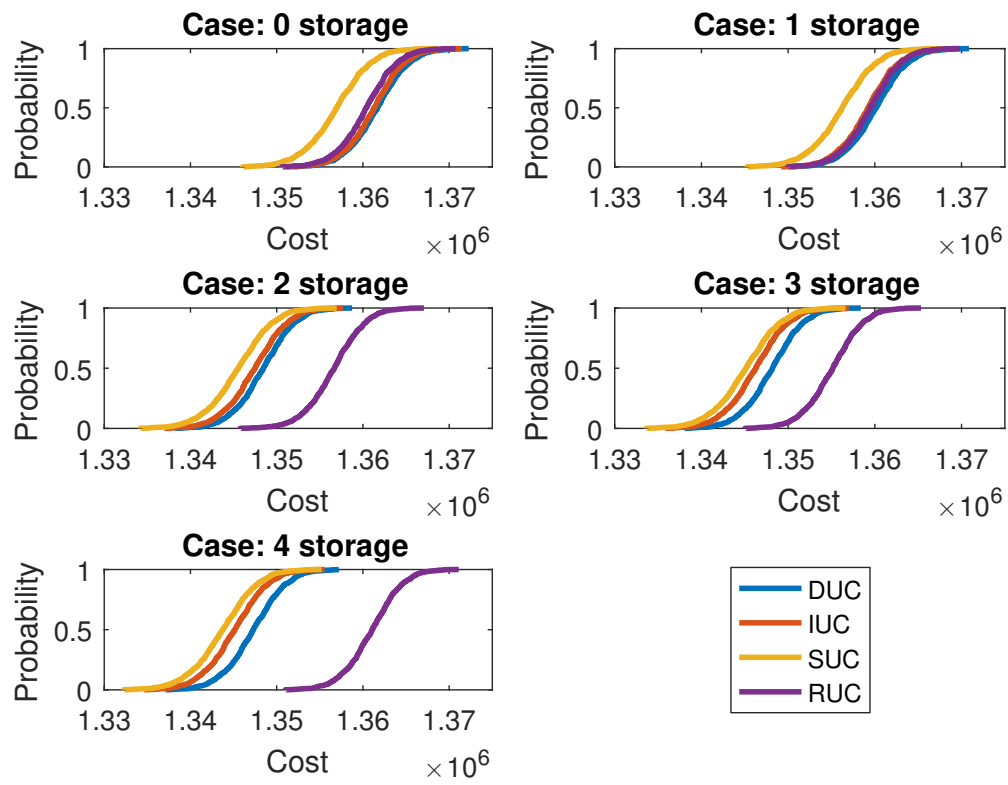


Figure 4.3: Comparison of the AOC cost distribution at Day 3

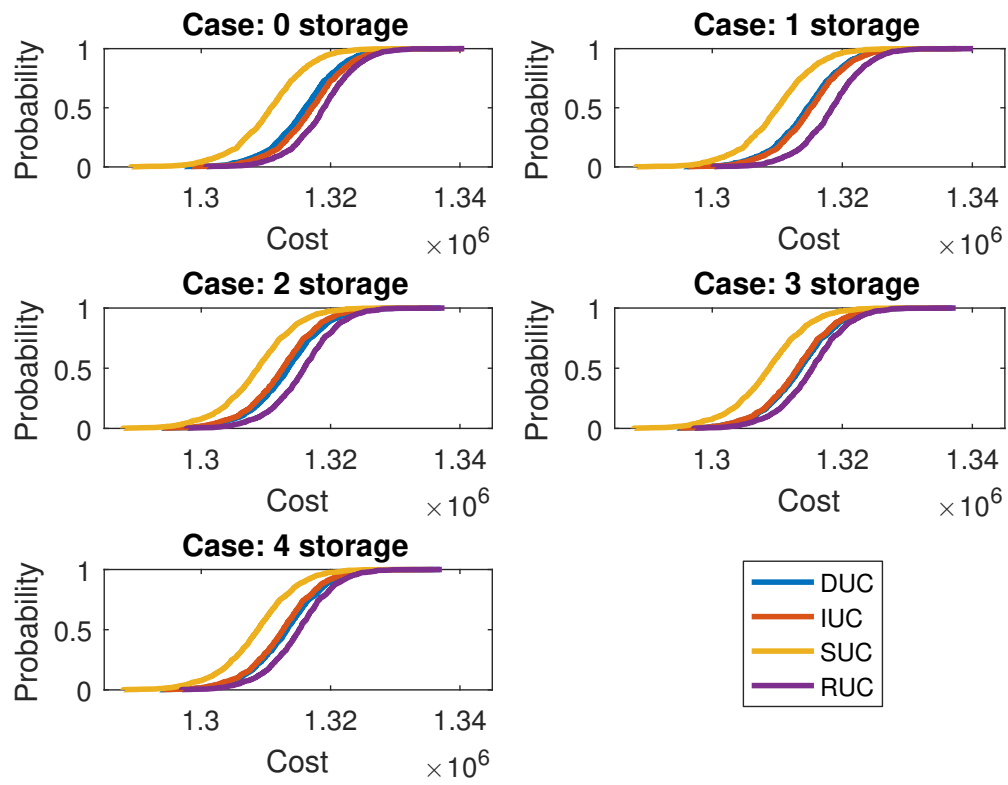


Figure 4.4: Comparison of the AOC cost distribution at Day 4

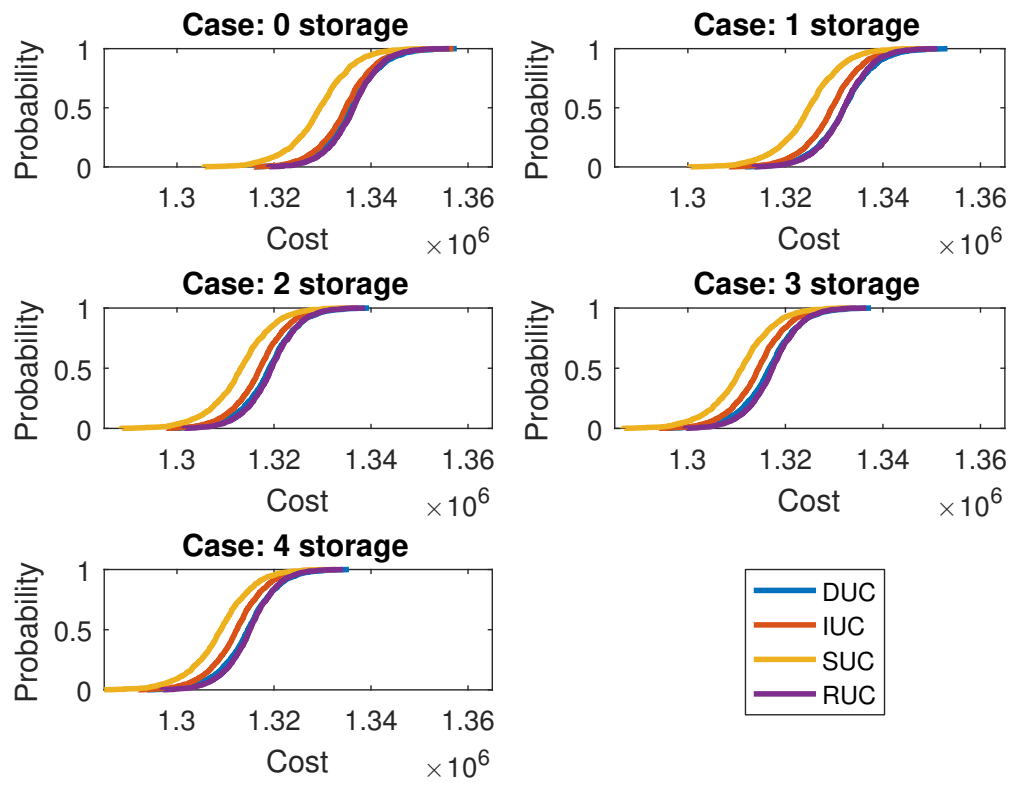


Figure 4.5: Comparison of the AOC cost distribution at Day 5

Table 4.18 presents the expected daily wind spillage for Day 5. The different UC formulations have almost the same effect on the utilization of wind. It is really the amount of storage that promotes the wind utilization. Wind curtailment was reduced by nearly 33 % from the reference cases without any energy storage. For other representative days, little wind spillage occurs in the reference case, and the installation of storage reduces it to zero.

Table 4.18: Expected daily wind spillage (MWh) for Day 5 with 45 % wind generation

Number of Storage	DUC	IUC	SUC	RUC
0	2748	2748	2747	2748
1	2704	2704	2704	2704
2	2038	2038	2038	2038
3	1933	1932	1934	1933
4	1831	1831	1831	1831

4.3 Summary

In this section, we presented a an analysis of the benefits of integrating energy storage in the power system. The deterministic unit commitment with reserve, interval unit commitment, stochastic scenario-based unit commitment and robust unit commitment are compared in terms of their day-ahead projected cost, day-ahead computing time, expected actual operating cost and wind utilization. Energy storage devices effectively reduce the system cost as well promote more wind utilization. The stochastic scenario-based unit commitment formulation consistently produces the least cost schedule but with the highest computational burden. The IUC and RUC require less computational resources yet do not achieve the same economic benefits. The deterministic UC serves as the benchmark, and “3+5” rule is over-conservative especially for a system with energy storage.

Chapter 5

ENERGY STORAGE OPERATION FROM MERCHANT OWNER PERSPECTIVE

5.1 Motivation

Due to its declining installation and operation cost, energy storage systems are likely to play an increasing role not only in power system operation but also in electricity markets. Because of its flexibility, storage can be used for energy arbitrage [14], transmission congestion relief [167], reserves [18], frequency regulation [16], post-contingency corrective actions [141] and other services. In order to encourage the deployment of energy storage, market rules are being revised to support products and pricing schemes better suited to the technical characteristics and constraints of storage systems [138].

In a vertically-integrated environment, energy storage is used to minimize the system operating cost. For example, Pandžić *et al.* [14], Qiu *et al.* [168] and Fernández-Blanco *et al.* [169] investigated the optimal siting and sizing of storage systems used for spatio-temporal arbitrage. The authors of [170] proposed a bilevel storage planning strategy that guarantees the recovery of the investment costs. Wogrin *et al.* [171] proposed to allocate storage for load-shifting and regulation services. Wen *et al.* [16] demonstrated the effectiveness of storage for post-contingency corrective actions. However, minimizing the operating cost, or more generally maximizing the benefits to the system, typically does not lead to a charging/discharging schedule that maximizes the profits of a merchant storage owner.

Instead of being fully controlled by the ISO, energy storage could participate actively in the wholesale electricity market. Private entities could buy and install their own storage devices to capture their revenue potential [137]. For these entities, reducing system cost is not the primary concerns. Instead, their goal is to earn as much profit as possible from

the market. In this sense, letting the ISO fully control the storage is probably not the best approach to maximize profits. Storage owners have two additional options to avoid centralized ISO mode, operating storage on their own or finding an aggregator to coordinate privately-owned storage. This aggregator is responsible for providing bidding and offering price/quantity pairs to the system, and the market cleared schedule is used to guide the storage state-of-charge movements.

Operation and planning of storage in a competitive market environment is therefore gaining increasing attention. Mohsenian-Rad [172] schedules the storage based on the demand and supply bidding optimization which takes the price uncertainty into account. Outer and inner problems are solved to obtain a globally optimal solution. The same authors extend their work by considering the energy and reserve market together with stochastic programming methods [173]. Xu [174] considers the participation of storage in the frequency regulation market including its degradation cost. Similarly, He *et al.* [175] determined the participation of storage in regulation services considering the storage performance score and the degradation from the storage cycling. Shafiee [156] proposed to use information gap decision theory (IGDT) to optimize compressed air energy storage (CAES). Mohsenian-Rad [176], Pandžić [177] and Hartwig [178] all proposed to use bilevel programming to model storage operation.

ESS can also coordinate with other resources in the market. Ding [179] proposed to bid and offer storage and a wind farm together with probabilistic forecast information in a rolling real-time market. A modified gradient descent is developed to solve the optimization problem. In the same authors' extended work [180], linear decision rules are applied to formulate the problem, and solved efficiently with commercial solvers. Thatte [181] proposed to use robust optimization to control the risk when storage is bidding together with wind farms. In addition to coordination with wind, virtual power plants (VPP), which aggregate demand response, small gas turbines, diesel generators, enter the market [182]. There is an abundant literature on the bidding strategies for electric vehicles (EV) which can be viewed as another storage resource [183]. Due to their special characteristics, EVs are excluded from

this discussion.

Information about rival generators' offers are also vital in optimizing storage bids and offers. In [184], strategic power producers calculated optimal offers considering the competing generators' participation. An inverse optimization was used to reveal the marginal offers from rival generators in [185]. Kazemi *et al.* [186] proposed a robust bidding model considering the uncertainty on competing generators' price and quantity offers. Equilibrium problems with equilibrium constraints (EPEC) have also been used to assess the interactions between multiple competing energy companies in [187, 188, 189].

Although many papers have discussed the operation storage, there are still some areas not covered in the literature. In this chapter, we first present a deterministic bilevel formulation to determine the optimal storage bidding and offering price/quantity pair. Then, we extend this strategy by a look-ahead setting with stochastic future information. The final stage of charge is also optimized, which is commonly neglected in the literature. The impacts of thermal generator ramping constraints on storage profitability are also evaluated.

5.2 Impact of Local Transmission Congestion on Energy Storage Arbitrage Opportunities

5.2.1 Introduction

Castillo *et al.* [190], Shafiee *et al.* [156], and Mohsenian-Rad [172] proposed techniques to maximize the operating profits of storage in a decentralized environment. Ding *et al.* [191, 192] coordinated the operation of storage and wind farms in a rolling real-time market. Xu *et al.* [138] analyzed the batteries' profits for regulation services considering its degradation costs [193]. Khani *et al.* [142] proposed to perform arbitrage while relieving transmission congestion. Relative to these works, this section's models impacts of local and system-wide transmission congestion and bidding of other market participants.

Several authors [194, 195] have shown that complementarity modeling is an effective tool to analyze the interactions between two or more market participants. In power systems, bilevel programming has been applied in vulnerability analysis [196], price-based market

clearing [197], robust unit commitment [166], trading wind power in futures market [198], generation investment [199], investments in wind power [200], bidding strategies for wind power [201], operation of storage [176] and investments in storage [170].

A large number of papers discuss how network constraints and transmission congestion affect electricity markets and influence the bidding strategy of market participants, *e.g.* [202, 203, 204]. However, previous work focused on system-wide transmission congestion, while this section examines how storage can affect or be affected by local transmission congestion.

A network-constrained market-clearing mechanism with storage participation is analyzed under perfect and imperfect competition. Perfect competition is modeled as an economic dispatch problem, while imperfect competition is modeled using a bilevel approach where the storage owner behaves strategically. The effect of local transmission congestion on the annual operating profit of storage is quantified to show how it affects the behavior of storage owners. The importance of optimizing siting and sizing decisions on the storage profitability is discussed.

5.2.2 Mathematical Formulation

This section provides the mathematical formulation of a network-constrained market clearing with storage participation under perfect and imperfect competition.

5.2.2.1 Market Clearing under Perfect Competition

Under perfect competition, each market participant bids or offers at its marginal cost because none of them is assumed to be able to exercise market power. The ISO then clears the market in a way that maximizes the social welfare. Storage owners participate in the market by submitting bids and offers that reflect their willingness to charge and discharge. These bids and offers should take into account how battery cycling affects the life of the battery, *i.e.* its incremental degradation cost. Market clearing takes the form of a network-constrained

economic dispatch problem. Storage owners and other market participants pay or are paid based on their locational marginal prices (LMPs).

$$\max \quad Obj^{SO} = \sum_{b,t} C_b^D d_{b,t} + \sum_{h,t} C_h^{bid} q_{h,t}^{chs} \quad (5.1)$$

$$- \sum_{i,t} C_i^G p_{i,t}^G - \sum_{h,t} C_h^{dis} q_{h,t}^{dis}$$

$$0 \leq q_{h,t}^{chs} \leq \overline{chs}_h x_{h,t}^{chs}, \forall h, \forall t \quad (5.2)$$

$$0 \leq q_{h,t}^{dis} \leq \overline{dis}_h x_{h,t}^{dis}, \forall h, \forall t \quad (5.3)$$

$$x_{h,t}^{dis} + x_{h,t}^{chs} \leq 1, \forall h, \forall t \quad (5.4)$$

$$x_{h,t}^{dis}, x_{h,t}^{chs} \in \{0, 1\}, \forall h, \forall t \quad (5.5)$$

$$SoC_{h,t} = SoC_{h,t-1} + q_{h,t}^{chs} \eta_h^{chs} - q_{h,t}^{dis} / \eta_h^{dis}, \forall t > 1, \forall h \quad (5.6)$$

$$SoC_{h,t} = SoC_h^{init} + q_{h,t}^{chs} \eta_h^{chs} - q_{h,t}^{dis} / \eta_h^{dis}, t = 1, \forall h \quad (5.7)$$

$$\underline{SoC}_h \leq SoC_{h,t} \leq \overline{SoC}_h, \forall h, \forall t \quad (5.8)$$

$$SoC_{h,N_T} = SoC_h^{init}, \forall h \quad (5.9)$$

$$0 \leq p_{i,t}^G \leq \bar{P}_i, \forall i, \forall t \quad (5.10)$$

$$0 \leq p_{w,t}^W \leq WF_{w,t}, \forall w, \forall t \quad (5.11)$$

$$\underline{D}_{b,t} \leq d_{b,t} \leq \bar{D}_{b,t}, \forall b, \forall t \quad (5.12)$$

$$d_{b,t} + \sum_{b|f(l)=b} p_{f,l,t} - \sum_{b|t(l)=b} p_{f,l,t} = \quad (5.13)$$

$$\sum_i p_{b(i),t}^G + \sum_w p_{b(w),t}^W + \sum_h (q_{b(h),t}^{dis} - q_{b(h),t}^{chs}), \forall b, \forall t$$

$$p_{f,l,t} = \frac{1}{X_l} (\theta_{f(l),t} - \theta_{t(l),t}), \forall l, \forall t \quad (5.14)$$

$$-\bar{F}_l \leq p_{f,l,t} \leq \bar{F}_l, \forall l, \forall t \quad (5.15)$$

$$-\bar{\theta} \leq \theta_{b,t} \leq \bar{\theta} \forall b \neq \text{ref}, \forall t \quad (5.16)$$

$$\theta_{b,t} = 0, b = \text{ref}, \forall t. \quad (5.17)$$

The objective function (5.1) maximizes the social welfare which includes generation and storage discharging offers as well as the consumers and storage charging bids. Constraints (5.2)–(5.3) set the limits on the storage bid/offer quantities. Eqs. (5.4)–(5.5) prevent simultaneous charging and discharging by enforcing constraints on binary variables. Constraints (5.6) – (5.8) track the state of charge (SoC) and enforce the limits on their operating range. Constraint (5.9) forces the final SoC to be identical to the initial SoC. Constraints (5.10) – (5.12) enforce the lower and upper bounds on the thermal generation, wind farms, and consumers. Constraint (5.13) is the nodal power balance. Constraint (5.14) calculates the power flows using a DC load flow model. Constraints (5.15)–(5.17) enforce the limits on the line flows and the voltage angles.

5.2.2.2 Market Clearing under Imperfect Competition

In an imperfectly competitive market, storage can achieve larger profits through strategic bidding and offering. This strategic behavior can be modeled as a bilevel program that captures the interactions between the merchant storage and the ISO. In the upper level, storage maximizes its operating profits and determines the price/quantity bids and offers to be submitted to the ISO. The lower level represents a network-constrained market clearing as described in the previous section. The accepted bids, offers and market-clearing locational marginal prices are fed back to the upper level where they are used to calculate the profits of storage.

$$\max Obj^{ESS} = \sum_{h,t} [\lambda_{b(h),t}(q_{h,t}^{dis} - q_{h,t}^{chs}) - C_h^{dis} q_{h,t}^{dis} - C_h^{chs} q_{h,t}^{chs}] \quad (5.18)$$

$$\rho_{h,t}^{dis}, \rho_{h,t}^{chs} \geq 0, \forall h, \forall t \quad (5.19)$$

$$0 \leq chs_{h,t} \leq \overline{chs}_h x_{h,t}^{chs}, \forall h, \forall t \quad (5.20)$$

$$0 \leq dis_{h,t} \leq \overline{dis}_h x_{h,t}^{dis}, \forall h, \forall t \quad (5.21)$$

$$(5.4) - (5.9) \quad (5.22)$$

$$\lambda_{b,t}, q_{h,t}^{dis}, q_{h,t}^{chs} \in \arg \max \left\{ SW = \sum_{b,t} C_b^D d_{b,t} \right. \\ \left. + \sum_{h,t} \rho_{h,t}^{chs} q_{h,t}^{chs} - \sum_{h,t} \rho_{h,t}^{dis} q_{h,t}^{dis} - \sum_{i,t} C_i^G p_{i,t}^G \right. \quad (5.23)$$

$$0 \leq q_{h,t}^{chs} \leq chs_{h,t}, \forall h, \forall t \quad (5.24)$$

$$0 \leq q_{h,t}^{dis} \leq dis_{h,t}, \forall h, \forall t \quad (5.25)$$

$$(5.10) - (5.17) \left. \right\}. \quad (5.26)$$

The upper-level objective function (5.18) maximizes the storage profits based on the LMPs and the cleared charging/discharging quantities. Constraint (5.19) enforces the non-negativity of bid and offer prices submitted by the storage. This constraint could be relaxed in systems with high renewable penetration where negative electricity prices can occur. Constraints (5.20)–(5.21) set the limits on the quantity bids and offers. Other constraints on storage operation are identical to those described in the previous section. LMPs and cleared quantities are obtained from the lower-level problem. This lower-level problem maximizes the social welfare (5.23). Partial-bids and offers for storage charging and discharging are accepted in constraints (5.24)–(5.25). Other constraints on market clearing (5.26) are again identical to those described in the previous section.

This bilevel formulation is nonlinear and non-convex. However, under the assumption of convexity of the lower level, it can be parametrized using the Karush–Kuhn–Tucker optimality conditions. The complementary slackness conditions are further linearized using the Fortuny-Amat and McCarl transformation [196]. The nonlinear terms in the objective function are also linearized using the strong duality condition. The resulting single-level equivalent is a mixed-integer linear program that can be solved with commercial solvers. Interested readers are referred to [205, 206] for further details.

