Optimal Load And Energy Storage Management In Electricity Markets

Zhongyang Zhao
Wayne State University

Follow this and additional works at: https://digitalcommons.wayne.edu/oa_dissertations

Part of the Electrical and Computer Engineering Commons

Recommended Citation
Zhao, Zhongyang, "Optimal Load And Energy Storage Management In Electricity Markets" (2020). Wayne State University Dissertations. 2461.
https://digitalcommons.wayne.edu/oa_dissertations/2461

This Open Access Embargo is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.
OPTIMAL LOAD AND ENERGY STORAGE MANAGEMENT IN ELECTRICITY MARKETS

by

ZHONGLANG ZHAO

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: ELECTRICAL ENGINEERING

Approved By:

Adviser

Date
DEDICATION

This thesis is dedicated to my parents and all my teachers.
ACKNOWLEDGMENTS

I have had a wonderful Ph.D. journey thanks to the people around me. Above all, I would like to express my most sincere gratitude to my advisor, Dr. Caisheng Wang. His detailed guidance and tremendous support for my search have led me to the way of scientific research and prepared me well for my future career.

I am grateful to my co-advisor Dr. Masoud H. Nazari. His personal integrity and expectations of excellence always encourage me to advance higher in the study and research.

I would also like to express my sincere gratitude to other members of the dissertation committee, Dr. Le Yi Wang, Dr. Feng Lin, and Dr. Carol J. Miller, who gave me many professional ideas and stimulated my thinking in different ways so that I could continuously improve my research.

I would like to thank Dr. Matthew Nokleby for his professional guidance on my data analysis skills and research. I also want to thank Dr. Huaiwei Liao for his insightful suggestions of power market modeling. I do appreciate Dr. Guangyi Liu for offering me the opportunities to work on the projects based on my interest during the internship.

During my study and research at Wayne State University, I’m pleased to have collaborated with Dr. Yang Wang, Dr. Chang Fu, Tingli Hu, Chen Duan, Mahdi Rouholamini, Wasseem Al-Rousan and Nima Abdolmaleki for their valuable discussions and interesting ideas.

Last but not least, I want to thank my parents and my friends. Love and encouragement coming from them always inspire me.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>CHAPTER 1: <strong>Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Electricity Market in the U.S.</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Independent System Operator</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 Two-Settlement System in a Wholesale Market</td>
<td>2</td>
</tr>
<tr>
<td>1.1.3 Locational Marginal Price</td>
<td>2</td>
</tr>
<tr>
<td>1.1.4 Power Plants</td>
<td>3</td>
</tr>
<tr>
<td>1.1.5 Battery Energy Storage System</td>
<td>3</td>
</tr>
<tr>
<td>1.1.6 Aggregation of Electric Vehicles</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Motivation</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Literature Review</td>
<td>9</td>
</tr>
<tr>
<td>1.3.1 Prediction of Electricity Price and Fuel Cost Distributions</td>
<td>9</td>
</tr>
<tr>
<td>1.3.2 Generation Dispatch and Load Management</td>
<td>10</td>
</tr>
<tr>
<td>1.3.3 Energy Storage Systems in Electricity Markets</td>
<td>11</td>
</tr>
<tr>
<td>1.3.4 EV Aggregators in Electricity Markets</td>
<td>13</td>
</tr>
<tr>
<td>1.4 Thesis Outline</td>
<td>13</td>
</tr>
<tr>
<td>CHAPTER 2: <strong>Predictions of Electricity Price and Fuel Cost Distributions in Electricity Markets</strong></td>
<td>15</td>
</tr>
<tr>
<td>2.1 Motivation and Contribution</td>
<td>15</td>
</tr>
</tbody>
</table>
2.2 Data Features of Electricity Price

2.2.1 Raw Data Analysis

2.2.2 Differential Series between the DALMP and RTLMP

2.3 Model Development for Electricity Price Prediction

2.3.1 Seasonal ARIMA Model

2.3.2 ARMA Model

2.3.3 Exogenous Variables

2.3.4 Autoregressive GARCH Model

2.3.5 Evaluation Methods

2.4 Prediction Performance of Electricity Price

2.4.1 Data Preprocessing

2.4.2 SARIMA and SARIMAX Model Selection

2.4.3 ARMA and ARMAX Model Selection

2.4.4 Results Comparison

2.5 Data Features of Power Plants’ Fuel Costs

2.5.1 Power Plants Data

2.5.2 Objective Data Selection

2.5.3 Natural Gas Fuel Cost Analysis

2.6 Model Development for Fuel Costs Prediction

2.6.1 ARIMA Model

2.6.2 Normal Distribution Fitting

2.6.3 Kullback-Leibler (KL) Divergence of Normal Distributions

2.7 Prediction Performance for Fuel Cost Distributions

2.7.1 Data Preprocessing
CHAPTER 3: Temporal and Spatial Load Management Methods for Cost and Emission Reduction

3.1 Motivation and Contribution

3.2 Load Management

3.2.1 Cost Model

3.2.2 Temporal and Spatial Load Management (TSLM)

3.2.3 Temporal Only Load Management (TLM)

3.2.4 Self-Optimizing Load Management (SOLM)

3.2.5 Improved Self-Optimizing Load Management

3.2.6 Total Cost and Emission

3.3 Algorithm Implementation And Simulation Studies

3.3.1 Cost Model

3.3.2 Base Case Formulation of IEEE 14-Bus System

3.3.3 Base Case without Load Management

3.3.4 Temporal and Spatial Load Management

3.3.5 Temporal Only Load Management

3.3.6 Self-Optimizing Load Management

3.3.7 Sliding Window Self-Optimizing Load Management

3.3.8 Day Ahead Self-Optimizing Load Management

3.3.9 Comparison Study of IEEE 14-Bus System

3.3.10 Case Studies on the IEEE 57-bus System

3.4 Summary
CHAPTER 4: Revenue Analysis and Optimal Placement of Stationary and Transportable Energy Storage Systems for Market Participants

4.1 Motivation and Contribution

4.2 Participation Models in Energy and Frequency Regulation Markets

4.2.1 Credit in Energy Market

4.2.2 Credit in Regulation Market

4.2.3 Cost of Storage Degradation

4.2.4 Participation of Energy Market

4.2.5 Participation of Energy and Frequency Regulation Markets

4.3 Revenue Analysis of BESSs

4.3.1 Data Preparation

4.3.2 Potential Revenue of Stationary BESS

4.3.3 Potential Revenue of Transportable BESS

4.4 Optimal Placement of BESSs

4.4.1 Prediction of LMP's Volatility

4.4.2 Clustering of Pricing Nodes

4.4.3 Optimal Placement Algorithm

4.4.4 Performance of the Proposed Algorithm

4.5 Summary

CHAPTER 5: A Two-Stage Optimal Bidding Algorithm for Incentive-based Aggregation of Electric Vehicles in Workplace Parking Lots

5.1 Motivation and Contribution

5.2 Modeling of EV Owners’ Behaviors

5.2.1 Initial Behaviors of EVs

5.2.2 EV’s Responses to Incentives
5.3 Two-Stage Optimal Bidding Algorithm ........................................... 94
  5.3.1 First Stage: Day-Ahead Planning of EV Aggregator .................. 94
  5.3.2 Second Stage: Real-Time Operation of EV Aggregator ............... 98
  5.3.3 Prediction of Market Data ..................................................... 105
5.4 Model Performance ................................................................. 107
  5.4.1 Incentives and EV Responses ............................................... 107
  5.4.2 First Stage: DA Planning ..................................................... 108
  5.4.3 Second Stage: RT Operation ............................................... 110
5.5 Summary .................................................................................. 114

CHAPTER 6: Conclusions and Future Work ........................................ 115

REFERENCES .................................................................................. 117

ABSTRACT ....................................................................................... 136

AUTOBIOGRAPHICAL STATEMENT ............................................. 138
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>PJM map with regions [1]</td>
<td>1</td>
</tr>
<tr>
<td>2.1</td>
<td>ACF and PACF plots of the MISO RTLMP between 1/5/2015 and 12/27/2015</td>
<td>17</td>
</tr>
<tr>
<td>2.2</td>
<td>ACF and PACF plots of $\Delta LMP$ of the MISO RTLMP between 1/5/2015 and 12/27/2015</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Improvement indices of ARMA and ARMAX models in 12-hour prediction</td>
<td>27</td>
</tr>
<tr>
<td>2.4</td>
<td>Boxplot of NG fuel cost of Texas in 2015</td>
<td>30</td>
</tr>
<tr>
<td>2.5</td>
<td>Boxplot of SUB fuel cost of Texas in 2015</td>
<td>31</td>
</tr>
<tr>
<td>2.6</td>
<td>$FC'$ forecasting process diagram</td>
<td>33</td>
</tr>
<tr>
<td>2.7</td>
<td>Fitted Normal Distributions for three-month delay fuel costs, forecasting fuel costs and real fuel costs</td>
<td>37</td>
</tr>
<tr>
<td>3.1</td>
<td>A typical 24-hour LMP curve of a node in PJM</td>
<td>46</td>
</tr>
<tr>
<td>3.2</td>
<td>SOLM algorithm flow chart</td>
<td>47</td>
</tr>
<tr>
<td>3.3</td>
<td>Management Process of the SW-SOLM</td>
<td>49</td>
</tr>
<tr>
<td>3.4</td>
<td>Management Process of the DA-SOLM</td>
<td>51</td>
</tr>
<tr>
<td>3.5</td>
<td>Basic load profile</td>
<td>54</td>
</tr>
<tr>
<td>3.6</td>
<td>LMP distribution before optimizing</td>
<td>55</td>
</tr>
<tr>
<td>3.7</td>
<td>LMP distribution after the TSLM</td>
<td>56</td>
</tr>
<tr>
<td>3.8</td>
<td>LMP distribution after TLM</td>
<td>56</td>
</tr>
<tr>
<td>3.9</td>
<td>LMP distribution after SOLM</td>
<td>57</td>
</tr>
<tr>
<td>3.10</td>
<td>LMP distribution after SW-SOLM</td>
<td>58</td>
</tr>
<tr>
<td>3.11</td>
<td>LMP distribution after DA-SOLM</td>
<td>59</td>
</tr>
<tr>
<td>3.12</td>
<td>24-h LMP of different cases in the IEEE 57-bus system</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>Comparison between $RegA$ and $RegD$</td>
<td>69</td>
</tr>
</tbody>
</table>
Figure 4.2: Comparison of LMPs in different seasons ........................................ 73
Figure 4.3: Comparison of LMPs at two different locations in PJM during the summer of 2018 .................................................. 74
Figure 4.4: Annual revenue percentiles in PJM ............................................. 76
Figure 4.5: Annual revenues by percentiles ............................................... 77
Figure 4.6: Daily revenue differences between ROCK and KINCA in 2018 .......... 78
Figure 4.7: LMP features of the top 20% daily revenue difference ................ 78
Figure 4.8: Output actions of node ROCK on 9/27 ..................................... 79
Figure 4.9: Output actions of node KINCA on 9/27 ..................................... 80
Figure 4.10: Revenue percentiles by season ............................................. 82
Figure 5.1: Arrival and departure EVs with different incentives .................. 108
Figure 5.2: Timeline of DA planning and RT operation ............................. 109
Figure 5.3: Average offers in energy and regulation markets by operating hours 111
Figure 5.4: Hourly average nonperformance offer and precision score .......... 112
LIST OF TABLES

Table 1.1: Review of the Load Management Literature

Table 2.1: BIC Values of the ARMA Model

Table 2.2: Comparison Results of all ARIMA Models

Table 2.3: Comparison Results of all ARIMA Models with $GARCH(1,1)$

Table 2.4: Raw Data in Form EIA-923

Table 2.5: Fuel cost data of Texas

Table 2.6: Mean and corresponding Error of Fitted Distributions

Table 2.7: Symmetric KL Divergence of Estimated Distributions

Table 3.1: Generator Cost Model Parameters

Table 3.2: Carbon Dioxide Emission Parameters

Table 3.3: Generator Fuel Types and Emission Factors

Table 3.4: Total Cost of Each Loop in DA-SOLM

Table 3.5: Total cost of each loop in extended DA-SOLM

Table 3.6: Bus#4 24-h LMP Mean Value and Standard Deviation

Table 3.7: Total Cost and Emission of Regular Cost Model

Table 3.8: Carbon Dioxide Emission with Different CO$_2$ Prices

Table 3.9: Generator Fuel Types and Emission Factors in the IEEE 57-bus System

Table 3.10: Total Cost of Each Loop of the DA-SOLM in the IEEE 57-bus System

Table 3.11: Total Cost and Emission of the IEEE 57-bus System

Table 4.1: Key Parameters of Energy Storage

Table 4.2: Seasonal Revenues of 1L1Y and 4L4S

Table 4.3: Revenues of 12L12M
Table 4.4: Performances of the Base Case and the Proposed Algorithm · · · · · · · 89
Table 5.1: Parameters of EV Original Behaviors’ Distributions · · · · · · · · · · · · 92
Table 5.2: Orders of Seasonal ARIMA Model · · · · · · · · · · · · · · · · · · · · · · · · 110
Table 5.3: Average Daily Performances in Different Cases · · · · · · · · · · · · · · · · · · · 113
CHAPTER 1: INTRODUCTION

1.1: Electricity Market in the U.S.
1.1.1: Independent System Operator

In the United States, there are more than 8,800 power plants, nearly 160,000 miles of high-voltage power lines, and millions of low-voltage power lines and distribution transformers in the power grid serving for 145 million customers. To efficiently and economically manage large power grids for their safe and reliable operations, wholesale and retail markets have been formed throughout the country. Several competitive wholesale markets in the United States are operated by the Independent System Operators (ISOs) or Regional Transmission Organizations (RTOs), such as Midcontinent Independent System Operator (MISO), California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), New York ISO, New England ISO, and PJM Interconnection LLC (PJM). Among these markets, the wholesale market administrated by PJM, as shown in Fig. 1.1, is the largest.

![Figure 1.1: PJM map with regions](image)

PJM regulates all or parts of 13 states and the District of Columbia and serves for
about 65 million people \(^3\). With 21% of the U.S. Gross Domestic Product (GDP) produced in its region, the PJM’s summer peak load can hit 165.49 GW with a generation capacity of 178.56 GW \(^3\).

1.1.2: Two-Settlement System in a Wholesale Market

Since the supply and demand of the electricity in the power system must be matched in real time, a two-settlement system including a day-ahead (DA) market and a real-time (RT) market is designed and operated by RTOs and ISOs. The two-settlement system can better arrange the generating resources to meet the anticipated and instantaneous power demands \(^4\) and determine the financial charges and credits to the market participants. In the DA market, the demand bids and supply offers submitted by the market participants are used to determine an economic dispatch and calculate the next-day hourly locational marginal prices (LMPs), also called day-ahead LMPs \(^5\). As the instantaneous generation and load always deviate from the day-ahead scheduled, the real-time balancing of supply and demand is performed by RTOs and ISOs in the RT market and 5-min LMPs are calculated simultaneously based on the actual operating conditions and economic dispatch. Afterward, the 5-min LMPs will be integrated into hourly LMPs to determine the financial settlement. \(^4\)

1.1.3: Locational Marginal Price

Given the demand bids and supply offers, the LMPs can be obtained via the programs of least-cost security-constrained unit commitment, which determine the start-up and shut-down schedules of generators, and the security-constrained economic dispatch programs which make the system demands be served at the least cost. The LMPs are combined with three pricing components: system energy price, transmission congestion cost, and marginal loss cost \(^6\). These components make the LMP reflect the marginal
cost of delivering an additional unit of energy to a specific location in the power system.

In the DA market, the 24 hours DA LMPs are determined based on the generation expectations and the anticipated loads for the next day. Compared to the RT LMPs, lower price volatility of DA LMPs is expected while the RT LMPs are influenced in the RT market by the updated bids, deviations between anticipated and actual loads, transmission and generation outages, etc.

1.1.4: Power Plants

Power plants are one of the most indispensable and influential elements in the electricity markets. According to Form EIA 923 published by the U.S. Energy Information Administration (EIA), there are over 8,800 power plants annually reporting their monthly generation outputs, heat contents, fuel types, etc. The most common energy sources in the U.S. are natural gas (NG), coal, nuclear, and renewable energy, which contribute 38%, 23%, 20%, and 17%, respectively, to the electricity generation in 2019. The power plants with different fuel types could have various heat rates, fuel costs, marginal costs, emission factors, etc., which can lead to different generation characteristics in the electricity markets.

1.1.5: Battery Energy Storage System

The battery energy storage system (BESS) analyzed in this dissertation is grid-scale. A BESS can be characterized by its power rating, energy capacity, storage duration, round trip efficiency, etc. The power rating of a BESS operated by the utility companies could range from 0.1 MW to 100 MW. In modern power systems, the BESS, as an important market participant, can provide a variety of functions, including energy arbitrage, ancillary services, generation capacity deferral, ramping, and accommodation of intermittent renewable energy. To benefit the operations of power systems from the
implementation of energy storage, Federal Energy Regulatory Commission (FERC) has issued Orders 755 and 841 in 2011 and 2018, respectively, which have further paved the way for market participation of energy storage. Order 755, titled “Frequency Regulation Compensation in the Organized Wholesale Power Markets,” requires the ISOs and RTOs to pay for the capacity offers and regulation performances provided by energy storage. Order 841, called “Electric Storage Participant in Markets Operated by Regional Transmission Organizations and Independent System Operators,” guides ISOs and RTOs toward lowering the barriers for the participation of electric storage resources in the capacity, energy, and ancillary service markets. Furthermore, FERC Order 841 establishes the minimum size requirement of not exceeding 100 kW for energy storage participation in the electricity markets. This order encourages more BESSs to participate in the wholesale market. The grid-scale BESS applications have experienced exponential growth in the past couple of years worldwide. For example, the BESS capacity in the U.S. has increased from 144.8 MW in 2012 to 1032.8 MW in 2019.

1.1.6: Aggregation of Electric Vehicles

As a type of fast responding and flexible load, electric vehicles (EVs) have been investigated to provide system-wide services and participate in electricity markets. It has been shown that EVs are capable of following the regulation signals from PJM. While a single EV with a limited capacity is not satisfied to participate in the wholesale markets, an EV aggregator, as an essential role between EV owners and the wholesale markets, can be implemented to integrate and coordinate a fleet of EVs to enter the markets. Since the global sales of EVs reached 2.26 million in 2019, the large amount of EVs has challenged the management of EV aggregation. The EV aggregator, as a third-party service, is commissioned to provide the EV owners benefits and operate
in competitive wholesale markets profitably.

1.2: Motivation

Load management (also called demand response or demand side management) techniques have been important tools in improving voltage profile, system efficiency, frequency regulation and stability, and for matching the stochastic output power of renewable sources [21,22]. Several load management programs have already been implemented in the real power grids, for example, 1) the New York Power Authority (NYPA)’s Peak Load Management (PLM) program [23] and 2) Texas load management program [24] which was designed to mitigate the peak amount of electric demand. As a part of the Smart Grid subprogram of the Advanced Digital Sciences Center (ADSC), a Demand-Side Management (DSM) [25] project has started in Singapore. This project applied an aggregator-level demand response mechanism to schedule controllable loads and to realize dynamic pricing and load curtailment. According to the Annual Electric Power Industry Report [26] published by EIA, there are over 9.75 million customers enrolled in the demand response programs in the U.S., which achieves a total of 1,426 GWh energy saving in 2018.

A majority of load management programs are based on electricity price signals [27]. In addition to the programs following real-time price signals, there are also many demand response programs using price forecasts to do load scheduling and management. The publicly available electricity price prediction is the DALMP published by ISOs or RTOs. However, day-ahead and real-time markets are basically two different markets, and there can be a large difference between the DALMP and RTLMP. In order to provide a more accurate RTLMP prediction for effective system operation and power demand management, an algorithm is proposed in Chapter 2 to improve short-term...
electricity price forecasting on the order of a few hours.

Prior to the applications of load management, the proposed methods must be validated with test data and/or historical conditions using validated, realistic, high fidelity power system network models. Important input variables to the development of such models include fuel costs (FCs), heat rates, generation costs, and emission rates associated with each class of prime mover and fuel type combination. The FC distributions can be used to generate large and realistic power system network models without compromising the confidentiality of the utilities. However, Form EIA-923 updates the fuel cost data with a three-month delay. The information available is the data three months before. **Due to the three months update lagging of the form, the most current data is not available for the timely study.** To address this issue, a forecasting algorithm is also developed in Chapter 2 to provide a more accurate FC characterization and distribution estimation instead of relying on the data with a three-month delay. The refined fuel cost distribution produced by the proposed method in Chapter 2 will be able to better serve for a marginal emission estimation model called Locational Emission Estimation Methodology (LEEM) [28]. LEEM can automatically track, analyze and report locational marginal emissions information for utility companies and energy users to make better demand and emission management decisions.

In the U.S., about 40% of all carbon dioxide emissions are attributed to electricity generation [29]. In order to reduce the negative impacts on the environment, strict emission constraint policies have already been implemented in the U.S. [30]. As a result, carbon dioxide emissions in 2018 was seen 13% decrease compared to the emission in 2005 [31]. To achieve sustainable energy development, one of the important targets in future smart grid implementation is to further decrease pollutant emissions
due to electric power generation and consumption. Besides the management of generation, load management can also help decrease emissions and costs, especially in future smart grids, where customers will have more flexibility in controlling their electricity usages. To further reduce the cost and emission via load management, several new load management methods, such as the temporal and spatial load management, are proposed for the market operator in Chapter 3.

While the developed load management methods for the market operator in Chapter 3 are proven to be capable of reducing the cost and emission in the system operation, the benefits for market participants to respond to the regulation signal and to manage their demands are worthy of study. As the battery technologies advance and the cost of battery decreases, BESSs have emerged as one of the popular and cost-effective energy storage technologies to participate in the electricity market. However, having access to the wholesale market and the new profitable opportunities do not mean to arbitrarily build more energy storage projects to dive into the markets while many factors can influence the profits. Due to the volatility and variations of LMP, Regulation Market Clearing Price (MCP), and regulation signal of different locations and seasons in the electricity market, the investment return can be vastly different. A poor selection of installation location for a BESS project in the electricity market can make the project struggle to compete with other market participants and cause detrimental impacts on the project. For the sake of supporting the energy storage market participants to better invest and manage BESS projects, a comprehensive analysis for finding the locations with the best potential revenues in the energy and frequency regulation markets is carried out in Chapter 4.

In addition, the BESS can also be formed by aggregating a fleet of EVs that have the
vehicle to grid (V2G) capabilities \[32\]. One of the key implementations is to aggregate the EVs at public parking lots \[33, 34\]. The peak hours of market conditions in both energy and regulation markets, such as locational marginal price (LMP) and regulation market-clearing prices, are between 7:00 and 21:00 \[35–37\], which largely overlap with many people’s working schedule. Hence, the aggregator in workplace parking lots has great potential to compete in the wholesale market during the peak hours of electricity prices.

