Electric Vehicle Charging Network Design and Control Strategies

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Fei Wu, M.S.

Graduate Program in Industrial and Systems Engineering

The Ohio State University

2016

Dissertation Committee:

Dr. Ramteen Sioshansi, Advisor

Dr. Güzin Bayraksan

Dr. Lixin Ye
Abstract

Electric vehicles hold great promise in improving transportation energy consumption efficiency. However, battery-related “range-anxiety” hampers large-scale EV adoption. To relieve range anxiety and accelerate wider EV adoption, two issues should be addressed. The first is how to develop and optimize the layout of a network of public EV charging stations to keep pace with increasing EV charging demand. Second, the strains of added EV charging loads on the electricity distribution infrastructure should be managed. This dissertation will address these two issues through new optimization models and simulations. Specifically, we formulate a stochastic flow-capturing charging station location model to solve the first issue and obtain an optimal charging station network layout. For the second issue, stochastic control models are created to schedule EV charging loads. This model assumes some flexibility to defer EV charging demands, to minimize distribution infrastructure investment and degradation costs. We expand this model to allow the charging station to participate in ancillary service markets, either by modulating EV charging or by using distributed resources. The outcome of this dissertation will help EV infrastructure planners develop an adequate charging station network and facilitate wider EV adoption. It will also help system operators and energy planners better understand the challenges of vehicle-grid integration and prepare appropriate strategies to address these issues.
This is dedicated to my family and friends.
Acknowledgments

This work is a product of my four year dedication and devotion. But these alone would always have fallen short without the support, guidance, inspiration, and advice of many others. This document will be incomplete without my mentioning their names.

I will be forever indebted to my advisor, Ramteen Sioshansi, for being a constant source of knowledge, encouragement, support, and inspiration over the past 4 years. I feel deeply grateful that he encourages me to explore the research world and allows me to start, stop, and fail along the way.

I would like to acknowledge the members of my dissertation committee, Prof. Bayraksan and Prof. Ye, for their comments and encouragements and also for the hard questions which encouraged me to improve my research from various perspectives.

I am also grateful to the Integrated Systems Engineering Department at the Ohio State University for the academic and technical supports, to the staff and faculty members in the department, especially Dr. Marc Posner and Dr. Theodore Allen, for their guidance, knowledge, and support. My sincere thank also goes to Dr. Josh Xiaomin Xi who enlightens me the first glance of the research.

Last, I would like to thank my friends and family for supporting me spiritually throughout writing this thesis.
Vita

2012 .......................... B.Eng. Industrial Engineering

2015 .......................... M.S. Industrial & Systems Engineering

2013-2016 ...................... Graduate Research Associate
                          The Ohio State University

2016 .......................... Graduate Teaching Associate
                          The Ohio State University