5.2.3 Case Study

The proposed market models have been tested on the modified version of the IEEE Reliability Test System (RTS) [119]. This test system consists of 24 buses, 32 generators, 38 transmission

lines, and 5 wind farms. The day-ahead wind power forecasts were generated based on the NREL Eastern Wind Dataset [47] and allow us to assess the proposed models over a one-year simulation horizon. We assume one storage with the optimized parameters from [14]: $\overline{dis}_h = \overline{chs}_h = 93$ MW, $\overline{SoC}_h = 629$ MWh, $\eta_h^{dis} = \eta_h^{chs} = 0.9$, $SoC_h^{init} = 315$ MWh, $C_h^{dis/chs} = 0$ \$/MWh, $C_h^{bid} = 30$ \$/MWh.

All simulations were carried out using GAMS 23.7 and CPLEX 12.5 on a Intel Xenon 2.55 GHz processor with 32 GB RAM. The computation time for each single day is less than or equal to 5 s with the MILP gap set to 0.005 %.

To examine the effect of local congestion on the transmission network, we first scale up the original line capacities by 50% to ensure that there is no congestion in the network, and that the LMP is the same at all buses. We then gradually reduce the capacities of the lines connected to the bus where the storage is located. Other line capacities remain unchanged in order to avoid congestion across the system. We adopted this approach because we wanted to focus on what happens when storage is not able to deliver its flexibility. By comparing operating profits when locating storage at different buses, we show how siting decisions affect storage profitability.

Figure 5.1 shows how the annual profit collected by the storage changes as the local line capacities are reduced when this storage is located at four different buses. Decreasing these line capacities tends to increase the LMP, which in turn tends to improve the profitability of storage for all four locations. However, these figures show that different patterns are possible. For example, if the storage is located at buses 2, 18 or 19, at some point the increase in LMP resulting from a reduction in transmission capacity is offset by a reduction in the amount of energy that the storage can physically deliver to the rest of the system. On the other hand, if the storage is located at bus 14, profitability increases monotonically as the local transmission capacity decreases. Factors such as the location of conventional and wind generation, the size and location of the loads, the topology of the network as well as the transmission capacity at nearby buses determine how local congestion affects storage profitability. While patterns for perfect and imperfect competition are similar, strategic

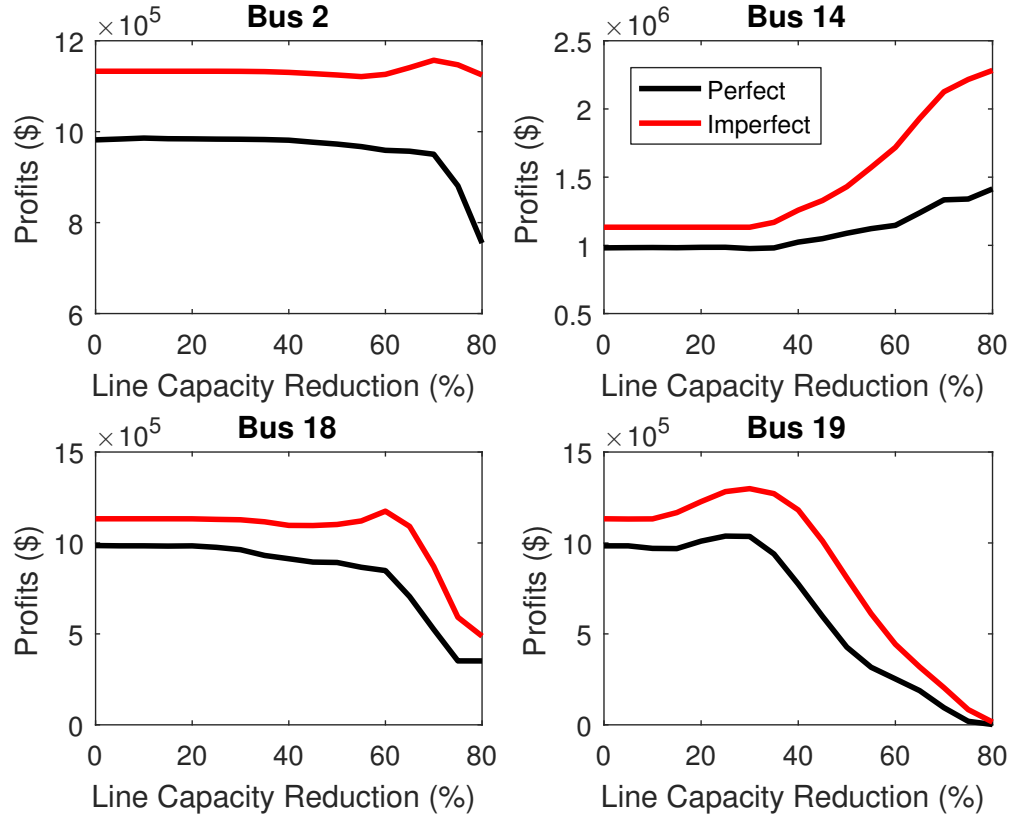


Figure 5.1: Typical patterns of annual storage profits.

bidding significantly increases profitability particularly when local congestion is significant.

Figure 5.2 shows the annual profit and social welfare for all possible storage locations under a 50% line capacity reduction. Imperfect competition enhances storage profitability for all locations but causes a reduction in the social welfare. Compared with the base case, storage can collect more or less profits depending on the location. This pattern is relatively consistent regardless of whether competition is perfect or imperfect. In terms of social welfare, the shapes of both cases are reasonably similar. For buses 15–19, social welfare is more sensitive to the congestion level. Strategic behavior actually does not result in a significant loss of social welfare. These observations motivate the search for the optimal storage locations because such locations would ensure that storage collects enough revenue



Figure 5.2: Annual profits at various buses under a 50% reduction in line transmission capacity.

while being less affected by congestion in the system.

Figure 5.3 shows the LMP duration curve at buses 14 and 19 based on a year-long simulation for the base case and for a 50% reduction on line capacity. The prices in perfect and imperfect competition are relatively close, but slightly more favorable for merchant storage in case of imperfect competition. On the other hand, comparing the prices for the base case shows that local congestion increases the average nodal price. The top left area of both plots suggest that high prices happen more frequently than in the base case, while the bottom right areas show that low prices happen less frequently. Significant penetration of renewables causes periods of zero or negative prices occur and are made more frequent by

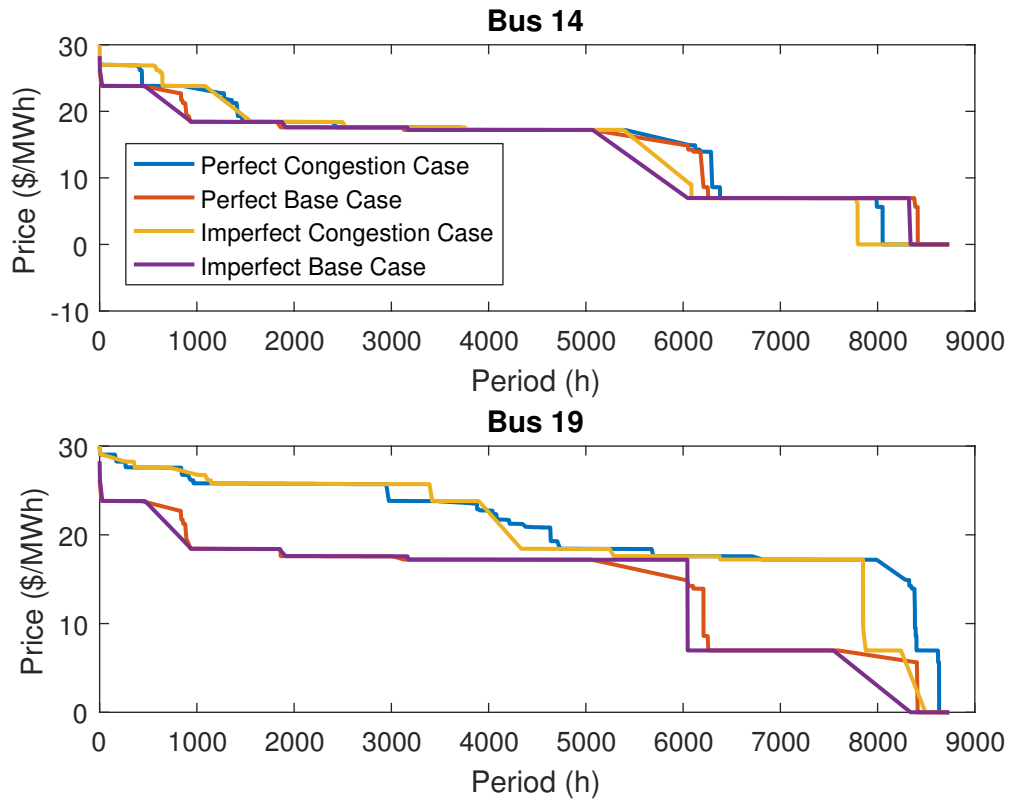


Figure 5.3: Locational marginal price duration curve with a 50% line capacity reduction.

local congestion as shown in the plot for bus 14.

Figure 5.4 shows a three-day snapshot of LMPs at bus 14. During the first 28 hours of this period, local congestion does not affect the LMPs and perfect competition actually produces a slightly higher average LMP. From about hour 28 onwards, strategic bidding can be used to create arbitrage opportunities by driving the LMPs down and up. Local congestion cases enhance these opportunities by extending the periods of lower and higher LMPs. Such price differences are valuable for storage because they make it possible to charge at low prices and discharge at high prices. Nontrivial profit differences between perfectly and imperfectly competitive markets stem from such infrequent periods with substantial price differences.

Figure 5.5 shows the histogram of daily profits for a storage located at bus 19. In the

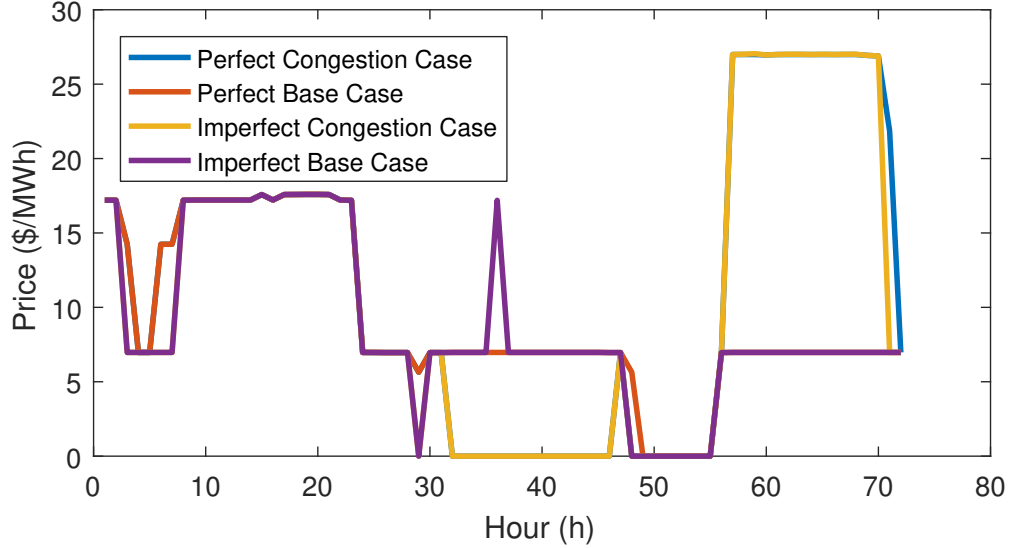


Figure 5.4: LMP at bus 14 on days 100–102 with a 50% line capacity reduction.

perfect competition case, the histogram does not exhibit a clear pattern. Local congestion increases the number of higher profit days and, to a lesser extent, the number of lower profit days, resulting in a higher total annual profit. The histogram for imperfect competition case is smoother and more evenly distributed. There are much fewer low profit days and many more high profit days. Here again, congestion increases the number of high profit days.

5.3 Look-Ahead Bidding Strategy for Energy Storage

5.3.1 Introduction

Large-scale deployment of energy storage (ES) is becoming an economically viable option due to technological advances in chemistry, material science and chemical engineering [15]. NREL [207] projects that about 152 GW of ES capacity will be installed by 2050 in the US. The California Independent System Operator (CAISO) has published a roadmap for ES, indicating the revenue potential that ES could reach [208].

ES can be used for spatio-temporal arbitrage [14], to relieve transmission congestion [167],

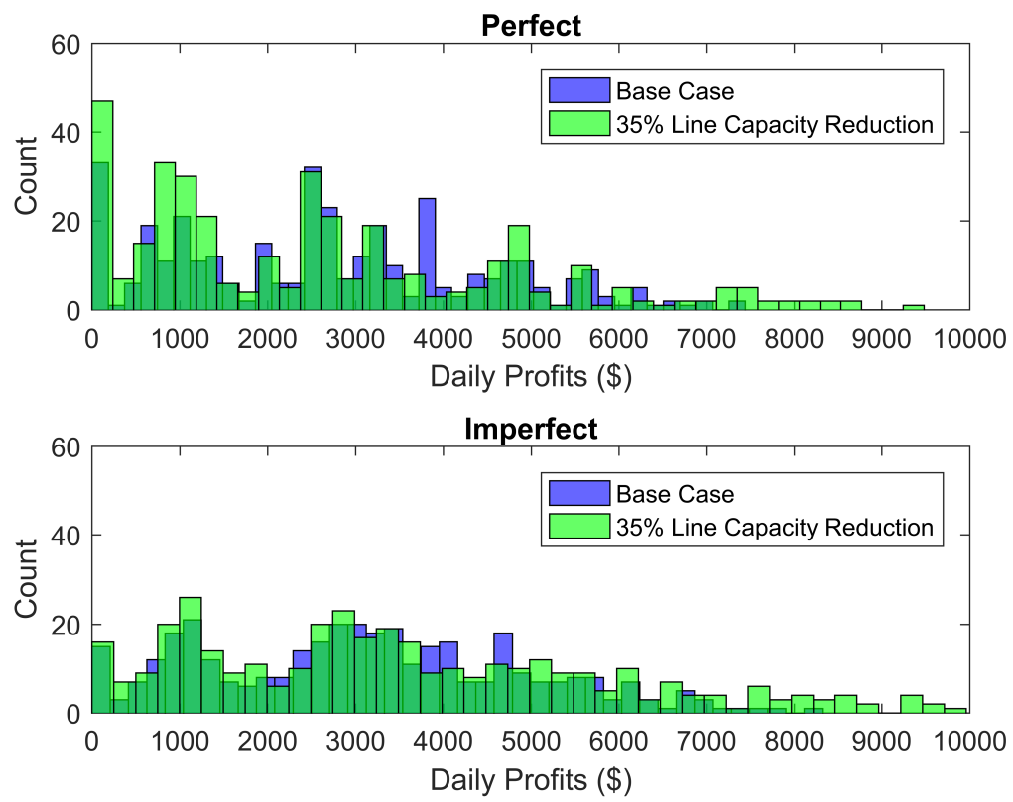


Figure 5.5: Histogram of daily storage profits at bus 19.

and to provide reserves [18] and corrective actions during contingencies [141]. In a vertically integrated environment, these activities are deployed by the system operator in a way that minimizes the system operating cost or maximizes the social welfare. These goals typically do not align with a merchant storage owner's objective to maximize its profits.

It is commonly assumed that energy storage should return to its initial state-of-charge (SoC) at the end of the last period of the scheduling horizon. In a vertically integrated environment, this practice provides system operators with the flexibility that they might need to deal with an unexpected situation during the next operating period. On the other hand, a merchant storage operator may want to adjust this final state of charge (fSoC) to maximize its profits. However, this adjustment should balance profit opportunities during the current market window against potential opportunities in the subsequent market window.

In this section, we propose a bilevel model that can be used by merchant energy storage operators to determine optimal price and quantity offers and bids in a day-ahead self-scheduling electricity market. The upper level accounts for future profit opportunities by considering both the current day-ahead (D_0) and the next day-ahead (D_1) market windows. The risk associated with the next day-ahead market is also constrained by co-optimizing the conditional value-at-risk (CVaR) for D_1 to incorporate the forecast uncertainty. This allows it to adjust the SoC at the end of the first day to optimize arbitrage opportunities over both days.

The lower level simulates the market clearing using a ramp-constrained DC optimal power flow and feeds the locational marginal prices (LMP) to the upper-level problem where they are used to calculate the profits. The uncertainty on the renewable forecast for D_1 is modeled using scenarios to represent different market clearing outcomes under different possible wind realizations.

The proposed hybrid bilevel approach maximizes the profit that a merchant storage operator can collect through spatio-temporal arbitrage by optimizing its bidding strategy not only for the day ahead but also for the following day by adjusting the storage state-of-charge at the end of the first day. The optimization for the day-ahead uses a deterministic

forecast, while scenarios are used to emphasize the increase uncertainty on the forecasts for the following day.

A risk-constrained profit term is added in the objective function to regularize the combined two-day profits and optimize the state-of-charge set points. With CVaR, this approach not only considers the potential future profits but also the future risks.

The case study shows how the results of this optimization are affected by the discount factor for the next day profit, by the length of the look-ahead window, by the forecast uncertainty, by the competing units' offer price, by the storage operating cost, by the system transmission capacity and by the ramp rate of the conventional generators.

Rather than conservatively constraining the state of charge at the end of the day ahead at a fixed value, the proposed approach determines the fSoC that maximizes the merchant storage's profits. Our method specifically focus on this storage time-coupling feature, which is commonly neglected. The way the future profit regularization term takes into account both the profit expectation and the risk is also novel.

5.3.2 Mathematical Formulation

Figure 5.6 illustrates the proposed bilevel formulation used to optimize the look-ahead bidding strategy of a merchant energy storage operator. The upper-level (UL) problem takes the perspective of this storage operator who aims to maximize the profits that it can obtain through spatio-temporal arbitrage [209] in a day-ahead electricity market. We consider the profit Π^{D_0} achieved on day D_0 using a deterministic forecast as well as the expected profit from $\mathbb{E}_{\omega_1} \Pi_{\omega_1}^{D_1}$ and a risk term $CVaR^{D_1}$ on day D_1 . Since this profit clearly depends on the state-of-charge at both the beginning and the end of the optimization horizon, we optimize the state-of-charge at the end of D_0 to maximize the sum of Π^{D_0} and Γ^{D_1} (a weighted sum that balances profits and risks on day D_1). In essence, we look ahead to the next day to balance profit opportunities between days D_0 and D_1 . However, because the forecasts upon which this optimization relies are less accurate for D_1 than for D_0 and because the optimization will be repeated a day later for D_1 , the anticipated profit for D_1 are discounted by a

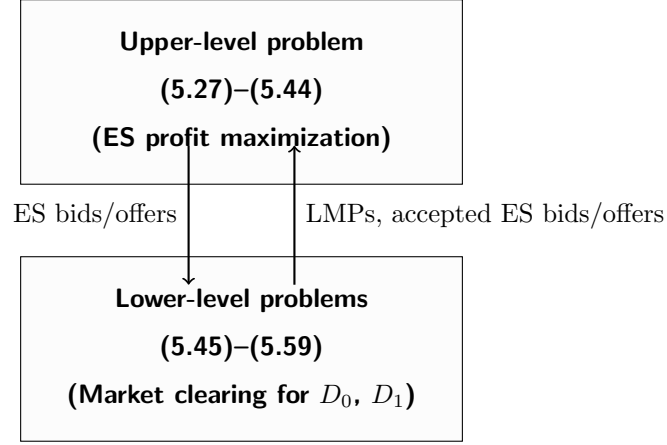


Figure 5.6: An illustration of the proposed bilevel model and the interfaces between the upper- and lower-level problems.

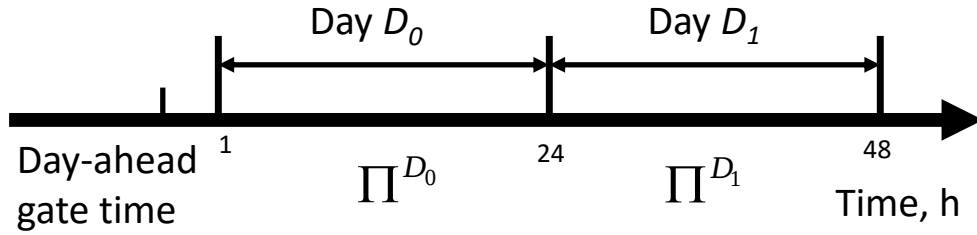


Figure 5.7: Time-frame for market-clearing on days D_0 and D_1 .

factor γ^{D_1} .

The UL problem outputs the optimal quantity and price storage bids and offers for each market interval of days D_0 and D_1 . These quantities are then used in the lower-level (LL) problem. The UL problem accounts for storage variable operating costs in its profit calculation. The LL problem clears the market separately for days D_0 and D_1 . Each market clearing maximizes the daily social welfare and accepts offers from conventional and renewable generation, bids from elastic loads, and bids and offers from the storage operator. This storage operator is assumed to be the only market participant that can act strategically in the day-ahead electricity market. Other market participants bid at their marginal costs. The LL problem produces the LMPs that the UL problem uses to calculate the profits.