Compared to the other demand response methods, such as the real-time pricing method \[38\], the incentive-based methods have the advantages of allowing the customers to participate voluntarily. Furthermore, the incentive information is not required to be always broadcasted to the customers. In other words, an incentive-based EV aggregation program can be activated only when it is beneficial for both the EV aggregator and owners. Meanwhile, the incentive is able to not only compensate the EV owners but also attract more EVs to join the aggregation program. The incentive-based demand response programs introduced in \[39–42\] have proven the ability to improve the demand flexibility in retail customers. In addition, according to the revenue analysis of battery in energy and regulation markets in \[36, 43\], the revenue obtained from the regulation market far exceeds that from the energy market. Most importantly, it can raise more regulation resources for power system operations. \textbf{However, there are still highly challenging issues, such as the uncertainties of EV owners’ behaviors and market conditions, which can derail the aggregator’s performance.} The EV owners’ behaviors, including their arrival/departure times, the EV batteries’ State of Charge (SOC) at arrival/departure, and their responses to the aggregator’s control signals, expose the EV aggregator to a complicated and risky situation. To address these challenges, a two-stage
optimal bidding algorithm for an aggregator of EVs in workplace parking lots to participate in RT energy and regulation markets, coordinated with the incentive to adjust the EVs’ behaviors, is designed to benefit both EV aggregator and EV owners in Chapter 5.

1.3: Literature Review
1.3.1: Prediction of Electricity Price and Fuel Cost Distributions

In time series analysis, autoregressive integrated moving average (ARIMA) model is a great tool for forecasting a stationary time series and a non-stationary time series that can be treated as “stationary” by differencing and other possible transformation techniques. The Integrated (I) part in an ARIMA model is used to account for non-stationary elements in a time series. ARIMA models have been analyzed and evaluated for forecasting electricity price [44–47]. An ARIMA model is used to analyze the time series with Box and Jenkins method [48] and the next-day MCP was predicted using the ARIMA model while considering explanatory variables, such as demand [44]. Furthermore, a wavelet transform was employed to decompose the ill-behaved historical price time series to a better-behaved constitutive series set [45]. After different ARIMA models were used for the better-behaved set, the inverse wavelet transform was then used to forecast the price [45]. A hybrid model using not only ARIMA but also an artificial neural network (ANN) was proposed for electricity spot price forecast [46,47]. A seasonal ARIMA (SARIMA) model was proposed to predict the time series for handling the weekly and daily periodical fluctuations [49,50]. The SARIMA with exogenous data, such as power production [49] and temperature [50], was developed to forecast the day-ahead spot electricity price in Sweden by considering weekly seasonal effects [49], and to produce short-term PV generation forecasting [50]. The ARIMA models using standard errors based on Hessian, Cochrane-Orcutt, Prais-Winstern and Hildreth-Lu have been developed to forecast the price of natural gas in [51]. For forecasting the Henry
Hub weekly natural gas spot price, the ARIMA model aiming for the approximation components associated with wavelet decomposition was implemented in [52].

1.3.2: Generation Dispatch and Load Management

A review of the literature on generation dispatch and load management to reduce cost and emission is tabulated in Table 1.1 [53–76]. These studies show the current load management programs are mostly from the perspective of time shifting [68,70,72,75] with a few on spatial shifting [74]. A set of methods were proposed in [70] to optimize demand resources with the supply and demand side considerations over day-ahead and real-time adjustments. A load forecasting based load management method was developed in [71] by considering multiple factors, such as distributed generation, bidding, and demand response on price. Since electricity markets become more popular, bidding curves and actual LMP have also become the focus of optimization. The estimation of residential demand response was used under an electricity price structure to shave electricity price peaks [72]. An interesting approach using self-adaptive global-based harmony search algorithm for optimal double-sided bidding strategy was developed in [73] to solve the bi-level problem in a market-based power system.

As more and new controllable loads such as EVs become available and popular, load management in a spatial frame can also be considered. In [74], a load distribution method in the spatial frame from the demand side was proposed. The method utilized LMP and/or locational marginal emission (LME) information in load management. Another integrated economic/emission/load profile management dispatch algorithm [75] aimed at the load management on the temporal frame where the emission issues are also considered. Different load management methods for cost and emission reduction were analyzed and compared in [76].
Table 1.1: Review of the Load Management Literature

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Reference number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Emission Dispatch (EED)</td>
<td>53–76</td>
</tr>
<tr>
<td>Multi-objective EED</td>
<td>56, 58, 60, 62, 67, 70</td>
</tr>
<tr>
<td>EED with transmission constraints</td>
<td>59, 60, 73, 76</td>
</tr>
<tr>
<td>EED with renewable source: Hydro thermal system, Wind</td>
<td>61, 63, 66</td>
</tr>
<tr>
<td>EED on temporal domain</td>
<td>68, 70, 72, 75, 76</td>
</tr>
<tr>
<td>EED with demand response</td>
<td>68, 73, 76</td>
</tr>
<tr>
<td>EED with bidding mechanism</td>
<td>68, 73</td>
</tr>
<tr>
<td>Algorithm: Abductive reasoning network</td>
<td>58</td>
</tr>
<tr>
<td>Algorithm: Strength Pareto evolutionary algorithm</td>
<td>60</td>
</tr>
<tr>
<td>Algorithm: Simulated annealing-based goal-attainment method</td>
<td>61</td>
</tr>
<tr>
<td>Algorithm: Evolution algorithm</td>
<td>64, 66</td>
</tr>
<tr>
<td>Algorithm: Opposition-based gravitational search algorithm</td>
<td>65</td>
</tr>
<tr>
<td>Algorithm: Fuzzy satisfaction-maximizing decision approach</td>
<td>67</td>
</tr>
<tr>
<td>Algorithm: Bi-level harmony search algorithm</td>
<td>73</td>
</tr>
</tbody>
</table>

1.3.3: Energy Storage Systems in Electricity Markets

To analyze the potential income of BESS in the energy and regulation markets, several studies have been done [77-83]. The energy storage technologies, potential applications, and comparison of policies on the participation of storage in different markets were investigated and studied in [77-78]. The analyses of potential revenues in MISO [79], PJM [80], CAISO [82], ERCOT [81], and New York ISO [83] have been carried out. Although the potential revenues of energy storage in energy and frequency regulation markets were estimated in [79, 80], the obtained estimations were only based on a single node, such as
In [81], the possible arbitrage and regulation revenues for BESS were estimated only for the load zones of ERCOT. Because of the insufficient study from the perspective of the whole system, the characteristics of the potential revenues throughout the entire power market have not been revealed for the BESS participants yet.

Instead of competing in both energy and frequency regulation markets at the same time, the revenues of just participating in one of the energy and frequency markets were surveyed in [82, 84–86]. The energy arbitrage revenue of BESS at all nodes in CAISO [82] and PJM [84] energy markets were studied. Considering the price uncertainty, the maximum arbitrage revenue in the day-ahead and real-time energy markets were discussed in [85, 86]. However, it does not provide meaningful guidance for market participants of energy storage since there is a lack of comparison between the possible profits in both the energy and frequency regulation markets when the potential revenue comes only from the energy arbitrage.

BESSs can be transportable by installing on a mobile platform such as a truck or by aggregating EVs with V2G capability. While a transportable BESS has the mobility to be implemented at different locations during operation, it is also imperative for the market participants to have an estimation on the current and future status of the power market and deploy the energy storage projects at the appropriate sites. Therefore, in addition to the system-wide revenue estimation for different times, a study to investigate the nodes’ profitability with considering the mobility of BESS is beneficial for the market participants as well. However, most of the existing studies on portable BESSs focus on the operations of power system, such as load shifting [87, 88], transmission congestion relief [89, 90], enhancing the resilience of distribution system [91], and the reduction of
wind curtailment [92]. Other than the spatiotemporal energy arbitrage for grid congestion relief investigated in [90], an optimal placement algorithm to maximize revenue in energy and frequency regulation markets for the market participants of transportable BESS has not been fully investigated yet.

1.3.4: EV Aggregators in Electricity Markets

Several studies have been carried out for helping EV aggregators to more effectively participate in the wholesale markets [38, 93–98]. An optimal bidding strategy based on the two-stage stochastic linear programming was developed in [94] for an EV aggregator to participate in DA energy and regulation markets with considering the uncertainties of market data and EV fleet. In [38], an EV aggregator model was proposed to study the operational behaviors of a parking lot with different demand response programs by changing the energy demand of EVs. In [96], a stochastic optimization model considering the conditional value at risk was proposed for the EV aggregator to participate in the frequency regulation market. Based on the transactive control, a two-stage optimal charging scheme was implemented in [97] to minimize the total cost of EV aggregator by managing the energy procurement in the DA market and scheduling EV charging. An optimal bidding strategy considering each individual vehicle’s behaviors was proposed in [98] to bid in the DA market for minimizing the charging cost of the aggregator. Nevertheless, the effectiveness and implementation of incentives to induce the EV owners to adjust their behaviors for providing more resources for the EV aggregator to participate in the wholesale markets have rarely been studied.

1.4: Thesis Outline

Chapters 2–5 are organized as follows:

- Chapter 2: Data analytics techniques, including the short-term electricity predic-
tion and the fuel cost distributions forecast, are proposed for improving the load and energy storage management.

- Chapter 3: New centralized and decentralized load management methods for the market operators are developed to reduce the cost and emission.

- Chapter 4: The revenue analysis and optimal placement of stationary and transportable BESSs are carried out for the market participants.

- Chapter 5: A two-stage optimal bidding algorithm is proposed for an incentive-based EV aggregator in workplace parking lots to participate in the energy and regulation markets.
CHAPTER 2: PREDICTIONS OF ELECTRICITY PRICE AND FUEL COST DISTRIBUTIONS IN ELECTRICITY MARKETS

2.1: Motivation and Contribution

The prediction for critical market data has become important for demand side management and power generation scheduling. Especially as the electricity market becomes more competitive, a more accurate prediction of market data will benefit participants in the market by improving load and demand managements.

For the sake of improving the short-term electricity price forecasting, this chapter first investigates the features of the DALMP and RTLMP data of MISO between January 2015 and December 2015 and forms a differential series between DALMP and RTLMP for increasing the forecasting accuracy. Four ARIMA models are then proposed and compared for electricity price forecasting in Section 2.3. Meanwhile, a generalized autoregressive conditional heteroskedastic (GARCH) model is proposed to implement on the ARIMA models for the volatility present in the LMP series. In Section 2.4, the MISO LMPs between 1/4/2015 and 1/31/2016 are used to validate the proposed models. The results show that the ARMAX-GARCH model produces the best result for short-term electricity price prediction in the real-time electricity market. It can improve the forecasting accuracy by more than 27% for the one-hour-ahead predictions.

Following the prediction of electricity price, this chapter also proposes an algorithm to forecast the fuel cost distributions against the three months update lagging of Form EIA-923. Via the exploration of monthly data features in Form EIA-923 from January 2013 to December 2016, the NG fuel cost of Texas is determined as the objective target to predict and estimate distributions for validation of the proposed forecasting method. The differential series between the plant’s fuel cost and NG hub spot price is formulated and implemented on the ARIMA model for a three-step-ahead prediction. As an enhancement
to the fuel cost prediction, the corresponding local natural gas hub spot price is utilized and predicted by one-step-ahead. In Section 2.7, the fuel cost data between January 2013 and June 2016 are used to train the formulated ARIMA model. The next six months data are adopted to validate the proposed forecasting algorithm. The results show the proposed method is able to achieve outstanding performance when the overall market price is not volatile. Besides, this method shows the capability to forecast the price after the price turning point appears.

2.2: Data Features of Electricity Price

In the following study, two types of LMP data, namely the DALMP and RTLMP, are analyzed and utilized for improving the accuracy of short-term electricity price forecasting.

2.2.1: Raw Data Analysis

The 24-h RTLMP between 1/5/2015 and 12/27/2015 in MISO is analyzed by observing the sample autocorrelation function (ACF) and partial autocorrelation function (PACF), shown in Fig. 2.1. The ACF and PACF plots are commonly used for ARIMA model selection using Box-Jenkins method [48]. The PACF plot is especially useful in identifying the order of an autoregressive model. In Fig. 2.1, a particular seasonal pattern is evident by the spikes appearing at approximately a 24-hour cycle. Hence, a SARIMA model should be investigated and tested to see whether a seasonal model might perform better. However, as discussed later in Section 2.4, a seasonal model does not necessarily guarantee a better performance if the seasonal pattern is weak.
According to the Box-Jenkins method, the time series is assumed to be stationary in using ARMA model \cite{48}. Hence, forming a stationary series is an important and necessary step before using an ARMA model.

The main idea of the proposed method is to utilize the DALMP data published in the day-ahead market by ISOs to improve the electricity price forecasting performance in short term. Both the historical RTLMP and DALMP series are used for improving the real-time LMP prediction. The series representing the differential between the past DALMP and RTLMP is extracted by (2.1):

$$\Delta LMP_t = DALMP_t - RTLMP_t$$  \hspace{1cm} (2.1)

where $\Delta LMP_t$ is the differential value between the DALMP and RTLMP at time $t$.

When $\Delta LMP_{t+i}$ is predicted by implementing the ARMA model, the forecasting $RTLMP_{t+i}$ is obtained by (2.2),

$$RTLMP'_{t+i} = DALMP_{t+i} - \Delta LMP_{t+i}, \hspace{1cm} i = 1, 2, \ldots$$  \hspace{1cm} (2.2)
where $DALMP_{t+i}$ is the DALMP published by the ISO for time $t+i$ and the $\Delta LMP_{t+i}$ is the forecasting differential LMP series at time $t+i$.

The stationarity of $\Delta LMP$ is revealed by the sample ACF and PACF plots in Fig. 2.2 because ACF decays quickly. Hence, the data of $\Delta LMP$ is appropriate to fit in the ARMA model. The three different LMP series, namely DALMP, RTLMP and $\Delta LMP$, introduced in this section, are to be used in the ARIMA models developed to forecast the electricity price in Section 2.3.

![Sample Autocorrelation Function](image1)

![Sample Partial Autocorrelation Function](image2)

Figure 2.2: ACF and PACF plots of $\Delta LMP$ of the MISO RTLMP between 1/5/2015 and 12/27/2015

### 2.3: Model Development for Electricity Price Prediction

Given the RTLMP with a seasonal pattern, a set of seasonal ARIMA models are proposed for short term forecasting, while an ARMA model is developed for $\Delta LMP$. Furthermore, exogenous data such as DALMP and weekday/weekend indicators are also incorporated into the ARIMA models to improve the forecasting accuracy.
2.3.1: Seasonal ARIMA Model

The first model used in this study is a seasonal ARIMA model. The seasonal ARIMA model for RTLMP is a standard \( ARIMA(p, d, q) \times (P, D, Q)_s \) model described as follows [48, 99]:

\[
\phi_p(B) \Phi_P(B^s) \nabla^d \nabla^D_s y_t = \mu + \theta_q(B) \Theta_Q(B^s) \varepsilon_t \tag{2.3}
\]

where \( B \) is the backward shift operator, i.e. \( B^h y_t = y_{t-h} \); \( p \) is the non-seasonal auto-regression (AR) order; \( d \) is the non-seasonal differencing order; \( q \) is the non-seasonal moving-average (MA) order; \( P \) is the seasonal AR order; \( D \) is the seasonal differencing; \( Q \) is the seasonal MA order; and \( S \) is the time span of repeating seasonal pattern. The error terms \( \varepsilon_t \) are generally assumed to be the independent and identically distributed noise with zero mean and finite variance, i.e. Gaussian white noise; \( \mu \) is a constant term.

\[
\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \cdots - \phi_p(B^p) \tag{2.4}
\]

\[
\Phi_P(B^s) = 1 - \Phi_1(B^s) - \Phi_2(B^{2s}) - \cdots - \Phi_P(B^{Ps}) \tag{2.5}
\]

\[
\nabla^d = (1 - B)^d \tag{2.6}
\]

\[
\nabla^D_s = (1 - B^s)^D \tag{2.7}
\]

\[
\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \cdots - \theta_q(B^q) \tag{2.8}
\]

\[
\Theta_Q(B^s) = 1 - \Theta_1(B^s) - \Theta_2(B^{2s}) - \cdots - \Theta_Q(B^{Qs}) \tag{2.9}
\]

Furthermore, \( \phi_1 \ldots \phi_p, \Phi_1 \ldots \Phi_P, \theta_1 \ldots \theta_q \) and \( \Theta_1 \ldots \Theta_Q \) are the coefficients of autoregressive, seasonal autoregressive, moving average and seasonal moving average polynomials, respectively.
2.3.2: ARMA Model

Aiming at predicting the $\Delta LMP$ series, the ARMA model is used for estimating the electricity price in short-term. The $ARMA(p,q)$ model can be described as \[48,99\]:

$$
\phi_p(B) y_t = \mu + \theta_q(B) \varepsilon_t
$$

(2.10)

where similar to the seasonal ARIMA model, $B$ is the backward shift operator; the $\phi_p(B)$, $\theta_q(B)$ and $\varepsilon_t$ have the same definitions as in the seasonal ARIMA model; $\mu$ is a constant term.

2.3.3: Exogenous Variables

Since the purpose of this study is to improve the short term electricity price forecast compared to the DALMP which can be obtained from the ISO before the day, it is useful to accurately capture the relationship between DALMP and RTLMP. The regressor embedded into the SARIMA and ARMA models for handling the exogenous variables are given in the following SARIMAX model (2.11) and ARMAX model (2.12).

$$
\phi_p(B) \Phi_P(B^s) \nabla^d \nabla^D_s y_t = \mu + \theta_q(B) \Theta_Q(B^s) \varepsilon_t + u_t^\gamma
$$

(2.11)

$$
\phi(B) y_t = \mu + \theta(B) \varepsilon_t + u_t^\gamma
$$

(2.12)

$$
u_t^\gamma = \sum_{k=1}^{r} \gamma_k u_{tk}
$$

(2.13)

where $u_t^\gamma$ is the vector of exogenous variables $[u_{t1} \ u_{t2} \ldots \ u_{tr}]$ and $\gamma$ is the coefficients vector for exogenous variables, such as DALMP and weekday/weekend indices.

2.3.4: Autoregressive GARCH Model

Due to the high volatility and price spikes in the time series, models that allow for heteroskedastic data are needed. In the generalized autoregressive conditional heteroskedastic $GARCH(p,q)$ models \[100,102\], the conditional variance is dependent on the past values of the time series and a moving average of the past conditional variance,
shown in (2.14) and (2.15):

\[ \varepsilon_t = \sigma_t z_t \]  

(2.14)

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \]  

(2.15)

where \( \varepsilon_t \) is the error term or the residual return at time \( t \); \( z_t \) is basically a white noise process; \( \alpha_0 > 0, \alpha_i > 0 \) and \( \beta_j > 0 \).

Additionally, it is concluded in [102] that ARIMA-GARCH can outperform generic ARIMA when the time series is volatile. Hence, a GARCH model is also developed to deal with varying nodal electricity prices for a more accurate prediction.

2.3.5: Evaluation Methods

A percentage index \( I_i \) is formulated in (2.16) for evaluating the performance of different prediction models. The percentage index \( I_i \) displays how much the forecast accuracy improves when compared to the DALMP.

\[ I_i = 1 - \frac{1}{n} \sum_{t=1}^{n} \left( \frac{|RTLMP_{t+i} - RTLMP'_{t+i}|}{|RTLMP_{t+i} - DALMP_{t+i}|} \right) \]  

(2.16)

where \( n \) is the sample number of the testing set; \( i \) means the \( i \)-th prediction hour; \( RTLMP_{t+i} \) is the real time LMP at time \( t+i \); \( RTLMP'_{t+i} \) is the \( i \)-th hour forecasting LMP at time \( t+i \); \( DALMP_{t+i} \) is the day-ahead LMP at time \( t+i \).

The models developed and discussed in this section, namely SARIMA, ARMA, SARI-MAX, ARMAX, ARMA-GARCH, and ARMAX-GARCH models will be evaluated and compared in the next section.

2.4: Prediction Performance of Electricity Price

The proposed forecasting models have been applied to predict the electricity prices in the region of MISO [103]. The training data set, from 1/5/2015 (Monday) to 12/27/2015
(Sunday), including 51 weeks hourly DALMPs and RTLMPs are used to obtain the parameters of the ARIMA models. The parameters estimation is based on the maximum likelihood estimation for the available training data. In this research, the parameter estimation is accomplished by the estimate tool in MATLAB \[104\]. After the models are developed, the hourly RTLMP and DALMP data, from 1/4/2016 (Monday) to 1/31/2016 (Sunday), are employed for the model validation. The forecasting results show the developed models make more accurate short-term predictions than the DALMP provided by the ISO in the real-time market.

2.4.1: Data Preprocessing

A preprocessing scheme is proposed as (2.17) for eliminating large price spikes in the LMP series:

\[
P_t = \begin{cases} 
UB \$/MWh & \text{if } RTLMP_t > UB \$/MWh \\
LMP_t & \text{otherwise} \\
LB \$/MWh & \text{if } RTLMP_t < LB \$/MWh
\end{cases} \tag{2.17}
\]

where \(UB\) and \(LB\) are the upper bound and the lower bound and the values of \(UB\) and \(LB\) are defined according to the processed time series.

Furthermore, a natural logarithm transformation, given in (2.18), is taken for suppressing larger fluctuations based on \[99\], because abnormal data points present in the observed time series could contribute to non-stationary and biased model fitting:

\[
y_t = \log{(P_t + c)} \tag{2.18}
\]

where \(P_t\) is the LMP value and \(c\) is a positive constant offset adding on \(P_t\) to guarantee the logarithm transformation.
2.4.2: SARIMA and SARIMAX Model Selection

The minimum RTLMp value in the training set is $-28.7$/MWh, thus the offset \( c \) in (2.18) is chosen to be 30 when RTLMp series from 1/5/2015 to 12/27/2015 is preprocessed by (2.18). The transformed series \( y_{RT,t} \) can then be obtained. Based on the plots in Fig. 2.1, the seasonality in the SARIMA model is set to 24. As a result, the seasonal ARIMA model is established as \( ARIMA(2, 0, 1) \times (1, 1, 1)_24 \) in (2.19):

\[
\phi_2(B) \Phi(B^{24}) \nabla_{24} y_{RT,t} = \mu + \theta(B) \Theta(B^{24}) \varepsilon_t. \tag{2.19}
\]

In the SARIMAX model, the DALMP time series is used as the exogenous data for improving prediction performance.

Before incorporating the DALMP into the SARIMAX model, the DALMP series is preprocessed by (2.18) to obtain \( y_{DA,t} \). Considering the minimum (negative) price in the DALMP time series and following the aforementioned RTLMp preprocessing, the offset \( c \) for preprocessing the DALMP is also chosen to be 30 in (2.18). Hence, the SARIMAX model is determined as (2.20)

\[
\phi_2(B) \Phi(B^{24}) \nabla_{24} y_{RT,t} = \mu + \theta(B) \Theta(B^{24}) \varepsilon_t + y'_{DA,t} \gamma \tag{2.20}
\]

2.4.3: ARMA and ARMAX Model Selection

Based on the sample ACF and PACF plots of \( \Delta LMP \) shown in Fig. 2.2, it is revealed that \( \Delta LMP \) is stationary. Thus, the integrated part for reducing non-stationary term in the ARIMA model is not necessary for dealing with \( \Delta LMP \). In other words, the initial differencing step is not required in the models for predicting the \( \Delta LMP \).

Thus, ARMA and ARMAX models are developed to predict the \( \Delta LMP \) series. Afterwards, the \( RTLMP_{t+i} \) forecast can be obtained by (2.2).

Since the minimum value of \( \Delta LMP \) series is $-457.45$/MWh which is beyond a normal range, (2.17) is utilized to shape the series for a better model fitting. In (2.17), the UB
and \( LB \) are set to 100 and -100, respectively. Then \( P_{\Delta LMP,t} \) series is processed by \((2.18)\) to become \( y_{\Delta LMP,t} \) when the offset \( c \) is set to 1000 for the transformation.