Fields of Study

Major Field: Industrial and Systems Engineering

Specialization: Operations Research

Minor Field: Economics

Minor Field: Statistics
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Dissertation Scope and Contributions</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1 Charging Network Optimal Layout Design</td>
<td>5</td>
</tr>
<tr>
<td>1.2.2 Charging Station Control Strategies</td>
<td>6</td>
</tr>
<tr>
<td>2. A Stochastic Flow-Capturing Model to Optimize the Location of Fast-Charging Stations with Uncertain Electric Vehicle Flows</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Modeling and Solution Approach</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Stochastic Flow-Capturing Location Model Formulation and Required Data</td>
<td>15</td>
</tr>
<tr>
<td>2.2.2 Input Data and Data Processing</td>
<td>18</td>
</tr>
<tr>
<td>2.2.3 Solution Technique</td>
<td>21</td>
</tr>
<tr>
<td>2.3 Case Study</td>
<td>35</td>
</tr>
<tr>
<td>2.4 Case Study Results</td>
<td>39</td>
</tr>
<tr>
<td>2.4.1 Stochastic Flow-Catching Location Model Results</td>
<td>40</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.4.2 Effect of Recharging Deviation Tolerance</td>
<td>43</td>
</tr>
<tr>
<td>2.4.3 Benefit of Public Charging Network to EV Owners</td>
<td>46</td>
</tr>
<tr>
<td>2.4.4 Charging Station Sizing</td>
<td>50</td>
</tr>
<tr>
<td>2.4.5 Comparison of Stochastic and Deterministic Flow-Catching Location Models</td>
<td>52</td>
</tr>
<tr>
<td>2.5 Conclusions and Future Work</td>
<td>57</td>
</tr>
<tr>
<td>3. A Stochastic Optimization Charging Station Control Model to Relieve Transformer Constraints</td>
<td>59</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>59</td>
</tr>
<tr>
<td>3.2 Electric Vehicle Charging Station-Control Model</td>
<td>64</td>
</tr>
<tr>
<td>3.2.1 Model Parameters</td>
<td>65</td>
</tr>
<tr>
<td>3.2.2 State Parameters</td>
<td>67</td>
</tr>
<tr>
<td>3.2.3 Decision Variables</td>
<td>69</td>
</tr>
<tr>
<td>3.2.4 Constraints</td>
<td>70</td>
</tr>
<tr>
<td>3.2.5 Objective Function</td>
<td>75</td>
</tr>
<tr>
<td>3.3 Solution Technique</td>
<td>76</td>
</tr>
<tr>
<td>3.3.1 Sample Average Approximation Problem</td>
<td>77</td>
</tr>
<tr>
<td>3.3.2 Solution of SAA method</td>
<td>78</td>
</tr>
<tr>
<td>3.3.3 Gap and Variance Estimators</td>
<td>81</td>
</tr>
<tr>
<td>3.3.4 Stopping Criterion and Sample Sizes</td>
<td>83</td>
</tr>
<tr>
<td>3.3.5 Convergence Properties of SAA</td>
<td>87</td>
</tr>
<tr>
<td>3.3.6 Asymptotic Validity and Finite Termination</td>
<td>94</td>
</tr>
<tr>
<td>3.4 Case Study</td>
<td>96</td>
</tr>
<tr>
<td>3.4.1 Electric Vehicle Usage</td>
<td>98</td>
</tr>
<tr>
<td>3.4.2 Non-Electric Vehicle Load</td>
<td>101</td>
</tr>
<tr>
<td>3.4.3 Photovoltaic Generation</td>
<td>101</td>
</tr>
<tr>
<td>3.4.4 Electricity Price</td>
<td>102</td>
</tr>
<tr>
<td>3.4.5 Battery Energy Storage System Characteristics</td>
<td>103</td>
</tr>
<tr>
<td>3.4.6 Transformer Characteristics</td>
<td>103</td>
</tr>
<tr>
<td>3.4.7 Temporal Resolution</td>
<td>104</td>
</tr>
<tr>
<td>3.4.8 Fractional Factorial Experimental Design</td>
<td>104</td>
</tr>
<tr>
<td>3.5 Case-Study Results</td>
<td>105</td>
</tr>
<tr>
<td>3.5.1 Station-Control Model Results</td>
<td>106</td>
</tr>
<tr>
<td>3.5.2 Parameter Sensitivity Analysis</td>
<td>114</td>
</tr>
<tr>
<td>3.6 Conclusion</td>
<td>117</td>
</tr>
<tr>
<td>4. Stochastic Control Strategies for EV Charging Station Control and Frequency Regulation</td>
<td>120</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>120</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 EV Charging Levels Defined by SAE J1772-2011 and Charging Times</td>
<td>3</td>
</tr>
<tr>
<td>Estimated on a 50 kWh Li-ion Battery</td>
<td></td>
</tr>
<tr>
<td>3.1 Fractional Factorial Experimental Design</td>
<td>105</td>
</tr>
<tr>
<td>3.2 Average Peak Load and Number of Hours Over the Year That the</td>
<td>113</td>
</tr>
<tr>
<td>Transformer is Overloaded With Station-Control Model and First-</td>
<td></td>
</tr>
<tr>
<td>Come, First-Served and Uniform Charging Heuristics</td>
<td></td>
</tr>
<tr>
<td>3.3 Average Unserved Electric Vehicle-Charging Load Over the Year</td>
<td>113</td>
</tr>
<tr>
<td>With Constrained First-Come, First-Served Heuristic</td>
<td></td>
</tr>
<tr>
<td>4.6 Test Cases</td>
<td>140</td>
</tr>
<tr>
<td>4.7 Control Model Frequency Regulation Performance in Case 1</td>
<td>143</td>
</tr>
<tr>
<td>4.8 Daily Operation Cost, Transformer Overload Time and Positive</td>
<td>144</td>
</tr>
<tr>
<td>Load Time Sensitivity Analysis</td>
<td></td>
</tr>
<tr>
<td>4.9 Quality of Frequency Regulation Service</td>
<td>144</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Case study simulation flowchart.</td>
<td>38</td>
</tr>
<tr>
<td>2.2</td>
<td>Optimal charging station locations as a function of ( p ).</td>
<td>41</td>
</tr>
<tr>
<td>2.3</td>
<td>Expected number of EVs caught as a function of ( p ).</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>95% confidence interval on relative optimality gap of solution given by SFCLM as a function of ( p ).</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>Comparison of Optimal Station Network with ( r = 1.6 ) and 6.4 km.</td>
<td>45</td>
</tr>
<tr>
<td>2.6</td>
<td>Expected number of EVs with typical daily driving tours longer than 86 miles, and percentage of those EVs that can complete their tours with a six-station EV-charging network if drivers can deviate at most a given distance.</td>
<td>46</td>
</tr>
<tr>
<td>2.7</td>
<td>Expected number of EVs with typical daily driving tours longer than a given EV range, and percentage of those EVs that can complete their tours with a six-station EV-charging network.</td>
<td>48</td>
</tr>
<tr>
<td>2.8</td>
<td>Expected number of EVs able to complete their daily tours with and without a six-station EV-charging network.</td>
<td>49</td>
</tr>
<tr>
<td>2.9</td>
<td>Cumulative distribution of the number of EVs in each of six stations with 100-kW chargers assuming 85-mile EV driving ranges and 20% SOC thresholds.</td>
<td>51</td>
</tr>
<tr>
<td>2.10</td>
<td>Cumulative distribution of the number of EVs in each of six stations with 200-kW chargers assuming 85-mile EV driving ranges and 20% SOC thresholds.</td>
<td>52</td>
</tr>
</tbody>
</table>
2.11 Optimal charging station locations from SFCLM and DFCLM with 
\( p = 6 \) ......................................................... 54