The problem can then be formulated as follows:

$$\max \quad Obj = \Pi^{D_0} + \gamma^{D_1} \Gamma^{D_1} \quad (5.27)$$

subject to:

$$\Gamma^{D_1} = \gamma^{CVaR} CVaR^{D_1} + (1 - \gamma^{CVaR}) \sum_{\omega_1} (\pi_{\omega_1} \Pi_{\omega_1}^{D_1}) \quad (5.28)$$

$$loss_{\omega_1}^{D_1} \geq VaR^{D_1} - \Pi_{\omega_1}^{D_1}, \forall \omega_1 \quad (5.29)$$

$$loss_{\omega_1}^{D_1} \geq 0, \forall \omega_1, \quad (5.30)$$

$$CVaR^{D_1} = VaR^{D_1} - \frac{1}{1 - \beta_0} \sum_{\omega_1} (\pi_{\omega_1} loss_{\omega_1}^{D_1}) \quad (5.31)$$

$$\begin{aligned} \Pi^{D_0} = \sum_{h,t_0} [\lambda_{b(h),t_0} (q_{h,t_0}^{dis} - q_{h,t_0}^{chs}) \\ - C_h^{ES} (q_{h,t_0}^{dis} + q_{h,t_0}^{chs})] \end{aligned} \quad (5.32)$$

$$\begin{aligned} \Pi_{\omega_1}^{D_1} = \sum_{h,t_1} [\lambda_{b(h),t_1,\omega_1} (q_{h,t_1,\omega_1}^{dis} - q_{h,t_1,\omega_1}^{chs}) \\ - C_h^{ES} (q_{h,t_1,\omega_1}^{dis} + q_{h,t_1,\omega_1}^{chs})] \end{aligned} \quad (5.33)$$

$$0 \leq chs_{h,t} \leq \overline{chs}_h (1 - x_{h,t}^{dis}), \forall h, \forall t \quad (5.34)$$

$$0 \leq dis_{h,t} \leq \overline{dis}_h x_{h,t}^{dis}, \forall h, \forall t \quad (5.35)$$

$$x_{h,t}^{dis} \in \{0, 1\}, \forall h, \forall t \quad (5.36)$$

$$\rho_{h,t}^{dis}, \rho_{h,t}^{chs} \geq 0, \forall h, \forall t \quad (5.37)$$

$$\begin{aligned} SoC_{h,t_0} = & SoC_h^{init} + q_{h,t_0}^{chs} \eta_h^{chs} - q_{h,t_0}^{dis} / \eta_h^{dis}, \\ & \forall t_0 = 1, \forall h \end{aligned} \quad (5.38)$$

$$\begin{aligned} SoC_{h,t_0} = & SoC_{h,t_0-1} + q_{h,t_0}^{chs} \eta_h^{chs} - q_{h,t_0}^{dis} / \eta_h^{dis}, \\ & \forall 1 < t_0 \leq N_{t_0}, \forall h \end{aligned} \quad (5.39)$$

$$\underline{SoC}_h \leq SoC_{h,t_0} \leq \overline{SoC}_h, \forall h, \forall t_0 \quad (5.40)$$

$$\begin{aligned} SoC_{h,t_1,\omega_1} = & SoC_{h,N_{t_0}} + q_{h,t_1,\omega_1}^{chs} \eta_h^{chs} \\ & - q_{h,t_1,\omega_1}^{dis} / \eta_h^{dis}, \forall t_1 = N_{t_0} + 1, \forall h, \forall \omega_1 \end{aligned} \quad (5.41)$$

$$\begin{aligned} SoC_{h,t_1,\omega_1} = & SoC_{h,t_1-1,\omega_1} + q_{h,t_1,\omega_1}^{chs} \eta_h^{chs} \\ & - q_{h,t_1,\omega_1}^{dis} / \eta_h^{dis}, \forall N_{t_0} + 1 < t_1 \leq N_{t_1}, \forall h, \forall \omega_1 \end{aligned} \quad (5.42)$$

$$\underline{SoC}_h \leq SoC_{h,t_1,\omega_1} \leq \overline{SoC}_h, \forall h, \forall t_1, \forall \omega_1 \quad (5.43)$$

$$SoC_{h,N_{t_1},\omega_1} \geq SoC_h^{init}, \forall h, \forall \omega_1 \quad (5.44)$$

$$\begin{aligned} \lambda_{b,t,\omega}, q_{h,t,\omega}^{dis}, q_{h,t,\omega}^{chs} \in \arg \max \Bigg\{ & SW_\omega = \sum_{b,t} C_b^D d_{b,t,\omega} \\ & + \sum_{h,t} \rho_{h,t}^{chs} q_{h,t,\omega}^{chs} - \sum_{i,t} C_i^G p_{i,t,\omega}^G - \sum_{h,t} \rho_{h,t}^{dis} q_{h,t,\omega}^{dis} \end{aligned} \quad (5.45)$$

subject to:

$$d_{b,t,\omega} + \sum_{f(l)=b} p_{f,l,t,\omega} - \sum_{t(l)=b} p_{f,l,t,\omega} = \sum_{b(i)=b} p_{i,t,\omega}^G + \quad (5.46)$$

$$\sum_{b(w)=b} p_{w,t,\omega}^W + \sum_{b(h)=b} (q_{h,t,\omega}^{dis} - q_{h,t,\omega}^{chs}) : \lambda_{b,t,\omega}, \forall b, \forall t$$

$$0 \leq p_{i,t,\omega}^G \leq \bar{P}_i : \underline{\mu}_{i,t,\omega}^G, \bar{\mu}_{i,t,\omega}^G, \forall i, \forall t \quad (5.47)$$

$$0 \leq p_{w,t,\omega}^W \leq WF_{w,t,\omega} : \underline{\mu}_{w,t,\omega}^W, \bar{\mu}_{w,t,\omega}^W, \forall w, \forall t \quad (5.48)$$

$$\underline{D}_{b,t} \leq d_{b,t,\omega} \leq \overline{D}_{b,t} : \underline{\mu}_{b,t,\omega}^D, \bar{\mu}_{b,t,\omega}^D, \forall b, \forall t \quad (5.49)$$

$$0 \leq q_{h,t,\omega}^{dis} \leq dis_{h,t} : \alpha_{h,t,\omega}^{dis}, \bar{\alpha}_{h,t,\omega}^{dis}, \forall h, \forall t \quad (5.50)$$

$$0 \leq q_{h,t,\omega}^{chs} \leq chs_{h,t} : \alpha_{h,t,\omega}^{chs}, \bar{\alpha}_{h,t,\omega}^{chs}, \forall h, \forall t \quad (5.51)$$

$$p_{i,t,\omega}^G - p_{i,t-1,\omega}^G \leq RU_i : \beta_{i,t,\omega}^u, \forall t > 1, \forall i \quad (5.52)$$

$$p_{i,t-1,\omega}^G - p_{i,t,\omega}^G \leq RD_i : \beta_{i,t,\omega}^d, \forall t > 1, \forall i \quad (5.53)$$

$$p_{i,t,\omega}^G - p_{i,0}^G \leq RU_i : \beta_{i,t,\omega}^u, t = 1, \forall i \quad (5.54)$$

$$p_{i,0}^G - p_{i,t,\omega}^G \leq RD_i : \beta_{i,t,\omega}^d, t = 1, \forall i \quad (5.55)$$

$$pf_{l,t,\omega} = \frac{1}{X_l}(\theta_{f(l),t,\omega} - \theta_{t(l),t,\omega}) : \nu_{l,t,\omega}, \forall l, \forall t \quad (5.56)$$

$$-\bar{F}_l \leq pf_{l,t,\omega} \leq \bar{F}_l : \varphi_{l,t,\omega}, \bar{\varphi}_{l,t,\omega}, \forall l, \forall t \quad (5.57)$$

$$-\bar{\theta} \leq \theta_{b,t,\omega} \leq \bar{\theta} : \underline{\delta}_{b,t,\omega}, \bar{\delta}_{b,t,\omega}, \forall b \neq \text{ref}, \forall t \quad (5.58)$$

$$\theta_{b,t,\omega} = 0 : \delta_t^{ref}, b = \text{ref}, \forall t \quad \left. \vphantom{\theta_{b,t,\omega} = 0} \right\} \quad (5.59)$$

Equation (5.27) is the objective function of the UL problem which aims to maximize the total profit of the storage owner over D_0 and D_1 . The discount factor $\gamma^{D_1} \in [0, 1]$ reflects the storage owner's willingness to balance profits between days D_0 and D_1 , considering that the market clearing for D_1 is more speculative than for D_0 . Varying γ^{D_1} affects the storage schedule for D_0 , including the state of charge at the end of that day, as well as the schedule for D_1 .

Constraints (5.28)–(5.31) define the look-ahead profit regularization term Γ^{D_1} with expected profits from $\Pi_{\omega_1}^{D_1}$ and risk $CVaR^{D_1}$. γ^{CVaR} is used to control the risk positions of storage owner due to the D_1 forecast uncertainty. A larger γ^{CVaR} indicates a more risk-averse position. This profit regularization term Γ^{D_1} determines the optimized risk trade-off based on a given discount factor and risk-aversion attitude.

Constraints (5.32)–(5.33) define the profits Π^{D_0} and $\Pi_{\omega_1}^{D_1}$ as the difference between the net revenue from spatio-temporal arbitrage and the operating cost. While the variable cost is negligible for pumped hydro and compressed air energy storage (CAES), cycling degradation must be taken into account for an accurate assessment of the profitability of batteries [193].

In this section, we adopt 4 \$/MWh from a PNNL report [155], which updates the results from an EPRI handbook [210]. We will show how these numbers affect storage revenues and profits in the case study. Locational marginal prices $\lambda_{h(b),t,w}$ in (5.32)–(5.33) and cleared storage capacity offers and bids $q_{h,t,w}^{dis}/q_{h,t,w}^{chs}$ are provided by the LL problems. Constraints (5.34)–(5.36) relate storage bids and offers with the power ratings. The price offers and bids are defined in (5.37). Constraints (5.38)–(5.43) keep track of the state-of-charge of the storage and enforce the limits. Constraint (5.44) ensures that the state-of-charge at the end D_1 returns to the level at the beginning of D_0 . This ensures that the storage cycling is energy neutral over this two-day horizon.

The LL problems maximize the social welfare as defined by (5.45). The nodal power balance constraint is enforced by (5.46). Constraints (5.47)–(5.51) impose the minimum and maximum power limits on bids and offers of conventional generators, wind power producers, storage, and demand. Note that the market can accept partial bids and offers submitted by storage as expressed in (5.50) and (5.51). The inter-hour cycling of conventional generators is constrained by their ramp rate limits in (5.52) – (5.55). Since initial positions of conventional units are also affected by various markets and signals, we relax constraints (5.54)–(5.55) for the D_1 market for simplicity. We assume that the initial D_0 generator positions and marginal offers are known. The impact of generators' offers is discussed in Section 5.3.4. Constraints (5.56)–(5.57) compute the power flows and enforce the power flow limits by using a dc load flow model. The voltage phase angles are constrained by (5.58) and the phase angle of the reference bus is set in (5.59).

5.3.3 Solution Technique

5.3.3.1 Single-level Equivalent

The single-level equivalent of the bilevel formulation in Section 5.3.2 can be obtained using the KKT-based approach [196]. In this approach, the LL problem is replaced by its primal and dual feasibility constraints and its complementary slackness conditions. For each market

clearing process ω_0 for D_0 and ω_1 for D_1 markets, we have following Karush-Kahn-Tucker conditions:

$$C_i^G - \lambda_{b(i),t,\omega} - \underline{\mu}_{i,t,\omega}^G + \bar{\mu}_{i,t,\omega}^G + \beta_{i,t,\omega}^u - \beta_{i,t,\omega}^d + \beta_{i,t+1,\omega|t < N_{T_1}}^d - \beta_{i,t+1,\omega|t < N_{T_1}}^u = 0, \forall i, \forall t \quad (5.60)$$

$$- \lambda_{b(w),t,\omega} - \underline{\mu}_{w,t,\omega}^W + \bar{\mu}_{w,t,\omega}^W = 0, \forall w, \forall t \quad (5.61)$$

$$- C_b^D + \lambda_{b,t,\omega} - \underline{\mu}_{b,t,\omega}^D + \bar{\mu}_{b,t,\omega}^D = 0, \forall b, \forall t \quad (5.62)$$

$$\rho_{h,t}^{dis} - \lambda_{b(h),t,\omega} - \underline{\alpha}_{h,t,\omega}^{dis} + \bar{\alpha}_{h,t,\omega}^{dis} = 0, \forall h, \forall t \quad (5.63)$$

$$- \rho_{h,t}^{chs} + \lambda_{b(h),t,\omega} - \underline{\alpha}_{h,t,\omega}^{chs} + \bar{\alpha}_{h,t,\omega}^{chs} = 0, \forall h, \forall t \quad (5.64)$$

$$\sum_b A_{l,b} \lambda_{b,t,\omega} - \varphi_{l,t,\omega} + \bar{\varphi}_{l,t,\omega} - \nu_{l,t,\omega} = 0, \forall l, \forall t \quad (5.65)$$

$$\sum_l \frac{A_{l,b} \nu_{l,t,\omega}}{X_l} - \delta_{b,t,\omega} + \bar{\delta}_{b,t,\omega} = 0, \forall b \neq \text{ref}, \forall t \quad (5.66)$$

$$\sum_l \frac{A_{l,b} \nu_{l,t,\omega}}{X_l} - \delta_{t,\omega}^{ref} = 0, b = \text{ref}, \forall t \quad (5.67)$$

$$\underline{\mu}_{i,t,\omega}^G p_{i,t,\omega}^G = 0, \forall i, \forall t \quad (5.68)$$

$$\bar{\mu}_{i,t,\omega}^G (\bar{P}_i - p_{i,t,\omega}^G) = 0, \forall i, \forall t \quad (5.69)$$

$$\underline{\mu}_{w,t,\omega}^W p_{w,t,\omega}^W = 0, \forall w, \forall t \quad (5.70)$$

$$\bar{\mu}_{w,t,\omega}^W (W F_{w,t,\omega} - p_{w,t,\omega}^W) = 0, \forall w, \forall t \quad (5.71)$$

$$\underline{\mu}_{b,t,\omega}^D (d_{b,t,\omega} - \underline{D}_{b,t}) = 0, \forall b, \forall t \quad (5.72)$$

$$\bar{\mu}_{b,t,\omega}^D (\bar{D}_{b,t} - d_{b,t,\omega}) = 0, \forall b, \forall t \quad (5.73)$$

$$\underline{\alpha}_{h,t,\omega}^{dis} q_{h,t,\omega}^{dis} = 0, \forall h, \forall t \quad (5.74)$$

$$\bar{\alpha}_{h,t,\omega}^{dis} (dis_{h,t} - q_{h,t,\omega}^{dis}) = 0, \forall h, \forall t \quad (5.75)$$

$$\underline{\alpha}_{h,t,\omega}^{chs} q_{h,t,\omega}^{chs} = 0, \forall h, \forall t \quad (5.76)$$

$$\bar{\alpha}_{h,t,\omega}^{chs} (chs_{h,t} - q_{h,t,\omega}^{chs}) = 0, \forall h, \forall t \quad (5.77)$$

$$\varphi_{l,t,\omega} (\bar{F}_l + p f_{l,t,\omega}) = 0, \forall l, \forall t \quad (5.78)$$

$$\bar{\varphi}_{l,t,\omega}(\bar{F}_l - pf_{l,t,\omega}) = 0, \forall l, \forall t \quad (5.79)$$

$$\underline{\delta}_{b,t,\omega}(\bar{\theta} + \theta_{b,t,\omega}) = 0, \forall b \neq \text{ref}, \forall t \quad (5.80)$$

$$\bar{\delta}_{b,t,\omega}(\bar{\theta} - \theta_{b,t,\omega}) = 0, \forall b \neq \text{ref}, \forall t \quad (5.81)$$

$$\beta_{i,t,\omega}^u(RU_i - (p_{i,t,\omega}^G - P_{i,0,\omega}^G)) = 0, t = 1, \forall i \quad (5.82)$$

$$\beta_{i,t,\omega}^d(RD_i - (P_{i,0,\omega}^G - p_{i,t,\omega}^G)) = 0, t = 1, \forall i \quad (5.83)$$

$$\beta_{i,t,\omega}^u(RU_i - (p_{i,t,\omega}^G - p_{i,t-1,\omega}^G)) = 0, t = 2 \dots N_T, \forall i \quad (5.84)$$

$$\beta_{i,t,\omega}^d(RD_i - (p_{i,t-1,\omega}^G - p_{i,t,\omega}^G)) = 0, t = 2 \dots N_T, \forall i \quad (5.85)$$

$$\begin{aligned} & \mu_{i,t,\omega}^G, \bar{\mu}_{i,t,\omega}^G, \mu_{w,t,\omega}^W, \bar{\mu}_{w,t,\omega}^W, \mu_{b,t,\omega}^D, \bar{\mu}_{b,t,\omega}^D, \\ & \beta_{i,t,\omega}^u, \beta_{i,t,\omega}^d, \alpha_{h,t,\omega}^{dis}, \bar{\alpha}_{h,t,\omega}^{dis}, \alpha_{h,t,\omega}^{chs}, \bar{\alpha}_{h,t,\omega}^{chs}, \\ & \varphi_{l,t,\omega}, \bar{\varphi}_{l,t,\omega}, \underline{\delta}_{b,t,\omega}, \bar{\delta}_{b,t,\omega}, \geq 0 \end{aligned} \quad (5.86)$$

$$\lambda_{b,t,\omega}, \delta_{t,\omega}^{ref}, \nu_{l,t,\omega} : \text{free} \quad (5.87)$$

The dual feasibility constraints of the LL problem are given in (5.60) – (5.67). Constraints (5.68) – (5.85) enforce the complementary slackness conditions of the LL problem. The dual variables of the LL problem are defined in (5.86) and (5.87).

5.3.3.2 Linearization

The single-level equivalent (5.27)–(5.87) is non-linear due to the products of continuous decision variables in constraints (5.32) – (5.33) and in the complementary slackness conditions (5.68) – (5.85).

5.3.3.3 Linearization of Eqs. (5.32) – (5.33)

Using the complementary slackness conditions in (5.63) – (5.64), the first term of eq. (5.32) and (5.33) can be reformulated as follows:

$$\begin{aligned}
 R &= \sum_{h,t} \lambda_{b(h),t,\omega} (q_{h,t,\omega}^{dis} - q_{h,t,\omega}^{chs}) \\
 &= \sum_{h,t} (\rho_{h,t}^{dis} - \underline{\alpha}_{h,t,\omega}^{dis} + \bar{\alpha}_{h,t,\omega}^{dis}) q_{h,t,\omega}^{dis} \\
 &\quad - \sum_{h,t} (\rho_{h,t}^{chs} + \underline{\alpha}_{h,t,\omega}^{chs} - \bar{\alpha}_{h,t,\omega}^{chs}) q_{h,t,\omega}^{chs}
 \end{aligned} \tag{5.88}$$

Next, complementary conditions (5.74) – (5.77) are used to recast (5.88) as follows:

$$\begin{aligned}
 R &= \sum_{h,t} (\rho_{h,t}^{dis} q_{h,t,\omega}^{dis} - \rho_{h,t,\omega}^{chs} q_{h,t,\omega}^{chs} \\
 &\quad + \bar{\alpha}_{h,t,\omega}^{dis} dis_{h,t} + \bar{\alpha}_{h,t,\omega}^{chs} chs_{h,t})
 \end{aligned} \tag{5.89}$$

Since the LL problem is convex, the strong duality theorem can be invoked to equate the LL primal and dual objective functions:

$$\begin{aligned}
 SW_{\omega} &= - \sum_{i,t} \bar{\mu}_{i,t,\omega}^G \bar{P}_i - \sum_{w,t} \bar{\mu}_{w,t,\omega}^W W F_{w,t,\omega} \\
 &\quad + \sum_{b,t} (\underline{\mu}_{b,t,\omega}^D \underline{D}_{b,t} - \bar{\mu}_{b,t,\omega}^D \bar{D}_{b,t}) \\
 &\quad - \sum_{h,t} (\bar{\alpha}_{h,t,\omega}^{dis} dis_{h,t} + \bar{\alpha}_{h,t,\omega}^{chs} chs_{h,t}) \\
 &\quad - \sum_{l,t} \bar{F}_l (\underline{\varphi}_{l,t,\omega} + \bar{\varphi}_{l,t,\omega}) \\
 &\quad - \sum_{t,b \neq \text{ref}} \bar{\theta} (\delta_{b,t,\omega} + \bar{\delta}_{b,t,\omega}) \\
 &\quad - \sum_{i,t} (\beta_{i,t,\omega}^u R U_i + \beta_{i,t,\omega}^d R D_i) \\
 &\quad - \sum_i P_{i,0}^G (\beta_{i,1,\omega}^u - \beta_{i,1,\omega}^d)
 \end{aligned} \tag{5.90}$$

Using (5.89) and (5.90), non-linear constraints (5.32) and (5.33) can be recast as the

following equivalent linear expressions:

$$\begin{aligned}
\Pi_{\omega_0}^{D_0} = & - \sum_{i,t_0} \bar{\mu}_{i,t_0}^G \bar{P}_i - \sum_{w,t_0} \bar{\mu}_{w,t_0}^W W F_{w,t_0} \\
& + \sum_{b,t_0} (\underline{\mu}_{b,t_0}^D \underline{D}_{b,t_0} - \bar{\mu}_{b,t_0}^D \bar{D}_{b,t_0}) \\
& - \sum_{l,t_0} \bar{F}_l (\underline{\varphi}_{l,t_0} + \bar{\varphi}_{l,t_0}) - \sum_{t_0, b \neq \text{ref}} \bar{\theta} (\underline{\delta}_{b,t_0} + \bar{\delta}_{b,t_0}) \\
& - \sum_{i,t_0} (\beta_{i,t_0}^u R U_i + \beta_{i,t_0}^d R D_i) \\
& - \sum_i P_{i,0}^G (\beta_{i,1}^u - \beta_{i,1}^d) \\
& - \sum_{i,t_0} C_i^G p_{i,t_0}^G + \sum_{b,t_0} C_b^D d_{b,t_0} \\
& - \sum_{h,t_0} C_h^{ES} (q_{h,t_0}^{dis} + q_{h,t_0}^{chs})
\end{aligned} \tag{5.91}$$

$$\begin{aligned}
\Pi_{\omega_1}^{D_1} = & - \sum_{i,t_1} \bar{\mu}_{i,t_1,\omega_1}^G \bar{P}_i - \sum_{w,t_1,\omega_1} \bar{\mu}_{w,t_1,\omega_1}^W W F_{w,t_1,\omega_1} \\
& + \sum_{b,t_1} (\underline{\mu}_{b,t_1,\omega_1}^D \underline{D}_{b,t_1} - \bar{\mu}_{b,t_1,\omega_1}^D \bar{D}_{b,t_1}) \\
& - \sum_{l,t_1} \bar{F}_l (\underline{\varphi}_{l,t_1,\omega_1} + \bar{\varphi}_{l,t_1,\omega_1}) \\
& - \sum_{t_1, b \neq \text{ref}} \bar{\theta} (\underline{\delta}_{b,t_1,\omega_1} + \bar{\delta}_{b,t_1,\omega_1}) \\
& - \sum_{i,t_1 > N_{T_0} + 1} (\beta_{i,t_1,\omega_1}^u R U_i + \beta_{i,t_1,\omega_1}^d R D_i) \\
& - \sum_{i,t_1} C_i^G p_{i,t_1,\omega_1}^G + \sum_{b,t_1} C_b^D d_{b,t_1,\omega_1} \\
& - \sum_{h,t_1} C_h^{ES} (q_{h,t_1,\omega_1}^{dis} + q_{h,t_1,\omega_1}^{chs})
\end{aligned} \tag{5.92}$$

5.3.3.4 Linearization of Eqs. (5.68) – (5.85)

Every non-linear complementary slackness condition (5.68) – (5.85) is linearized using the Fortuny-Amat and McCarl transformations [211] as described below.