In addition, although the DALMP information has already been utilized in the prediction of \( \Delta LMP \) when using the ARMA model, the additional weekday and weekend information is also valuable for consideration to further improve the prediction accuracy. In this study, the weekday’s hours are designed to be represented by “1” while weekend’s hours are represented by “0”. Therefore, a repeating 120 “1”s and 48 “0”s sequence (Note: 120 hours in the weekday and 48 hours in the weekend) is applied as the exogenous data to develop the ARMAX model.

Then Bayesian information criterion (BIC) \([105]\) is utilized to handle the model order selection. As the BIC values of ARMA shown in Table \[2.1\], the model order is selected to \( ARMA(1, 2) \) due to the lowest BIC value appearing at \( p=1 \) and \( q=2 \) in Table \[2.1\]. Similarly, the order of the ARMAX model is determined to be \( ARMAX(1, 1) \) according to the BIC values.

<table>
<thead>
<tr>
<th>( p )</th>
<th>( q )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>-68875.1</td>
<td>-69085.5</td>
<td>-69067.3</td>
<td>-68905.7</td>
<td>-69056.9</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>-69073</td>
<td>-69065.4</td>
<td>-69082.9</td>
<td>-68292.3</td>
<td>-69008</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-69063.2</td>
<td>-69050.2</td>
<td>-68988.3</td>
<td>-69069.4</td>
<td>-69067.6</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>-69084.7</td>
<td>-69024.4</td>
<td>-69042.1</td>
<td>-69034.2</td>
<td>-68958.5</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>-69078.4</td>
<td>-68944.7</td>
<td>-69062.7</td>
<td>-69031.2</td>
<td>-69014</td>
</tr>
</tbody>
</table>
2.4.4: Results Comparison

When the $ARIMA(2,0,1) \times (1,1,1)_{24}$, $ARIMAX(2,0,1) \times (1,1,1)_{24}$, $ARMA(1,2)$ and $ARMAX(1,1)$ models are built, the GARCH model aiming at handling price volatility is applied. The $GARCH(1,1)$ model minimizes the BIC value, so it is implemented for every proposed model to account for volatility.

Table 2.2: Comparison Results of all ARIMA Models

<table>
<thead>
<tr>
<th>Model</th>
<th>$I_1$ (%)</th>
<th>$I_2$ (%)</th>
<th>$I_3$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA</td>
<td>-7.09</td>
<td>-23.16</td>
<td>-29.59</td>
</tr>
<tr>
<td>ARMA</td>
<td>26.64</td>
<td>13.03</td>
<td>6.79</td>
</tr>
<tr>
<td>ARMAX</td>
<td>26.50</td>
<td>12.75</td>
<td>6.53</td>
</tr>
</tbody>
</table>

Table 2.3: Comparison Results of all ARIMA Models with $GARCH(1,1)$

<table>
<thead>
<tr>
<th>Model</th>
<th>$I_1$ (%)</th>
<th>$I_2$ (%)</th>
<th>$I_3$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA – GARCH</td>
<td>-6.10</td>
<td>-20.66</td>
<td>-27.27</td>
</tr>
<tr>
<td>SARIMAX – GARCH</td>
<td>-29.53</td>
<td>-57.55</td>
<td>-66.52</td>
</tr>
<tr>
<td>ARMA – GARCH</td>
<td>27.14</td>
<td>16.85</td>
<td>10.70</td>
</tr>
<tr>
<td>ARMAX – GARCH</td>
<td>27.21</td>
<td>17.10</td>
<td>11.35</td>
</tr>
</tbody>
</table>

For validating and comparing the proposed models, the hourly LMP data set from 1/4/2016 (Monday) to 1/31/2016 (Sunday) is implemented. The short-term forecast results of every model are listed in Table 2.2 and Table 2.3. In Table 2.2, the improvement indices of the SARIMA and SARIMAX are negative and getting worse when the predicting hour increases, which means the SARIMA and SARIMAX models perform poorly.
compared to the DALMP series. The SARIMA and SARIMAX add two extra parameters to the model; the additional model complexity increases the risk of overfitting, but it does not help substantially in predicting RTLMP, which is only weakly seasonal.

In contrast, the one-hour-ahead prediction (i.e. $I_1$) percentage indices of the ARMA and ARMAX models are 26.6% better than the DALMP and while it is 13% better for the two-hour-ahead prediction indices. In other words, the ARMA models can provide a better prediction than the DALMP in short-term. The results of the different models with $GARCH(1, 1)$ are listed in Table 2.3 in which we can find the GARCH model is able to increase the prediction accuracy of the ARMA and ARMAX models by about 0.7% in the first hour and 3.5% in the second hour. Actually, the DALMP has predicted very well in January 2016 with mean absolute error (MAE) of 2.06 $/MWh while the MAE of the SARIMA, SARIMAX, ARMA and ARMAX with the GARCH model at the one-hour ahead prediction are 2.19 $/MWh, 2.67 $/MWh, 1.50 $/MWh and 1.50 $/MWh, respectively.

ARMA-GARCH and ARMAX-GARCH models have outperformed the other models on the next-hour and the next-couple-hour predictions. Both of them have done well and performed close to each other. For further comparing and analyzing these two models, the improvement indices $I$ of the next 12-hour prediction are plotted in Fig. 2.3.

According to Fig. 2.3, the ARMAX-GARCH beats ARMA-GARCH by about 0.9% after the 4th step prediction and it shows the model with exogenous variables could predict better than the one without the weekday/weekend index while both of the models have enough capability to outperform DALMP in the next 12-hour prediction. That means the inclusion of the weekday/weekend information as the exogenous variables is useful in improving the prediction accuracy in long term. At the same time, even though
the forecasting capabilities of the models in Fig. 2.3 have little differences in first couple of hours, the models with $GARCH(1,1)$ demonstrate the advantages over a longer-term prediction.

![Figure 2.3: Improvement indices of ARMA and ARMAX models in 12-hour prediction](image)

**2.5: Data Features of Power Plants’ Fuel Costs**

In addition to the electricity price prediction, the forecast of fuel cost distributions is introduced in the following study. The public fuel cost data extracted from Form EIA-923 [8] are analyzed and utilized for achieving more accurate probability density estimation using ARIMA model forecasting.

**2.5.1: Power Plants Data**

According to Form EIA-923, covering the monthly fuel cost data from 2013 January to 2016 December, the Date (Year and Month), Plant ID, Plant State, Energy Source, Quantity, Average Heat Content and Fuel Cost are archived and shown in Table 2.4 for reference.
Table 2.4: Raw Data in Form EIA-923

<table>
<thead>
<tr>
<th>Date</th>
<th>Plant ID</th>
<th>Plant State</th>
<th>Energy Source</th>
<th>Q (ton)</th>
<th>AHC (MMBtu/ton)</th>
<th>FC ($/MMBtu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>476</td>
<td>17.7</td>
<td>1.10</td>
</tr>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>28866</td>
<td>17</td>
<td>2.29</td>
</tr>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>43824</td>
<td>16.8</td>
<td>2.08</td>
</tr>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>86254</td>
<td>17</td>
<td>2.12</td>
</tr>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>43256</td>
<td>17</td>
<td>2.10</td>
</tr>
</tbody>
</table>

As shown in Table 2.4, multiple fuel cost data information for same Energy Source of a plant in a single month can be found. Hence, the fuel cost data are required to be grouped together by Date, Plant ID and Energy Source for obtaining an updated overall fuel cost for each Energy Source in a month. The updated fuel cost $F_{C'}$ are calculated by (2.21), (2.22) and (2.23):

\[
\text{Total Cost}_{\text{Plant ID}=i, \text{Date}=ym, \text{Source}=\text{type}} = \sum_{\text{Plant ID}=i, \text{Date}=ym, \text{Source}=\text{type}} Q \times AHC \times FC \tag{2.21}
\]

\[
\text{Total Heat}_{\text{Plant ID}=i, \text{Date}=ym, \text{Source}=\text{type}} = \sum_{\text{Plant ID}=i, \text{Date}=ym, \text{Source}=\text{type}} Q \times AHC \tag{2.22}
\]

\[
F_{C'}_{\text{Plant ID}, \text{Date}, \text{Source}} = \frac{\text{Total Cost}_{\text{Plant ID}, \text{Date}, \text{Source}}}{\text{Total Heat}_{\text{Plant ID}, \text{Date}, \text{Source}}} \tag{2.23}
\]

where $\text{Total Cost}_{\text{Plant ID}=i, \text{Date}=ym, \text{Source}=\text{type}}$ represents the total fuel cost calculated by the Quantity ($Q$), the Average Heat Content ($AHC$), and the Fuel Cost ($FC$) of the plant with ID = $i$ (e.g., 127) when Data = $ym$ (e.g., 201301) for Energy Source = type (e.g., SUB). Similarly, $\text{Total Heat}$ stands for the total heat of the corresponding plant with respect to the specified date and energy source.
2.5.2: Objective Data Selection

While the gross quantity is aggregated to each state in Form EIA-923 from 2013 to 2016, Texas is found to be the top state in the energy consumption sources and is taken as an example for the study. Meanwhile, by combining the quantity of each energy source, NG, Bituminous coal (BIT), Sub-bituminous coal (SUB) are discovered to be the top 3 consumed energy sources from 2013 to 2016. Hence, the NG, SUB and BIT in Texas are aimed to be analyzed subsequently.

After the $FC'$ are updated and associated Plant ID, Energy Source and Date by (2.21), (2.22) and (2.23), the NG, SUB and BIT fuel cost data in Texas are obtained and part of the data is shown in Table 2.5 as an example.

Based on the boxplot of NG shown in Fig. 1, the fuel cost distributions of NG varies month by month. Furthermore, there is a three-month delay for EIA to publish the fuel cost information for the plants. In other words, the most updated fuel cost information can be obtained for estimating the fuel cost distribution in Nov. 2015 is three months ago, i.e., the data of Aug. 2015. As shown in Fig. 2.4, the distributions between Aug. 2015 and Nov. 2015 have large differences. A more accurate NG distribution estimation is crucial and necessary.

<table>
<thead>
<tr>
<th>Date</th>
<th>Plant ID</th>
<th>Plant State</th>
<th>Energy Source</th>
<th>$FC'$ ($/MMBtu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>201301</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>2.13</td>
</tr>
<tr>
<td>201302</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>2.10</td>
</tr>
<tr>
<td>201303</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>2.09</td>
</tr>
<tr>
<td>201304</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>2.10</td>
</tr>
<tr>
<td>201305</td>
<td>127</td>
<td>TX</td>
<td>SUB</td>
<td>2.13</td>
</tr>
</tbody>
</table>
In contrast, the fuel cost distributions of SUB stays fairly stable, according to Fig. 2.5. In other words, it is still sufficient and effective to estimate a proper SUB distribution by using a 3-month delayed data. Meanwhile, BIT, as another energy source of coal, has a similar pattern of fuel cost distributions as SUB.

Therefore, the probability density prediction and estimation will focus on NG instead of SUB and BIT in this work.
2.5.3: Natural Gas Fuel Cost Analysis

According to the previous data exploration and analysis, the primary goal is to study the utility fuel cost of natural gas and estimate the probability density function for the next month in lack of updated information from EIA.

In order to forecast and estimate the NG fuel cost distribution of all 37 NG consuming plants in Texas, the NG monthly fuel costs in each plant are considered as a time series. However, due to 8 of those plants miss more than two months of fuel cost data during 2013-2016, the data from the rest 29 plants are investigated for forecasting. For those plants only miss one data point, the new data is interpolated by

\[ FC_{k,new} = \frac{FC'_{k-1} + FC'_{k+1}}{2}, \quad k = 1, 2, \ldots \]  

(2.24)
where $FC'_{k,\text{new}}$ is the new interpolated fuel cost at time $k$; $FC'_{k-1}$ and $FC'_{k+1}$ are the available fuel cost before and after time $k$, respectively. With the monthly fuel cost series of 29 plants, the ARIMA model is capable of forecasting the fuel cost three-step ahead for each plant.

In addition to the plant’s fuel cost information, the more recent electricity used NG spot price at the objective region, which is related to the corresponding plants’ fuel cost, is able to be obtained as well. For example, the monthly aggregated NG price of Texas can be found as Henry Hub Natural Gas Spot Price [107].

Both the plant’s fuel cost and the aggregated monthly Henry Hub NG price series are used for forecasting the next month fuel cost and estimating the probability distribution. Since the aggregated NG price is able to provide a more recently updated fuel cost data, it is considered to enhance plant’s fuel cost prediction by taking the difference between the aggregated price at the hub and the plant’s fuel cost. The series representing the difference between the aggregated NG price and plants’ fuel cost is extracted by (2.25):

$$\Delta FC'_t = FC'_t - FC'^{\text{hub}}_t$$

where $\Delta FC'_t$ is the differential value between the $FC'_t$ and $FC'^{\text{hub}}_t$ at time $t$.

With the differential monthly fuel cost series $\Delta FC'_t$ of 29 plants, the ARIMA model is implemented to forecast the fuel cost three-step ahead. As shown in Fig. 2.6, the next month differential fuel cost $\Delta FC'_{t+1}$ is required to be forecasted from the three-month delayed fuel cost $\Delta FC'_{t-2}$. Meanwhile, the next month aggregated fuel cost $FC'^{\text{hub}}_{t+1}$ at the NG hub is also required a one-step ahead prediction. When $\Delta FC'_{(t-2)+3}$ and $FC'^{\text{hub}}_{t+1}$ are predicted by implementing the ARIMA model, the next month forecasting $FC'_{t+1}$ can be obtained by (2.26).

$$FC'_{t+1} = \Delta FC'_{(t-2)+3} + FC'^{\text{hub}}_{t+1}$$
where $FC_{t+1}^{hub}$ is the one-step ahead forecasting NG fuel cost of Henry Hub at time $t+1$ and the $\Delta FC'_{(t-2)+3}$ is the three-step ahead forecasting differential fuel cost series at time $t+1$.

![Diagram](image)

Figure 2.6: $FC'$ forecasting process diagram.

The three fuel cost time series $FC'$, $FC^{hub}$ and $\Delta FC'$, will be used in the ARIMA model to forecast the next month fuel cost of each plant. While these forecasted fuel costs of the plants are obtained by each corresponding ARIMA model, the costs will be fitted into Normal Distribution to estimate the next month fuel cost distribution. The numerical results will be given in Section 2.7.

2.6: Model Development for Fuel Costs Prediction

Given the differential fuel cost $\Delta FC'$ and $FC^{hub}$ without any seasonal pattern, a standard ARIMA model is proposed for forecasting the NG fuel cost series of each plant to estimate the distribution for the next month.

2.6.1: ARIMA Model

Different the seasonal ARIMA model and ARMA model introduced in Section 2.3, the model used for fuel cost prediction is a standard ARIMA model, which is popular for forecasting a stationary time series or a non-stationary time series that can be made to be “stationary” by differencing. The ARIMA model for predicting $\Delta FC'$ and $FC^{hub}$ is
a standard ARIMA\((p, d, q)\) model described as follows:

\[
\phi_p (B) \nabla^d y_t = \mu + \theta_q (B) \varepsilon_t
\] (2.27)

where \(B\) is the backward shift operator, i.e. \(B^h y_t = y_{t-h}\); \(p\) is the auto-regression (AR) order, which determines how many past values are used for regression; \(d\) is the differencing order, which is often used when the stationary assumption is not met; \(q\) is the moving-average (MA) order, which determines how many previous error term \(\varepsilon_t\) of the process should be considered. The error terms \(\varepsilon_t\) are generally assumed to be the independent and identically distributed noise with zero mean and finite variance, i.e., Gaussian white noise; \(\mu\) is a constant term.

\[
\phi_p (B) = 1 - \phi_1 (B) - \phi_2 (B^2) - \cdots - \phi_p (B^p)
\] (2.28)

\[
\nabla^d = (1 - B)^d
\] (2.29)

\[
\theta_q (B) = 1 - \theta_1 (B) - \theta_2 (B^2) - \cdots - \theta_q (B^q)
\] (2.30)

Furthermore, \(\phi_1 \ldots \phi_p\) and \(\theta_1 \ldots \theta_q\) are the coefficients of the autoregressive and the moving average polynomials, respectively.

2.6.2: Normal Distribution Fitting

Since the normal distribution in (2.31) has already been clarified and proven to be a good selection for estimating the fuel cost distribution in [110]. This study keeps using the standard normal distribution on fitting the data with maximum likelihood method to estimate the probability density function.

\[
F (X) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\] (2.31)

where \(\sigma\) and \(\mu\) represent the standard deviation and mean of the fitted density function, respectively.
2.6.3: Kullback-Leibler (KL) Divergence of Normal Distributions

To evaluate the forecasting performance with fitted normal distribution, the symmetric KL divergence [111] between two distributions is formulated (2.32):

\[ D_{p,q} = KL(p||q) + KL(q||p) \] (2.32)

where \( p \) and \( q \) are two probability distributions and \( D_{p,q} \) stands for the divergence between \( p \) and \( q \); \( KL(p,q) \) for the continuous probability distributions is expressed as (2.33):

\[ KL(p||q) = - \int p(x) \log q(x) \, dx + \int p(x) \log p(x) \, dx \] (2.33)

Since the fitted probability distribution is normal (Gaussian) distribution, which means \( p(x) = N(\mu_1, \sigma_1) \) and \( q(x) = N(\mu_2, \sigma_2) \), \( KL(p||q) \) can be derived as:

\[ KL(p||q) = \log \frac{\sigma_2}{\sigma_1} + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2} \] (2.34)

2.7: Prediction Performance for Fuel Cost Distributions

The proposed prediction model has been applied to forecast the NG fuel cost in Texas. The training data set, from Jan. 2013 to June 2016, including 42 months of fuel cost data and NG hub spot price data are used to obtain the parameters of the ARIMA model. The maximum likelihood estimation method is used to fit the model parameters with the training dataset. In this study, the toolboxes in Python [112] are used to estimate the model parameters and forecast.

Six months of NG fuel cost data from July 2016 to Dec. 2016 are employed for validating the model. Then the normal distribution is used to fit the data to obtain an estimated distribution of the fuel cost. Meanwhile, the symmetric KL divergence is used to evaluate the distribution forecasting algorithm performance. The forecasting results show the developed algorithm makes more accurate predictions than using the three-month delayed EIA-923 data when the NG market does not have high volatility.
2.7.1: Data Preprocessing

In order to reduce the data fluctuation before fitting the ARIMA model, the natural logarithm transformation for both the Henry hub NG price and plants’ fuel cost, given in (2.35) and (2.36), are implemented:

\[
FC_{log}^{hub} = \log (FC_{t}^{hub})
\]  
(2.35)

\[
\Delta FC_{log}' = \log (\Delta FC_{t}' + c)
\]  
(2.36)

where \(c\) is a positive constant offset adding on \(\Delta FC_{t}'\) to guarantee positive value for the logarithm transformation, because there exists negative differential price value between plant’s fuel cost and NG hub spot price obtained by (2.25). Aiming at forecasting \(\Delta FC_{(t-2)+3}'\) of each plant and the NG hub fuel cost \(FC_{t+1}^{hub}\), the ARIMA models for these two types time series are described as (2.37) and (2.38):

\[
\phi_{p}(B) \nabla^{d} FC_{log}'_{t} = \mu + \theta_{q}(B) \varepsilon_{t}
\]  
(2.37)

\[
\phi_{p}(B) \nabla^{d} FC_{log}^{hub}_{t} = \mu + \theta_{q}(B) \varepsilon_{t}
\]  
(2.38)

By applying Bayesian Information Criterion (BIC) [113] and observing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots [109], the model order is selected to be \(ARIMA(2,1,1)\) for predicting the processed Henry hub NG spot price \(FC_{log}^{hub}_{t}\). Similarly, either \(ARIMA(2,1,1)\) or \(ARIMA(2,0,1)\) is selected in order to predict the differential fuel costs \(\Delta FC_{log}'_{t}\) of the objective plants in Texas. Since the \(\Delta FC_{log}'_{t}\) series of some plants are stationary, the series differencing is not necessary and the \(ARIMA(2,0,1)\) is selected for them. Otherwise, \(ARIMA(2,1,1)\) is implemented.

2.7.2: Results Comparison

When the prediction results between July 2016 and Dec. 2016 are obtained, they are fitted to normal distributions, as shown in Fig. [2.7]. At the same time, the corresponding
fitted three-month delayed distribution and real fuel cost distribution are plotted in Fig. 2.7 as well.

By observing these distributions between July 2016 and Sept. 2016, the proposed distribution forecasting method using the proposed ARIMA model achieves much better results than the method estimating the distribution by using the plants’ fuel cost data from three months ago. In Oct. and Nov., 2016, both the forecasted distribution and the delayed distribution are very close to the real one. However, in Dec. 2016, the forecasting
method gives a worse result than the delayed data.

The mean of each fitted normal distribution are presented in Table 2.6. It can be found that the proposed ARIMA forecasting method is capable of reducing the error by about 10 to 20% when the overall NG market price is steadily increasing from 2.97 \$/MMBtu in July 2016 to 3.27 \$/MMBtu in Oct. 2016. However, when NG market price unexpectedly dropped to 2.95 \$/MMBtu in Nov. 2016 and bounced back to 4.2 \$/MMBtu, the forecasting model does not perform well. The reason is that forecasting algorithm turns the prediction price lower for Dec. 2016 to catch up the decreasing price while it receives the price dropping signal from Nov. 2016. Hence, when the overall NG market prices are volatile without any pattern, the forecasting algorithm does not achieve a better result than just using the delayed data.

<table>
<thead>
<tr>
<th>Month</th>
<th>$Mean_{real}$</th>
<th>$Mean_{delay}$</th>
<th>Error</th>
<th>$Mean_{forecast}$</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>201607</td>
<td>2.97</td>
<td>2.29</td>
<td>22.9%</td>
<td>3.01</td>
<td>1.3%</td>
</tr>
<tr>
<td>201608</td>
<td>2.98</td>
<td>2.26</td>
<td>24.2%</td>
<td>3.29</td>
<td>10.4%</td>
</tr>
<tr>
<td>201609</td>
<td>3.24</td>
<td>2.72</td>
<td>16.0%</td>
<td>3.12</td>
<td>3.7%</td>
</tr>
<tr>
<td>201610</td>
<td>3.27</td>
<td>2.97</td>
<td>9.1%</td>
<td>3.27</td>
<td>0%</td>
</tr>
<tr>
<td>201611</td>
<td>2.95</td>
<td>2.98</td>
<td>1%</td>
<td>3.26</td>
<td>10.5%</td>
</tr>
<tr>
<td>201612</td>
<td>4.20</td>
<td>3.24</td>
<td>22.9%</td>
<td>2.8</td>
<td>33%</td>
</tr>
</tbody>
</table>
Since the data listed in Table 2.6 only considers the mean, KL divergence formulated in (2.35) is implemented to check the divergence for the distributions of three-month delay fuel costs, forecasting fuel cost and the actual fuel cost. The KL divergence numerical results are listed in Table 2.7. In information theory, the KL divergence is used to measure the difference between two probability distributions over the same variable. The higher value of $D_{p,q}$ indicates the larger diversity between distribution $p$ and $q$. In Table 2.7, the forecasting algorithm is able to improve the estimation by 99.3%, 74.6%, 85.8% and 41.6% while estimation becomes worse in Nov. and Dec. of 2016.