2.12 Optimal charging station locations from SFCLM and DFCLM with 
\( p = 31 \) ......................................................... 55

2.13 Expected number of EVs caught with charging station locations chosen 
by SFCLM and DFCLM as a function of \( p \) ............................................. 56

3.1 Modeled PV Availability From a 200 kW Panel on 5 December, 2013: 
Actual and Forecast Mean, Three-Standard Deviation Band, and Scenario-
1 Forecast as of 15:11 ...................................................... 102

3.2 Transformer-Load Profile From Station-Control Model and Using First-
Come, First-Served Heuristic ................................................. 107

3.3 Expected Operational Life of 500 kVA Transformer Using a First-
Come, First-Served Heuristic With Different Electric Vehicle Penetration 
Levels and Charging Thresholds ........................................... 109

3.4 Load-Duration Curve of 1000 Hours of the Year With Highest Trans-
former Loads Under Station-Control Model and Three Heuristics ... 112

3.5 Average Load-Duration Curve of 1000 Hours of the Year With Highest 
Transformer Loads With and Without Battery Energy Storage System 115

3.6 Average Load-Duration Curve of 1000 Hours of the Year With Highest 
Transformer Loads With and Without Photovoltaic System .......... 116

3.7 Average Load-Duration Curve of 1000 Hours of the Year With Highest 
Transformer Loads With 120 kW and 50 kW Electric Vehicle Chargers 118

4.1 Comparison of Realized Non-EV Load and ARIMA Forecasting Non-
EV Load ................................................................. 135

4.2 Comparison of Realized Frequency Regulation Capacity and ARIMA 
Forecasting Regulation Capacity .......................................... 137

4.3 Comparison of Realized One-Hour Average Photovoltaic Generation 
and ARIMA Forecasting Generation ..................................... 138
4.4 Frequency Regulation Capacity and Fulfillment and BESS State of Charge Profiles .................................................. 141

4.5 Transformer Positive Load Duration of Case 1 ................................. 142
Chapter 1: Introduction

1.1 Background

Electric Vehicles (EVs) are one possible solution to improving transportation energy consumption efficiency. This is because of two main characteristics of an electric motor, which is suitable for current road transportation conditions. First, electric motors have high torque at low rotational speed when compared to internal combustion engines, which has to stay at idle at temporary stop. Second, a electric motor has wide high torque output range, which simplify the vehicle transmission system and reduces the energy losses in complex transmission devices. Miller et al. (2011) report that the average EV energy conversion efficiency is between 59% and 62% from the grid to power at the wheel, compared to 17-21% for conventional gasoline vehicles.