Assume that $a \geq 0$ and $b \geq 0$ are such that $ab = 0$. As explained in [211], the product ab can be replaced using an auxiliary binary variable z constrained as follows:

$$0 \leq a \leq Mz \quad (5.93)$$

$$0 \leq b \leq M(1 - z), \quad (5.94)$$

where M is a big enough constant. Thus, if $z = 0$, eq. (5.93) returns $a = 0$ and eq. (5.94) ensures $b \leq M$. If $z = 1$, eq. (5.94) returns $b = 0$ and eq. (5.93) ensures $a \leq M$. If M is chosen large enough, a and b reach the optimal solution. Given the transformation in (5.93) and (5.94), the complementary slackness conditions in (5.68) – (5.85) are replaced by the equivalent constraints given below in matrix form for the sake of compactness:

$$Ex + Fy \leq G, \quad (5.95)$$

where x is the vector of the LL primal and dual variables, y is the vector of auxiliary binary variables and E, F, G denote matrices of parameters.

5.3.4 Case Study

5.3.4.1 Test System & Data

The proposed MPEC formulation has been tested on a modified version of the one-area IEEE RTS, which consists of 24 buses, 32 generators with a total capacity of 3,105 MW, 38 transmission lines and five wind farms with a total installed capacity of 2,100 MW [14]. The day-ahead wind power forecasts for days D_0 and D_1 were generated based on the NREL Eastern Wind Dataset [47] sequentially for 7 consecutive days to evaluate the proposed MPEC model over a week. A deterministic forecast is used for D_0 . For day D_1 , we first use an ensemble approach to generate 1000 training scenarios. Then, we apply a submodular

Table 5.1: Parameters of the storage devices

Storage	$\overline{chs} = \overline{dis}$ (MW)	$\overline{SoC}$ (MWh)	$\eta^{\text{dis}} = \eta^{\text{ch}}$	Bus location
1	41	303	0.9	16
2	17	117	0.9	17
3	36	247	0.9	19
4	93	629	0.9	21

scenario reduction to reduce this set to 10 scenarios. Fig. 5.8 shows the wind and load profiles for this simulated week. Four storage devices with the parameters given in Table 5.1 were sited and sized using the procedure presented in [14]. The initial state-of-charge on day D_0 is set at 50% of $\overline{SoC}_h$ for all storage devices.

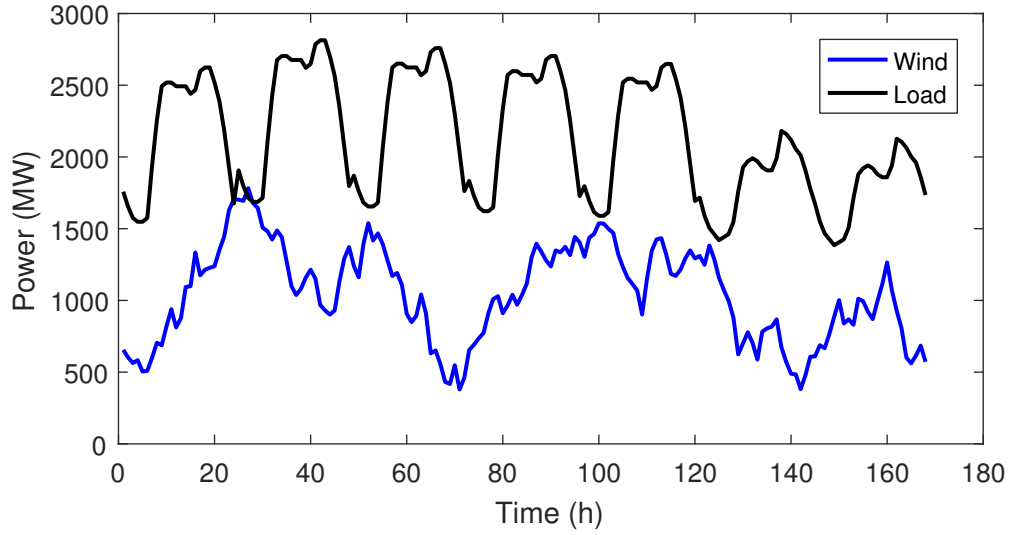


Figure 5.8: Wind and load profiles for the test week

All simulations were carried out using GAMS v.23.7 [212] and CPLEX v.12.5 with optimality gap 0.1% on a Intel Xenon 2.55 GHz processor with at least 32 GB RAM [213].

5.3.4.2 Inter-day Profit Arbitrage

This section analyzes the inter-day profit arbitrage by the storage devices and its sensitivity to the final state-of-charge (fSoC) of day D_0 and to the discount factor γ^{D_1} and γ^{CVaR} . Table 5.2 itemizes the optimized fSoC of every storage device for every day of the week studied and different values of the discount factor γ^{D_1} . The economic dispatch (ED) is the reference case where storage is economically dispatched by the system operator and does not behave strategically. The benchmark is the case where the fSoC of each storage device at the end of each day must return to 50% $\overline{SoC}_h$. Transmission and ramp rate constraints are omitted to obtain the worst-case profit opportunities for storage. As the discount factor γ^{D_1} increases, the optimization puts more weight on Γ^{D_1} . In particular, if $\gamma^{D_1} = 1$, the profits terms Π^{D_0} and Γ^{D_1} are weighted equally. Compared to the benchmark case, the total ES profit increases for all values of discount factor γ^{D_1} . However, the maximum profit is achieved when $\gamma^{D_1} = 0.75$, which demonstrates the usefulness of discounting potential future profits because of the larger forecast errors that might affect these profits. Overly conservative discounting, *i.e.* $\gamma^{D_1} < 0.5$, overestimates the effect of forecast errors and reduces the total actual profits collected by the storage devices. On the other hand, underestimating these forecast errors by setting $\gamma^{D_1} > 0.75$, reduces this total profit even more. The weight given to the risk γ^{CVaR} affects the amount energy stored at end of the day. For a small γ^{D_1} , the risk impact is less obvious than for the high γ^{D_1} cases. Although this risk aversion has a smaller effect on the total fSoC at the end of D_0 , it tends to make small adjustments to the offers submitted by the various storage devices. Fig. 5.9 shows the total profits for the test week with different γ^{D_1} and γ^{CVaR} values.

5.3.4.3 Effect of the Level of Forecasting Error

In order to assess its effect on storage profitability, the amount of forecasting error on Day D_1 is progressively increased. We first compute the Day D_1 day-ahead forecast errors from the realizations. Then to reflect a larger forecast error level, these errors are proportionally

Table 5.2: Final state of charge (MWh) and profits (\$) without transmission and ramp rate constraints

Model	ES	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total Profits
Economic Dispatch	ES1	151	151	151	151	151	151	151	23230
	ES2	58	58	58	58	58	58	58	
	ES3	123	123	123	123	123	123	123	
	ES4	314	314	314	314	314	314	314	
Benchmark	ES1	151	151	151	151	151	151	151	26800
	ES2	58	58	58	58	58	58	58	
	ES3	123	123	123	123	123	123	123	
	ES4	314	314	314	314	314	314	314	
$\gamma^{D_1} = 0$	ES1	0	0	0	0	0	0	0	39060
	ES2	0	0	0	0	0	0	0	
	ES3	0	0	0	0	0	0	0	
	ES4	0	0	0	0	0	0	0	
$\gamma = 0.50$ $\gamma^{CVaR} = 0.75$	ES1	2	0	0	0	0	0	0	39379
	ES2	8	0	0	0	0	0	0	
	ES3	1	0	0	0	0	0	0	
	ES4	15	0	0	0	0	0	0	
$\gamma = 0.75$ $\gamma^{CVaR} = 0.75$	ES1	49	48	0	45	24	0	0	41404
	ES2	25	15	0	10	10	0	0	
	ES3	44	41	0	20	0	0	0	
	ES4	69	84	0	43	23	0	0	
$\gamma = 0.75$ $\gamma^{CVaR} = 1$	ES1	65	50	0	45	27	0	0	41240
	ES2	25	20	0	10	10	0	0	
	ES3	20	33	0	20	18	0	0	
	ES4	58	84	0	43	7	0	0	
$\gamma = 1$ $\gamma^{CVaR} = 0$	ES1	52	58	0	74	74	0	144	36084
	ES2	10	19	0	31	31	0	41	
	ES3	53	43	0	53	64	0	89	
	ES4	90	103	0	167	167	0	169	
$\gamma = 1$ $\gamma^{CVaR} = 0.75$	ES1	57	78	0	74	74	0	166	35947
	ES2	25	38	0	10	30	0	59	
	ES3	53	39	0	63	63	0	81	
	ES4	112	128	0	167	167	5	140	
$\gamma = 1$ $\gamma^{CVaR} = 1$	ES1	77	119	30	46	74	0	181	33991
	ES2	10	25	4	31	31	0	100	
	ES3	26	44	0	53	64	0	129	
	ES4	90	171	0	108	167	0	158	

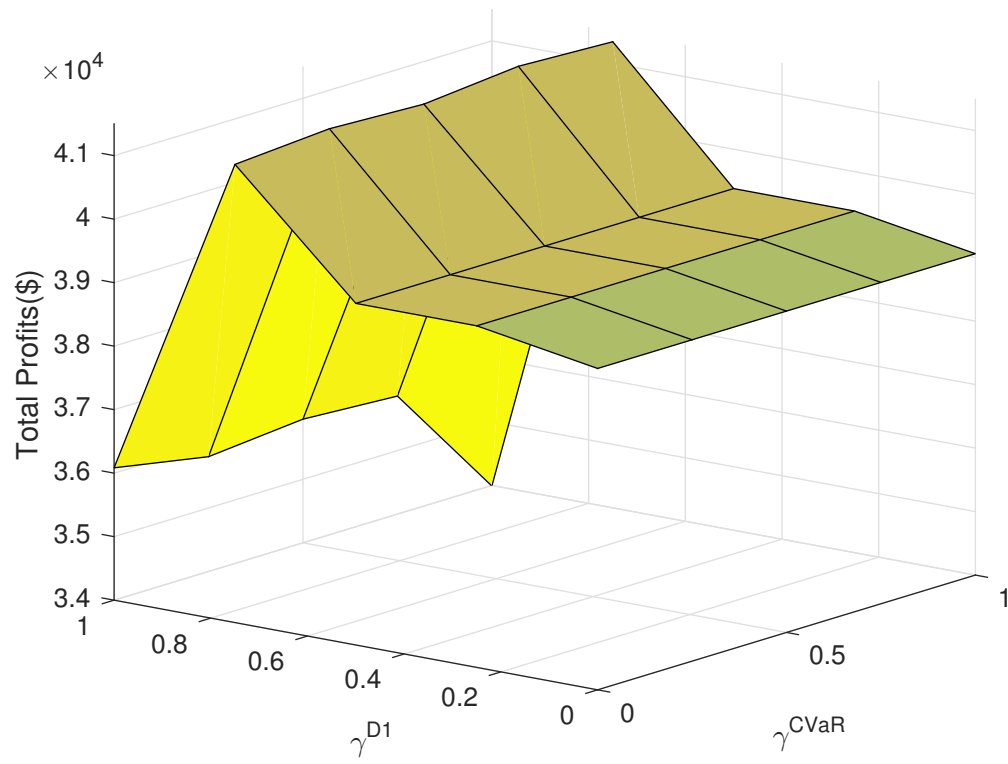


Figure 5.9: Storage profits with different γ^{D1} and γ^{CVaR}

increased to represent larger forecast errors during the ensemble training stage for scenario generation.

Fig. 5.10 shows how the total profit collected by the storage devices changes as a function of the wind power forecast error levels for discount factor $\gamma^{D_1} = 0.75$ with different γ^{CVaR} . The total ES profit decreases as the wind power forecast error increases and this reduction is less sensitive for large values of γ^{CVaR} . This confirms that considering the risk using Γ^{D_1} when the wind power forecast are inaccurate is necessary. Well-tuned γ^{D_1} and γ^{CVaR} values ensure higher profits for all cases of wind power forecast errors. In the next subsections, we will use a value of $\gamma^{CVaR} = 0.75$.

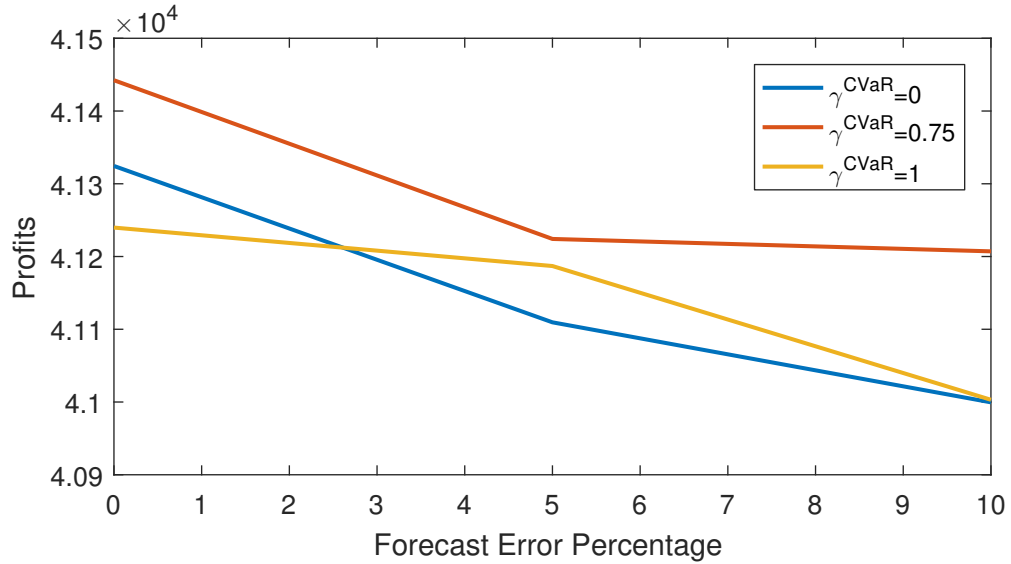


Figure 5.10: Profits for different D+1 forecasting error levels for $\gamma^{D_1} = 0.75$

Fig. 5.11 shows the optimized price and quantity offers that storage device 4 would submit on day 4 for different forecasting error levels using a discount factor $\gamma^{D_1} = 0.75$ and $\gamma^{CVaR} = 0.75$. Due to the forecast error, the ES owners change their offering strategies according to the forecast that they have available. The discount factor results in a relatively consistent offering strategy. While their offer prices do not change much with different

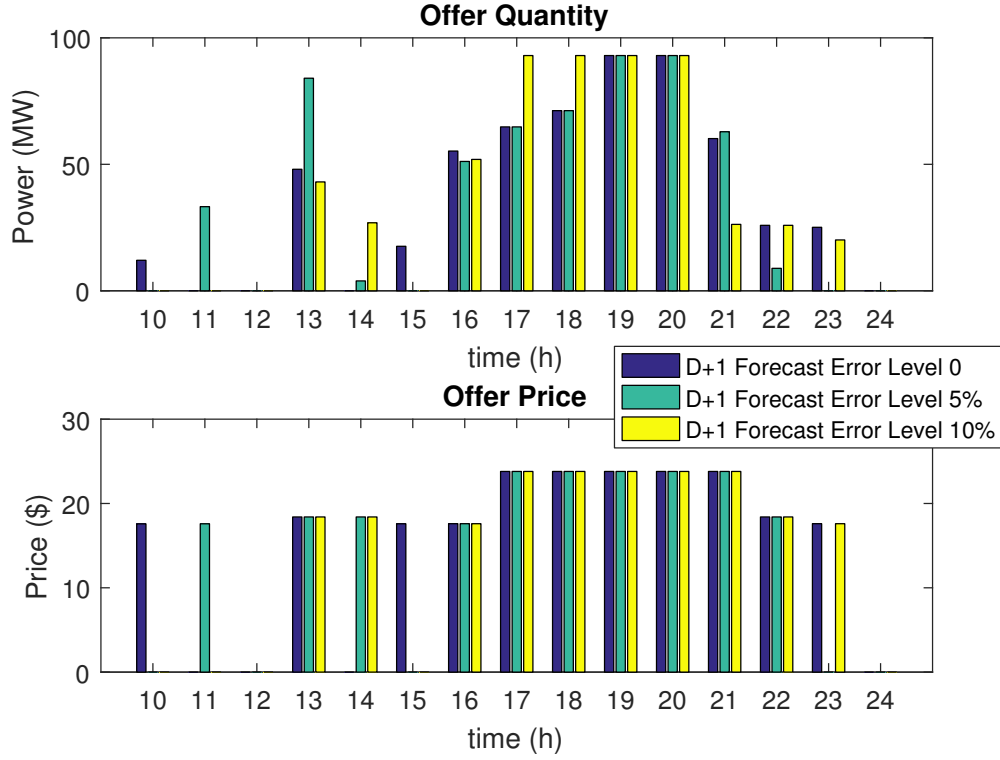


Figure 5.11: Storage 4 Offer in Day 4

forecast error levels, the offer quantities change a lot because of possible future arbitrage opportunities. $D + 1$ serves as a profit estimator, but this estimator must be discounted unless we have full confidence in the accuracy of the forecast.

5.3.4.4 Effect of the Look-ahead Window Length

Fig. 5.12 shows how storage profits change with different lengths of D_1 market. We compare the 6-, 12-, 24-hour look-ahead window lengths with ED, benchmark and $\gamma^{D_1} = 0$, which means no look-ahead. We set $\gamma^{CVaR} = 0.75$ as it does not affect results as significantly as γ^{D_1} . Two γ^{D_1} values 0.75 and 1 are shown to illustrate two distinct patterns. When $\gamma^{D_1} = 0.75$, the optimization properly discounts the future. The more information about D_1 market is available, the more profits storage can realize in the future by setting the fSoC accordingly.

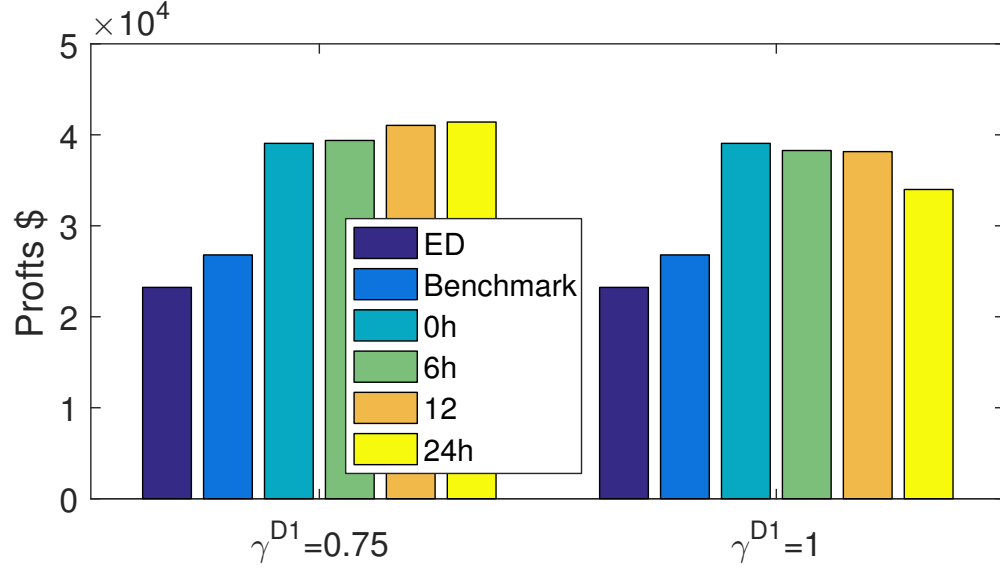


Figure 5.12: Storage profits with different look-ahead window lengths

On the contrary, when $\gamma^{D1} = 1$, the optimization is over-optimistic about its ability to predict the future and charges more at the end of the D_0 markets. When the look-ahead window is shortened, the algorithm anticipates less profit opportunities and charges less, which actually improves the profitability. This figure shows that using a 24-hour look-ahead window is appropriate for this short-term day-ahead operation. For pumped hydro units with medium-term (weekly or monthly) schedules, additional constraints could be added in the upper-level problem to keep enough energy for future use.

5.3.4.5 Effect of the Storage Variable Operating Costs

To illustrate the effect of storage variable costs on their bidding outcomes, we consider 4 typical values:

- \$4/MWh represents the base case for the operating cost of relatively large scale storage, such as pumped hydro or CAES [155].

- \$7/MWh is a typical value for a battery according to [155].
- \$9/MWh is a typical value a combustion turbine [155].
- \$0/MWh is used as a reference.

Fig. 5.13 shows the total revenue and profits storage collected from temporal arbitrage where revenue is defined as the income from arbitrage without considering the storage variable operating cost. When this operating cost varies from 0 to 7, we can observe that the revenue does not change much. Since the storage cost in this range is lower than the offer from base units, the market clearing results are mostly unchanged. Only when storage is exposed to a high price, does it changes its bidding strategy to avoid a loss of profit. The resulting revenue also decrease significantly because storage needs large price differences to compensate for the operating costs. As expected, storage profits decrease significantly with higher costs and this rate of change is nonlinear.