The reason why the three-month delay estimated distributions outperform the forecasting distributions is price bouncing back to the previous level after a suddenly decreasing, which causes an imprecise NG hub spot price prediction. The price data from three months ago was still at a similar level, while the forecasting algorithm determines that is a turning point in Nov. 2016 and provides a declining price prediction in Dec. 2016.
2.8: Summary

In this chapter, various ARMA and ARIMA models (seasonal, non-seasonal, and with or without exogenous data included) were developed and compared for short-term electricity price prediction using the published DALMP and historical RTLMP data in an electricity market. A GARCH model was developed to handle price volatility for all the developed ARMA and ARIMA models. All the SARMA, SARMAX, ARMA and ARMAX models and the models with the GARCH model were developed and tested using the actual day-ahead and historical real-time LMPs data in MISO. The results show that adding seasonal data does not help improve the prediction accuracy. The results also show that the ARMAX-GARCH model with the weekday/weekend information as the exogenous data, has the best performance compared to all models. It achieves 27.21% improvement for the one-hour-ahead predictions, 17.10% improvement for the two-hour-ahead predictions and 11.35% improvement for the three-hour-ahead predictions. The ARMAX-GARCH model also shows over 5% improvement for predictions 12 hours ahead.

In addition, an algorithm for forecasting the fuel cost was developed to obtain more accurate distribution estimations. Based on the data exploration, the NG fuel cost in Texas was selected as an example to develop and validate the proposed fuel cost forecast algorithm. The results show the proposed forecasting algorithm has a superior performance over the method that uses the three-month delayed data. Especially when the low volatility presents in the forecasting time domain, the proposed method is capable of achieving a very accurate estimation on the distributions mean with errors of 1.3%, 10.4%, 3.7% and 0%. The results also indicate the good capability of the method in handling the prediction after the turning point. The proposed algorithm is able to be extended and implemented in other states and for other types of fuel costs when there
are sufficient data.
CHAPTER 3: TEMPORAL AND SPATIAL LOAD MANAGEMENT METHODS FOR COST AND EMISSION REDUCTION

3.1: Motivation and Contribution

To address the challenges of climate change, reducing emissions due to electric power generation and consumption has received increasing attention worldwide. The previous research efforts have been focused on emission reduction on the generation side, but the positive impacts through demand side load management are often ignored. In this chapter, load shifting in both space and time frames is investigated by extending the preliminary results reported in [76]. A full load management method called Temporal and Spatial Load Management (TSLM) is proposed, which is able to dispatch load and generation simultaneously. The proposed method is then compared with a temporal only load management method which can only shift load on timeline. Moreover, a self-optimizing load management method based on day-ahead LMP is presented. To overcome the limitation of the basic self-optimizing load management (which will be discussed in detail later in this chapter), two improved individual load management methods, namely sliding window self-optimizing load management and day ahead self-optimizing load management, are proposed. In the sliding window self-optimizing load management scheme, customers can adjust their loads in a sliding time frame from 24-h to 2-h in a day. In the day ahead self-optimizing load management scheme, customers will use the 24-h day ahead LMPs to do load management. The customers can communicate with the ISO for several rounds to continuously optimize the load management scheme before executing the final optimized 24-hour load scheduling plan. Simulation studies on the IEEE 14-bus system are used to verify the proposed temporal and spatial load-generation management method and to compare the different load management methods discussed in the chapter. A more practical power system, i.e. the IEEE 57-bus system, is also employed to further test the
progressive self-optimizing load management methods.

3.2: Load Management
3.2.1: Cost Model

The cost models used in this chapter are developed based on the models in [73–75]. One model only considers generation cost and the other one has emission cost added to the whole cost model, as shown in (3.1) - (3.3).

\[ G_i = a_i P_{G,i}^2 + b_i P_{G,i} + c_i , \quad i = 1, 2 \ldots N_G \] (3.1)

where \( P_{G,i} \) is the output power of generator \( i \); \( a_i, b_i \) and \( c_i \) are the corresponding parameters in the quadratic cost model; \( N_G \) is the total number of generators in the system.

\[ e_i = \gamma_{CO_2} \times EF_{CO_2,i}, \quad i = 1, 2 \ldots N_G \] (3.2)

\[ E_i = a_i P_{G,i}^2 + b_i P_{G,i} + c_i + e_i P_{G,i}, \quad i = 1, 2 \ldots N_G \] (3.3)

where \( e_i \) represents the generation emission cost factor; \( \gamma_{CO_2} \) is the emission price of \( CO_2 \) and \( EF_{CO_2,i} \) represents the emission factor of generator \( i \); \( e_i P_{G,i} \) represents the emission cost part in the combined cost model.

3.2.2: Temporal and Spatial Load Management (TSLM)

This load management method is formed to optimize the load distribution on both the temporal and spatial frame for a whole control area. The objective function of the TSLM is:

\[ C = \min \sum_{i}^{N_G} G_i \] (3.4)

where \( G_i \) represents the cost of generator \( i \), and \( C \) is the total cost of generators.

If the emission is taken into account in the object, the cost model of (3.1) will be
taken place by the combined cost model of (3.3). The objective function then becomes:

$$C = \min \sum_{i} E_i$$  \hspace{1cm} (3.5)

As each node in the area has a certain amount of controllable load capacity, the load at each bus may vary between a maximum value and a minimum value for every hour as shown in (3.6). All the generators’ outputs should be within their limits in (3.7). Without considering the transmission loss, the total generators’ output power should be equal to the total load demand at each hour (3.8). The total amount of optimized load demand in one day is kept constant as given in (3.9).

$$P_{L,i,h}^{0,min} \leq P_{L,i,h} \leq P_{L,i,h}^{0,max}, \ i = 1, 2 \ldots N_B$$  \hspace{1cm} (3.6)

$$P_{G,i,h}^{0,min} \leq P_{G,i,h} \leq P_{G,i,h}^{0,max}, \ i = 1, 2 \ldots N_G$$  \hspace{1cm} (3.7)

$$\sum_{i=1}^{N_B} P_{L,i,h} = \sum_{i=1}^{N_G} P_{G,i,h}$$  \hspace{1cm} (3.8)

$$\sum_{h=1}^{H} \sum_{i=1}^{N_B} P_{L,i,h} = \sum_{h=1}^{H} \sum_{i=1}^{N_B} P_{L,i,h}^{0}$$  \hspace{1cm} (3.9)

where $P_L$ and $P_G$ represent the bus power and generator power respectively meanwhile $P_L^0$ and $P_G^0$ are the power of the loads and generators in the base case (i.e., without load management). $N_B$ is the number of buses in the simulation case. The $h$ represents the hour and $H$ serves as the total hours for optimization. The $\min$ and $\max$ superscripts define the lower bound and upper bound of the power.

Besides the constraint conditions in (3.6)-(3.9), the constraints in DC power flow
(DCPF) are also implemented:

\[ P_i = \sum_{i=1, i \neq j}^{N_B} \frac{1}{x_{ij}} (\theta_i - \theta_j), \quad i, j = 1, 2 \ldots N_B \]  

\[ \frac{1}{x_{ij}} (\theta_i - \theta_j) \leq F_{ij}^{\text{max}}, \quad i, j = 1, 2 \ldots N_B \]  

\[ \frac{1}{x_{ij}} (\theta_i - \theta_j) \geq F_{ij}^{\text{min}}, \quad i, j = 1, 2 \ldots N_B \]  

\[ -\pi \leq \theta_i \leq \pi, \quad i = 1, 2 \ldots N_B \]  

where \( P_i \) represents the active power injection at bus \( i \) and \( x_{ij} \) is the reactance of the line from bus \( i \) to bus \( j \) with line resistance \( R=0 \). \( \theta \) is bus voltage angle vector. \( F_{ij}^{\text{min}} \) and \( F_{ij}^{\text{max}} \) mean the lower bound and upper bound of the branch flow. Equation (3.13) defines the limits for each bus voltage angle: between \(-\pi\) and \(\pi\).

3.2.3: Temporal Only Load Management (TLM)

This load management method considers that the load can only be shifted in time and the total load consumption should be kept unchanged in the period of study (one day in this study). In other words, each bus in the control area has to reach its settled load level in one day and there is no load shifting between different buses in the area. With this consideration, the constraint of (3.14) is added. This is the main difference between the TSLM and the TLM schemes.

\[ \sum_{h=1}^{H} P_{L,i,h} = \sum_{h=1}^{H} P_{L,i,h}^0, \quad i = 1, 2 \ldots N_B \]  

where \( P_{L,i,h}^0 \) is the load at each bus in the base case and \( P_{L,i,h} \) is the load at each bus after optimization.

The objective function for TLM can be (3.4) or (3.5) depending on whether the emission cost is included or not. Equation (3.14) is an additional constraint in TLM and all the other constraints of TLM are the same as those in the TSLM, namely (3.6), (3.7), (3.8), (3.9), (3.10), (3.11), (3.12) and (3.13).
3.2.4: Self-Optimizing Load Management (SOLM)

Based on the day ahead LMP information as shown in Fig. 3.1, customers will have their own strategies to arrange their load distributions in order to reduce the cost. The LMP profile in Fig. 3.1 shows a typical 24-hour day ahead LMP curve of a price node in PJM [115].

![LMP Curve](image)

Figure 3.1: A typical 24-hour LMP curve of a node in PJM.

The SOLM is a kind of autonomous scheme, which means the customers do not have the communication with each other and all the decisions are made on the basis of the given LMP. The objective function of SOLM is given in (3.15).

\[
C_i = \min \sum_{h=1}^{H} (LMP_{i,h} \times P_{L,i,h}) , \quad i = 1, 2 \ldots N_B
\]  

(3.15)

where \( LMP_{i,h} \) is the day ahead LMP for bus \( i \) at hour \( h \), \( P_{L,i,h} \) is the power demand of bus \( i \) at hour \( h \) and \( C_i \) is the cost after self-administration at bus \( i \).

The constraints in SOLM are (3.14) and (3.16)

\[
P_{L,i,h}^{0,\min} \leq P_{L,i,h} \leq P_{L,i,h}^{0,\max}
\]

(3.16)

Equation (3.16) is the inequality constraint to represent the level of controllable load at each bus and (3.14) in SOLM is to keep the total amount of power demand unchanged.
If the emissions are considered, the generators cost models will become \( (3.3) \). At the same time, the emission cost will be a part of the given LMPs delivered from the ISO to customers.

When the customers have the day ahead LMP from the ISO, they can manage their loads separately with the objective function of \( (3.15) \) under the constraints of \( (3.14) \) and \( (3.16) \). Eventually, a new load distribution would be formed by combining each customer’s scheduled load. The load management flow chart is shown in Fig. 3.2.

![Figure 3.2: SOLM algorithm flow chart.](image)

**3.2.5: Improved Self-Optimizing Load Management**

As the results of SOLM (which will be discussed later) show that the SOLM will cause peak LMP re-strike at different hours and could not reduce the total cost in an effective
way, two improved SOLMs, namely sliding window self-optimizing load management and
day-ahead self-optimizing load management, are developed to achieve a better demand
side load management than the original SOLM.

**Sliding Window Self-Optimizing Load Management (SW-SOLM)**

In the SW-SOLM, the customers and the ISO would have a real time communication
from 0:00 to 22:00 in a day. When the customers get the 24-h LMP from the ISO at 0:00,
they would arrange the 24-h load and execute the first hour load scheduling plan by
sending the next 23-h load arrangement back to the ISO. After receiving the new 23-h
load management plan from the customers, the ISO would use it to analyze and obtain
a new set of LMPs for the next 23-h and forward the updated LMP information to the
customers accordingly. The customers would then adjust their loads based on the updated
23-h LMP information and execute the second hour load management at 1:00 and send
back a 22-h load arrangement to the ISO again. Hence, the individual load management
by each customer would be accomplished based on a temporal sliding window, such as
24-h, 23-h ... 3-h and 2-h. The real time communication process between the ISO and
individual customers is shown in Fig. 3.3.
During the whole exchange processes between the customers and the ISO, the customers will do the individual load management following (3.14), (3.15), (3.16) and (3.17):

\[
|P_{L,i,h} - P'_{L,i,h}| \leq R \times P_{L,i,h}^{\text{Controllable}}
\]  

(3.17)

where \(P'_{L,i,h}\) is the previously scheduled hourly load at each bus and \(P_{L,i,h}\) is the arranged hourly load by the customers. \(P_{L,i,h}^{\text{Controllable}}\) represents the hourly controllable load at each bus of the based case and \(R\) is a scale ratio. Following constraint (3.17), all the customers would be prohibited from adjusting too much controllable loads to the low LMP hours at one time, which prevents some unexpected LMP peaks. It could generate new price spikes if too much controllable loads are shifted to the same low LMP hours and could convert the original valley hours to the peak hours.
Day Ahead Self-Optimizing Load Management (DA-SOLM)

As shown in Fig. 3.4 in the DA-SOLM, it is assumed that the ISO would publish the day ahead LMP at 16:00 on the current day for customers to arrange their hourly loads in the next day. The customers could use the SOLM by following (3.14), (3.15), (3.16) and (3.17) to adjust their 24-h loads and send back their load management plans for the next day to the ISO at 16:20 assuming that the process can be done in 20 min. After receiving the updated information from the customers, the ISO will carry out a new round of calculation and optimization and release a set of new LMPs to the customers. It is assumed that the process at the ISO side can be finished also in 20 min. Therefore, the customers can receive a new set of LMPs at 16:40 on the current day. This process will continue until the time reaches the next day. After the ISO receives the last round of load management information from the customers at 23:20, the ISO will do the final round of LMP update and send the updated LMP to the customers. The customers will choose the best results among all the trial arounds of load management then finalize their load management plans right before the next day.
3.2.6: Total Cost and Emission

It is expected that the five different load management methods would produce improved load distribution from the base case. The new LMP at each bus can then be calculated by using the optimal power flow (OPF) in MATPOWER [116] from hourly loads in the optimized cases. Therefore, the overall cost $C$ and the total carbon dioxide emission $E_{CO_2}$ can be calculated based on the optimization results of each method by
and (3.19), respectively. The five methods are compared based on the above two parameters, i.e., the total cost and the total emission. It should be noted that the total cost is calculated based on the LMP values.

\[
C = \sum_{h=1}^{H} \sum_{i=1}^{NB} LMP_{i,h} \times P_{L,i,h}
\]

(3.18)

\[
E_{CO_2} = \sum_{h=1}^{H} \sum_{i=1}^{NG} P_{G,i,h} \times EF_{CO_2,i}
\]

(3.19)

3.3: Algorithm Implementation And Simulation Studies

3.3.1: Cost Model

Based on the generator cost data from the U.S. Energy Information Administration [117], the parameters of the generators quadratic cost model can be obtained as given in Table 3.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fuel Type</th>
<th>Coal</th>
<th>Natural Gas</th>
<th>Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td>0.035</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td>31</td>
<td>46</td>
<td>240</td>
</tr>
<tr>
<td>c</td>
<td></td>
<td>200</td>
<td>300</td>
<td>400</td>
</tr>
</tbody>
</table>

According to the social cost of carbon (SCC) estimated by the U.S. Environmental Protection Agency (EPA) in 2013 [118], the price of carbon dioxide is set at 10 $/ton. Hence the price of \(CO_2\) could be obtained by (3.20). Based on the EIA’s data, the emission factors of different types of generators are listed in Table 3.2. Due to the various emission factors of natural gas, the emission factors for natural gas generators are set to two typical values, as shown in Table 3.2. Therefore, the emission cost factor \(e\) can be calculated...
by \((3.20)\). The fuel types and emission factors of each generator in the IEEE-14 bus test system are set as shown in Table 3.3.

\[
\gamma_{CO_2} = 10\$/ton \times 0.00045\text{ton/lb} = 0.0045\$/lb
\]

where \(10\$/ton\) is the emission price of \(CO_2\) and the conversion rate of \(CO_2\) from \(lb\) to \(ton\) is 0.00045.

Table 3.2: Carbon Dioxide Emission Parameters

<table>
<thead>
<tr>
<th>Fuel Type</th>
<th>Coal</th>
<th>Natural Gas</th>
<th>Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>(EF_{CO_2}) (lbs/MWh)</td>
<td>2159</td>
<td>934/1450</td>
<td>1911</td>
</tr>
<tr>
<td>(c) ($/MWh)</td>
<td>9.72</td>
<td>4.20/6.53</td>
<td>8.60</td>
</tr>
</tbody>
</table>

Table 3.3: Generator Fuel Types and Emission Factors

<table>
<thead>
<tr>
<th>Generator Number</th>
<th>Bus Number</th>
<th>Fuel Type</th>
<th>(EF_{CO_2}) (lbs/MWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Coal</td>
<td>2159</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Natural gas</td>
<td>934</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Natural gas</td>
<td>1450</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>Oil</td>
<td>1911</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>Oil</td>
<td>1911</td>
</tr>
</tbody>
</table>

3.3.2: Base Case Formulation of IEEE 14-Bus System

The aforementioned five optimization methods are tested on the IEEE 14-Bus system. The basic 24-hour load profile in the case is developed based on the data from one of the PJM daily hourly loads [115], shown in Fig. 3.5.
Figure 3.5: Basic load profile.

If all the nodes follow the identical load ratio for 24 hours, the loads at all the buses have same type of load profile over the 24-hour period. With considering the randomness in consumers’ activities and their different behaviors, the hourly load at each node can vary from the load curve shown in Fig. 3.5 following a normal distribution between -3% and 3%. It is assumed that ±10% of controllable load can be adjusted at each bus.

3.3.3: Base Case without Load Management

The LMPs of the base case is obtained via MATPOWER, shown in Fig. 3.6. As shown in the figure, without load management there are two sets of LMP peaks appearing at around 10:00 and 20:00 which are consistent with the load peaks in Fig. 3.5. If there is no transmission congestion in the control area, the LMPs in this area should be the same. But in Fig. 3.6 the LMPs at buses 1, 2 and 3 are lower than the other buses’ around 10:00 and 20:00, indicating that some congestions occurred.
3.3.4: Temporal and Spatial Load Management

The result of TSLM is shown in Fig. 3.7. By comparing Figs. 6 and 7, it is clearly seen that at around 10:00 and 20:00 the LMPs in Fig. 3.6 are almost 4 times of the LMPs before implementing the TSLM (shown in Fig. 3.7). This means that the TSLM has eliminated congestion and reduced the peaks in the load demands, resulting in a flatter LMP distribution as well.
3.3.5: Temporal Only Load Management

The optimization result of the TLM is shown in Fig. 3.8 which looks similar to Fig. 3.7. However, the amplitudes in Fig. 3.8 are slightly higher than those of LMPs in Fig. 3.7 since the TSLM has more capability in controlling loads to achieve a better result.
3.3.6: Self-Optimizing Load Management

SOLM is a distributed load management by individual customers while both the TSLM and TLM are central management methods to reach the optimization goal. When the customers only have the 24-hour LMP to make their self-governed decision in the SOLM, an interesting result can be obtained as shown in Fig. 3.9. The LMP distribution after the SOLM displays some new peak levels of LMPs, when compared to the LMP distribution figures of TSLM and TLM. By comparing the base case and the SOLM, it can be found that the high level of LMPs peaks re-struck at around 8:00, 16:00 and 23:00 in the SOLM, which are different from the peaks happened at 10:00 and 20:00 in the base case. Hence, some unexpected load peaks would be generated, if the customers only consider their own behaviors individually.

![Figure 3.9: LMP distribution after SOLM.](image)

3.3.7: Sliding Window Self-Optimizing Load Management

For enhancing the SOLM, the SW-SOLM is designed based on a sliding time frame for customers to accomplish the individual load arrangement hour by hour in 24 hours. In
the simulation study, the ratio $R$ in (3.17) is chosen to be 0.618 and the controllable load $P_{\text{Controllable}}^{L,i,h}$ is assigned to 10% of the original load $P_{L,i,h}^0$. The result of the SW-SOLM is shown in Fig. 3.10.

![Figure 3.10: LMP distribution after SW-SOLM.](image)

As shown in Fig. 3.10, the LMP peaks re-struck back at 10:00, 15:00, 19:00 and 21:00. Nevertheless, comparing the result of SOLM in Fig. 3.9 and the result of the SW-SOLM given in Fig. 3.10, it can be seen that less LMP spikes are generated in the SW-SOLM than in the basic SOLM. This indicates that the SW-SOLM has a better performance in cost reduction.

3.3.8: Day Ahead Self-Optimizing Load Management

Another enhanced self-optimizing load management, DA-SOLM, via using multiple rounds of communication between ISO and customers, has also been developed to improve the performance of distributed self-optimizing load management.

Table 3.4 shows the result of DA-SOLM after 11 rounds of communication (i.e., LMP and load information exchanges between the ISO and the customers). According to the
total cost listed, the 10th round gives the best result, i.e. the minimum cost among all the trial rounds. In other words, the total cost is minimum cost among all the records of the total cost. Hence, the 10th scheme will be chosen by the customers for the next day load management.

Table 3.4: Total Cost of Each Loop in DA-SOLM

<table>
<thead>
<tr>
<th>Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($ \times 10^5$)</td>
<td>7.0133</td>
<td>5.8810</td>
<td>6.2667</td>
<td>6.2443</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($ \times 10^5$)</td>
<td>6.2551</td>
<td>6.2450</td>
<td>6.6297</td>
<td>5.5057</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($ \times 10^5$)</td>
<td>6.6285</td>
<td>4.7676</td>
<td>6.6354</td>
</tr>
</tbody>
</table>

Figure 3.11: LMP distribution after DA-SOLM.

By using the 10th load arrangement scheme, the LMP distribution of the DA-SOLM has been acquired, shown in Fig. 3.11. There are two LMP peaks at 13:00 and 14:00.
after the DA-SOLM is implemented. By comparing the result in Fig. 3.11 and the results shown in Figs. 3.9 (SOLM) and 3.10 (SW-SOLM), it can be seen that the day ahead communication can help achieve a better distributed demand side load management. Besides, a better result of the DA-SOLM can be achieved if more rounds of trials can be done before the next-day actual load management plan implementation. As shown in Table 3.5 if the number of rounds can reach 18, a better result can be achieved. In the extended result of the DA-SOLM, the 16th round produces an even lower cost than the 10th around which was chosen when only 11 rounds of communication were done. The more number of rounds of information exchange, the more opportunity to discover a better load management scheme with a lower total cost.

<table>
<thead>
<tr>
<th>Number</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($ 	imes 10^5$)</td>
<td>4.7657</td>
<td>7.0039</td>
<td>4.3951</td>
<td>7.0016</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($ 	imes 10^5$)</td>
<td>4.0276</td>
<td>7.3757</td>
<td>4.3955</td>
</tr>
</tbody>
</table>

**3.3.9: Comparison Study of IEEE 14-Bus System**

In Table 3.6 the LMP mean value and standard deviation of bus #4 over 24 hours in the IEEE 14-bus system for each load management method are listed. The TSLM and TLM have the best performance and they are able to reduce cost and flatten the LMP distribution in the time domain with a small standard deviation. For the SOLM methods, they are good for enhancing the demand side management because of the lower mean value and standard deviation than the base case although they are not as powerful as the TSLM and TLM, which are centralized and need more system resources for optimization.
<table>
<thead>
<tr>
<th>Case</th>
<th>Mean($)</th>
<th>Standard deviation($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>110.70</td>
<td>75.00</td>
</tr>
<tr>
<td>TSLM</td>
<td>54.13</td>
<td>1.95</td>
</tr>
<tr>
<td>TLM</td>
<td>57.06</td>
<td>1.67</td>
</tr>
<tr>
<td>Basic SOLM</td>
<td>92.30</td>
<td>66.89</td>
</tr>
<tr>
<td>SW-SOLM</td>
<td>80.91</td>
<td>56.76</td>
</tr>
<tr>
<td>DA-SOLM</td>
<td>69.11</td>
<td>41.85</td>
</tr>
</tbody>
</table>

The total cost and emission in the different cases are summarized in Table 3.7 for comparison when the emission cost is not considered. It should be noted that, compared to the base case, the TSLM, TLM and SOLM can all reduce both the cost and emission, i.e., about 48.24%, 46.01% and 15.29% reduction in cost and $171 \times 10^3$ lbs, $4 \times 10^3$ lbs and $18 \times 10^3$ lbs reduction in emissions, respectively. Additionally, the SW-SOLM and the DA-SOLM can produce a better result compared to the original SOLM. According to the cost of the base case, the SW-SOLM and the DA-SOLM cut down the cost by 25.94% and 36.07%, respectively.
Table 3.7: Total Cost and Emission of Regular Cost Model

<table>
<thead>
<tr>
<th>Case</th>
<th>Cost ($)</th>
<th>Emission (lbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Case</td>
<td>$7.4571 \times 10^5$</td>
<td>$1.3721 \times 10^7$</td>
</tr>
<tr>
<td>TSLM</td>
<td>$3.8601 \times 10^5$</td>
<td>$1.3550 \times 10^7$</td>
</tr>
<tr>
<td>TLM</td>
<td>$4.0261 \times 10^5$</td>
<td>$1.3717 \times 10^7$</td>
</tr>
<tr>
<td>SOLM</td>
<td>$6.3169 \times 10^5$</td>
<td>$1.3703 \times 10^7$</td>
</tr>
<tr>
<td>SW-SOLM</td>
<td>$5.5224 \times 10^5$</td>
<td>$1.3713 \times 10^7$</td>
</tr>
<tr>
<td>DA-SOLM</td>
<td>$4.7676 \times 10^5$</td>
<td>$1.3716 \times 10^7$</td>
</tr>
</tbody>
</table>

Furthermore, in order to analyze the impact of $CO_2$ price on emission reduction, the simulations with different $CO_2$ price have been carried out, too. As shown in Table 3.8, the TSLM method has the maximum carbon dioxide emission reduction. Meanwhile, with the increasing $CO_2$ price from $10$ to $25$, the average emission reduction of all the load management methods is about 0.74 million lbs. In general, the higher emission price is implemented to the combined cost model, the more emission reduction could be achieved in all the methods.

Table 3.8: Carbon Dioxide Emission with Different $CO_2$ Prices

<table>
<thead>
<tr>
<th>$\gamma_{CO_2} ($/ton)$</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
</table>
3.3.10: Case Studies on the IEEE 57-bus System

In a large power system, it is hard to collect the whole system’s information to implement optimal load management. While a centralized load management in a wide power system requires overwhelming system resources, the DSM will become more useful to reduce the cost and emissions than a centralized load regulation scheme.

The two proposed advanced self-optimization load management methods have been tested on the IEEE 57-bus power system. The test system has seven generators including three coal generators, three natural gas generators and one oil generator \[119\]. The fuel type and emission factor of each generator are listed as in Table 3.9 according to \[117-119\].

Table 3.9: Generator Fuel Types and Emission Factors in the IEEE 57-bus System

<table>
<thead>
<tr>
<th>Generator Number</th>
<th>Bus Number</th>
<th>Fuel Type</th>
<th>$EF_{CO_2}$ (lbs/MWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Coal</td>
<td>2159</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Coal</td>
<td>2159</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Coal</td>
<td>2159</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>Natural gas</td>
<td>934</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>Natural gas</td>
<td>934</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>Natural gas</td>
<td>1450</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>Oil</td>
<td>1911</td>
</tr>
</tbody>
</table>

By using the same 24-h load ratio profile with a base load of 1520.03MW and a load deviation between -3% and 3% among all the buses as in the base case formulation of the IEEE 14-bus system, the base case load profile of the IEEE 57-bus system has been formed. The 24-h LMPs can be obtained as shown in Fig. 3.12. From the figure, LMP peaks of the base case are seen around 10:00 and 20:00, which occur at about the same
time points as in the previously studied IEEE 14-bus case.

When the SW-SOLM is used to manage the load in the IEEE 57-bus case, a new LMP curve can be obtained, also shown in Fig. 3.12. There are three peak restrikes at 13:00, 19:00 and 21:00, but less LMP peaks occur after executing the SW-SOLM compared to the base case. The total costs are listed in Table 3.11. The result shows that the SW-SOLM is able to cut down the cost by 21.6% and reduce $2.11 \times 10^5$ lbs of CO$_2$ emission.

In the DA-SOLM process, the total cost after each communication between ISO and customers in the IEEE 57-bus system are listed in Table 3.10, which has the lowest total cost at the 8th step. Based on the proposed DA-SOLM, customers will follow the 8th load management plan.

<table>
<thead>
<tr>
<th>Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($\times 10^6$)</td>
<td>4.4600</td>
<td>3.7244</td>
<td>3.4855</td>
<td>3.4877</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($\times 10^6$)</td>
<td>3.7237</td>
<td>3.4848</td>
<td>3.9658</td>
<td>3.0119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost ($\times 10^6$)</td>
<td>4.1978</td>
<td>3.2510</td>
<td>4.1913</td>
</tr>
</tbody>
</table>

The 24-h LMPs after executing the DA-SOLM is also shown in Fig. 3.12. It can be seen from the figure that the LMP peaks have been eliminated by the DA-SOLM. As listed in Table 3.11, the DA-SOLM has cut down the total cost by 36.7% and the CO$_2$ emission by $2.70 \times 10^5$ lbs.
Table 3.11: Total Cost and Emission of the IEEE 57-bus System

<table>
<thead>
<tr>
<th>Case</th>
<th>Cost ($)</th>
<th>Emission (lbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Case</td>
<td>$4.7589 \times 10^6$</td>
<td>$5.8929 \times 10^7$</td>
</tr>
<tr>
<td>SW-SOLM</td>
<td>$3.7304 \times 10^6$</td>
<td>$5.8718 \times 10^7$</td>
</tr>
<tr>
<td>DA-SOLM</td>
<td>$3.0119 \times 10^6$</td>
<td>$5.8659 \times 10^7$</td>
</tr>
</tbody>
</table>

Figure 3.12: 24-h LMP of different cases in the IEEE 57-bus system.

3.4: Summary

Five different load management methods have been developed and compared in this chapter. Among them, the TSLM allows dispatching loads in both space and time frames while the TLM only shifts load in time domain. The TSLM and TLM are the central control methods which need the overall system information, such as power system topology, customers’ power demand, generator cost model, etc. The SOLMs, a distributed load management method self-governed by individual users, were also formed and analyzed. The weakness, such as multiple occurrences of new LMP peaks, of the basic SOLM
was revealed in the simulation studies and two other types of enhanced SOLM methods, namely, SW-SOLM and DA-SOLM, have been developed. The simulation studies have been conducted on the IEEE 14-bus system and the IEEE 57-bus system to verify and compare the five different load management methods. The simulation results show that while all the load management methods can reduce cost and emissions the TSLM can achieve the best result with a $48.24\%$ reduction in the total cost and cut down $1.71 \times 10^5$ lbs CO$_2$ emission. The results also show that the communication between the ISO and customers is vital for improving the performance of demand side load management where DA-SOLM is able to reduce $36.7\%$ of the total cost and crush certain LMP peaks. The sensitivity analysis of carbon emission prices over the emission reduction shows that raising the carbon emission price in the future would make a significant impact on reducing greenhouse gas emissions.
CHAPTER 4: REVENUE ANALYSIS AND OPTIMAL PLACEMENT OF STATIONARY AND TRANSPORTABLE ENERGY STORAGE SYSTEMS FOR MARKET PARTICIPANTS

4.1: Motivation and Contribution

A comprehensive revenue analysis of BESSs is critical for market participants to install such systems in a market-based power system. Taking PJM as an example, this chapter carries out a thorough revenue analysis for the entire system. Considering LMPs, regulation MCP and regulation signals, and battery degradation, the potential revenues for BESSs to participate in both the energy and regulation markets are estimated for every node throughout the entire power market, and the features of profit differences among different nodes are characterized. The potential revenue in the energy and regulation markets is analyzed from the temporal and spatial domains by considering transportable BESSs to grasp more profitable opportunities. According to the characteristics discovered at the most profitable node, an optimal placement algorithm is proposed in this chapter for market participants to find the profitable locations for better installing and operating the BESS, particularly when it is transportable.

4.2: Participation Models in Energy and Frequency Regulation Markets

In order to find out the nodes where the BESS projects are most likely to have better revenues, this study focuses on analyzing the BESS in PJM. The following models for participating in PJM energy and frequency regulation markets are used to estimate the potential profit at each commercial pricing node which is valid for auction.

4.2.1: Credit in Energy Market

The arbitrage credit of energy storage from the energy market during a period of time $T$ is calculated by [4.1]. The time interval is assumed to be one hour throughout this
chapter unless otherwise specified.

\[
Credit_E = \sum_{t=1}^{T} P_{t}^{E,\text{dis}} \times LMP_{t} - \sum_{t=1}^{T} P_{t}^{E,\text{ch}} \times LMP_{t}
\]  

(4.1)

where \( P_{t}^{E,\text{ch}} \) and \( P_{t}^{E,\text{dis}} \) represent the charging and discharging power in the energy market at hour \( t \); \( LMP_{t} \) is the LMP at hour \( t \).

4.2.2: Credit in Regulation Market

In addition to the energy market, energy storage can participate in the regulation market, or both the markets simultaneously. Since this study focuses on analyzing the electric storage projects competing in PJM, the credit calculation method of participating in the frequency regulation market follows the PJM’s rules. According to the section about regulation credits in the PJM manual [120], the regulation remuneration is obtained from two parts: capability and performance, as shown in (4.2).

\[
Credit_{R} = Credit_{\text{cap}} + Credit_{\text{perf}}
\]  

(4.2)

Given an offered capacity \( P_{t}^{R} \), Regulation Market Capability Clearing Price \( RMCCP_{t} \), and performance score \( \rho_{t} \), the \( Credit_{\text{cap}} \) can be calculated by (4.3). Performance score \( \rho_{t} \) is used to evaluate how well the resource is following the regulation signal [121].

\[
Credit_{\text{cap}} = \sum_{t=1}^{T} P_{t}^{R} \times RMCCP_{t} \times \rho_{t}
\]  

(4.3)

In PJM’s frequency regulation market, there are two types of regulation signals generated every two seconds for participants to follow. One is the Regulation D signal (RegD), which is a fast and dynamic signal for fast-responding resources such as a BESS. The other one, called Regulation A signal (RegA), is a slower signal for conventional resources like hydropower [121].
To show the difference between these two types of signals, a series of 40-minute regulation test signals provided by PJM \cite{122} are plotted in Fig. 4.1. According to the plots of these two signals, it is clear to observe that the \textit{RegD} has a much higher volatility than \textit{RegA}, which means \textit{RegD} is more suitable for the fast ramping energy resources to follow. Meanwhile, compared to the \textit{RegA} resources, which typically do not have a duration limit, the \textit{RegD} resources are expected to have a short period duration for Area Control Error (ACE) control \cite{123}. Hence, the BESS discussed in this chapter is only considered to submit the offer for following \textit{RegD} signals in the regulation market. In the frequency regulation market, the performance credit $\text{Credit}_{\text{perf}}$ is calculated by \begin{equation} Credit_{\text{perf}} = \sum_{t=1}^{T} P_t^R \times \text{RMPCP}_t \times \beta_t \times \rho_t \end{equation} \label{eq:4.4} In the equation, $\text{RMPCP}_t$ is the Regulation Market Performance Clearing Price at hour $t$ and the Mileage Ratio $\beta_t$ of \textit{RegD} over \textit{RegA} can be obtained by \begin{equation} \beta_t \end{equation}. The mileage of a regulation signal indicates the total movements of that signal in a given interval. The mileages of \textit{RegD} and \textit{RegA} ($\text{Mileage}_{t}^{\text{RegD}}$ and $\text{Mileage}_{t}^{\text{RegA}}$) are given in \begin{equation} \text{Mileage}_{t}^{\text{RegD}} \end{equation} and \begin{equation} \text{Mileage}_{t}^{\text{RegA}} \end{equation}, respectively. $n$ in those two equations is the number of time intervals during hour $t$. For example, to calculate the mileage of a 2-second signal in one hour, $n$ is 1800.
\( \beta_t = \frac{\text{Mileage}_{t}^{\text{RegD}}}{\text{Mileage}_{t}^{\text{RegA}}} \) \hspace{1cm} (4.5)

\[ \text{Mileage}_{t}^{\text{RegD}} = \sum_{i=0}^{n} |\text{RegD}_i - \text{RegD}_{i-1}| \] \hspace{1cm} (4.6)

\[ \text{Mileage}_{t}^{\text{RegA}} = \sum_{i=0}^{n} |\text{RegA}_i - \text{RegA}_{i-1}| \] \hspace{1cm} (4.7)

### 4.2.3: Cost of Storage Degradation

Based on [124, 125], the degradation cost of a BESS is assumed to be a linear function of BESS output as given in (4.8) and (4.9):

\[ \text{Cost}_D = \sum_{t=1}^{24} (P_{t}^{\text{ch}} \times \eta_c + P_{t}^{\text{dis}} \times \eta_d^{-1}) \times \text{DEG}_{\text{rate}} \] \hspace{1cm} (4.8)

\[ \text{DEG}_{\text{rate}} = \left( \upsilon \times \pi_{\text{ES}} \times 0.5 \right) / (1 - \sigma_{\text{EOL}}) \] \hspace{1cm} (4.9)

where \( P_{t}^{\text{ch}} \) and \( P_{t}^{\text{dis}} \) are the overall charging and discharging power at hour \( t \), respectively; the \( \eta_c \) and \( \eta_d \) are the corresponding charging and discharging efficiency of BESS; \( \text{DEG}_{\text{rate}} \) represents the cost of each MWh charge/discharge of the BESS; \( \upsilon \) represents the degradation speed of energy storage; \( \pi_{\text{ES}} \) stands for the cost of energy storage; \( \sigma_{\text{EOL}} \) means the storage state at the end of life (EOL).

### 4.2.4: Participation of Energy Market

When the electric storage only participates in the energy market, the objective of energy arbitrage is to maximize the revenue in each day at a pricing node described by (4.10). For the sake of calculating the maximum profit, the forecast of electricity price in a day is assumed to be perfect and the participation model is expected to be price-taking. Therefore, the model implemented for participating in the energy market can be described as:

\[ \text{Max}(\text{Credit}_E - \text{Cost}_D) \] \hspace{1cm} (4.10)

s.t. (4.1), (4.8), (4.9), (4.11) - (4.18).
\[ P_t^{ch} = P_t^{E,ch} \] (4.11)
\[ P_t^{dis} = P_t^{E,dis} \] (4.12)
\[ 0 \leq P_t^{E,ch} \leq P_{max} \] (4.13)
\[ 0 \leq P_t^{E,dis} \leq P_{max} \] (4.14)
\[ P_t^{E,ch} \times P_t^{E,dis} = 0 \] (4.15)
\[ 0 \leq S_t \leq 100\% \] (4.16)
\[ S_t = S_{t-1} + (P_t^{ch} \times \eta_c - P_t^{dis} \times \eta_d^{-1})/E_{max} \times 100\% \] (4.17)
\[ S_0 = S_{24} \] (4.18)

The \( P_t^{ch} \) and \( P_t^{dis} \) in the degradation cost are determined by (4.11) and (4.12). \( P_{max} \) and \( E_{max} \) are the power rating and energy capacity of the storage device. \( S_t \) represents the state of charge (SOC) at hour \( t \). The operation limits of the storage device are set by (4.13)-(4.16). \( S_0 \) and \( S_{24} \) are the SOC at the beginning \( t = 0 \) and the end \( t = 24 \) of each day. \( S_t \) is obtained from the previous SOC \( S_{t-1} \) and the hourly output as given in (4.17). (4.18) is implemented to ensure the conditions of daily optimization to be consistent. (4.15) is used to force the optimization model not to make charging and discharging offers at the same time since a BESS cannot be physically charging and discharging simultaneously. Meanwhile, (4.15) is essential for the market participation model to deal with negative LMPs.

4.2.5: Participation of Energy and Frequency Regulation Markets

When the BESS participates in the energy and frequency regulation markets at the same time, the objective function is formulated as (4.19) to achieve the maximum daily profit at a node.

\[ \text{Max}(Credit_E + Credit_R - Cost_D) \] (4.19)
Different from (4.11) and (4.12) for the participation only in energy market, the $P_{t}^{ch}$ and $P_{t}^{dis}$ of the degradation cost for both the energy and regulation markets are calculated by (4.20) and (4.21), where $RegD_{t}^{up}$ and $RegD_{t}^{down}$ are the absolute values of hourly accumulations of the regulation up and down signals. In PJM, a unit offering in the regulation market is required to be able to provide the same amount of positive capacity and negative capacity. The constraints of output offers are given in (4.22)-(4.24).

\[
\begin{align*}
    P_{t}^{ch} &= P_{t}^{E, ch} + P_{t}^{R} \times RegD_{t}^{down} \\
    P_{t}^{dis} &= P_{t}^{E, dis} + P_{t}^{R} \times RegD_{t}^{up}
\end{align*}
\]

\[
\begin{align*}
    0 &\leq P_{t}^{R} \leq P_{\text{max}} \\
    0 &\leq P_{t}^{E, ch} + P_{t}^{R} \leq P_{\text{max}} \\
    0 &\leq P_{t}^{E, dis} + P_{t}^{R} \leq P_{\text{max}}
\end{align*}
\]

Implementing the introduced participation models with the real market data, such as LMP, RMCCP, RMPCP, and $RegD$, the maximum daily revenue at each node can be obtained and the nodes with the best profitability in different periods are able to be identified, as discussed in Section 4.3.

### 4.3: Revenue Analysis of BESSs

#### 4.3.1: Data Preparation

In PJM, there are over 12,000 pricing nodes, and about 7000 of them are valid for auction [126]. This study focuses on analyzing the potential revenue of BESS at the valid nodes for auction. According to the market participation models, the LMP, regulation signals, regulation mileage, RMCCP, and RMPCP are required for estimating the potential revenue of participating in the energy and regulation markets. The corresponding 12-month market data in 2018 are collected via PJM Data Miner 2 [127]. Meanwhile, to
investigate the potential profit of transportable BESS in different seasons, 12 months are separated and grouped into four seasons as spring: 3, 4, 5, summer: 6, 7, 8, autumns: 9, 10, 11, and winter: 12, 1, 2.

In the energy market, LMP reflects the value of electricity at different locations or nodes, which is defined as the cost of supplying the next 1 MW at the node. LMP is a combination of three components: energy component, congestion component, and loss component. Since these components are affected by load, fuel cost, system topology, etc., the LMPs can have different characteristics and values at different periods and locations, which can cause significant changes in the BESS revenue. As an example, the boxplots of hourly LMPs in 2018 spring and summer are plotted in Fig. 4.2. For better clarification, the outliers are not displayed in the boxplots. Fig. 4.2 clearly shows the 24-hour LMPs of PJM RTO in summer have higher volatility than spring, which indicates the storage project can have more energy arbitrage opportunities in summer.

![Image](image_url)

**Figure 4.2:** Comparison of LMPs in different seasons.

In addition to the features in different seasons, the LMPs at different locations can have significant differences, as well. In Fig. 4.3, the LMP boxplots of two nodes:
BAYVIEW and 9 JOLIET, in the summer are plotted. The hourly averages of the LMPs at node BAYVIEW are found to be higher than the LMPs at node 9 JOLIET. Meanwhile, the boxplots indicate the difference of price uncertainties at these nodes, even though they are collected from the same market period.

Figure 4.3: Comparison of LMPs at two different locations in PJM during the summer of 2018.

Unlike the energy market, the regulation market is a pool-based market, which means there is only one market clearing price for each MW supplied in the regulation market. Hence, all resource owners in the regulation market are credited by the same price for providing the scheduled regulation output in PJM. To calculate the potential profit of participating in the regulation market, the hourly LMP, RMCCP, RMPCP, RegA mileage, and RegD mileage in 2018 are collected from PJM Data Miner 2 [127]. Meanwhile, to determine the hourly regulation output and the hourly SOC, the 2-second regulation signals of 2018 provided by PJM [129] are also collected and aggregated to the hourly regulation up signal and regulation down signal.

To estimate the potential revenue, several key parameters are listed in Table 4.1 and
utilized to characterize a 10 MW/10 MWh BESS studied in this chapter.

<table>
<thead>
<tr>
<th>$P_{max}$</th>
<th>$E_{max}$</th>
<th>$\eta_c$</th>
<th>$\eta_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 MW</td>
<td>10 MWh</td>
<td>0.95</td>
<td>0.95</td>
</tr>
</tbody>
</table>

This example BESS can achieve charge/discharge ($\eta_c/\eta_d$) efficiency at 95%. To estimate the degradation cost based on the CostD, the degradation speed of energy storage $v$, the cost of energy storage $\pi_{ES}$, and the storage state at the end of life $\sigma_{EOL}$ are set as $3 \times 10^{-5}$, $1 \times 10^5$ $$/MWh$, and 0.8, respectively, [124, 125]. In the case study, the initial SOC $S_0$ of each day is set to be 50%. Given these BESS parameters, the degradation rate $DEG_{rate}$ in (4.9) can be obtained as $7.5$ $$/MWh$. This degradation cost indicates that the charging action is worthless for the battery when the profit of 1 MWh charge cycle is less than $15.

### 4.3.2: Potential Revenue of Stationary BESS

The first case studied in this chapter is to place the energy storage at one location for an entire year (1L1Y), which is the most common setup investigated in the previous studies [77, 86]. The proposed energy storage location algorithm finds a profitable location in the system to build the BESS and obtains the credit settled by the corresponding LMP and regulation MCP for a long period. In the view of this case, the 1L1Y utilizes the real market data, including the RTLMP, RMCCP, RMPCP, and the regulation signal, to estimate the potential annual revenues at different settlement price nodes in PJM, which provides the valuable information for implementing storage projects from the perspective
Given the 2018 market data collected from PJM, the maximum potential revenue and the corresponding generation outputs of each node in PJM can be obtained by the optimization models proposed in Section 4.2. As shown in Fig. 4.4, the percentiles of all nodes’ potential revenues are grouped and mapped into five different colors, which indicate the percentiles of 0-20%, 20-40%, 40-60%, 60-80%, and 80-100%, respectively. According to the distributions of the colors on the map, the nodes of the top 20% (i.e., 80-100%) annual revenue are found to concentrate in some regions. In other words, the nodes, which have high potential revenue for stationary BESS, are not randomly distributed in the power market. It is feasible for market participants to locate some sites to enhance their competitiveness in the power market. Furthermore, the annual revenues of each node in the five percentile intervals are plotted in Fig. 4.5, which clearly indicates that there is a sharp increase in revenue for the nodes falling in the 80-100% percentile interval. Notably, the annual revenue at the most lucrative node, named ROCKWLKN69KV LS-EGRET (ROCK), is $2,783,529.66, which is $529,389.02 higher than the annual revenue of market participants.

Figure 4.4: Annual revenue percentiles in PJM.
at the least lucrative node, called 21 KINCA20KV KN-1 (KINCA). In other words, the
BESS at ROCK could earn 23.5% more profit than the BESS at KINCA in 2018.

Figure 4.5: Annual revenues by percentiles.