5.3.4.6 *Effect of the Competing Generators Price Offers*

In a perfectly competitive market, every participant is expected to submit its true marginal cost. Fig. 5.14 shows how the profit of storage operators changes when conventional units adjust their offers by discrete multipliers [0.25, 0.5, 1.0, 2.0, 3.0, 4.0]. We select 4 generators out of 32 to represent different generator types (base units, peakers or mid-range generators). When one unit changes its offer, we assume that the remaining 31 units continue to submit competitive offers. Gen 22 is the cheapest unit in the system. When it offers below \$14/MWh, the less it offers, the more profits storage achieves. For a system with high renewable penetration, a lower offer from base units creates more low-LMP time intervals suitable for storage to charge. When this base unit gradually increase offers, it increases the LMPs and storage earns more from discharging. By contrast, Gen 18 is the most expensive unit in the system and rarely generates. As the almost flat line after the multiplier reaches 1 shows, when Gen 18 increases its offer it does not affect market clearing much, and

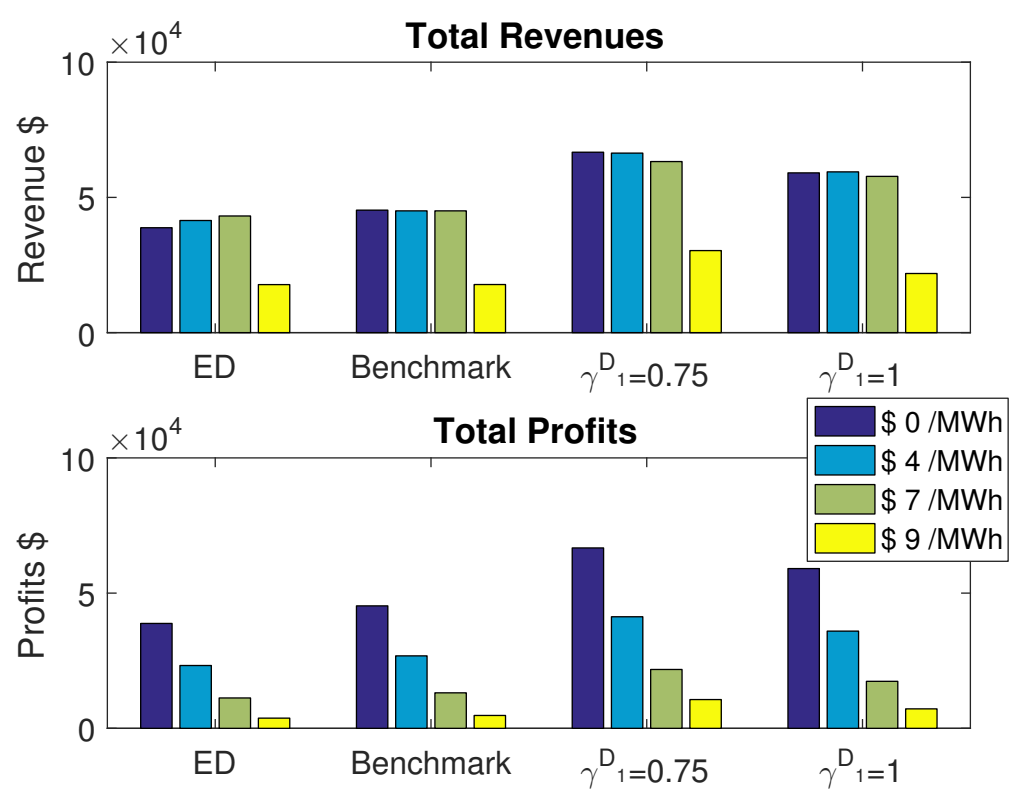


Figure 5.13: Storage revenue and profits with typical cost numbers

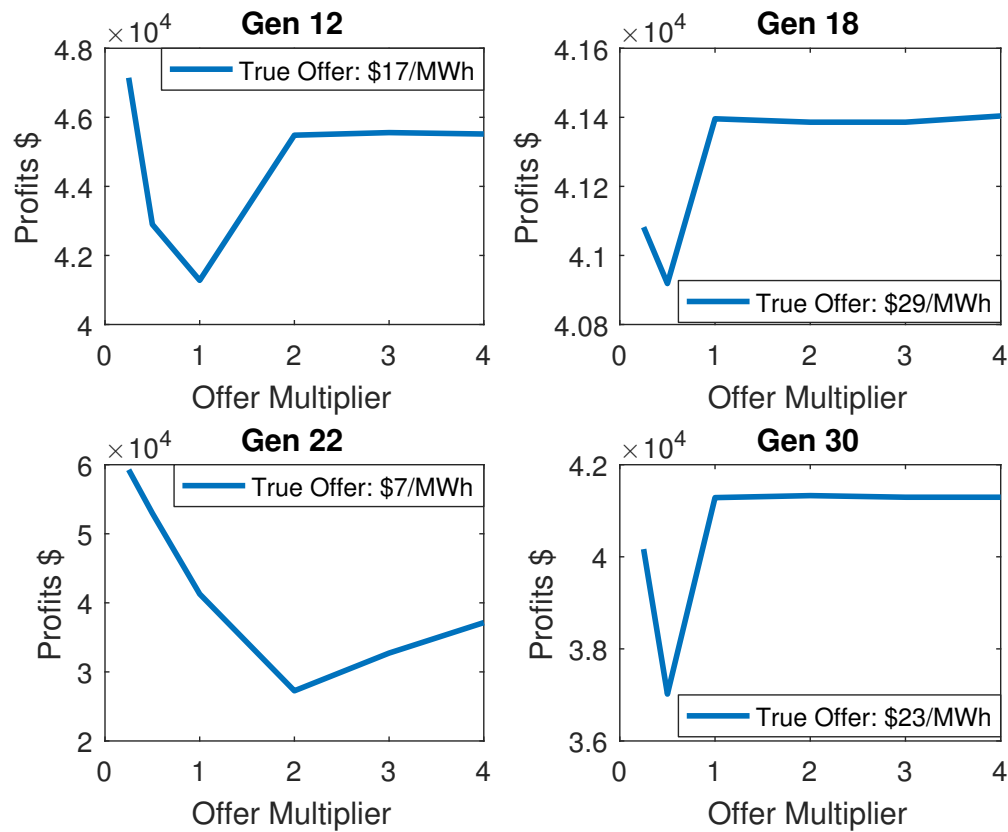


Figure 5.14: Storage profits with competing generators offers

storage profits remain relatively unchanged. When it reduces its price offer, the electricity market is more competitive as can be observed from the sharp change between multipliers 0.5 and 1, and storage loses profits. When the offer is reduced below \$15/MWh, the profits of storage increases as explained before. For mid-range generators Gen 18 and Gen 30, we clearly see a mixed behavior when they vary the offer price. We can observe that the most competitive market for storage occurs when generators offer in the range between \$11/MWh to \$23/MWh. Analyze how the generator offer uncertainty affects the profits of storage when both conventional units' price and quantity offer are random variables is an interesting question but is outside the scope of this paper.

Table 5.3: Total profits (\$) with different system conditions

Network	Ramp	ED	Benchmark	$\gamma = 0$	$\gamma = 0.50$	$\gamma = 0.75$	$\gamma = 1$
Y	N	26058	33547	42647	43687	43935	38793
N	Y	29072	37209	43976	50067	50238	45075
Y	Y	26488	38066	46543	47122	47483	42008

5.3.4.7 Effects of ramp rate and power flow limits on the profitability of storage devices

Tables 5.3 summarizes the effects of the ramp rate limits of the conventional generators and of the power flow limits on the scheduling of the storage devices and their profits. The last row gives these results for the case where both types constraints are considered in the optimization. In general, both types of constraints increase the profit opportunities for the storage devices. Quantitatively, ramp rate limits have a larger effect on the profits than the power flow limits, because thermal generation ramping capabilities are limited and congestion on the lines connected to an ES limits its ability to perform spatio-temporal arbitrage. In all cases presented, the proposed look-ahead optimization yields greater storage profits than the ED and benchmark cases where the state of charge is returned to its initial value at the end of the optimization. Storage profitability remains sensitive to the value of the discount factor γ^{D_1} when ramping and transmission constraints are taken into consideration. However, the optimal value of γ^{D_1} is then in the 0.50–0.75 range. These constraints also affect the optimal values of the fSOC as shown in Fig. 5.15. As in the case without network and ramp limits, γ^{CVaR} has a smaller impact on the results, and we use a value of 0.75 here for illustration.

Fig. 5.16 shows the total storage profits when the ramping limits of conventional generators are gradually increased using discrete multipliers [1.0, 1.5, 2.0, 2.5, 3.0]. If the multiplier is set to 1, the ramp limits are the same as above. Higher multipliers lead to more flexibility from conventional generators, which in turn leads to lower total profits collected by the storage devices regardless of the value of the discount factor γ^{D_1} . When $\gamma^{D_1} = 0$, the profit term

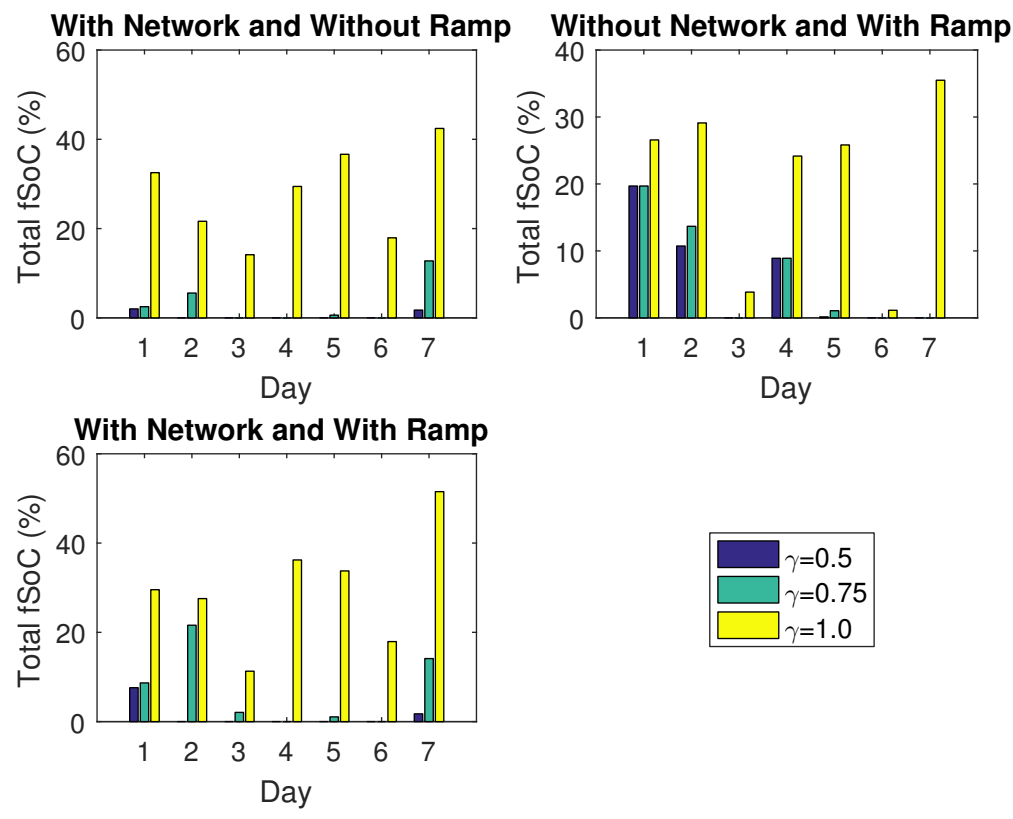


Figure 5.15: Storage fSoC level with different system conditions

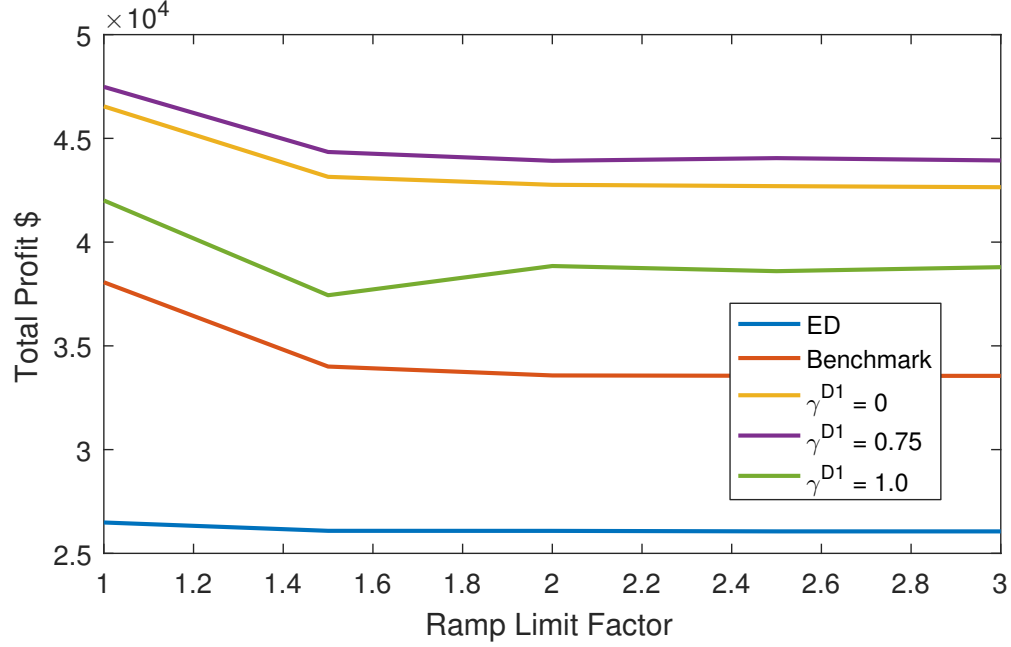


Figure 5.16: Storage profits for different ramp limits with normal transmission capacities

$\Pi_{\omega_1}^{D_1}$ is discarded and the state of charge is not optimized at the end of day D_0 . However, tuning γ^{D_1} increases storage profitability as shown for $\gamma^{D_1} = 0.75$. Increasing γ^{D_1} decreases (e.g. $\gamma^{D_1} = 1$) the storage profitability due to the high charging cost for the fSoC and the uncertainty on day D_1 .

Fig. 5.17 shows the storage profits for different power flow limits. In these simulations the power flow limits were adjusted using multipliers $[0.85, 0.95, 1.0, 1.05, 1.15, 1.5]$ relative to the base values used above. For any value of the discount factor γ^{D_1} , the total storage profit decreases for the highly constrained cases (*i.e.* multipliers 0.85, 0.95). This is due to the inability of the storage devices to perform spatio-temporal arbitrage because of the congestion. However, if congestion is eliminated (*i.e.* multiplier 1.5), all ES become available for providing spatio-temporal arbitrage and they can collect almost as much profit as in the base case. Even though more congestion may result in higher prices, congestion may also decrease the quantity that storage gets cleared in the market. The trade-off between these

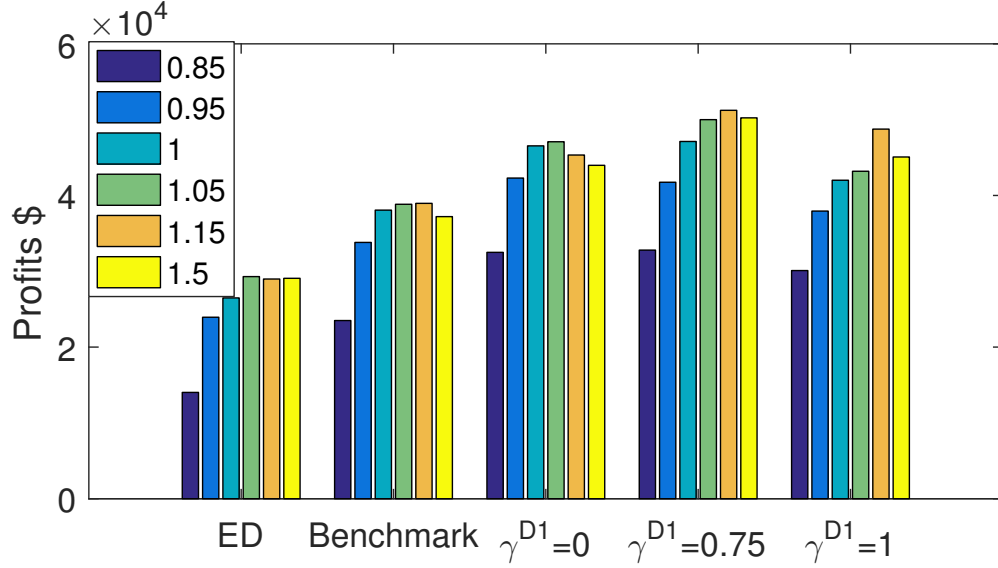


Figure 5.17: Profits under different transmission capacities with normal ramp limits

two factors determines whether or not it achieves a larger profit. For a merchant storage owner with several geographically dispersed devices, a moderately constrained case (*e.g.* a multiplier of 1.15) is more profitable.

5.4 Summary

In this section, we propose a risk-constrained, look-ahead bilevel technique to determine the optimal bids and offers for energy storage owners participating in a day-ahead electricity market. We maximize the total operating profits over two consecutive days. This look-ahead technique determines the energy state-of-charge at the end of the first day that best positions the storage device for taking advantage of arbitrage opportunities over the following day. This bilevel formulations emulates the interactions between the storage owner and the system operator. The upper-level problem maximizes the profits collected by the storage devices considering the operating constraints on the storage as well as the market clearing results from the lower level. In the lower-level problem, the system operator maximizes the

social welfare to clear the market. The resulting bilevel model is non-linear and non-convex and thus hard to solve. Karush-Kuhn-Tucker conditions and linearization techniques are used to transform this model into a single-level, mixed-integer linear equivalent, which can be solved efficiently with current commercial solvers. Test results demonstrate that this look-ahead bidding strategy improves the profitability of energy storage by setting a better initial position for the next market window. We also show that taking into account limits on the ramping capacity of conventional generators and transmission constraints increases the profitability of storage.

Chapter 6

CONCLUSIONS AND SUGGESTION FOR THE FUTURE WORKS

6.1 *Conclusions*

To address the operation challenges from wind uncertainty, this dissertation proposed an ensemble based wind modeling tool to generate wind scenarios. This ensemble provides a better point forecast and scenario input for the stochastic optimization. Due to limited computation power, current power system scheduling and market clearing engines cannot support operation with a large number of scenarios. To relieve such issue, we need to select a subset of scenarios to represent important operating conditions. By comparing current scenario reduction techniques, Fast Forward Selection algorithm outperforms other algorithms. Motivated with FFS, we propose a submodular scenario reduction algorithm to optimally rank and select scenarios. The scalable performance suggests its strong capabilities to handle the huge number of scenarios with multiple uncertainty sources.

To investigate the benefits from energy storage systems to the system and merchant participants. We first analyze ESS operation strategies from ISO perspective. Stochastic programming based strategies demonstrate a better ex-post economic and reliable performances comparing with deterministic, interval and robust strategies. In addition, the energy storage also assists the system operator to reach a better operating position to hedge against wind uncertainty. On the other hand, to incentive ESS investments from merchant participants, we propose a look-ahead bilevel bidding strategies to exploit energy arbitrage opportunities. The current intra-day scheme is extended to a inter-day setting with a more efficient energy allocation for the future usage.

6.2 *Suggestion for the Future Work*

Recently, deep learning or deep neural networks (DNN) has shown great progress and success in the machine learning filed. Many computer vision, natural language processing, speech recognition benchmarks are led with deep learning methods. Power system itself is a complex network with huge data processing needs. Naturally, we can improve the system with DNN in proper applications.

Forecasting is one direct application for DNN. Currently, one specific recurrent neural network (RNN) structure called Long short-term memory (LSTM) is suitable for time series predictions. Actually, load/price/wind/solar time series are very structured data. Directly applying DNN without any domain knowledge adjustments does not necessarily results in a significant performance improvements over existing ensemble methods. However, DNN is expected to show the full potentials if we, power system engineers, provide the tools with power system understandings to carefully select features and design the algorithms.

In addition to regression, DNN would be helpful for sequential decision-making processes as well with the deep reinforcement learning tools. It is also not surprising to see how recent progress in semi-supervised learning helps analyzing power systems data dependencies.

Submodular optimization is another nice tool to be further studied in the power systems. Works have been done with submodular optimization for PMU placement [214, 215], scenario reduction [165] and voltage control [216]. Since our power systems are full of discrete decisions to make, submodularity will be a great addition to derive efficient algorithms under big-data contexts. However, the physical network flows can create problems when defining the problem submodularity. More progress is expected in this field.

Back to energy storage, in this work, we mainly studied the ESS applications in energy arbitrage, but ESS can do much more than that. However, when providing multiple services, it is an interesting question to investigate the optimal portfolios for entering different markets. Moreover, the nonlinear degradation effects with BESS further complicates this problem when developing optimization formulations.

BIBLIOGRAPHY

- [1] Daniel Kirschen and Goran Strbac. *Fundamentals of Power System Economics*. John Wiley & Sons, Ltd, Chichester, UK, mar 2004.
- [2] Allen J. Wood and Bruce F. Wollenberg. *Power Generation, Operation, and Control*. Wiley, 2 edition, 1996.
- [3] M.A. Ortega-Vazquez and D.S. Kirschen. Estimating the Spinning Reserve Requirements in Systems With Significant Wind Power Generation Penetration. *IEEE Transactions on Power Systems*, 24(1):114–124, feb 2009.
- [4] G Giebel, R Brownsword, G Kariniotakis, Denhard. M, and C Draxl. The State of the Art in Short-Term Prediction of Wind Power, A Literature Overview, 2nd Edition. Technical report, ANEMOS.plus, 2011.
- [5] Yao Zhang, Jianxue Wang, and Xifan Wang. Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32:255–270, apr 2014.
- [6] Anthony Papavasiliou and Shmuel S. Oren. Multiarea Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network. *Operations Research*, 61(3):578–592, jun 2013.
- [7] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, and INESC Porto. Wind power forecasting : state-of-the-art 2009. Technical report, Argonne National Laboratory (ANL), nov 2009.
- [8] Miguel A. Ortega-Vazquez and Daniel S. Kirschen. Assessing the Impact of Wind Power Generation on Operating Costs. *IEEE Transactions on Smart Grid*, 1(3):295–301, Dec. 2010.
- [9] Zhi Zhou and Audun Botterud. Dynamic scheduling of operating reserves in co-optimized electricity markets with wind power. *IEEE Transactions on Power Systems*, 29(1):160–171, 2014.
- [10] Anthony Papavasiliou, Shmuel S. Oren, and Richard P. O’Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197–2206, nov 2011.

- [11] Peter Meibom, Rüdiger Barth, Bernhard Hasche, Heike Brand, Christoph Weber, and Mark O'Malley. Stochastic Optimization Model to Study the Operational Impacts of High Wind Penetrations in Ireland. *IEEE Transactions on Power Systems*, 26(3):1367–1379, aug 2011.
- [12] Qipeng P. Zheng, Jianhui Wang, Panos M. Pardalos, and Yongpei Guan. A decomposition approach to the two-stage stochastic unit commitment problem. *Annals of Operations Research*, 210(1):387–410, nov 2013.
- [13] Nicole Gröwe-Kuska, Holger Heitsch, and Werner Römisch. Scenario reduction and scenario tree construction for power management problems. *2003 IEEE Bologna PowerTech - Conference Proceedings*, 3:152–158, 2003.
- [14] Hrvoje Pandzic, Yishen Wang, Ting Qiu, Yury Dvorkin, and Daniel S. Kirschen. Near-Optimal Method for Siting and Sizing of Distributed Storage in a Transmission Network. *IEEE Transactions on Power Systems*, 30(5):2288–2300, sep 2015.
- [15] B. Dunn, H. Kamath, and J.-M. Tarascon. Electrical Energy Storage for the Grid: A Battery of Choices. *Science*, 334(6058):928–935, nov 2011.
- [16] Yunfeng Wen, Wenyuan Li, Gang Huang, and Xuan Liu. Frequency Dynamics Constrained Unit Commitment With Battery Energy Storage. *IEEE Transactions on Power Systems*, 31(6):5115–5125, nov 2016.
- [17] Yunfeng Wen, Chuangxin Guo, Hrvoje Pandzic, and Daniel S Kirschen. Enhanced Security-Constrained Unit Commitment With Emerging Utility-Scale Energy Storage. *IEEE Transactions on Power Systems*, 31(1):652–662, jan 2016.
- [18] David Pozo, Javier Contreras, and Enzo E Sauma. Unit Commitment With Ideal and Generic Energy Storage Units. *IEEE Transactions on Power Systems*, 29(6):2974–2984, nov 2014.
- [19] Ruiwei Jiang, Jianhui Wang, and Yongpei Guan. Robust unit commitment with wind power and pumped storage hydro. *IEEE Transactions on Power Systems*, 27(2):800–810, 2012.
- [20] Dimitris Bertsimas, Eugene Litvinov, Xu Andy Sun, Jinye Zhao, and Tongxin Zheng. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems*, 28(1):52–63, 2013.
- [21] Yang Wang, Qing Xia, and Chongqing Kang. Unit commitment with volatile node injections by using interval optimization. *IEEE Transactions on Power Systems*, 26(3):1705–1713, 2011.