To gain a more comprehensive understanding of the factors causing the revenue dis-
pparity, the differences in daily revenues in 2018 between these two nodes are calculated
and plotted in Fig. 4.6. It can be seen that most of the daily revenue differences are
low and the significant revenue differences are found to concentrate in May, October, and
November. Moreover, when the daily revenue differences are sorted into ascending order,
the accumulation of the sorted daily revenue differences can be obtained. According to
the accumulated revenue difference, 80% of the total annual revenue difference is found
to be contributed by the top 19.45% of the daily revenue difference. This means the
potential profit at node ROCK does not overwhelm the potential profit at node KINCA
on every single day. Most of the annual revenue difference is just accumulated in a short
period.
Furthermore, to obtain more details about the revenue difference, the hourly LMPs of these two nodes on the days of the top 19.45% daily revenue differences are extracted. The boxplots of the extracted 24-hour LMPs in Fig. 4.7 clearly shows the hourly LMPs at ROCK have much higher volatility than the LMPs at KINCA. At node ROCK, the LMPs from 6:00 to 9:00, and from 13:00 to 23:00, are distributed on a larger price range than the other hours, which enables the BESS to have more arbitrage opportunities in the energy market. Notably, a large amount of negative prices is found at node ROCK, and the negative rates make the charging action become profitable.
According to the optimization results on 9/27, the battery at node ROCK can earn $17,131.46 while the battery at node KINCA can just earn $5,922.84. There is an $11,208.61 revenue difference, which is more than 189% of the daily revenue at node KINCA on 9/27. It is the most significant difference in daily revenue between these two nodes in 2018. To better illustrate the energy storage behaviors, their optimized generation outputs on 9/27 are extracted and displayed in Figs. 4.8 and 4.9. Given the optimized outputs, their hourly SOCs on 9/27 can be calculated by (4.17). As a reference, the corresponding hourly state of energy (SOE) derived from SOC is plotted with the LMPs in Figs. 4.8 and 4.9 as well.

Figure 4.8: Output actions of node ROCK on 9/27.

In Figs. 4.8 and 4.9 the left y-axis indicates the hourly output power and the SOE of the BESS. The y-axis on the right shows the prices of LMP at the node. The hour 24 means the first hour of the next day. The bars of different colors at each hour display charging power, discharging power, and regulation output, respectively. The solid line and dash line show the hourly SOEs and LMPs in 24 hours, correspondingly.

As shown in Fig. 4.8, the LMP at node ROCK is very volatile, and the peak prices on 9/27/2018 happen at 3:00, 6:00, 11:00, and 23:00 when the discharging actions are
also taken for arbitraging in the energy market. When the LMPs are low, the charging actions are clearly found at 5:00, 8:00, 13:00, and 19:00.

![Output actions of node KINCA on 9/27.](image)

In Fig. 4.9, the standard deviation of 24-hour LMPs is $4.67/MWh. Compared to the LMPs in Fig. 4.8 with $128.11/MWh standard deviation, the LMPs at KINCA is less volatile, which indicates fewer arbitrage opportunities. Besides, the degradation cost implemented in the market participation models forces the battery not charging or discharging any single MWh when the arbitrage revenue cannot cover the degradation cost. Due to the degradation cost, it can be seen in Fig. 4.9 that most of the output power at node KINCA is offered to the regulation market. From 12:00 to 15:00, the BESS at KINCA does not even provide full output capacity since the earnings from energy and regulation markets during these hours are not able to cover the degradation cost.

Comparing the only discharging actions at node KINCA to the multiple charging and discharging actions at node ROCK, the factor that leads to an $11,208.61 daily revenue difference on 9/27 can be easily discovered. It is the volatility of LMP at node ROCK, which makes the energy storage system produce more revenue at the node. As aforementioned, the LMP is a combination of three components. Since there is only one
energy component for all the buses in the system and the loss component is relatively small compared to the other two components [6], the congestion component becomes the dominating factor that influences the potential revenue for the BESS of participating in the energy and regulation markets. In other words, to achieve more revenue in the market, the storage project should search for a place having a highly volatile congestion component.

An interesting point that can be observed in Fig. 4.9 is the advantage of participating in both the energy and regulation markets. According to the optimization results, the battery at node KINCA can still make $5,922.84 profit on 9/27, which means the battery can always guarantee a certain amount of profit by participating in the regulation market even though there is a lack of arbitrage opportunity in the energy market.

4.3.3: Potential Revenue of Transportable BESS

As aforementioned, a significant part of the annual revenue difference is just accumulated in a short period. If the storage system is on a transportable platform (such as trucks) that can be installed and commissioned at different places in different seasons or months, it is expected to capture more revenue opportunities in the market and deal with the risks of market changes. A transportable BESS can also be formed by aggregating a fleet of electric vehicles that have the vehicle to grid (V2G) capabilities. To investigate the potential revenue of the transportable BESS, the cases of settling the BESS at 4 Locations in 4 Seasons (4L4S), and 12 Locations in 12 Months (12L12M) are studied in the following.
Based on the annual motor carrier operations cost report [130] provided by American Transportation Research Institute (ATRI), the average carrier cost, including driver costs, fuel, insurance, permits, tolls, etc., is set to $66.65/hour. Similar to [90], the size of BESS on each truck is assumed to be 2.5MW/2.5MWh. Therefore, the 10MW/10MWh BESS analyzed in this chapter can be carried by four trucks when transported. Meanwhile, the maximum transportation distance in PJM is expected to be 900 miles, which can cover the distance from Chicago, IL, to Virginia Beach, VA. The average speed of the truck is assumed to be 60 miles/hour, so the maximum transportation time is about 15 hours and transportation cost is $3,999. In addition to the cost during transportation, it is assumed that the wage of a professional electrician is $65/hour and the integration work will take 10 electricians 4 hours to complete. Then the labor cost for integration is estimated at $3,000, including the electrician cost, insurance, etc. Meanwhile, according to the PJM transmission tariff [131,132], the cost for transmission interconnection of a 10MW/10MWh BESS is assumed to be $2,200, which includes the costs for application,
feasibility study, and non-refundable deposit. Hence, the maximum cost for each re-location of the 10MW/10MWh BESS is $9,199. It should be noted that the above re-location cost does not apply to a virtual transportable BESS consisting of aggregated EVs.

As aforementioned, the months in 2018 are assigned to four seasons base on the regulation requirement of PJM. According to the optimization results, all nodes’ daily revenues can be calculated and aggregated to the seasonal revenues. Similar to the analysis of annual revenue percentiles, seasonal revenue percentiles are grouped into five intervals and mapped to five different colors. As shown in Fig. 4.10, the seasonal revenue percentiles are plotted by spring, summer, autumn, and winter, correspondingly. According to the contour maps in different seasons, it can be observed that nodes having the top 20% revenue distribute on different areas in four seasons, which reveals a positive sign for the transportable BESS. The sign indicates the portable battery project can earn more profit by moving the storage systems from one comparably profitable place to another in different seasons. In other words, based on the seasonal revenue analysis for the whole system, the mobile BESS can always chase a better profitable node through the entire market in different seasons to achieve a higher potential of earning.

More specifically, the most profitable nodes and their seasonal earnings are obtained based on the daily maximum revenues, as listed in Table 4.2. As a comparison, the seasonal revenues of 1L1Y are shown in Table 4.2 as well. Comparing the seasonal revenues of 1L1Y with the revenues of 4L4S, the seasonal profits of 4L4S are found to surpass the seasonal profits of 1L1Y by $27,819.10, $33,706.69, and $117,746.47 in spring, summer, and winter, respectively. In autumn, the nodes having the best potential revenue of 1L1Y and 4L4S are the same. Therefore, if the BESS is transportable, the best possible
Table 4.2: Seasonal Revenues of 1L1Y and 4L4S

<table>
<thead>
<tr>
<th>Season</th>
<th>Node</th>
<th>1L1Y Profit</th>
<th>Node</th>
<th>4L4S Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>ROCK</td>
<td>$688,138.22</td>
<td>STRASBUR</td>
<td>$715,957.32</td>
</tr>
<tr>
<td>Summer</td>
<td>ROCK</td>
<td>$488,798.10</td>
<td>CEDARCRE</td>
<td>$522,504.79</td>
</tr>
<tr>
<td>Autumn</td>
<td>ROCK</td>
<td>$769,195.04</td>
<td>ROCK</td>
<td>$769,195.04</td>
</tr>
<tr>
<td>Winter</td>
<td>ROCK</td>
<td>$837,398.29</td>
<td>GARDNERS</td>
<td>$955,144.76</td>
</tr>
</tbody>
</table>

earning for 4L4S is $2,962,801.91, which is $179,272.25 higher than the annual revenue at the best fixed location (i.e. node ROCK). Even considering the $36,796 transportation cost for 4L4S, the transportable BESS is still more profitable than stationary BESS in 1L1Y. Furthermore, according to the price node mapping by state, zip code, and transmission zone provided by PJM [133], these nodes are roughly marked on the contour maps as shown in Fig. 4.10. It can be found that all the nodes having the best seasonal revenues are in the eastern region of PJM. In addition to the more profitable chances, these nearby nodes can lower the moving and relocation cost for mobile BESSs.

In addition to the 4L4S, the 12L12M scheme is studied for transportable BESS as well. In the 12L12M case, the BESS are transported every month to maximize the revenue in the energy and regulation markets. According to the monthly revenue listed in Table 4.3, the potential revenue of 12L12M is $3,213,195.73. Even considering with $110,388 transportation cost in a year, the maximum profit of 12L12M is $176,801.82 higher than the revenue of 4L4S and $319,278.07 more than the revenue of 1L1Y.

4.4: Optimal Placement of BESSs

The previous revenue analysis is based on historical data. However, the most profitable nodes identified using historical data may not still be the appropriate locations for BESSs
in the future. Hence, an algorithm is needed for placing and operating the BESSs. This algorithm is particularly important for the transportable BESSs since they can move to different locations in the next month or season for more profitable opportunities. To achieve this goal, an optimal placement algorithm to identify highly profitable locations in the power market based on the time series prediction is proposed in this section. Without loss of generality, the algorithm that aims at the optimal placement for the next month can be readily extended to other future times, such as next season or year.

4.4.1: Prediction of LMP’s Volatility

According to the characteristics discovered at the most profitable node in the previous section, the revenue difference among the nodes is found to be caused by the volatility of LMP in the energy market. Hence, to identify the nodes which have more profitable opportunities in the next month, the prediction of LMP volatility becomes a critical task. To accomplish this task, a standard ARIMA model \[ \phi_p(B) \nabla^d y_t = \mu + \theta_q(B) \varepsilon_t \] (4.25) is used. An ARIMA model can be described by (4.25). 

Table 4.3: Revenues of 12L12M

<table>
<thead>
<tr>
<th>Season</th>
<th>Spring</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Profit (k)</td>
<td>190.85</td>
<td>262.09</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Profit (k)</td>
<td>184.96</td>
<td>178.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Season</th>
<th>Autumn</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Profit (k)</td>
<td>218.41</td>
<td>317.28</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Profit (k)</td>
<td>220.29</td>
<td>726.09</td>
</tr>
</tbody>
</table>
where $B$ is the backward shift operator; $p$ is the auto-regression (AR) order, which determines how many past values are used for regression; $d$ is the differencing order, which is often used for making the training series stationary; $q$ is the moving-average (MA) order to determine how many previous error terms $\varepsilon_t$ should be considered. The error terms $\varepsilon_t$ are generally assumed to be the independent and identically distributed noise with zero mean and finite variance.

Given the hourly LMP, the monthly standard deviations of LMPs at each node can be calculated. For reducing the price volatility caused by the system energy component and formulating a stationary series for the ARIMA model, the time series $y_{i,t}$ is obtained by a natural logarithm transformation of a differential series, as shown in (4.26).

$$y_{i,t} = \log(\sigma_{i,t} - \sigma_{PJM,t} + c)$$

(4.26)

where $\sigma_{i,t}$ is the standard deviation of LMPs at node $i$ in month $t$; $\sigma_{PJM,t}$ represents the standard deviation of system’s LMPs in month $t$; $c$ is a positive constant offset to guarantee the logarithm transformation; $y_{i,t}$ is the transformed time series in month $t$, which is used by the ARIMA model to forecast the LMP volatility at node $i$. In this study, the LMPs from 1/1/2016 to 12/31/2018 of each node are used for training the ARIMA model to predict the corresponding standard deviation in the next month. The market data between 1/1/2019 and 6/30/2019 are used to validate the proposed algorithm.

### 4.4.2: Clustering of Pricing Nodes

In addition to the estimation of LMPs’ volatility, the clustering of nodes is also important for the optimal BESS placement. After the nodes are grouped according to the LMPs’ features in the past month, the algorithm can place the transportable BESSs at the nodes in different clusters for the purpose of reducing the risk and seeking for more profitable opportunities throughout the market. The K-means [137], as one of the
most widely used clustering algorithms, is implemented to cluster the nodes into different
groups via minimizing the objective function described in (4.27):

\[ \text{Min} \left( \sum_{k=1}^{K} \sum_{x \in G_k} \| x - \mu_k \|^2 \right) \quad (4.27) \]

where \( K \) is the total number of clusters, which is determined by the elbow method \cite{138} in
this study; \( G_k \) represents the \( k \)-th group of the clustered data \( x \); \( \mu_k \) is the mean, also called
centroid, of \( x \) in \( G_k \). In other words, the objective of K-means is to minimize the sum
of the Euclidean distance between the data points and the centroid of the corresponding
group.

In order to deal with high-dimensional hourly LMP data for K-means clustering,
the hourly LMP in the previous month is grouped by each day and a daily standard
deviation of the LMP is calculated first. Given the daily standard deviations of each
node, the Principle Component Analysis (PCA) \cite{139} is implemented to further reduce
the dimension of the LMP data. The PCA is a linear transformation to represent the
original data by a set of orthogonal variables. Via PCA, the original data can be described
with a low dimensional subspace with keeping much of the variance in the dataset. In
this study, the \( x \) in (4.27) is the PCA components of normalized daily standard deviation
of the LMPs at each node. With the implementation of PCA and K-means for the LMPs
in the previous month, the clusters of different nodes can be obtained.

4.4.3: Optimal Placement Algorithm

Given the prediction of LMP’s monthly standard deviations and the clusters of pricing
nodes, an algorithm shown in (4.28)-(4.31) is designed to assist the market participants
with placing the BESSs for the next month. The objective of this algorithm is to maximize
the sum of the predicted LMP’s volatility of the selected nodes, as shown in (4.28).

\[
Max \left( \sum_{i=1}^{N} \delta_i \sigma'_{i,t+1} \right) \tag{4.28}
\]

s.t. (4.29)-(4.31).

\[
\sum_{i=1}^{N} \delta_i = N_{ES} \tag{4.29}
\]

\[
\sum_{i \in G_k} \delta_i \leq N_{C_{lu}} \tag{4.30}
\]

\[
\delta_i \delta_j d_{i,j}^{Geo} \geq D_{Geo}, \ \forall i < j \tag{4.31}
\]

where \(\sigma'_{i,t+1}\) is the forecast monthly standard deviation of the LMPs at node \(i\); \(\delta_i\) is a binary decision variable; when \(\delta_i = 1\), it means the node \(i\) is selected for placing transportable BESS in the next month. (4.29) indicates that the algorithm needs to find \(N_{ES}\) locations for placing the BESSs; (4.30) enforces that no more than \(N_{C_{lu}}\) BESSs can be placed in the same cluster \(G_k\). Since the pricing nodes in the same cluster could have the similar characteristics of LMPs, (4.30) enables the market participants to place the BESSs in different clusters for the risk diversification. In (4.31), the restriction about geographical distance is implemented. Given the geographical information of the pricing nodes, (4.31) enforces that the geographical distance \(d_{i,j}^{Geo}\) between nodes \(i\) and \(j\) must be larger than \(D_{Geo}\), which is also beneficial for the risk management of market participants. It is worth pointing out that, when more system information is available to the market participants, the geographical distance implemented in this study can be extended to other types of distance, such as the number connections between the nodes or electrical distance. Furthermore, (4.31) can be transformed to linear constraints as shown in (4.32)-(4.34), where \(\delta_{i,j}\) is a binary variable and \(M\) is a large positive number. Therefore, the optimal placement algorithm with the objective function (4.28), subject to
\[ \delta_{i,j} \geq \delta_i + \delta_j - 1, \quad \forall i < j \quad (4.32) \]
\[ \delta_{i,j} \leq \delta_i, \quad \delta_{i,j} \leq \delta_j, \quad \forall i < j \quad (4.33) \]
\[ \delta_{i,j} \delta_{i,j}^{Geo} + (1 - \delta_{i,j}) M \geq D_{Geo}, \quad \forall i < j \quad (4.34) \]

4.4.4: Performance of the Proposed Algorithm

In this study, it is assumed that the market participant needs to place five 10MW/10MWh BESSs in the market for the next month. \( N_{Clu} \) is set to 1, which means no more than 1 BESS can be placed in the same cluster of nodes. The minimum distance \( D_{Geo} \) between two BESSs is chosen to be 50 miles. With the implementation of the proposed optimal placement algorithm for the first 6 months in 2019, the selected nodes for placing the BESSs in each month can be obtained. Meanwhile, the market data, such as RTLMP, RMCCP, RMPCP, and the regulation signals, from 1/1/2019 to 6/31/2019, are collected to verify the proposed algorithm in finding the optimal placement for the 5 BESSs under study. For comparison, a base case for placing the BESSs in the market based on the historical revenue is included, where the market participants use the 5 most profitable nodes in the current month as the placements for the next month.

Table 4.4: Performances of the Base Case and the Proposed Algorithm

<table>
<thead>
<tr>
<th>Mo.</th>
<th>Base</th>
<th>Proposed</th>
<th>Mo.</th>
<th>Base</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$550650.97</td>
<td>$629201.43</td>
<td>4</td>
<td>$560359.15</td>
<td>$618180.98</td>
</tr>
<tr>
<td>2</td>
<td>$512690.76</td>
<td>$558914.55</td>
<td>5</td>
<td>$434092.99</td>
<td>$509234.69</td>
</tr>
<tr>
<td>3</td>
<td>$673090.96</td>
<td>$733176.52</td>
<td>6</td>
<td>$592339.61</td>
<td>$475226.81</td>
</tr>
</tbody>
</table>

As shown in Table 4.4, the proposed optimal placement algorithm can outperform the base case in 5 out of 6 months. In the first 5 months, the proposed algorithm can
earn 11.4% more revenue than the base case. Even though the base case beats the proposed algorithm by $117,112.80 in the last month, the optimal placement algorithm can still make $194,410.55 more profit than the base case. In other words, the proposed optimal placement algorithm is capable of benefiting market participant by identifying the potentially profitable sites to deploy the BESSs.

4.5: Summary

This chapter carried out a comprehensive and system-wide analysis of potential revenues of BESSs in a market-based power system. The PJM market has been used as an example for the study, and the features of PJM market data of different locations and seasons have been characterized for the analyses of potential revenue. According to the geographic distributions of energy storage revenue, a set of candidate sites have been discovered for installing a BESS in the market to achieve the highest revenue of BESS. By comparing the best and worst profitable nodes, LMP volatility has been found to be the dominating factor that influences the potential revenue. Meanwhile, most of the annual revenue difference was observed to be accumulated in a short period, which indicates the potential opportunity of transportable BESS for capturing more profits from different locations in several periods. Through the comprehensive study of different seasons and months at all the nodes in the whole system, the results indicate a transportable BESS is capable of yielding a higher revenue than a corresponding stationary BESS in the energy and regulation markets. Furthermore, an algorithm has been developed and verified for optimal placement of BESSs. The results reported and the method developed in this chapter can be helpful to the energy storage market participants for investing and managing BESS projects in the energy and frequency regulation markets.
CHAPTER 5: A TWO-STAGE OPTIMAL BIDDING ALGORITHM FOR INCENTIVE-BASED AGGREGATION OF ELECTRIC VEHICLES IN WORKPLACE PARKING LOTS

5.1: Motivation and Contribution

In this chapter, a two-stage optimal bidding algorithm for an aggregator of EVs in workplace parking lots to participate in RT energy and regulation markets, coordinated with the incentive to adjust the EVs’ behaviors, is designed to benefit both EV aggregator and EV owners. To secure more resources for the aggregator when needed and address the uncertainty of EV owners’ behaviors, the EV’s responses to the incentives are modeled and considered in the proposed two-stage optimal bidding algorithm for the EV aggregator to participate in both RT energy and regulation markets. In the first stage, a DA planning model based on the predictions of the seasonal ARIMA model is proposed for the EV aggregator to determine if the EV aggregation program should be activated for the next day and what level of the incentive should be broadcasted to optimally change the EVs’ behaviors. Based on the incentive determined in the previous stage, a RT optimal bidding strategy is designed for the EV aggregator to maximize the profit in the RT energy and regulation markets, complying with the energy requirements of EVs.

5.2: Modeling of EV Owners’ Behaviors

5.2.1: Initial Behaviors of EVs

The EV aggregator in this study is designed to provide two types of incentives to the EVs. The first one is a fixed reward $\hat{\pi}$. Once the EV aggregation program is activated, there is a fixed reward provided for each EV owner. This reward is designed for attracting EV owners to participate in the EV aggregation program. The other one is the incentive $\pi$ designed to attract the EV owners to change their behaviors for providing more resources for the aggregator to participate in the energy and regulation markets when needed.

In order to study the responses of an individual EV owner to different incentives, the
truncated Gaussian distributions are used to model the original EV behaviors, including the arrival time, departure time, SOC when arriving, and desired SOC when leaving. For example, the distribution of arrival time can be modeled by using a truncated Gaussian distribution ($f_{TG}$):

$$t_{arv} = f_{TG}(x; \mu_{t_{arv}}, \sigma_{t_{arv}}, t_{min}^{t_{arv}}, t_{max}^{t_{arv}})$$  \hspace{1cm} (5.1)

Based on [38, 93, 94], the parameters of the distributions for modeling the EV owners’ behaviors are listed in Table 5.1. For the arrival and departure times, the mean values indicate that people would mostly like to go to work at around 8:30 AM and leave at around 5:30 PM; min values show the earliest time when people arrive/depart and max values represent the latest time of arrival/departure. Similarly, the parameters of SOC at arrival/departure represent the randomness of the EV owners’ behaviors in charging their EVs. Given the parameters listed in Table 5.1, the initial behaviors of each EV owner can be modeled by the corresponding truncated Gaussian distributions. Since the EV aggregator is proposed to participate in the RT markets on the hourly basis, the EV arriving between 8:00 and 8:59 is considered as arriving within the hour of 8.
5.2.2: EV’s Responses to Incentives

With the initial behaviors of each EV owner, the step function for modeling the EV owner’s response to different levels of incentives can be built. As formulated in (5.2), the arrival time $t_{arv}^i$ of EV $i$ with different incentives are described with three intervals. It shows the EV $i$ initially arrives at hour 9 and is willing to come at hour 7 when the incentive $\pi$ is over $\pi_1^i$. Similarly, the step response functions of departure time $t_{dep}^i$, arrival SOC $SOC_{arv}^i$ and departure SOC $SOC_{dep}^i$ for EV $i$ can be developed. In these step functions, the minimum and maximum values of the different responses are the same as the initial behavioral parameters as shown in Table 5.1. The advantage of estimating the response behavior individually is the EV owners can be characterized with the diversity of response elasticity by different step functions.

$$t_{arv}^i = \begin{cases} 9, & 0 \leq \pi < \pi_1^i \\ 8, & \pi_1^i \leq \pi < \pi_2^i \\ 7, & \pi_2^i \leq \pi < \pi_3^i. \end{cases}$$ (5.2)

Given the step functions of each EV’s responses, the individual responses can be aggregated to the responses of the whole EV fleet with different incentives. For example, the number ($N_{t,arv}^{EV}$) of newly arriving EV available to the aggregator at hour $t$ with different incentives can be obtained as (5.3):

$$N_t^{EV,arv} = \begin{cases} N_{t,1}^{EV,arv}, & 0 \leq \pi < \pi_1 \\ \vdots & \vdots \\ N_{t,\omega}^{EV,arv}, & \pi_{\omega-1} \leq \pi < \pi_\omega \\ \vdots & \vdots \\ N_{t,\Omega}^{EV,arv}, & \pi_{\Omega-1} \leq \pi < \pi_\Omega \end{cases}$$ (5.3)
where $\omega$ indicates the $\omega$th interval of the incentive $\pi$ and $\Omega$ is the number of all intervals in the response functions of the whole EV fleet; $N_{t,\omega}^{EV,arv}$ means the number of EV newly available for the EV aggregator at hour $t$ if the incentive $\pi$ is between $\pi_{\omega-1}$ and $\pi_{\omega}$. It should be noted that $t_{arv}^i$ indicates EV $i$ arrives at the parking lot within that hour, which means if the EVs arrive at hour $t$, they are not available for EV aggregator to operate until hour $t+1$. Hence, the first hour for $N_{t,\omega}^{EV,arv}$ at the aggregator is $t = 7$ while the earliest arrival time of EVs is hour 6. Similarly, the number of EV leaving parking lot $N_{t,\omega}^{EV,dep}$ can be obtained and the last operation hour for the aggregator to participate in the markets is set as $t = 19$ for the EV owners who leave the parking lot at hour 20. In addition to $N_{t,\omega}^{EV,arv}$ and $N_{t,\omega}^{EV,dep}$, the arrival energy $E_{t,\omega}^{EV,arv}$ and desired departure energy $E_{t,\omega}^{EV,dep}$ of the EV fleet at hour $t$ with different incentives can be derived from individual SOC response functions and EV capacity $E_{max}^{EV}$.

5.3: Two-Stage Optimal Bidding Algorithm
5.3.1: First Stage: Day-Ahead Planning of EV Aggregator

As aforementioned, the EV aggregator is designed to provide two types of incentives to the EV owners. The incentive $\pi$ is for inducing the EV owners to change their behaviors. The incentive information is required to be sent to the EVs one day ahead in order to change the EV behaviors of the next day. Therefore, a DA planning model considering the market information is important for the EV aggregator to determine the optimal incentive signal $\pi$ delivered to the EV owners.

In the following DA planning model, the objective function (5.4), subject to (5.5)-(5.25), is to maximize the profit for EV aggregator to participate in the energy and regulation markets, which are represented by $Credit_{E,\omega}$ and $Credit_{R,\omega}$, with distributing the fixed reward $\hat{\pi}$ and the incentive $\pi$ to EV owners. $W_{\omega}$ is a binary weight related to the intervals of $\pi$. (5.5) enforces the DA planning model to maximize the profit within
only one interval of \( \pi \). Each interval of \( \pi \) is defined as \([5.7]\). For example, \( W_1 = 1 \) and \( W_\omega = 0 \) for \( \omega \neq 1 \) when \( \pi_0 \leq \pi < \pi_1 \).

\[
Max \sum_{\omega=1}^{\Omega} W_\omega (Credit_{E,\omega} + Credit_{R,\omega} - (\hat{\pi} + \pi)) \quad (5.4)
\]

s.t. \((5.5)-(5.25)\).

\[
\sum_{\omega=1}^{\Omega} W_\omega = 1 \quad (5.5)
\]

\[
W_\omega \in \{0, 1\} \quad (5.6)
\]

\[
\pi_{\omega-1} \leq \pi < \pi_\omega \quad (5.7)
\]

**Credits in Markets:** When \( \pi \in [\pi_{\omega-1}, \pi_\omega) \), the credit of energy market \( Credit_{E,\omega} \) and the credit of regulation market \( Credit_{R,\omega} \) are determined by \((5.8)\) and \((5.9)\), where \( P_{E,ch}^{t,\omega}, P_{E,dis}^{t,\omega} \), and \( P_{R}^{t,\omega} \) respectively represent the charging power, discharging power, and regulation power offered in the markets. These power offers are the decision variables determined by the DA planning model to maximize the profit of the EV aggregator based on the estimated market information, including the LMP, \( LMP_t \), Regulation Market Capability Clearing Price, \( RMCCP_t \), Regulation Market Performance Clearing Price, \( RMPCP_t \), performance score \( \overline{\rho}_t \), and mileage ratio \( \overline{\beta}_t \), estimated for the next day (Note: The method to estimate the market information for the next day is introduced in Section \([5.3.3]\)). The performance score \( \overline{\rho}_t \) is used to evaluate how well the regulation resources follow the regulation signal and penalize the resources that fail to follow. In the PJM’s regulation market, there are two types of regulation signals: Regulation A and Regulation D, ranging between 0 and 1. Compared to the fast and dynamic Regulation D signal, Regulation A is a slower signal. Considering the battery is capable of fast responding to the dispatch signals, the EV aggregator is proposed to follow Regulation D. The mileage
ratio in the regulation market measures the mileages of Regulation D over Regulation
A, which is used to evaluate the relative movement of following the different two signals.
For further clarifications of the market data, please refer to the PJM Manual \[140\] and
related studies \[36, 43, 141\].

\[
Credit_{E,\omega} = \sum_{t \in T} (P_{E,\omega}^{E,\text{dis}} - P_{E,\omega}^{E,\text{ch}}) \times LMP_t
\] (5.8)

\[
Credit_{R,\omega} = \sum_{t \in T} (P_{R,\omega} \times RMPCCP_t \times \bar{\rho}_t + P_{R,\omega} \times RMPCP_t \times \hat{\beta}_t \times \bar{\rho}_t)
\] (5.9)

**Power Constraints:** By aggregating the \( N_{EV,\text{arv}}^{EV} \) and \( N_{EV,\text{dep}}^{EV} \) from the first operation
hour to hour \( t \), the number of EVs at the aggregator \( N_{EV}^{EV} \) can be obtained, and the max
power of the aggregator \( P_{\text{max},t,\omega}^{Agg} \) at hour \( t \) can be calculated by (5.10). (5.11)-(5.13)
enforces the offers in the energy and regulation markets could not exceed the maximum
power of each hour. Since the aggregator cannot charge and discharge in the energy
market at the same time, a binary decision variable \( \delta_{t,\omega} \) is implemented in (5.14) and
(5.15) to enforce the aggregator to only offer either charging or discharging power in a
given hour segment. Meanwhile, the total power offered to the energy and regulation
markets cannot exceed the maximum power of the aggregator at hour \( t \) as described by
(5.17) and (5.18). Given the \( P_{E,\omega}^{E,\text{ch}}, P_{E,\omega}^{E,\text{dis}}, \text{ and } P_{R,\omega} \), the total charging power \( P_{t,\omega}^{Agg,\text{ch}} \)
and discharging power \( P_{t,\omega}^{Agg,\text{dis}} \) of the aggregator can be obtained by (5.19) and (5.20), where
the \( \overline{\text{RegD}}_{\text{down}}^t \) and \( \overline{\text{RegD}}_{\text{up}}^t \) are the estimated Regulation D sent by the ISO or RTO to
conduct the charge and discharge outputs in the regulation market, respectively.
Energy Constraints: In addition to the constraints (5.10)-(5.20) on the power outputs of the aggregator, there are also constraints on the EV aggregator’s energy amount in the DA planning model. (5.21) specifies the aggregator’s energy $E_{t,\omega}^{Agg}$ at hour $t$, which is obtained according to the aggregator’s energy $E_{t-1,\omega}^{Agg}$ of the previous hour, the aggregator’s departure energy $E_{t,\omega}^{Agg,dep}$, the energy $E_{t,\omega}^{EV,arv}$ provided by the new arrival EVs, and the total aggregated output at hour $t$. In (5.21), the $\eta_c$ and $\eta_d$ are the charging and discharging efficiencies, respectively. (5.22) shows the $E_{t,\omega}^{Agg,dep}$ should be not only larger than desired departure energy $E_{t,\omega}^{EV,dep}$ of the EVs, which is obtained from response functions, but also smaller than maximum energy of the departure EVs obtained by the number $N_{t,\omega}^{EV,dep}$ of EVs leaving the parking lot at hour $t$ and the capacity $E_{max}^{EV}$ of each EV. In other words, the aggregator can leave more departure energy than the demand.
of EVs when they depart. Therefore, in order to make the aggregators’ operation more flexible, the $E_{t,\omega}^{Agg,dep}$ is implemented in (5.21) instead of $E_{t,\omega}^{EV,dep}$. As shown in (5.23), the $E_{t,\omega}^{Agg}$ should be able to fulfill the energy demand $E_{t+1,\omega}^{Agg,dep}$ of the EVs departure at the next hour and less than the maximum energy $E_{max,t,\omega}^{Agg}$ of the aggregator which is obtained by (5.24). In (5.23), $\lambda$ is a percentage (e.g., 5%) and a margin of $\lambda E_{max,t,\omega}^{Agg}$ in the lower and upper bounds is set up against the uncertainty of regulation signals. The constraint (5.25) indicates the rest of the aggregator’s energy after the EVs depart at the next hour should not exceed the energy capacity of the rest EVs with considering the energy margin set up by $\lambda$.

\[
E_{t,\omega}^{Agg} = E_{t-1,\omega}^{Agg} - E_{t,\omega}^{Agg,dep} + E_{t,\omega}^{EV,arv} + \]

(5.21)

\[
P_{t,\omega}^{Agg,ch} \times \eta_c - P_{t,\omega}^{Agg,dis} / \eta_d
\]

\[
E_{t,\omega}^{EV,dep} \leq E_{t,\omega}^{Agg,dep} \leq N_{t,\omega}^{EV,dep} \times E_{max}^{EV}
\]

(5.22)

\[
E_{t+1,\omega}^{Agg,dep} + \lambda E_{max,t,\omega}^{Agg} \leq E_{t,\omega}^{Agg} \leq (1 - \lambda) E_{max,t,\omega}^{Agg}
\]

(5.23)

\[
E_{max,t,\omega}^{Agg} = N_{t,\omega}^{EV} \times E_{max}^{EV}
\]

(5.24)

\[
E_{t,\omega}^{Agg} - E_{t+1,\omega}^{Agg,dep} \leq (N_{t,\omega}^{EV} - N_{t+1,\omega}^{EV,dep})(1 - \lambda) E_{max}^{EV}
\]

(5.25)

According to the DA planning model (5.4)-(5.25), the aggregator can estimate the revenue and determine the incentives for the next day. If the aggregator finds out there is a profitable opportunity to operate in the next day, it can activate the EV aggregation program and broadcast the incentives, which are obtained via the DA planning model, to the EV owners.

### 5.3.2: Second Stage: Real-Time Operation of EV Aggregator

After the aggregation program is activated and the optimal incentive $\pi$ has been delivered to the EV owners, the aggregator needs to participate in the RT energy and regulation markets for gaining revenues. Similar to [41][42], the RT market is assumed
to be an hourly cleared market in this study and the market participants can submit
the bids before the hour \( t \) starts. In other words, the aggregator needs to keep bidding
for the next operation hour in the RT market until the end of the aggregation program.
To maximize the aggregator’s revenue in the RT market and generate the optimal bids
for the next operation hour, a RT optimal bidding algorithm for the incentive-based EV
aggregator is proposed.

Based on the DA planning model, a RT bidding model described in (5.26)-(5.47) is
proposed to maximize the profits for the rest of the operation hours \( \hat{T} \) of the day. Since
the incentive has already been determined via the DA planning model, the RT operation
model only considers maximizing the credits from the energy and regulation markets as
shown in (5.26). In (5.26)-(5.43), the \( LMP_t', RMCCP_t', RMPCP_t', \rho_t', \beta_t', RegD_t^{up} \), and
\( RegD_t^{down} \) are the market information predicted before the next operation hour of the
RT markets. The other variables and parameters in (5.26)-(5.43) are set up similar to
those in the DA planning model.

Maximize \( (Credit'_E + Credit'_R) \)  \hspace{1cm} \text{(5.26)}

s.t. \( (5.27)-(5.47) \).

\[
Credit'_E = \sum_{t \in \hat{T}} (P_{E,dis}^t - P_{E,ch}^t) \times LMP_t' \hspace{1cm} \text{(5.27)}
\]

\[
Credit'_R = \sum_{t \in \hat{T}} (P_R^t \times RMCCP_t' \times \rho_t^t + P_R^t \times RMPCP_t' \times \beta_t^t \times \rho_t^t) \hspace{1cm} \text{(5.28)}
\]

\[
P_{\text{max},t}^A = N_t^{EV} \times P_{\text{max}}^{EV} \hspace{1cm} \text{(5.29)}
\]

\[
0 \leq P_{t,\text{ch}}^E \leq P_{\text{max},t}^A \hspace{1cm} \text{(5.30)}
\]
\[ 0 \leq P_{t}^{E,dis} \leq P_{\text{max},t}^{\text{Agg}} \]  (5.31)

\[ 0 \leq P_{t}^{E,ch} \leq \delta_{t} \times M \]  (5.32)

\[ 0 \leq P_{t}^{E,dis} \leq (1 - \delta_{t}) \times M \]  (5.33)

\[ \delta_{t} \in \{0, 1\} \]  (5.34)

\[ 0 \leq P_{t}^{E,ch} + P_{t}^{R} \leq P_{\text{max},t}^{\text{Agg}} \]  (5.35)

\[ 0 \leq P_{t}^{E,dis} + P_{t}^{R} \leq P_{\text{max},t}^{\text{Agg}} \]  (5.36)

\[ P_{t}^{\text{Agg,\text{ch}}} = P_{t}^{E,ch} + P_{t}^{R} \times \text{RegD}_{t}^{\text{down}} \]  (5.37)

\[ P_{t}^{\text{Agg,\text{dis}}} = P_{t}^{E,dis} + P_{t}^{R} \times \text{RegD}_{t}^{\text{up}} \]  (5.38)

\[ E_{t}^{\text{Agg}} = E_{t-1}^{\text{Agg}} - E_{t}^{\text{Agg,dep}} + E_{t}^{\text{EV,arv}} + \]  (5.39)

\[ P_{t}^{\text{Agg,\text{ch}}} \times \eta_{c} - P_{t}^{\text{Agg,\text{dis}}} / \eta_{d} \]

\[ E_{t}^{\text{EV,dep}} \leq E_{t}^{\text{Agg,dep}} \leq N_{t}^{\text{EV,dep}} \times E_{\text{max}}^{\text{EV}} \]  (5.40)

\[ E_{t+1}^{\text{Agg,dep}} + \lambda E_{t}^{\text{Agg}} \leq E_{t}^{\text{Agg}} \leq (1 - \lambda) E_{\text{max},t}^{\text{Agg}} \]  (5.41)

\[ E_{\text{max},t}^{\text{Agg}} = N_{t}^{\text{EV}} \times E_{\text{max}}^{\text{EV}} \]  (5.42)

\[ E_{t}^{\text{Agg}} - E_{t+1}^{\text{Agg,dep}} \leq (N_{t}^{\text{EV}} - N_{t+1}^{\text{EV,dep}})(1 - \lambda) E_{\text{max}}^{\text{EV}} \]  (5.43)

**Additional Power Constraints:** In addition to the power constraints similar to the DA planning model, there are four constraints (5.44)-(5.47) in the RT operation model for the power offered to the next operation hour \( \tau \). During the RT operation, the aggregator needs to guarantee the EVs always leave the parking lot with their desired departure energy, which is important for the EV aggregator to keep operating the incentive-based aggregation program successfully: Having a satisfied customer. Therefore, the EV aggregator should schedule the charging/discharging power in the energy market to fulfill the energy demand of departure EVs against the uncertainty of regulation signals. \[5.44\]
indicates if the aggregator’s energy at the beginning of hour \( \tau \), which is obtained by
\[
(E^{Agg}_{\tau-1} + E_{\tau}^{EV,arv} - E^{Agg}_{\tau,dep})
\]
is not enough for the desired departure energy \( E^{EV,dep}_{\tau+1} \) at
hour \( \tau + 1 \), the aggregator must purchase a certain amount of energy from the energy
market at hour \( \tau \) to satisfy the energy demand of the departure EVs at hour \( \tau + 1 \).
Similarly, \( (5.45) \) restricts the aggregator to over discharge in the energy market at hour
\( \tau \). As aforementioned, the \( \lambda E_{max,\tau}^{Agg} \) is a margin for scheduling the aggregator’s charg-
ing/discharging power, which is used against the uncertainty of regulation signals and
improves the aggregator’s performance in the regulation market. \( (5.46) \) and \( (5.47) \) indi-
cate the power offered at hour \( \tau \) should be utilized to keep the aggregator from violating
the \((1 - \lambda)E_{max,\tau}^{Agg}\) at the end of hour \( \tau \), which is also for better handling the uncertainty
of the regulation signals at hour \( \tau \) with an energy margin.

\[
P^{E,ch}_{\tau} \times \eta_c \geq \max\{0, (E^{EV,dep}_{\tau+1} + \lambda E_{max,\tau}^{Agg})
- (E^{Agg}_{\tau-1} + E_{\tau}^{EV,arv} - E^{Agg}_{\tau,dep})\}
\]
\[
P^{E,dis}_{\tau} / \eta_d \leq \max\{0, (E^{Agg}_{\tau-1} + E_{\tau}^{EV,arv} - E^{Agg}_{\tau,dep})
- (E^{EV,dep}_{\tau+1} + \lambda E_{max,\tau}^{Agg})\}
\]
\[
P^{E,ch}_{\tau} \times \eta_c \leq \max\{0, (1 - \lambda)E_{max,\tau}^{Agg} - (E^{Agg}_{\tau-1}
+ E^{EV,arv}_{\tau} - E^{Agg}_{\tau,dep})\}
\]
\[
P^{E,dis}_{\tau} / \eta_d \geq \max\{0, (E^{Agg}_{\tau-1} + E_{\tau}^{EV,arv} - E^{Agg}_{\tau,dep})
- (1 - \lambda)E_{max,\tau}^{Agg}\}
\]

At the end of hour \( \tau \), the aggregator needs to calculate a provisional energy status
\( \hat{E}_{\tau}^{Agg} \) of the aggregator with the realized offers, including the \( P^{E,ch}_{\tau}, P^{E,dis}_{\tau}, \) and \( P^{R}_{\tau} \), and
the actual regulation signals represented by \( RegD^{up}_{\tau} \) and \( RegD^{down}_{\tau} \) as shown in \( (5.48) \).
It should be noted that the \( E^{Agg}_{\tau-1} \) and \( E^{Agg}_{\tau,dep} \) are obtained from the operation of hour
\( \tau - 1 \) and the process to obtain \( E_{\tau-1}^{Agg} \) and \( E_{\tau}^{Agg,dep} \) is introduced with Algorithm 5.1 subsequently.

\[
\hat{E}_{\tau}^{Agg} = E_{\tau-1}^{Agg} - E_{\tau}^{Agg,dep} + E^{EV,arv}_{\tau}
\]
\[
+ (P_{E,ch,\tau} + P_{R,\tau} \times \text{RegD}_{\tau}^{down}) \eta_c \]
\[
- (P_{E,dis,\tau} + P_{R,\tau} \times \text{RegD}_{\tau}^{up}) \eta_d
\] (5.48)

Due to the uncertainty of regulation signals, it is possible that the predicted regulation signals significantly deviate from the actual signals. Therefore, the \( \hat{E}_{\tau}^{Agg} \) may violate the energy constraints at the end of hour \( \tau \). In order to keep the aggregator operating within the energy requirements, three criteria corresponding to (5.41) and (5.43) are introduced in (5.49)-(5.51) for checking the \( \hat{E}_{\tau}^{Agg} \). If \( \hat{E}_{\tau}^{Agg} \) violates any one of these criteria, a post-process at the end of hour \( \tau \) has to be taken to adjust the \( \hat{E}_{\tau}^{Agg} \) and \( E_{\tau+1}^{Agg,dep} \).

\[
\hat{E}_{\tau}^{Agg} \geq E^{EV,dep}_{\tau+1}
\] (5.49)
\[
\hat{E}_{\tau}^{Agg} \leq E_{max,\tau}^{Agg}
\] (5.50)
\[
\hat{E}_{\tau}^{Agg} - E^{EV,dep}_{\tau+1} \leq (N_{\tau}^{EV} - N_{\tau+1}^{EV,dep}) \times E_{max}^{EV}
\] (5.51)

When \( \hat{E}_{\tau}^{Agg} < E^{EV,dep}_{\tau+1} \), the energy difference \( \Delta E_{\tau}^{Agg} \) obtained by (5.52) is utilized to calculate the nonperformance offer \( \Delta P_{\tau}^{R} \), which is a shortfall performance of the \( P_{\tau}^{R} \) in the regulation market, and the actual aggregator’s energy \( E_{\tau}^{Agg} \) after the adjustment via (5.53) and (5.54), respectively.

\[
\Delta E_{\tau}^{Agg} = \hat{E}_{\tau}^{Agg} - E^{EV,dep}_{\tau+1}
\] (5.52)
\[
\Delta P_{\tau}^{R} = \Delta E_{\tau}^{Agg} / (\text{RegD}_{\tau}^{down} \eta_c - \text{RegD}_{\tau}^{up} / \eta_d)
\] (5.53)
\[
E_{\tau}^{Agg} = \hat{E}_{\tau}^{Agg} - \Delta E_{\tau}^{Agg}
\] (5.54)
If $\hat{E}_{Agg}^{\tau} > E_{\text{max},\tau}^{Agg}$, the corresponding $\Delta E_{\tau}^{Agg}$, $\Delta P_{\tau}^{R}$, and $E_{\text{tau}}^{Agg}$ are calculated by (5.55), (5.53) and (5.54), respectively.

$$\Delta E_{\tau}^{Agg} = \hat{E}_{\tau}^{Agg} - E_{\text{max},\tau}^{Agg}$$  \hspace{1cm} (5.55)$$

After the $E_{\tau}^{Agg}$ has been obtained, the $E_{\tau+1}^{Agg,dep}$ can be calculated by (5.56) and (5.57) if the criterion (5.51) is violated.

$$\Delta E_{\tau+1}^{Agg,dep} = E_{\tau}^{Agg} - E_{\tau+1}^{EV,dep}$$

$$- (N_{\tau}^{EV} - N_{\tau+1}^{EV,dep}) \times E_{\text{max}}^{EV}$$

$$E_{\tau+1}^{Agg,dep} = E_{\tau+1}^{EV,dep} + \Delta E_{\tau+1}^{Agg,dep}$$  \hspace{1cm} (5.57)$$

The process of the second stage of the proposed bidding algorithm is illustrated in Algorithm 5.1. In the algorithm, $\tau$ indicates the next operation hour. Before the hour $\tau$, the aggregator needs to predict the market data for the rest operation hours $\hat{T}$ and solve the RT operation model in (5.26)-(5.47) to obtain the bids for hour $\tau$. At the end of hour $\tau$, the $\hat{E}_{\tau}^{Agg}$ is calculated by (5.48) and checked with the criteria (5.49)-(5.51). Then, the obtained $E_{\tau}^{Agg}$ and $E_{\tau+1}^{Agg,dep}$ are used in the RT operation model of the next operation hour $\tau + 1$.