- [22] Yury Dvorkin, Hrvoje Pandzic, Miguel A. Ortega-Vazquez, and Daniel S Kirschen. A Hybrid Stochastic/Interval Approach to Transmission-Constrained Unit Commitment. *IEEE Transactions on Power Systems*, 30:621–631, 2015.
- [23] Hrvoje Pandzic, Yury Dvorkin, Ting Qiu, Yishen Wang, and Daniel S Kirschen. Toward Cost-Efficient and Reliable Unit Commitment Under Uncertainty. *IEEE Transactions on Power Systems*, 31(2):970–982, mar 2016.
- [24] M. R. Sarker, Y. Dvorkin, and M. A. Ortega-Vazquez. Optimal participation of an electric vehicle aggregator in day-ahead energy and reserve markets. *IEEE Transactions on Power Systems*, 31(5):3506–3515, Sept 2016.
- [25] John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer New York, New York, NY, 2011.
- [26] Dimitris Bertsimas, David B. Brown, and Constantine Caramanis. Theory and Applications of Robust Optimization. *SIAM Review*, 53(3):464–501, jan 2011.
- [27] J. W. Chinneck and K. Ramadan. Linear Programming with Interval Coefficients. *The Journal of the Operational Research Society*, 51(2):209, feb 2000.
- [28] Yongpei Guan and Jianhui Wang. Uncertainty Sets for Robust Unit Commitment. *IEEE Transactions on Power Systems*, 29(3):1439–1440, may 2014.
- [29] Kjetil Høyland, Michal Kaut, and Stein W Wallace. A Heuristic for Moment-Matching Scenario Generation. *Computational Optimization and Applications*, 24(2/3):169–185, 2003.
- [30] Kjetil Høyland and Stein W Wallace. Generating scenario trees for multi-stage decision problems. Technical Report 2, 2001.
- [31] Colm Lowery and Mark O’Malley. Wind Power Scenario Tree Tool: Development and Methodology. In *Reliability and Risk Evaluation of Wind Integrated Power Systems*, pages 13–27. Springer, 2013.
- [32] J.M. Morales, R. Mínguez, and A.J. Conejo. A methodology to generate statistically dependent wind speed scenarios. *Applied Energy*, 87(3):843–855, mar 2010.
- [33] R. Billinton, Hua Chen, and R. Ghajar. A sequential simulation technique for adequacy evaluation of generating systems including wind energy. *IEEE Transactions on Energy Conversion*, 11(4):728–734, 1996.

- [34] R Karki, P Hu, and R Billinton. A Simplified Wind Power Generation Model for Reliability Evaluation. *IEEE Transactions on Energy Conversion*, 21(2):533–540, jun 2006.
- [35] Alexander Sturt and Goran Strbac. Efficient Stochastic Scheduling for Simulation of Wind-Integrated Power Systems. *IEEE Transactions on Power Systems*, 27(1):323–334, feb 2012.
- [36] Marcos S. Miranda and Rod W. Dunn. Spatially Correlated Wind Speed Modelling for Generation Adequacy Studies in the UK. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–6. IEEE, jun 2007.
- [37] G. Papaefthymiou and B. Klockl. MCMC for Wind Power Simulation. *IEEE Transactions on Energy Conversion*, 23(1):234–240, mar 2008.
- [38] George Papaefthymiou and Dorota Kurowicka. Using Copulas for Modeling Stochastic Dependence in Power System Uncertainty Analysis. *IEEE Transactions on Power Systems*, 24(1):40–49, feb 2009.
- [39] Ning Zhang, Chongqing Kang, Qian Yao Xu, Changming Jiang, Zhixu Chen, and Jun Liu. Modelling and Simulating the Spatio-Temporal Correlations of Clustered Wind Power Using Copula. *Journal of Electrical Engineering and Technology*, 8(6):1615–1625, nov 2013.
- [40] Pierre Pinson, Henrik Madsen, Henrik Aa Nielsen, George Papaefthymiou, and Bernd Klöckl. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy*, 12(1):51–62, jan 2009.
- [41] George Papaefthymiou and Pierre Pinson. Modeling of spatial dependence in wind power forecast uncertainty. In *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems*, pages 1–9, 2008.
- [42] Xi-Yuan Ma, Yuan-Zhang Sun, and Hua-Liang Fang. Scenario Generation of Wind Power Based on Statistical Uncertainty and Variability. *IEEE Transactions on Sustainable Energy*, 4(4):894–904, oct 2013.
- [43] Lei Wu, Mohammad Shahidehpour, and Zuyi Li. Comparison of Scenario-Based and Interval Optimization Approaches to Stochastic SCUC. *IEEE Transactions on Power Systems*, 27(2):913–921, may 2012.
- [44] Henrik Aalborg Nielsen, Torben Skov Nielsen, Henrik Madsen, Gregor Giebel, Jake Badger, Lars Landberg, Kai Sattler, Lars Voulund, and John Tofting. From wind

- ensembles to probabilistic information about future wind power production $\hat{p}$ results from an actual application. In *2006 International Conference on Probabilistic Methods Applied to Power Systems*, pages 1–8. IEEE, jun 2006.
- [45] James W. Taylor, Patrick E. McSharpy, and Roberto Buizza. Wind Power Density Forecasting Using Ensemble Predictions and Time Series Models. *IEEE Transactions on Energy Conversion*, 24(3):775–782, sep 2009.
 - [46] Emil M Constantinescu, Victor M Zavala, Matthew Rocklin, Sangmin Lee, and Mihai Anitescu. A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Transactions on Power Systems*, 26(1):431–441, 2011.
 - [47] NREL. NREL Eastern Wind Dataset.
 - [48] NREL. NREL Western Wind Dataset.
 - [49] Niya Chen, Zheng Qian, Ian T. Nabney, and Xiaofeng Meng. Wind Power Forecasts Using Gaussian Processes and Numerical Weather Prediction. *IEEE Transactions on Power Systems*, 29(2):656–665, mar 2014.
 - [50] Stylianos I. Vagropoulos, Evaggelos G. Kardakos, Christos K. Simoglou, Anastasios G. Bakirtzis, and João P.S. Catalão. ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electric Power Systems Research*, 134:9–18, may 2016.
 - [51] George Sideratos and Nikos D. Hatziaargyriou. Probabilistic wind power forecasting using radial basis function neural networks. *IEEE Transactions on Power Systems*, 27(4):1788–1796, 2012.
 - [52] Nima Amjady, Farshid Keynia, and Hamidreza Zareipour. Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization. *IEEE Transactions on Sustainable Energy*, 2(3):265–276, 2011.
 - [53] Song Li, Peng Wang, and Lalit Goel. Wind Power Forecasting Using Neural Network Ensembles With Feature Selection. *IEEE Transactions on Sustainable Energy*, 6(4):1447–1456, oct 2015.
 - [54] J. P. Catalao, H. Pousinho, and V. Mendes. Hybrid wavelet-pso-anfis approach for short-term wind power forecasting in portugal. *IEEE Transactions on Sustainable Energy*, 2(1):50–59, 2011.

- [55] Alexander Sturt and Goran Strbac. Time series modelling of power output for large-scale wind fleets. *Wind Energy*, 14(8):953–966, nov 2011.
- [56] Peiyuan Chen, Troels Pedersen, Birgitte Bak-Jensen, and Zhe Chen. ARIMA-Based Time Series Model of Stochastic Wind Power Generation. *IEEE Transactions on Power Systems*, 25(2):667–676, may 2010.
- [57] Yongqian Liu, Jie Shi, Yongping Yang, and Wei-jen Lee. Short-Term Wind-Power Prediction Based on Wavelet Transform&-Support Vector Machine and Statistic-Characteristics Analysis. *IEEE Transactions on Industry Applications*, 48(4):1136–1141, jul 2012.
- [58] Andrew Kusiak, Haiyang Zheng, and Zhe Song. Short-term prediction of wind farm power: A data mining approach. *IEEE Transactions on energy conversion*, 24(1):125–136, 2009.
- [59] Tao Hong, Pierre Pinson, and Shu Fan. Global energy forecasting competition 2012. *International Journal of Forecasting*, 30(2):357–363, 2014.
- [60] E. Mangalova and E. Agafonov. Wind power forecasting using the k-nearest neighbors algorithm. *International Journal of Forecasting*, 30(2):402–406, apr 2014.
- [61] Tao Hong, Pierre Pinson, Shu Fan, Hamidreza Zareipour, Alberto Troccoli, and Rob J. Hyndman. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *International Journal of Forecasting*, 32(3):896–913, jul 2016.
- [62] Pierre Pinson, Henrik Aa. Nielsen, Jan K. Møller, Henrik Madsen, and George N. Kariniotakis. Non-parametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 10(6):497–516, nov 2007.
- [63] Pierre Pinson, Christophe Chevallier, and George N. Kariniotakis. Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power. *IEEE Transactions on Power Systems*, 22(3):1148–1156, aug 2007.
- [64] R. J. Bessa, V. Miranda, A. Botterud, Jianhui Wang, and E. M. Constantinescu. Time adaptive conditional kernel density estimation for wind power forecasting. *IEEE Transactions on Sustainable Energy*, 3(4):660–669, 2012.
- [65] James W. Taylor and Jooyoung Jeon. Forecasting wind power quantiles using conditional kernel estimation. *Renewable Energy*, 80:370–379, aug 2015.

- [66] Wei Wu, Keyou Wang, Bei Han, Guojie Li, Xiuchen Jiang, and Mariesa L. Crow. A Versatile Probability Model of Photovoltaic Generation Using Pair Copula Construction. *IEEE Transactions on Sustainable Energy*, 6(4):1337–1345, oct 2015.
- [67] George E. P. Box, Gwilym M. Jenkins, and Gregory C. Reinsel. *Time Series Analysis*. John Wiley & Sons, Inc., Hoboken, NJ, jun 2008.
- [68] Leo Breiman. Random Forest. *Machine Learning*, 45(1):5–32, 2001.
- [69] Pu Wang, Bidong Liu, and Tao Hong. Electric load forecasting with recency effect: A big data approach. *International Journal of Forecasting*, 32(3):585–597, jul 2016.
- [70] Katrin Kirchhoff, Yuzong Liu, and Jeff Bilmes. Classification of developmental disorders from speech signals using submodular feature selection. In *Proceedings of Annual Conference of the International Speech Communication Association (INTER-SPEECH)*, 2013.
- [71] Hanchuan Peng, Fuhui Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, aug 2005.
- [72] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York, New York, NY, 2009.
- [73] Bradley Efron. Bootstrap Methods: Another Look at the Jackknife. pages 569–593. 1992.
- [74] Alex J. Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, aug 2004.
- [75] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27, apr 2011.
- [76] Leo Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, aug 1996.
- [77] Pan Li, Baosen Zhang, Yang Weng, and Ram Rajagopal. A sparse linear model and significance test for individual consumption prediction. *IEEE Transactions on Power Systems*, 2017. in press.
- [78] Mark Landry, Thomas P. Erlinger, David Patschke, and Craig Varrichio. Probabilistic gradient boosting machines for {GEFCom2014} wind forecasting. *International Journal of Forecasting*, 32(3):1061 – 1066, 2016.

- [79] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794. ACM, 2016.
- [80] Pierre Gaillard, Yannig Goude, and Raphaël Nedellec. Additive models and robust aggregation for {GEFCom2014} probabilistic electric load and electricity price forecasting. *International Journal of Forecasting*, 32(3):1038 – 1050, 2016.
- [81] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [82] Manuel a. Matos and R. J. Bessa. Setting the Operating Reserve Using Probabilistic Wind Power Forecasts. *IEEE Transactions on Power Systems*, 26(2):594–603, May 2011.
- [83] Audun Botterud, Zhi Zhou, Jianhui Wang, Jean Sumaili, Hrvoje Keko, Joana Mendes, Ricardo J. Bessa, and Vladimiro Miranda. Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: A case study of Illinois. *IEEE Transactions on Sustainable Energy*, 4(1):250–261, 2013.
- [84] Z. Zhou, A. Botterud, J. Wang, R.J. Bessa, H. Keko, J. Sumaili, and V. Miranda. Application of probabilistic wind power forecasting in electricity markets. *Wind Energy*, 16(3):321–338, Apr 2013.
- [85] Kenneth Bruninx, Kenneth Van den Bergh, Erik Delarue, and William D’haeseleer. Optimization and Allocation of Spinning Reserves in a Low-Carbon Framework. *IEEE Transactions on Power Systems*, 31(2):872–882, Mar 2016.
- [86] K. Bruninx and E. Delarue. Endogenous probabilistic reserve sizing and allocation in unit commitment models: cost-effective, reliable and fast. *IEEE Transactions on Power Systems*, 2016. in press.
- [87] Joshua D Lyon, Fengyu Wang, Kory W Hedman, and Muhong Zhang. Market Implications and Pricing of Dynamic Reserve Policies for Systems With Renewables. *IEEE Transactions on Power Systems*, 30(3):1593–1602, May 2015.
- [88] Yonghong Chen, Ryan Leonard, Marc Keyser, and Joe Gardner. Development of Performance-Based Two-Part Regulating Reserve Compensation on MISO Energy and Ancillary Service Market. *IEEE Transactions on Power Systems*, 30(1):142–155, jan 2015.

- [89] Miguel Angel Ortega Vazquez. *Optimizing the spinning reserve requirements*. PhD thesis, University of Manchester, 2006.
- [90] Y. Dvorkin, D. S. Kirschen, and M. A. Ortega-Vazquez. Assessing flexibility requirements in power systems. *IET Generation, Transmission Distribution*, 8(11):1820–1830, 2014.
- [91] Y. Dvorkin, M. A. Ortega-Vazquez, and D. S. Kirschen. Wind generation as a reserve provider. *IET Generation, Transmission Distribution*, 9(8):779–787, 2015.
- [92] X. Zhang, K. Tomsovic, and A. Dimitrovski. Security constrained multi-stage transmission expansion planning considering a continuously variable series reactor. *IEEE Transactions on Power Systems*, 2017. in press.
- [93] G. Liu, M. Starke, Xiaohu Zhang, and K. Tomsovic. A milp-based distribution optimal power flow model for microgrid operation. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, July 2016.
- [94] X. Zhang, K. Tomsovic, and A. Dimitrovski. Optimal investment on series facts device considering contingencies. In *2016 North American Power Symposium (NAPS)*, pages 1–6, Sept 2016.
- [95] G. Liu, M. Starke, B. Xiao, X. Zhang, and K. Tomsovic. Microgrid optimal scheduling with chance-constrained islanding capability. *Electric Power Systems Research*, 145:197 – 206, 2017.
- [96] X. Wang, D. Shi, Z. Wang, C. Xu, Q. Zhang, X. Zhang, and Z. Yu. Online calibration of phasor measurement unit using density-based spatial clustering. *IEEE Transactions on Power Delivery*, 2017. in press.
- [97] Anthony Papavasiliou, Shmuel S Oren, and Barry Rountree. Applying High Performance Computing to Transmission-Constrained Stochastic Unit Commitment for Renewable Energy Integration. *IEEE Transactions on Power Systems*, 30(3):1109–1120, May 2015.
- [98] Yishen Wang, Zhi Zhou, Cong Liu, and Audun Botterud. Evaluating stochastic methods in power system operations with wind power. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–6. IEEE, Apr 2016.
- [99] Kwok Cheung, Dinakar Gade, César Silva-Monroy, Sarah M. Ryan, Jean-Paul Watson, Roger J.-B. Wets, and David L. Woodruff. Toward scalable stochastic unit commitment. *Energy Systems*, 6(3):417–438, 2015.

- [100] P. A. Ruiz and R. Philbrick. Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems*, 24(2):642–651, 2009.
- [101] Lei Wu, Mohammad Shahidehpour, and Tao Li. Stochastic Security-Constrained Unit Commitment. *IEEE Transactions on Power Systems*, 22(2):800–811, May 2007.
- [102] Qipeng P. Zheng, Jianhui Wang, and Andrew L. Liu. Stochastic Optimization for Unit Commitment—A Review. *IEEE Transactions on Power Systems*, 30(4):1913–1924, Jul 2015.
- [103] Yury Dvorkin, Yishen Wang, Hrvoje Pandzic, and Daniel Kirschen. Comparison of scenario reduction techniques for the stochastic unit commitment. In *2014 IEEE PES General Meeting — Conference & Exposition*, pages 1–5. IEEE, Jul 2014.
- [104] Juan M. Morales, Salvador Pineda, Antonio J. Conejo, and Miguel Carrión. Scenario reduction for futures market trading in electricity markets. *IEEE Transactions on Power Systems*, 24(2):878–888, 2009.
- [105] Yonghan Feng and Sarah M Ryan. Scenario reduction for stochastic unit commitment with wind penetration. In *2014 IEEE PES General Meeting — Conference & Exposition*, pages 1–5. IEEE, Jul 2014.
- [106] Kenneth Bruninx and Erik Delarue. Scenario reduction techniques and solution stability for stochastic unit commitment problems. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–7, Apr 2016.
- [107] J. Dupačová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming. *Mathematical Programming*, 95(3):493–511, Mar 2003.
- [108] Nicole Gröwe-Kuska. GAMS/SCENRED Tutorial. Technical report.
- [109] Yonghan Feng and Sarah M. Ryan. Solution sensitivity-based scenario reduction for stochastic unit commitment. *Computational Management Science*, 13(1):29–62, Jan 2016.
- [110] J. Gomez-Martinez, M. Ortega-Vazquez, and L. F. Ochoa. Scenario selection of wind forecast errors for stochastic unit commitment: A uk case study. In *IEEE PES Innovative Smart Grid Technologies, Europe*, pages 1–6, Oct 2014.
- [111] J. Sumaili, H. Keko, V. Miranda, Z. Zhou, A. Botterud, and J. Wang. Finding representative wind power scenarios and their probabilities for stochastic models. In *2011 16th International Conference on Intelligent System Applications to Power Systems*, pages 1–6, Sept 2011.

- [112] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 1967.
- [113] David Arthur and Sergei Vassilvitskii. K-Means++: The Advantages of Careful Seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, 2007.
- [114] J F Benders. Partitioning procedures for solving mixed-variables programming problems. *Computational Management Science*, 2(1):3–19, Jan 2005.
- [115] Chao Li, Muhong Zhang, and Kory W Hedman. N-1 Reliable Unit Commitment via Progressive Hedging. *Journal of Energy Engineering*, 141(1), 2014.
- [116] Yu An and Bo Zeng. Exploring the Modeling Capacity of Two-Stage Robust Optimization: Variants of Robust Unit Commitment Model. *IEEE Transactions on Power Systems*, 30(1):109–122, Jan 2015.
- [117] Miguel Carrión and José Manuel Arroyo. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 21(3):1371–1378, 2006.
- [118] Morales-Espana, Jesus M. Latorre, and Andres Ramos. Tight and Compact MILP Formulation for the Thermal Unit Commitment Problem. *IEEE Transactions on Power Systems*, 28(4):4897–4908, Nov 2013.
- [119] The REAL Lab Library, Department of Electrical Engineering, University of Washington.
- [120] Hrvoje Pandzic, Ting Qiu, and Daniel S. Kirschen. Comparison of state-of-the-art transmission constrained unit commitment formulations. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5, 2013.
- [121] R. Baldick. Wind and energy markets: A case study of texas. *IEEE Systems Journal*, 6(1):27–34, March 2012.
- [122] K.K. Kariuki and R.N. Allan. Evaluation of reliability worth and value of lost load. *IEE Proceedings - Generation, Transmission and Distribution*, 143(2):171, 1996.
- [123] Bonneville Power Administration. Wind generation & Total Load in the BPA balancing authority.

- [124] M. S. Nazir and F. Bouffard. Intra-hour wind power characteristics for flexible operations. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE, July 2012.
- [125] Gerald J. Hahn and S.S. Shapiro. *Statistical Models in Engineering*. John Wiley and Sons Ltd, 1967.
- [126] Yuzong Liu, Kai Wei, Katrin Kirchhoff, Yisong Song, and Jeff Bilmes. Submodular feature selection for high-dimensional acoustic score space. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2013.
- [127] Yuzong Liu and Katrin Kirchhoff. Graph-based semi-supervised learning for phone and segment classification. In *Proceedings of Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2013.
- [128] Yuzong Liu and Katrin Kirchhoff. Graph-based semisupervised learning for acoustic modeling in automatic speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(11):1946–1956, 2016.
- [129] Yuzong Liu*, Kai Wei*, Katrin Kirchhoff, and Jeff Bilmes. Using document summarization techniques for speech data subset selection. In *HLT-NAACL*, pages 721–726, 2013.
- [130] Kai Wei, Yuzong Liu, Katrin Kirchhoff, and Jeff Bilmes. Unsupervised submodular subset selection for speech data. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 4107–4111. IEEE, 2014.
- [131] Kai Wei, Yuzong Liu, Katrin Kirchhoff, Chris Bartels, and Jeff Bilmes. Submodular subset selection for large-scale speech training data. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 3311–3315. IEEE, 2014.
- [132] Yuzong Liu, Rishabh Iyer, Katrin Kirchhoff, and Jeff Bilmes. Switchboard ii and fisver i: High-quality limited-complexity corpora of conversational english speech. In *Interspeech*, 2015.
- [133] Yuzong Liu, Rishabh Iyer, Katrin Kirchhoff, and Jeff Bilmes. Switchboard-ii and fisver-i: Crafting high quality and low complexity conversational english speech corpora using submodular function optimization. *Computer Speech & Language*, 42:122–142, 2017.
- [134] Yuzong Liu and Katrin Kirchhoff. A comparison of graph construction and learning algorithms for graph-based phonetic classification. *UWEE Technical Report, UWETR-2012-0005*, 2012.

- [135] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, Dec. 1978.
- [136] Michel Minoux. Accelerated greedy algorithms for maximizing submodular set functions. In *Optimization Techniques*, volume XXXIII, pages 234–243. Springer-Verlag, 2014.
- [137] IEA. Technology roadmap energy storage. Technical report, International Energy Agency, 2014.
- [138] Bolun Xu, Yury Dvorkin, Daniel S. Kirschen, C. A. Silva-Monroy, and Jean-Paul Watson. A comparison of policies on the participation of storage in U.S. frequency regulation markets. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, jul 2016.
- [139] Yang-Wu Shen, De-Ping Ke, Yuan-Zhang Sun, Kirschen Daniel, Yi-Shen Wang, and Yuan-Chao Hu. A novel transient rotor current control scheme of a doubly-fed induction generator equipped with superconducting magnetic energy storage for voltage and frequency support. *Chinese Physics B*, 24(7):070201, Jul. 2015.
- [140] Hongbo Sun, Yishen Wang, Daniel Nikovski, and Jinyun Zhang. Flex-Grid: A dynamic and adaptive configurable power distribution system. In *2015 IEEE Eindhoven PowerTech*, pages 1–6, Jun. 2015.
- [141] Yunfeng Wen, Chuangxin Guo, Daniel S Kirschen, and Shufeng Dong. Enhanced Security-Constrained OPF With Distributed Battery Energy Storage. *IEEE Transactions on Power Systems*, 30(1):98–108, Jan. 2015.
- [142] Hadi Khani, Mohammad R. Dadash Zadeh, and Amir H Hajimiragha. Transmission Congestion Relief Using Privately Owned Large-Scale Energy Storage Systems in a Competitive Electricity Market. *IEEE Transactions on Power Systems*, 31(2):1449–1458, Mar. 2016.
- [143] Trishna Das, Venkat Krishnan, and James D. McCalley. High-Fidelity Dispatch Model of Storage Technologies for Production Costing Studies. *IEEE Transactions on Sustainable Energy*, 5(4):1242–1252, Oct. 2014.
- [144] Trishna Das, Venkat Krishnan, and James D. McCalley. Assessing the benefits and economics of bulk energy storage technologies in the power grid. *Applied Energy*, 139:104–118, Feb. 2015.

- [145] Kenneth Bruninx, Yury Dvorkin, Erik Delarue, Hrvoje Pandzic, William Dhaeseleer, and Daniel S Kirschen. Coupling Pumped Hydro Energy Storage With Unit Commitment. *IEEE Transactions on Sustainable Energy*, 7(2):786–796, Apr. 2016.
- [146] Zhe Yu, Shanjun Li, and Lang Tong. Market dynamics and indirect network effects in electric vehicle diffusion. *Transportation Research Part D: Transport and Environment*, 47:336 – 356, 2016.
- [147] Z. Yu and L. Tong. Demand response via large scale charging of electric vehicles. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, July 2016.
- [148] Z. Yu, S. Chen, and L. Tong. An intelligent energy management system for large-scale charging of electric vehicles. *CSEE Journal of Power and Energy Systems*, 2(1):47–53, March 2016.
- [149] Zhe Yu. *Large scale charging of electric vehicles: Technology and economy*. PhD thesis, Cornell University, 2017.
- [150] M. R. Sarker, H. Pandžić, and M. A. Ortega-Vazquez. Optimal operation and services scheduling for an electric vehicle battery swapping station. *IEEE Transactions on Power Systems*, 30(2):901–910, March 2015.
- [151] Ciara O’Dwyer and Damian Flynn. Using Energy Storage to Manage High Net Load Variability at Sub-Hourly Time-Scales. *IEEE Transactions on Power Systems*, 30(4):2139–2148, jul 2015.
- [152] Nan Li and Kory W. Hedman. Economic Assessment of Energy Storage in Systems With High Levels of Renewable Resources. *IEEE Transactions on Sustainable Energy*, 6(3):1103–1111, jul 2015.
- [153] Nan Li, Canan Uckun, Emil M Constantinescu, John R Birge, Kory W Hedman, and Audun Botterud. Flexible Operation of Batteries in Power System Scheduling With Renewable Energy. *IEEE Transactions on Sustainable Energy*, 7(2):685–696, apr 2016.
- [154] Canan Uckun, Audun Botterud, and John R Birge. An Improved Stochastic Unit Commitment Formulation to Accommodate Wind Uncertainty. *IEEE Transactions on Power Systems*, 31(4):2507–2517, jul 2016.
- [155] V. Viswanathan, M. Kintner-Meyer, P. Balducci, and C. Jin. National assessment of energy storage for grid balancing and arbitrage phase ii: Volume 2 cost and performance characterization. Technical report, PNNL, 2013.

- [156] Soroush Shafiee, Hamidreza Zareipour, Andrew M. Knight, Nima Amjady, and Behnam Mohammadi-Ivatloo. Risk-Constrained Bidding and Offering Strategy for a Merchant Compressed Air Energy Storage Plant. *IEEE Transactions on Power Systems*, 32(2):946–957, 2017.
- [157] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, and D. Kirschen. Modeling of lithium-ion battery degradation for cell life assessment. *IEEE Transactions on Smart Grid*, 2016. in press.
- [158] Y. Li and Y. Han. A new perspective on battery cell balancing: Thermal balancing and relative temperature control. In *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 1–5, Sept 2016.
- [159] Y. Li and Y. Han. A module-integrated distributed battery energy storage and management system. *IEEE Transactions on Power Electronics*, 31(12):8260–8270, Dec 2016.
- [160] Y. Li and Y. Han. Evaluation of a module-integrated distributed battery energy storage system. In *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 1351–1358, Sept 2015.
- [161] Y. Li and Y. Han. Modelling and control of single-input and multiple-output modular converter systems. In *2015 IEEE 16th Workshop on Control and Modeling for Power Electronics (COMPEL)*, pages 1–8, July 2015.
- [162] Y. Li and Y. Han. Used-battery management with integrated battery building block system. In *2015 IEEE Applied Power Electronics Conference and Exposition (APEC)*, pages 3177–3182, March 2015.
- [163] Y. Li and Y. Han. Power electronics integration on battery cells. In *2014 IEEE Applied Power Electronics Conference and Exposition - APEC 2014*, pages 3318–3322, March 2014.
- [164] B. Dong, Y. Li, and Y. Han. Parallel architecture for battery charge equalization. *IEEE Transactions on Power Electronics*, 30(9):4906–4913, Sept 2015.
- [165] Yishen Wang, Yuzong Liu, and Daniel S. Kirschen. Scenario reduction with submodular optimization. *IEEE Transactions on Power Systems*, 32(3):2479–2480, May 2017.
- [166] Cheng Wang, Feng Liu, Jianhui Wang, Feng Qiu, Wei Wei, Shengwei Mei, and Shunbo Lei. Robust Risk-Constrained Unit Commitment With Large-Scale Wind Generation: An Adjustable Uncertainty Set Approach. *IEEE Transactions on Power Systems*, 32(1):723–733, jan 2017.

- [167] L. S. Vargas, G. Bustos-Turu, and F. Larraín. Wind power curtailment and energy storage in transmission congestion management considering power plants ramp rates. *IEEE Transaction on Power Systems*, 30(5):2498–2506, Sept. 2015.
- [168] Ting Qiu, Bolun Xu, Yishen Wang, Yury Dvorkin, and Daniel S. Kirschen. Stochastic Multistage Coplanning of Transmission Expansion and Energy Storage. *IEEE Transactions on Power Systems*, 32(1):643–651, Jan. 2017.
- [169] R. Fernández-Blanco, Y. Dvorkin, B. Xu, Y. Wang, and D. S. Kirschen. Optimal energy storage siting and sizing: A wecc case study. *IEEE Transactions on Sustainable Energy*. early access.
- [170] Yury Dvorkin, Ricardo Fernandez-Blanco, Daniel S. Kirschen, Hrvoje Pandzic, Jean-Paul Watson, and Cesar A. Silva-Monroy. Ensuring Profitability of Energy Storage. *IEEE Transactions on Power Systems*, 32(1):611–623, jan 2017.
- [171] Sonja Wogrin, David Galbally, and Javier Reneses. Optimizing Storage Operations in Medium- and Long-Term Power System Models. *IEEE Transactions on Power Systems*, 31(4):3129–3138, jul 2016.
- [172] Hamed Mohsenian-Rad. Optimal Bidding, Scheduling, and Deployment of Battery Systems in California Day-Ahead Energy Market. *IEEE Transactions on Power Systems*, 31(1):442–453, jan 2016.
- [173] H. Akhavan-Hejazi and H. Mohsenian-Rad. Optimal operation of independent storage systems in energy and reserve markets with high wind penetration. *IEEE Transactions on Smart Grid*, 5(2):1088–1097, March 2014.
- [174] Bolun Xu, Alexandre Oudalov, Jan Poland, Andreas Ulbig, and Göran Andersson. BESS Control Strategies for Participating in Grid Frequency Regulation. *IFAC Proceedings Volumes*, 47(3):4024–4029, 2014.
- [175] Guannan He, Qixin Chen, Chongqing Kang, Pierre Pinson, and Qing Xia. Optimal Bidding Strategy of Battery Storage in Power Markets Considering Performance-Based Regulation and Battery Cycle Life. *IEEE Transactions on Smart Grid*, 7(5):2359–2367, Sep. 2016.
- [176] Hamed Mohsenian-Rad. Coordinated Price-Maker Operation of Large Energy Storage Units in Nodal Energy Markets. *IEEE Transactions on Power Systems*, 31(1):786–797, Jan. 2016.

- [177] Hrvoje Pandzic and Igor Kuzle. Energy storage operation in the day-ahead electricity market. In *2015 12th International Conference on the European Energy Market (EEM)*, number 13, pages 1–6. IEEE, May 2015.
- [178] Karl Hartwig and Ivana Kockar. Impact of Strategic Behavior and Ownership of Energy Storage on Provision of Flexibility. *IEEE Transactions on Sustainable Energy*, 7(2):744–754, Apr. 2016.
- [179] Huajie Ding, Pierre Pinson, Zechun Hu, and Yonghua Song. Integrated Bidding and Operating Strategies for Wind-Storage Systems. *IEEE Transactions on Sustainable Energy*, 7(1):163–172, Jan. 2016.
- [180] Huajie Ding, Pierre Pinson, Zechun Hu, and Yonghua Song. Optimal Offering and Operating Strategies for Wind-Storage Systems With Linear Decision Rules. *IEEE Transactions on Power Systems*, 31(6):4755–4764, Nov. 2016.
- [181] Anupam a. Thatte, Le Xie, Daniel E. Viassolo, and Sunita Singh. Risk Measure Based Robust Bidding Strategy for Arbitrage Using a Wind Farm and Energy Storage. *IEEE Transactions on Smart Grid*, 4(4):2191–2199, Dec. 2013.
- [182] Qianchuan Zhao, Ying Shen, and Mingyang Li. Control and Bidding Strategy for Virtual Power Plants With Renewable Generation and Inelastic Demand in Electricity Markets. *IEEE Transactions on Sustainable Energy*, 7(2):562–575, Apr. 2016.
- [183] Zhiwei Xu, Zechun Hu, Yonghua Song, and Jianhui Wang. Risk-Averse Optimal Bidding Strategy for Demand-Side Resource Aggregators in Day-Ahead Electricity Markets Under Uncertainty. *IEEE Transactions on Smart Grid*, 8(1):96–105, Jan 2017.
- [184] C. Ruiz and A. J. Conejo. Pool strategy of a producer with endogenous formation of locational marginal prices. *IEEE Transactions on Power Systems*, 24(4):1855–1866, Nov 2009.
- [185] C. Ruiz, A. J. Conejo, and D. J. Bertsimas. Revealing rival marginal offer prices via inverse optimization. *IEEE Transaction on Power Systems*, 28(3):3056–3064, Aug 2013.
- [186] Mostafa Kazemi, Hamidreza Zareipour, Mehdi Ehsan, and William D. Rosehart. A Robust Linear Approach for Offering Strategy of a Hybrid Electric Energy Company. *IEEE Transactions on Power Systems*, 2016. in press.
- [187] D. Pozo and J. Contreras. Finding multiple nash equilibria in pool-based markets: A stochastic epec approach. *IEEE Transactions on Power Systems*, 26(3):1744–1752, Aug 2011.

- [188] T. Dai and W. Qiao. Finding equilibria in the pool-based electricity market with strategic wind power producers and network constraints. *IEEE Transactions on Power Systems*, 32(1):389–399, Jan 2017.
- [189] C. Ruiz, A. J. Conejo, and Y. Smeers. Equilibria in an oligopolistic electricity pool with stepwise offer curves. *IEEE Transactions on Power Systems*, 27(2):752–761, May 2012.
- [190] Anya Castillo and Dennice F. Gayme. Profit maximizing storage allocation in power grids. In *52nd IEEE Conference on Decision and Control*, pages 429–435. IEEE, dec 2013.
- [191] Huajie Ding, Zechun Hu, and Yonghua Song. Rolling Optimization of Wind Farm and Energy Storage System in Electricity Markets. *IEEE Transactions on Power Systems*, 30(5):2676–2684, sep 2015.
- [192] Huajie Ding, Pierre Pinson, Zechun Hu, Jianhui Wang, and Yonghua Song. Optimal offering and operating strategy for a large wind-storage system as a price maker. *IEEE Transactions on Power Systems*. in press.
- [193] M.A. Ortega-Vazquez. Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty. *IET Transactions on Generation, Transmission & Distribution*, 8(6):1007–1016, 2014.
- [194] Stephan Dempe. *Foundations of bilevel programming*. Springer Science & Business Media, 2002.
- [195] S. Gabriel, A. Conejo, J. Fuller, B. F. Hobbs, and C. Ruiz. *Complementarity modeling in energy markets*, volume 180. Springer Science & Business Media, 2012.
- [196] J.M. Arroyo. Bilevel programming applied to power system vulnerability analysis under multiple contingencies. *IET Generation, Transmission & Distribution*, 4(2):178, 2010.
- [197] R. Fernández-Blanco, J. M. Arroyo, and N. Alguacil. A unified bilevel programming framework for price-based market clearing under marginal pricing. *IEEE Transactions on Power Systems*, 27(1):517–525, Feb. 2012.
- [198] M. Carrion, J. M. Arroyo, and A. J. Conejo. A bilevel stochastic programming approach for retailer futures market trading. *IEEE Transactions on Power Systems*, 24(3):1446–1456, Aug. 2009.

- [199] S. J. Kazempour, A. J. Conejo, and C. Ruiz. Strategic generation investment using a complementarity approach. *IEEE Transactions on Power Systems*, 26(2):940–948, May 2011.
- [200] L. Baringo and A. J. Conejo. Strategic wind power investment. *IEEE Transactions on Power Systems*, 29(3):1250–1260, May 2014.
- [201] L. Baringo and A. J. Conejo. Offering strategy of wind-power producer: A multi-stage risk-constrained approach. *IEEE Transactions on Power Systems*, 31(2):1420–1429, March 2016.
- [202] Y. Liu and F.F. Wu. Impacts of network constraints on electricity market equilibrium. *IEEE Transactions on Power Systems*, 22(1):126–135, Feb. 2007.
- [203] T. Peng and K. Tomsovic. Congestion influence on bidding strategies in an electricity market. *IEEE Transactions on Power Systems*, 18(3):1054–1061, Aug 2003.
- [204] F. Gao and G.B. Sheble. Electricity market equilibrium model with resource constraint and transmission congestion. *Electric Power Systems Research*, 80(1):9–18, 2010.
- [205] Yishen Wang, Yury Dvorkin, Ricardo Fernández-blanco, and Daniel Kirschen. Look-Ahead Strategic Energy Storage in Ramp-Constrained Markets. In *FERC Technical Conference Increasing Real-Time and Day-Ahead Market Efficiency Through Improved Software*, 2016.
- [206] Ehsan Nasrolahpour, Hamidreza Zareipour, William D Rosehart, and S Jalal Kazempour. Bidding strategy for an energy storage facility. In *2016 Power Systems Computation Conference (PSCC)*, pages 1–7, jun 2016.
- [207] Renewable electricity futures study. Technical report, NREL, 2012.
- [208] Advancing and maximizing the value of energy storage technology. Technical report, CAISO, 2014.
- [209] Yishen Wang, Yury Dvorkin, Ricardo Fernández-Blanco, Bolun Xu, and Daniel S. Kirschen. Impact of Local Transmission Congestion on Energy Storage Arbitrage Opportunities. In *2017 IEEE Power Engineering Society General Meeting*, pages 1–5. IEEE, Jun 2017.
- [210] A. Akhil and *et al.* Doe/epri 2013 electricity storage handbook in collaboration with nreca. Technical report, Sandia National Lab, 2013.

- [211] José Fortuny-Amat and Bruce McCarl. A representation and economic interpretation of a two-level programming problem. *Journ. of the Oper. Res. Soc.*, 32(9):783–792, 1981.
- [212] GAMS Development Corporation, Washington, DC, USA. *GAMS - A User's Guide, GAMS Release 24.2.1*, 2013.
- [213] University of Washington, Seattle, WA, USA. *Hyak Supercomputer*, 2016.
- [214] Q. Li, T. Cui, Y. Weng, R. Negi, F. Franchetti, and M. D. Ilic. An information-theoretic approach to pmu placement in electric power systems. *IEEE Transactions on Smart Grid*, 4(1):446–456, March 2013.
- [215] P. Yang, Z. Tan, A. Wiesel, and A. Nehorai. Placement of pmus considering measurement phase-angle mismatch. *IEEE Transactions on Power Delivery*, 30(2):914–922, April 2015.
- [216] Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. Submodular optimization for voltage control. *IEEE Transactions on Power Systems*, 2017. in press.
- [217] Lars Kai Hansen and Peter Salamon. Neural network ensembles. *IEEE transactions on pattern analysis and machine intelligence*, 12(10):993–1001, 1990.
- [218] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning internal representations by error propagation. Technical report, DTIC Document, 1985.
- [219] Mads Almassalkhi, Yury Dvorkin, Jennifer Marley, Ricardo Fernandez-Blanco, Ian Hiskens, Daniel Kirschen, Jonathon Martin, Hrvoje Pandzic, Ting Qiu, Mushfiqu Sarker, Maria Vrakopoulou, Yishen Wang, and Mengran Xue. Incorporating storage as a flexible transmission asset in power system operation procedure. In *2016 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, jun 2016.
- [220] Kyri Baker, Gabriela Hug, and Xin Li. Energy Storage Sizing Taking Into Account Forecast Uncertainties and Receding Horizon Operation. *IEEE Transactions on Sustainable Energy*, 8(1):331–340, jan 2017.
- [221] Abebe W. Bizuayehu, Agustin A. Sanchez de la Nieta, Javier Contreras, and Joao P. S. Catalao. Impacts of Stochastic Wind Power and Storage Participation on Economic Dispatch in Distribution Systems. *IEEE Transactions on Sustainable Energy*, 7(3):1336–1345, jul 2016.

- [222] Rabih A. Jabr, Izudin Dzafic, and Bikash C. Pal. Robust Optimization of Storage Investment on Transmission Networks. *IEEE Transactions on Power Systems*, 30(1):531–539, jan 2015.
- [223] Rabih a Jabr, Sami Karaki, and Joe Akl Korbane. Robust Multi-Period OPF With Storage and Renewables. *IEEE Transactions on Power Systems*, 30(5):2790–2799, sep 2015.
- [224] Nan Li and Kory Hedman. Enhanced Pumped Hydro Storage Utilization using Policy Functions. *IEEE Transactions on Power Systems*, 2016. in press.
- [225] Zhengshuo Li, Qinglai Guo, Hongbin Sun, and Jianhui Wang. Storage-like devices in load leveling: Complementarity constraints and a new and exact relaxation method. *Applied Energy*, 151:13–22, aug 2015.
- [226] Zhengshuo Li, Qinglai Guo, Hongbin Sun, and Jianhui Wang. Sufficient Conditions for Exact Relaxation of Complementarity Constraints for Storage-Concerned Economic Dispatch. *IEEE Transactions on Power Systems*, 31(2):1653–1654, mar 2016.
- [227] Alvaro Lorca and Xu Sun. Multistage Robust Unit Commitment with Dynamic Uncertainty Sets and Energy Storage. *IEEE Transactions on Power Systems*, 2016. in press.
- [228] Carsten Matke, Daniel Bienstock, Gonzalo Muñoz, Shuoguang Yang, David Kleinhans, and Sebastian Sager. Robust optimization of power network operation: storage devices and the role of forecast errors in renewable energies. In *Complex Networks & Their Applications V: Proceedings of the 5th International Workshop on Complex Networks and their Applications (COMPLEX NETWORKS 2016)*, volume 693, pages 809–820. Springer International Publishing, 2017.
- [229] Ciara O’Dwyer, Lisa Ryan, and Damian Flynn. Efficient Large-Scale Energy Storage Dispatch: Challenges in Future High Renewables Systems. *IEEE Transactions on Power Systems*, 2017. in press.
- [230] Junjie Qin, Insoon Yang, and Ram Rajagopal. Submodularity of energy storage placement in power networks. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 686–693. IEEE, dec 2016.
- [231] Sonja Wogrin and Dennice F Gayme. Optimizing Storage Siting, Sizing, and Technology Portfolios in Transmission-Constrained Networks. *IEEE Transactions on Power Systems*, 30(6):3304–3313, nov 2015.

- [232] Zechun Hu, Shu Zhang, Fang Zhang, and Haiyan Lu. SCUC with battery energy storage system for peak-load shaving and reserve support. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.
- [233] Peng Zou, Qixin Chen, Qing Xia, Guannan He, and Chongqing Kang. Evaluating the Contribution of Energy Storages to Support Large-Scale Renewable Generation in Joint Energy and Ancillary Service Markets. *IEEE Transactions on Sustainable Energy*, 7(2):808–818, apr 2016.
- [234] R. Fernandez-Blanco, J. M. Arroyo, and N. Alguacil-Conde. On the solution of revenue- and network-constrained day-ahead market clearing under marginal pricing—part i: An exact bilevel programming approach. *IEEE Transactions on Power Systems*. early access.
- [235] R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of Classical Reliability in Test Construction. *Journal of Risk*, 2(3):1, jan 2000.
- [236] Wei Wei, Jianhui Wang, and Shengwei Mei. Dispatchability Maximization for Co-Optimized Energy and Reserve Dispatch With Explicit Reliability Guarantee. *IEEE Transactions on Power Systems*, 31(4):3276–3288, jul 2016.
- [237] Jinye Zhao, Tongxin Zheng, and Eugene Litvinov. Variable Resource Dispatch Through Do-Not-Exceed Limit. *IEEE Transactions on Power Systems*, 30(2):820–828, mar 2015.
- [238] Bolong Cheng and Warren Powell. Co-optimizing Battery Storage for the Frequency Regulation and Energy Arbitrage Using Multi-Scale Dynamic Programming. *IEEE Transactions on Smart Grid*, 2016. in press.
- [239] Guannan He, Qixin Chen, Chongqing Kang, Qing Xia, and Kameshwar Poolla. Cooperation of Wind Power and Battery Storage to Provide Frequency Regulation in Power Markets. *IEEE Transactions on Power Systems*, 2016. in press.
- [240] Dheepak Krishnamurthy, Canan Uckun, Zhi Zhou, Prakash Thimmapuram, and Audun Botterud. Energy Storage Arbitrage Under Day-Ahead and Real-Time Price Uncertainty. *IEEE Transactions on Power Systems*, 2017. in press.
- [241] Ehsan Nasrolahpour, Jalal Kazempour, Hamidreza Zareipour, and William D. Rosehart. Impacts of Ramping Inflexibility of Conventional Generators on Strategic Operation of Energy Storage Facilities. *IEEE Transactions on Smart Grid*, 2016. in press.