It is worth pointing out that there will be no non-compliant issue and penalization for the aggregator in the energy market since the charging/discharging power cleared in the energy market has been enforced by (5.44)-(5.47) to satisfy the energy requirements of the coming operation hour. However, the predictions of the regulation signals can be significantly different from the actual values, the worst case for the aggregator is $\Delta P_{t}^{R} = P_{t}^{R}$, which means the aggregator is unable to perform any regulation services in the regulation market during the hour $t$. The departure energy required by the EVs can still be guaranteed under this circumstance.
Algorithm 5.1: Real-Time Operation

Initialization: $\tau = 7$, $E_{\tau-1}^{Agg} = 0$, $E_{\tau}^{Agg,dep} = E_{\tau}^{EV,dep}$; for $\tau = 7, 8, \ldots, 19$ do

\[ \hat{T} = \{\tau, \ldots, 19\}; \]

Estimate market data for $t \in \hat{T}$;

Solve the RT operation model (5.26)-(5.47);

Obtain $P_{\tau}^{E,ch}$, $P_{\tau}^{E,dis}$, and $P_{\tau}^{R}$;

Calculate $\hat{E}_{\tau}^{Agg}$ by (5.48);

if $\hat{E}_{\tau}^{Agg} < E_{\tau+1}^{EV,dep}$ then

| Obtain $\Delta P_{\tau}^{R}$ and $E_{\tau}^{Agg}$ by (5.52), (5.53), (5.54) |

else if $\hat{E}_{\tau}^{Agg} > E_{\tau}^{Agg, max, \tau}$ then

| Obtain $\Delta P_{\tau}^{R}$ and $E_{\tau}^{Agg}$ by (5.55), (5.53), (5.54) |

else

| $\Delta P_{\tau}^{R} = 0$, $E_{\tau}^{Agg} = \hat{E}_{\tau}^{Agg}$, |

end

if $E_{\tau}^{Agg} - E_{\tau+1}^{EV,dep} > (N_{\tau}^{EV} - N_{\tau+1}^{EV,dep}) \times E_{\tau}^{EV, max}$ then

| Obtain $E_{\tau+1}^{Agg,dep}$ by (5.56), (5.57) |

else

| $E_{\tau+1}^{Agg,dep} = E_{\tau+1}^{EV,dep}$ |

end

Prepare $E_{\tau}^{Agg}$, $E_{\tau+1}^{Agg,dep}$ for the next hour $\tau + 1$;

end

Result: $P_{\tau}^{E,ch}$, $P_{\tau}^{E,dis}$, $P_{\tau}^{R}$, $\Delta P_{\tau}^{R}$, $E_{\tau}^{Agg}$, and $E_{\tau+1}^{Agg,dep}$ for $t \in T$
As aforementioned, the performance score $\rho_t$ is used to penalize the market participants for failing to provide the cleared regulation capacity. In PJM, the performance score shown in (5.58) is a combination of the precision score ($\rho_t^{\text{precision}}$), the correlation score ($\rho_t^{\text{correlation}}$), and the delay score ($\rho_t^{\text{delay}}$) ranging between 0 and 1 \cite{141, 143}. The precision score defined by (5.59) is used to evaluate the error of a regulation response.

$$\rho_t = \frac{\rho_t^{\text{precision}} + \rho_t^{\text{correlation}} + \rho_t^{\text{delay}}}{3} \tag{5.58}$$

$$\rho_t^{\text{precision}} = 1 - \frac{(P_t^R - |\Delta P_t^R|)}{P_t^R} \tag{5.59}$$

The correlation score is implemented to measure the correlation between the regulation signals and resource’s responses. The delay score is to evaluate the delay of the response between the regulation signal and the change of resource’s output. Considering the outstanding ramping capability of the battery for following the regulation signal, both the correlation and delay scores can be approximated to be the same as the precision score at each hour for the simplification based on \cite{141, 143}. Therefore, $\rho_t$ is represented by $\rho_t^{\text{precision}}$.

Given the cleared offers, such as $P_t^{E,ch}$, $P_t^{E,dis}$, $P_t^R$, and the actual market data including $LMP_t$, $RMCCP_t$, $RMPCP_t$, $\beta_t$, and $\rho_t$, the credits of participation in the energy and regulation markets for the EV aggregator can be calculated based on (5.27) and (5.28).

### 5.3.3: Prediction of Market Data

According to the studies for market data in \cite{35, 37, 144}, the seasonality has been found in the LMP \cite{35}, regulation market clearing price \cite{144}, and regulation signals \cite{37}. Meanwhile, a similar feature of the mileage ratio collected from PJM has been discovered, as well. Hence, to predict the market data in the DA planning and RT operation models, the seasonal ARIMA model ($ARIMA(p, d, q) \times (P, D, Q)_s$) \cite{134, 145} described by (5.60).
is implemented.

$$\phi_p(B)\Phi_P(B^s)\nabla^d \nabla^D y_t = \mu + \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$  \hspace{1cm} (5.60)

where $y_t$ is the time series fitted by the seasonal ARIMA model; $p$ is the non-seasonal auto-regression (AR) order; $d$ is the non-seasonal differencing order; $q$ is the non-seasonal moving-average (MA) order; $P$ is the seasonal AR order; $D$ is the seasonal differencing; $Q$ is the seasonal MA order; $S$ represents the time steps of repeating seasonal pattern; The error term $\varepsilon_t$ is an independent and identically distributed noise with zero mean and finite variance; $\mu$ is a constant term for the model offset; $B$ is the backward shift operator, i.e. $B^h y_t = y_{t-h}$. The operators $\phi_p(B)$, $\Phi_P(B^s)$, $\nabla^d$, $\nabla^D$, $\theta_q(B)$, and $\Theta_Q(B^s)$ are described as (5.61)-(5.66).

\[
\begin{align*}
\phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \hspace{1cm} (5.61) \\
\Phi_P(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps} \hspace{1cm} (5.62) \\
\nabla^d &= (1 - B)^d \hspace{1cm} (5.63) \\
\nabla^D &= (1 - B^s)^D \hspace{1cm} (5.64) \\
\theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \hspace{1cm} (5.65) \\
\Theta_Q(B^s) &= 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs} \hspace{1cm} (5.66)
\end{align*}
\]

where $\phi_1 \ldots \phi_p$, $\Phi_1 \ldots \Phi_P$, $\theta_1 \ldots \theta_q$ and $\Theta_1 \ldots \Theta_Q$ are the coefficients of the corresponding operators.

In order to fit the market data with the seasonal ARIMA model, a preprocess scheme is implemented. First, the market data is preprocessed by (5.67) to eliminate the outliers.
in data series:

\[
\bar{y}_t = \begin{cases} 
\mu_y + 3\sigma_y, & \text{if } y_t > \mu_y + 3\sigma_y \\
y_t, & \text{otherwise} \\
\mu_y - 3\sigma_y, & \text{if } y_t < \mu_y - 3\sigma_y
\end{cases}
\]  

(5.67)

where \( \mu_y \) is the mean of \( y_t \) and \( \sigma_y \) is the standard deviation of \( y_t \). Then, based on [145], a natural logarithm transformation, given in (5.68), is taken to reduce the fluctuations and make the series stationary for fitting the seasonal ARIMA model.

\[
\hat{y}_t = \log(\bar{y}_t + c)
\]  

(5.68)

where \( \bar{y}_t \) is the series obtained by (5.67); \( c \) is a constant value to offset \( \bar{y}_t \) for the logarithm transformation. The \( \hat{y}_t \) acquired by (5.68) is utilized with the seasonal ARIMA model for the prediction.

5.4: Model Performance
5.4.1: Incentives and EV Responses

In order to study the performance of the proposed models and algorithms, 200 EVs are assumed to participate in the EV aggregation program. The power and capacity of each EV are set to 50 kW and 50 kWh [146]. Hence, the maximum power and energy capacity of the aggregator are 10 MW and 10 MWh, respectively. Based on the revenue analysis for energy storage systems in energy and regulation markets [36,43], the average daily revenue for a 10 MW/ 10 MWh energy storage system with considering the battery degradation and perfect price forecasting is about $6,300/day. Since the EV aggregator proposed in this chapter is operating between 7:00 and 19:00, the maximum credit obtained from energy and regulation markets during the 13 operating hours is expected to be about $3,500/day. Accordingly, the maximum reward for the EV owners is set to be $2500/day, including the fixed reward \( \hat{\pi} \) of $1000/day and the incentive \( \pi \) up to $1500/day. The
credit less the rewards is considered as the EV aggregator’s gross revenue, covering the operating cost.

The initial behaviors of each EV are sampled via the truncated Gaussian functions introduced in Section 5.2. Given the initial behaviors of each EV, the response step functions of the behaviors, including $t_{\text{arv}}^i$, $t_{\text{dep}}^i$, $SOC_{\text{arv}}^i$, and $SOC_{\text{dep}}^i$, can be generated with the incentive $\pi$ ranging from 0 to 1500. Consequently, the step functions of $N_{t_{\text{EV,arv}}}^i$, $N_{t_{\text{EV,dep}}}^i$, $E_{t_{\text{EV,arv}}}^i$, $E_{t_{\text{EV,dep}}}^i$ for the whole EV fleet corresponding to different incentive levels can be derived. As a reference, the $N_{t_{\text{EV,arv}}}^i$ and $N_{t_{\text{EV,dep}}}^i$ of $\pi = 0, 750,$ and 1500 are plotted in Fig. 5.1 where the horizontal axis indicates the operating hours and the vertical axis shows the number of arrival/departure EVs. It can be clearly seen that EVs tend to arrive earlier and depart later when the incentive $\pi$ increase from 0 to 1500.

![Figure 5.1: Arrival and departure EVs with different incentives.](image)

5.4.2: First Stage: DA Planning

The real market data, such as LMP, RMCCP, RMPCP, mileage ratio and regulation signals, between 1/1/2018 and 12/31/2018 are collected from PJM Data Miner 2 [115] to test the performance of the proposed models in the energy and regulation markets. To mitigate the impact of the transmission congestion on influencing the analysis of
model performance, the LMP of PJM-RTO, which is an aggregated LMP of the whole energy market, is utilized in the following study to analyze the overall performance of the proposed model in the markets first. Then, the proposed two-stage optimal bidding algorithm is evaluated with the LMPs of different cities located in PJM.

In this study, the EV aggregator is designed to broadcast the incentive information to the EV owners at 16:00 before the operating day, as shown in Fig. 5.2, in order to obtain more information on the RT market and let the EV owners get the incentive information before most of them leave the parking lot. Since the aggregation program starts at 7:00 on an operating day, the EV aggregator needs to predict the market data for the next 27 hours, including the rest 7 hours of the DA day and the 20 hours of the operating day before the operation ends.

![Timeline of DA planning and RT operation.](image)

The introduced forecasting method based on the seasonal ARIMA model has been implemented to predict the market data. According to [35, 134], the seasonality of the market data is found to be 24 via the analysis of autocorrelation function (ACF) and partial autocorrelation function (PACF). Meanwhile, the other orders of the seasonal ARIMA model are determined with the values listed in Table 5.2 depending on the stationarity of the forecasting time series and the analysis of ACF and PACF. In this study, the market data before 2/26/2018 (Monday) are used to train the corresponding seasonal ARIMA models. Therefore, the proposed models are tested during the period
between 2/26/2018 (Monday) and 12/30/2018 (Sunday).

Table 5.2: Orders of Seasonal ARIMA Model

<table>
<thead>
<tr>
<th>Orders</th>
<th>p</th>
<th>d</th>
<th>q</th>
<th>P</th>
<th>D</th>
<th>Q</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>2, 3</td>
<td>0, 1</td>
<td>0, 1</td>
<td>1</td>
<td>0, 1</td>
<td>0, 1</td>
<td>24</td>
</tr>
</tbody>
</table>

After the DA planning model is solved with the forecasting market data, the estimated revenue for the operating day can be calculated. Since the intention of the proposed EV aggregator at the parking lot is to benefit the EV owners with the incentives and conduct more regulation resources to the power system, the aggregation program is set to be activated whenever the estimated revenue is larger than $0. Meanwhile, it should be mentioned that the EV aggregation program is activated only when the operating day is not a weekend or holiday. Between 2/26/2018 and 12/30/2018, there are 213 working days valid for the EV aggregator at the parking lot to participate in the markets. According to the estimated revenue obtained via the DA planning model, the EV aggregation program is activated for 125 out of 213 working days. During these days, the EV owners can be rewarded with $229,625, which means the 200 EV owners can share $1,837/day and each EV owner is expected to get a reward of $9.19/day for participating in the EV aggregation program.

5.4.3: Second Stage: RT Operation

When the operating days and the corresponding incentives have been determined by the DA planning model, the EV aggregator needs to participate in the RT energy and regulation markets for the credits. According to the designed RT operation algorithm, the hourly offers, including $P_{t}^{E,ch}$, $P_{t}^{E,dis}$, and $P_{t}^{R}$, and the performance score can be obtained. In Fig. 5.3, the average offers by the operating hours are plotted. As shown
in Fig. 5.3, the aggregator offers most of its capacity to the regulation market, which indicates the operation model tends to obtain the credits by providing regulation service instead of energy arbitrage. The aggregator keeps bidding a certain amount of power charging in the energy market after hour 13, as the EVs start leaving at hour 13, and the aggregator needs to guarantee that the EVs leave with their desired energy level.

![Figure 5.3: Average offers in energy and regulation markets by operating hours.](image)

As aforesaid, the aggregator needs to adjust its actual output in the regulation market for complying with the energy constraints, especially when the actual regulation signals are quite different from the forecasting regulation signals. Therefore, the aggregator may fail to perform the committed regulation offer and get a penalty evaluated by the performance score for some hours. In Fig. 5.4, the average nonperformance power and the precision score at different operating hours are plotted. The left and right vertical indices in Fig. 5.4 indicate the MW of nonperformance offer and the performance score, respectively. As shown in Fig. 5.4, the performance scores at all operating hours are over 91% and the average performance score of the aggregator is over 95%, which means the aggregator performs well on complying with the cleared regulation offers. Since the aggregator bids most of its capacity in the regulation market as displayed in Fig. 5.3, it is very important that the aggregator’s profit is not affected by a poor performance in
the regulation market.

Figure 5.4: Hourly average nonperformance offer and precision score.

The total credit of the aggregator obtained from the energy and regulation markets during the 125 operating days is $360,833.73 and the average daily credit is $2,886.67. After issuing the incentives to EV owners, the aggregator’s revenue is $131,208.73. With a further analysis of the market credits, the average daily credits of energy and regulation markets are $46.41/day and $2,840.26/day, respectively. The credit from the energy market is only 1.63% of the regulation market. In other words, the energy market mainly helps the aggregator to comply with the energy requirements instead of providing a meaningful revenue. For comparison, a base case for participating in RT energy and regulation markets is included, where the aggregator maximizes its credits without using the incentives to change the EVs’ behaviors or implementing the energy margin $\lambda$ in the optimization models against the uncertainty of the regulation signals. It should be noted that the EV aggregation program without the procedure of DA planning is activated everyday. With this strategy, the EV aggregator can only obtain $1,456.66 per day with a performance score of 85.9%.

In addition to the 125 operating days determined via the DA planning model, the proposed models and algorithm have also been tested for the rest 88 working days in the
test period. The average daily credit obtained from the RT markets is $1,415.89, which is only 49% the average credit earned during the operating days. In other words, the profit margin would get shrunk if the aggregator kept activating the incentive-based EV aggregation program on those working days. It also means the aggregator could suffer a higher risk of loss without the proposed DA planning model by still providing blind incentives.

After the overall performances of the proposed algorithm have been evaluated with the LMP of PJM-RTO, the LMPs of the price nodes located in several large cities, including: Chicago, Cleveland, Pittsburgh, and Philadelphia, are selected for a further analysis of the bidding algorithm for the EV aggregator in workplace parking lots. As shown in Table 5.3, the average daily credits earned by the EV aggregation program in different cities are close to the credit of the overall performances, which indicates the proposed incentive-based EV aggregation program is capable of benefitting the EV owners and aggregators at different locations.

<table>
<thead>
<tr>
<th>Case</th>
<th>Avg. Credit</th>
<th>EV</th>
<th>Aggregator</th>
<th>Avg. $\rho_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>$1,456.66</td>
<td>$1,000.00</td>
<td>$456.66</td>
<td>85.9%</td>
</tr>
<tr>
<td>PJM-RTO</td>
<td>$2,886.67</td>
<td>$1,837.00</td>
<td>$1,049.67</td>
<td>95.6%</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>$2,869.32</td>
<td>$1,820.89</td>
<td>$1,048.43</td>
<td>95.7%</td>
</tr>
<tr>
<td>Cleveland, OH</td>
<td>$2,885.81</td>
<td>$1,838.08</td>
<td>$1,047.73</td>
<td>95.1%</td>
</tr>
<tr>
<td>Pittsburgh, PA</td>
<td>$2,879.95</td>
<td>$1,838.67</td>
<td>$1,041.28</td>
<td>95.6%</td>
</tr>
<tr>
<td>Philadelphia, PA</td>
<td>$2,943.88</td>
<td>$1,812.02</td>
<td>$1,131.86</td>
<td>96.2%</td>
</tr>
</tbody>
</table>
5.5: Summary

A two-stage optimal bidding algorithm for an incentive-based aggregator of EVs at workplace parking lots has been proposed in this chapter. Based on the features and incentive responses of individual vehicles’ behaviors, such as arrival/departure times, arrival energy, and desired departure energy, the behaviors of the whole EV fleet aggregator have been characterized. A DA planning model considering the incentive response functions has been proposed to determine whether or not to provide incentives and the incentive level for the aggregator. With a specified incentive, a RT operation algorithm has been designed for maximizing the aggregator’s profit in the RT energy and regulation markets while satisfying the EVs’ energy demand. The proposed models have been tested using the real PJM energy and regulation market data in 2018. The results show that 200 EV owners are expected to get an average daily reward of $1,837 in total, and the aggregator can have a profit of $1,049.67 per day. The revenue analysis of the overall performances has shown the proposed two-stage optimal bidding algorithm is effective in benefitting both the EV owners and the aggregator for participating in the energy and regulation markets.
CHAPTER 6: CONCLUSIONS AND FUTURE WORK

In this thesis, several algorithms and methods for optimal load and energy storage management have been proposed and studied. The proposed forecasting algorithms based on the ARIMA models have been proven to be capable of improving the electricity price prediction and fuel cost distributions forecasting, which can provide more accurate information for load management programs to make better decisions. The centralized and decentralized load management methods, including Temporal and Spatial Load Management (TSLM), Temporal Only Load Management (TLM), Sliding Window Self-Optimizing Load Management (SW-SOLM), and Day Ahead Self-Optimizing Load Management (DA-SOLM), have been proposed and studied using the IEEE 14-bus and IEEE 57-bus systems. The simulation results show that all the proposed load management methods can reduce the cost and emissions, and the TSLM can outperform the other methods.

Taking PJM as an example, the highly profitable nodes in the system have been revealed and characterized for battery energy storage system (BESS) installation. A comparison study of stationary and transportable BESSs has shown the transportable BESS can produce higher potential revenue in the energy and regulation markets. Meanwhile, an optimal placement algorithm has been developed for market participants to invest and manage BESS projects. Furthermore, a two-stage optimal bidding algorithm has been proposed for an incentive-based EV aggregator to participate in the energy and regulation markets. The simulation results indicate that the proposed optimal bidding algorithm is capable of handling the uncertainty of regulation signals. The revenue analysis has also shown that the incentive-based EV aggregator in workplace parking lots can have a stable performance when participating in energy and regulation markets at
different locations, such as the metropolitan areas.

In order to further improve the energy storage management for market participants, there can be several future extensions to the presented research work: 1) A more comprehensive study and modeling of parking behaviors and incentive responses of EV owners obtained via social studies can be beneficial for market participants to get a more sophisticated model of EV aggregator. 2) Since the incentive-based EV aggregator can be implemented at different sites day by day, it is possible for market participants to promptly activate the aggregation program around the upstream and downstream sides of a congested transmission line for more energy arbitrage opportunities. Meanwhile, the transmission congestion can be alleviated. 3) When multiple EV aggregators participate in the electricity market simultaneously, it is worth developing an effective bidding algorithm to coordinate the aggregators in different locations to manage the risks (e.g., the uncertainties of electricity prices and EVs’ behaviors) and maximize the overall profit. 4) Photovoltaic (PV) panels have been utilized in certain EVs. An incentive-based EV aggregator can cooperate with the management of onboard PVs. Based on the market conditions, EV owners’ behaviors, and the sunset/sunrise time, a special incentive scheme can be designed to maximize the aggregator’s profit and solar energy harvesting.
REFERENCES


[18] M. A. Ortega-Vazquez, F. Bouffard, and V. Silva, “Electric vehicle aggregator/sys-


[67] C.-M. Huang, H.-T. Yang, and C.-L. Huang, “Bi-objective power dispatch using


[82] R. H. Byrne, T. A. Nguyen, D. A. Copp, R. J. Concepcion, B. R. Chalamala, and


[89] Y. Sun, Z. Li, M. Shahidehpour, and B. Ai, “Battery-based energy storage trans-


ABSTRACT

OPTIMAL LOAD AND ENERGY STORAGE MANAGEMENT IN ELECTRICITY MARKETS

by

ZHONGYANG ZHAO

August 2020

Advisor: Dr. Caisheng Wang

Co-Advisor: Dr. Masoud H. Nazari

Major: Electrical and Computer Engineering

Degree: Doctor of Philosophy

In the U.S., several electricity markets have been formed throughout the country to efficiently and economically manage large power grids for their safe and reliable operations. In the electricity markets, the load management techniques have become important tools in improving voltage profile, frequency regulation, system efficiency, and stability. In this thesis, several algorithms and methods for optimal load and energy storage managements are proposed and studied. As the accurate prediction of market information is critical for demand side management and power generation scheduling, the algorithms based on the autoregressive integrated moving average (ARIMA) models are developed in this thesis to improve the predictions of electricity price and fuel cost distributions.

The load management can also help decrease emissions and costs, especially in future smart grids where customers will have more flexibility in controlling their electricity usages. Several new load management methods, such as the temporal and spatial load management, are proposed for the market operator in the thesis. The results show the proposed load management methods are capable of reducing the cost and emission, and mitigating the price spikes.
In modern power systems, the battery energy storage system (BESS), as an important market participant, can provide a variety of functions, such as energy arbitrage and ancillary services. A comprehensive revenue analysis of BESSs is carried out in this thesis. Based on the results of revenue analysis, an optimal placement algorithm is proposed for finding the profitable sites to install BESSs in the system. The effectiveness of the proposed algorithm is validated with real electricity market data.

Furthermore, the BESS can also be formed by aggregating a fleet of electric vehicles (EVs) that have the vehicle to grid (V2G) capabilities. A two-stage optimal bidding algorithm is proposed for an incentive-based EV aggregator to participate in the electricity markets. The simulation results show that the proposed bidding algorithm performs well on handling the uncertainty of market data, and the EV aggregator can have a stable revenue when participating in the electricity market at different locations.
AUTOBIOGRAPHICAL STATEMENT

ZHONGYANG ZHAO

Education
Ph.D. Wayne State University, Detroit, MI, USA 2020
M.S. Wayne State University, Detroit, MI, USA 2014
B.S. South China University of Technology, Guangzhou, Guangdong, China 2013

Selected Honors and Awards
Best Conference Paper Award in IEEE PES General Meeting 2018
Thomas C. University Graduate Rumble Fellowship 2017
The Dean’s Merit Scholarship 2012

Selected Publications