- [242] Ehsan Nasrolahpour, S. Jalal Kazempour, Hamidreza Zareipour, and William D. Rosehart. Strategic Sizing of Energy Storage Facilities in Electricity Markets. *IEEE Transactions on Sustainable Energy*, 7(4):1462–1472, oct 2016.
- [243] Soroush Shafiee, Payam Zamani-Dehkordi, Hamidreza Zareipour, and Andrew M. Knight. Economic assessment of a price-maker energy storage facility in the Alberta electricity market. *Energy*, 111:537–547, sep 2016.
- [244] Ramteen Sioshansi. Using Storage-Capacity Rights to Overcome the Cost-Recovery Hurdle for Energy Storage. *IEEE Transactions on Power Systems*, 2016. in press.
- [245] Ramteen Sioshansi, Paul Denholm, Thomas Jenkin, and Jurgen Weiss. Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects. *Energy Economics*, 31(2):269–277, mar 2009.
- [246] Yanli Wei, Hantao Cui, Xin Fang, and Fangxing Li. Strategic scheduling of energy storage for load serving entities in locational marginal pricing market. *IET Generation, Transmission & Distribution*, 10(5):1258–1267, apr 2016.
- [247] Ying Jun Zhang, Changhong Zhao, Wanrong Tang, and Steven H. Low. Profit Maximizing Planning and Control of Battery Energy Storage Systems for Primary Frequency Control. *IEEE Transactions on Smart Grid*, 2016. in press.

Appendix A

NOMENCLATURE

A.1 Chapter 3 and Chapter 4

A.1.1 Sets and Indices

B	Set of buses, indexed by b .
I	Set of conventional generators, indexed by i .
J	Set of conventional generator startup cost segments, indexed by j .
K	Set of conventional generator production cost blocks, indexed by k .
L	Set of transmission lines, indexed by l .
S	Set of scenarios, indexed by s .
T	Set of time intervals, indexed by t .
W	Set of wind generators, indexed by w .
H	Set of energy storage devices, indexed by h .
$f(l)$	Indices of sending buses of line l .
$t(l)$	Indices of receiving buses of line l .
$r(b)$	Indices of resources connected at buses b .

A.1.2 Binary Decision Variables

$x_{i,t}$	Commitment variable for generator i at time t .
$y_{i,t}$	Shut down indicator for generator i at time t .
$z_{i,t}$	Start-up indicator for generator i at time t .
$q_{i,j,t}$	Stepwise start-up cost indicator for generator i segment j at time t .

A.1.3 Continuous Non-negative Decision Variables

$csu_{i,t}$	Start-up cost for generator i at time t .
$cpg_{i,t,s}$	Energy production cost for generator i at time t for scenario s .
$p_{i,t,s}$	Dispatched power output for generator i at time t for scenario s .
$g_{i,k,t,s}$	Dispatched power output for generator i block k at time t for scenario s .
$r_{i,t,s}$	Scheduled reserve for generator i at time t for scenario s .
$ens_{b,t,s}$	Amount of not served energy at bus b time t for scenario s .
$w_{w,t}$	Scheduled wind generation output for wind farm w at time t .
$curt_{w,t,s}$	Wind dispatch curtailment for wind farm w at time t for scenario s .
$SoC_{h,t,s}$	Schedule Storage state of charge for storage h at time t for scenario s .
$q_{h,t,s}^{dis}$	Scheduled storage discharging power quantity for storage h at time t for scenario s .
$q_{h,t,s}^{chs}$	Scheduled storage charging power quantity for storage h at time t for scenario s .

A.1.4 Continuous Decision Variables

$pf_{l,t,s}$ Power flow on line l at time t for scenario s .

$\theta_{b,t,s}$ Voltage angle at bus b time t for scenario s .

A.1.5 Parameters

$PROB_s$ Probability of scenario s .

$VoLL$ Value of lost load.

$VoWS$ Value of wind spillage.

$D_{b,t}$ Load at bus b time t .

FL_l^{cap} Real power capacity of line l .

X_l Reactance of line l .

$SUC_{i,j}$ Startup cost of generator i cost segment j .

NLC_i No-load cost of generator i .

$MC_{i,k}$ Marginal cost of generator i at cost segment k .

MUT_i Minimum up time of generator i .

MDT_i Minimum down time of generator i .

P_i^{MIN} Minimum generation level of generator i .

P_i^{MAX} Maximum generation level of generator i .

$PR_{i,k}$ Power capacity of generator i on cost segment k .

$T_{i,j}^{MIN}$	Lower limit of stepwise start-up cost curve of generator i and segment j .
$T_{i,j}^{MAX}$	Upper limit of stepwise start-up cost curve of generator i and segment j .
RU_i	Hourly ramp up limit of generator i .
RD_i	Hourly ramp down limit of generator i .
PO_i	Initial power generation level of generator i .
$WF_{w,t,s}$	Forecasted wind generation of wind farm w at time t for scenario s (denoted as central forecast in DUC case).
$WF_{w,t}^{LB}$	Lower bound of forecasted wind generation of wind farm w at time t .
$WF_{w,t}^{UB}$	Upper bound of forecasted wind generation of wind farm w at time t .
$\eta_h^{chs}/\eta_h^{dis}$	Storage charging / discharging efficiency for storage h .
SoC_h^{min}/SoC_h^{max}	Minimum / Maximum SoC level for storage h , MWh.
SoC_h^{init}	Initial SoC for storage h , MWh.
Dis_h^{cap}/Chs_h^{cap}	Discharging/charging rate limit for storage h , MW.
α^{load}	Percentage of load to be reserved.
α^{wind}	Percentage of wind to be reserved.
MSR_i	Reserve rate per minute for generator i .
ΔT	Reserve time requirement.

A.2 Chapter 5

A.2.1 Sets and Indices

B	Set of buses, indexed by b .
I	Set of thermal generators, indexed by i .
L	Set of transmission lines, indexed by l .
H	Set of storage devices, indexed by h .
W	Set of wind generators, indexed by w .
Ω_1	Set of wind scenarios at D_1 , indexed by ω_1 .
Ω	Set of wind scenarios, indexed by ω . Note that ω_0 stands for D_0 deterministic forecast profiles.
T^{D_0}	Set of time intervals of the first market window indexed by $t_0 = 1 \dots N_{T_0}$
T^{D_1}	Set of time intervals of the second market window indexed by $t_1 = (N_{T_0} + 1) \dots N_{T_1}$.
T	Set of all time intervals, indexed by t . Note that $T = T^{D_0} \cup T^{D_1}$.
$f(l)/t(l)$	Indices of the from/to buses of line l .

A.2.2 Variables

$chs_{h,t}/dis_{h,t}$	Quantity of charging bid/discharging offer of storage h at time t , MW.
$\rho_{h,t}^{chs}/\rho_{h,t}^{dis}$	Price of charging bid/discharging offer of storage h at time t , \$/MWh.

$x_{h,t}^{dis}$	Binary status variable of storage h at time t . Equals to 1 if discharging and 0 otherwise.
$\Pi^{D_0}/\Pi_{\omega_1}^{D_1}$	Profit at D_0 / D_1 stage.
Γ^{D_1}	Risk-constrained look-ahead profits adjustments.
$CVaR^{D_1}$	Conditional value-at-risk of D_1 profits.
Var^{D_1}	Value-at-risk of D_1 profits.
$loss_{\omega_1}^{D_1}$	Auxiliary variables to compute D_1 CVaR.
$SoC_{h,t,\omega}$	State of charge of storage h at time t and scenario ω , MWh.
$d_{b,t,\omega}$	Cleared demand at bus b at time t and scenario ω , MW.
$p_{i,t,\omega}^G$	Cleared power output of conventional generator i at time t and scenario ω , MW.
$p_{w,t,\omega}^W$	Cleared power output of wind generator w at time t and scenario ω , MW.
$pf_{l,t,\omega}$	Power flow on line l at time t and scenario ω , MW.
$q_{h,t,\omega}^{dis/chs}$	Cleared discharging/charging rate for storage h at time t and scenario ω , MW.
$\theta_{b,t,\omega}$	Voltage phase angle at bus b at time t and scenario ω , rad.

A.2.3 Dual Variables

$\lambda_{b,t,\omega}$	Nodal power balance constraint at time t and scenario ω .
$\underline{\mu}_{i,t,\omega}^G / \bar{\mu}_{i,t,\omega}^G$	Min/max power output constraints of generator i at time t and scenario ω .
$\underline{\mu}_{w,t,\omega}^W / \bar{\mu}_{w,t,\omega}^W$	Min/max power output constraints for wind generator w at time t and scenario ω .
$\underline{\mu}_{b,t,\omega}^D / \bar{\mu}_{b,t,\omega}^D$	Min/max bounds on demand at bus b time t at scenario ω .
$\underline{\alpha}_{h,t,\omega}^{dis} / \bar{\alpha}_{h,t,\omega}^{dis}$	Min/max discharging rate constraints of storage h at time t and scenario ω .
$\underline{\alpha}_{h,t,\omega}^{chs} / \bar{\alpha}_{h,t,\omega}^{chs}$	Min/max charging rate constraints of storage h at time t and scenario ω .
$\nu_{l,t,\omega}$	Power flow constraints on line l at time t and scenario ω .
$\underline{\varphi}_{l,t,\omega} / \bar{\varphi}_{l,t,\omega}$	Min/max power flow constraints on line l at time t and scenario ω .
$\underline{\delta}_{b,t,\omega} / \bar{\delta}_{b,t,\omega}$	Min/max voltage phase angle constraints at bus b at time t and scenario ω .
$\delta_{b,t,\omega}^{ref}$	Reference bus voltage angle constraints at bus b at time t and scenario ω .
$\beta_{i,t,\omega}^u / \beta_{i,t,\omega}^d$	Up/down ramp limit constraints for generator i at time t and scenario ω .

A.2.4 Parameters

γ^{D_1}	Discount factor for the profit term of day D_1 .
γ^{CVaR}	Weight factor for the risk position of day D_1 .

C_h^{ES}	Variable operating cost for storage h , \$/MW.
C_i^G / C_w^W	Offer price for generator i / wind w , \$/MWh.
C_b^D	Bid price for demand b , \$/MWh.
$\overline{dis}_h / \overline{chs}_h$	Discharging/charging rate limit for storage h , MW.
$\eta_h^{dis} / \eta_h^{chs}$	Discharging/charging efficiency for storage h .
$\underline{D}_{b,t} / \overline{D}_{b,t}$	Min/max bounds on demand at bus b at time t , MW.
$\bar{F}_l$	Power flow limit on line l , MW.
$P_i / \bar{P}_i$	Min/max power output limits for generator i , MW.
$P_{i,0}^G$	Initial power output of generator i , MW.
RU_i / RD_i	Ramp up/down limits for generator i , MW.
$\underline{SoC}_h / \overline{SoC}_h$	Min/max state-of-charge for storage h , MWh.
SoC_h^{init}	Initial state-of-charge level for storage h , MWh.
X_l	Reactance of line l .
π_{ω_1}	Probability for wind scenario ω_1 on day D_1 .
β_0	Probability threshold for risk calculation.
$A_{l,b}$	Incidence matrix of transmission lines. 1/-1 indicates bus b is the from/to bus for line l , and 0 otherwise.
$WF_{w,t,\omega}$	Forecast wind w at time t and scenario ω , MW.

Appendix B

DIAGRAMS OF IEEE RTS TEST SYSTEMS

B.1 One-Area RTS

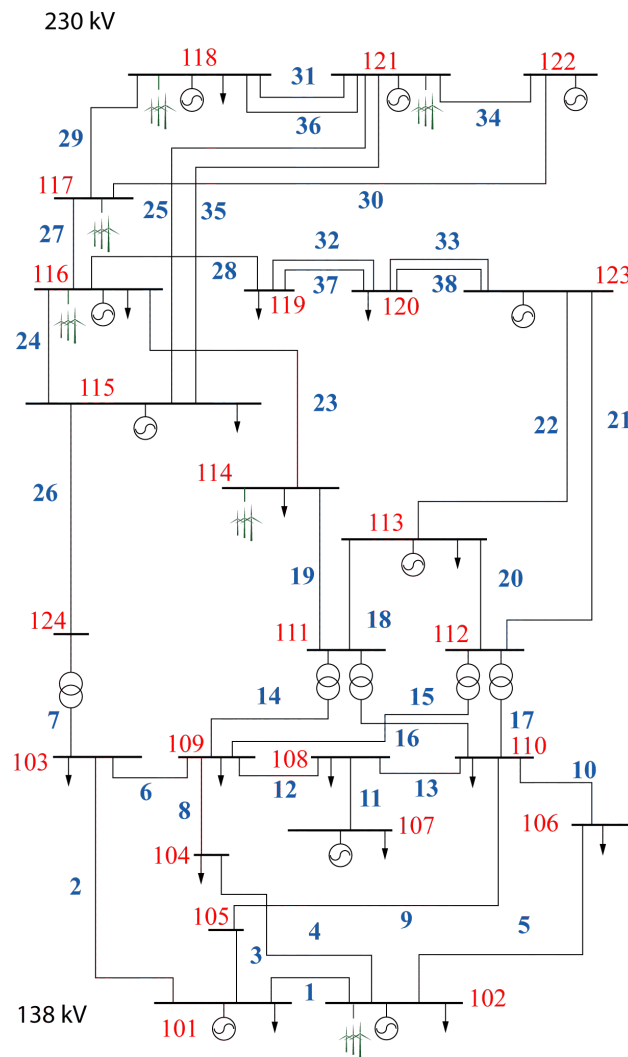


Figure B.1: One-Area RTS Diagram

B.2 Three-Area RTS

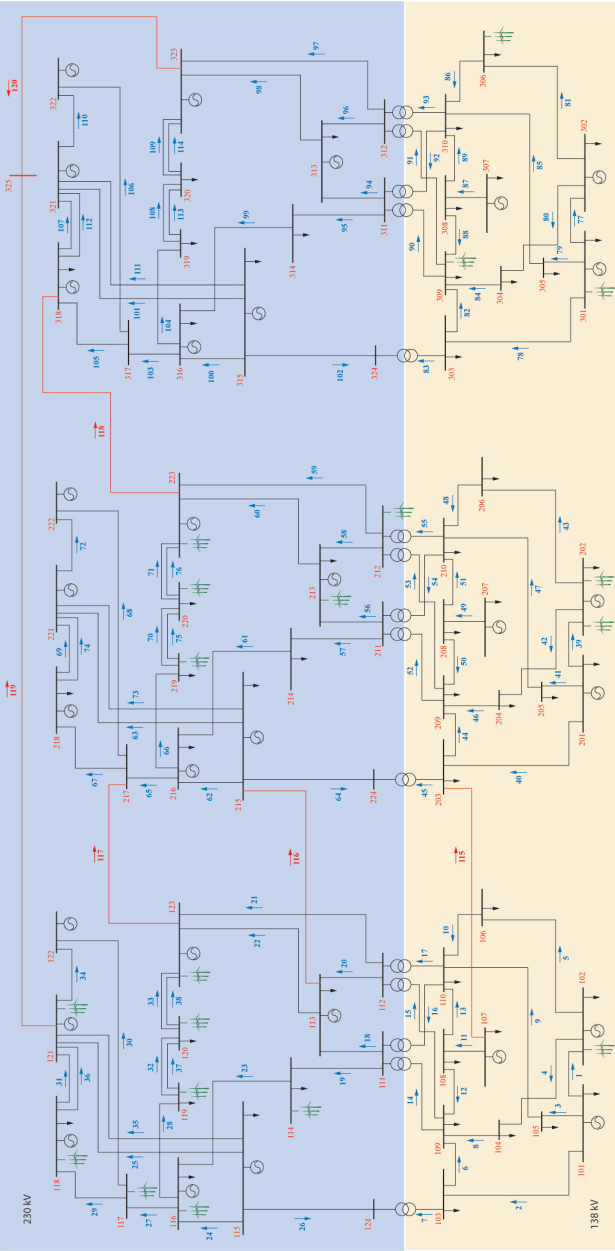


Figure B.2: Three-Area RTS Diagram

Appendix C

AUTHOR'S BIBLIOGRAPHY***Journal***

1. Y. Dvorkin, R. Fernández-Blanco, **Y. Wang**, B. Xu, D. S. Kirschen, H. Pandžić, J.-P. Watson, C. A. Silva-Monroy, “Co-planning of Investments in Transmission and Merchant Energy Storage,” in *IEEE Transactions on Power Systems*, in press, 2017. [pdf]
2. B. Xu, **Y. Wang**, Y. Dvorkin, R. Fernández-Blanco, D. S. Kirschen, C. A. Silva-Monroy, J.-P. Watson, “Scalable Planning of Energy Storage in Energy and Reserve Markets,” in *IEEE Transactions on Power Systems*, in press, 2017. [pdf]
3. **Y. Wang**, Y. Dvorkin, R. Fernández-Blanco, B. Xu, T. Qiu, D. S. Kirschen, “Look-Ahead Bidding Strategy for Energy Storage,” in *IEEE Transactions on Sustainable Energy*, in press, 2017. [pdf]
4. **Y. Wang**, Y. Liu, D. S. Kirschen, “Scenario Reduction with Submodular Optimization,” in *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 2479–2480, May 2017. [pdf]
5. R. Fernández-Blanco, Y. Dvorkin, B. Xu, **Y. Wang**, D. S. Kirschen, “Optimal Energy Storage Siting and Sizing: A WECC Case Study,” in *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 733–743, April 2017. [pdf]
6. T. Qiu, B. Xu, **Y. Wang**, Y. Dvorkin, D. S. Kirschen, “Stochastic Multi-Stage Co-Planning of Transmission Expansion and Energy Storage,” in *IEEE Transactions on*

- Power Systems*, vol. 32, no. 1, pp. 643–651, Jan. 2017. [pdf]
7. H. Pandžić, Y. Dvorkin, T. Qiu, Y. Wang, D. S. Kirschen, “Toward Cost-Efficient and Reliable Unit Commitment under Uncertainty,” in *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 970–982, March 2016. [pdf]
 8. Shen Y., Ke D., Sun Y. D. S. Kirschen, Wang Y., Hu Y., “A Novel Transient Rotor Current Control Scheme of a Doubly-fed Induction Generator Equipping with SMES for Voltage and Frequency Support,” in *Chinese Physics B*, vol. 24, no. 7, pp. 1–12, July 2015.
 9. H. Pandžić, Y. Wang, T. Qiu, Y. Dvorkin, D. S. Kirschen, “Near-Optimal Method for Siting and Sizing of Distributed Storage in a Transmission Network,” in *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2288–2300, Sept. 2015. [pdf]

Conference

1. J. E. Contereras-Ocaña, Y. Wang, M. A. Ortega-Vazquez, B. Zhang, “Energy Storage: Market Power and Social Welfare,” in *Proc. of 2017 IEEE PES General Meeting*, Chicago, IL, 2017, pp. 1-5.
2. Y. Wang, Y. Dvorkin, R. Fernández-Blanco, B. Xu, D. S. Kirschen, “Impact of Local Transmission Congestion on Energy Storage Arbitrage Opportunities,” in *Proc. of 2017 IEEE PES General Meeting*, Chicago, IL, 2017, pp. 1-5.
3. M. Almassalkhi, Y. Dvorkin, J. Marley, R. Fernández-Blanco, I. Hiskens, D. Kirschen, J. Martin, H. Pandzic, T. Qiu, M. Sarker, M. Vrakopoulouz, Y. Wang, M. Xue, “Incorporating Storage as a Flexible Transmission Asset in Power System Operation Procedure,” *2016 Power Systems Computation Conference (PSCC)*, Genoa, 2016, pp. 1-7. [pdf]

4. **Y. Wang**, Y. Dvorkin, R. Fernández-Blanco, D. S. Kirschen, “Look-Ahead Strategic Energy Storage in Ramp-Constrained Markets,” *FERC Technical Conference on Increasing Real-Time and Day-ahead Market Efficiency Through Improved Software*, Washington, DC, 2016. [pdf]
5. **Y. Wang**, Z. Zhou, C. Liu, A. Botterud, “Evaluating Stochastic Methods in Power Systems Operations with Wind Power,” *2016 IEEE International Energy Conference (ENERGYCON)*, Leuven, 2016, pp. 1-6. [pdf]
6. H. Sun, **Y. Wang**, D. Nikovski, J. Zhang, “Flex-Grid: A Dynamic and Adaptive Configurable Power Distribution System,” *2015 IEEE Eindhoven PowerTech*, Eindhoven, 2015, pp. 1-6. [pdf]
7. Y. Dvorkin, **Y. Wang**, H. Pandžić, D. S. Kirschen, “Comparison of Scenario Reduction Techniques for the Stochastic Unit Commitment,” *2014 IEEE PES General Meeting | Conference & Exposition*, National Harbor, MD, 2014, pp. 1-5. [pdf]
8. H. Pandžić, Y. Dvorkin, **Y. Wang**, T. Qiu, D. S. Kirschen, “Effect of Time Resolution on Unit Commitment Decisions in Systems with High Wind Penetration,” *2014 IEEE PES General Meeting | Conference & Exposition*, National Harbor, MD, 2014, pp. 1-5. [pdf]
9. D. S. Kirschen, Y. Dvorkin, **Y. Wang**, T. Qiu, H. Pandžić, “A Comparison of Various Unit Commitment Techniques Dealing with Uncertainty,” *FERC Technical Conference on Increasing Real-Time and Day-ahead Market Efficiency Through Improved Software*, Washington, DC, 2014. [pdf]

Appendix D

VITA

Yishen Wang received the B.Eng. degree from Department of Electrical Engineering at Tsinghua University, Beijing, China, in June 2011. He is currently a Ph.D. candidate in Department of Electrical Engineering at the Renewable Energy Analysis Laboratory at University of Washington, Seattle.

Previously, Yishen was a research intern at Mitsubishi Electric Research Laboratories (2014) and research aide at Argonne National Laboratory Energy Systems Division (2015). His research interests include power systems operation and planning with renewable generation, energy storage, power system economics, renewable forecasting and machine learning applications in power systems.