Recent aggressive policies in the U.S., Europe and China under greenhouse-gas emission-reduction pressures have lead to fast-growth of EV markets in these countries. In the U.S., federal government offers a tax credit for electric vehicle production and purchase. The successes of startup companies in the automotive industry, such as Tesla, also accelerate EV market’s growth. Since the year 2013 most automotive manufacturers have promoted their own models of plug-in electric vehicles, as well as hybrid electric vehicles that are modified from their mature vehicle models. Electric
Drive Transporation Association reports that 500,000 electric vehicles were sold in the year 2015 in U.S., within which 30% are pure electric vehicles. EV market has taken 2.87% of the whole automotive market in the U.S..

At the same time, public EV charging station network has also rapidly expanded. The Department of Energys Office of Energy Efficiency and Renewable Energy reports that 14,445 public charging stations have installed in all 50 states by the end of September 2016. Nineteen-hundred of these stations are DC fast charging stations. This broad charging network also contributes to EV market fast growth.

EVs, in general, can be categorized into two types: hybrid electric vehicles (HEVs) and pure plug-in electric vehicles (PEVs). HEVs are powered by a combustion engine, which usually uses gas or diesel, and an electric moter, which uses electricity as “fuel”. HEVs can further be categorized into three subcategories, parallel hybrid, series hybrid, and power-split hybrid, based on their powertrains. A parallel hybrid vehicle, such as the Honda Insight, use either a combustion engine or an electric motor to drive the vehicle. In combustion driving state, the electric motor is usually used as an energy collector which generates electricity to charge a battery in braking and idle modes. In electric driving state, the combustion engine stops working and the vehicle is only drove by the electric motor.

A series hybrid vehicle, such as the Chevrolet Volt, is also refered to extended range electric vehicles. Different from parallel hybrids, this type of vehicle is only driven by the electric motor. The combustion engines only acts as generators to charge the battery. Most series hybrids can be pluged-in and the battery can be charged from the power grid directly.
A power-split hybrid vehicle, such as the Toyota Pruis, uses a planetary gear to decouple the combustion engine and the electric motor. The two engines output power to the wheels through the same powertrain and drive a vehicle together.

A PEVs, different from a HEV, are only powered by electricity, which is installed with a powerful motor and a large capacity battery. Typical PEV models include Nissan Leaf, BMW i3, and Tesla Model S. A PEV is usually charged by power from grid through a charging port. PEVs’ driving ranges are usually between 80 to 300 miles, compared to the more than 400 miles range of conventional vehicles and HEVs.

EV charging standard in the U.S. is proposed in SAE J1772-2011 by Society of Automotive Engineers. It defines an EV charging connector type and 6 charging levels as shown in Table 1.1. The three AC charging levels are usually applied on home and workplace charging because of their long recharging time. This type of charging connectors is easy to install at home and parking lot. DC charging standards are used by commercial charging stations for en-route charging because EVs can be recharged within an acceptable time. The DC fast charging station network can support intercity and interstate travel.

Table 1.1: EV Charging Levels Defined by SAE J1772-2011 and Charging Times Estimated on a 50 kWh Li-ion Battery

<table>
<thead>
<tr>
<th>Charging level</th>
<th>Voltage level(V)</th>
<th>Power Capacity(kW)</th>
<th>Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC Level 1</td>
<td>120 VAC</td>
<td>1.4/1.9</td>
<td>17 h</td>
</tr>
<tr>
<td>AC Level 2</td>
<td>240 VAC</td>
<td>19.2</td>
<td>7 h/3.5 h</td>
</tr>
<tr>
<td>AC Level 3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DC Level 1</td>
<td>200-500 VDC</td>
<td>up to 40 kW</td>
<td>1.2 h</td>
</tr>
<tr>
<td>DC Level 2</td>
<td>200-500 VDC</td>
<td>up to 100 kW</td>
<td>20 min</td>
</tr>
<tr>
<td>DC Level 3</td>
<td>200-600 VDC</td>
<td>up to 240 kW</td>
<td>-</td>
</tr>
</tbody>
</table>
A survey by Franke and Krems (2013b) in Berlin Germany finds that 62% of EV drivers’ prefer to charge vehicles in public charging stations. However, more than 2/3 of the public charging stations in the U.S. are located on the east and west coast. Twenty-two percent of 1,900 DC fast charging stations are installed in California. Because of this situation, vehicle range is an important concern for driving EVs in places lacking a DC fast charging network. Planning a charging station network is very necessary in these places to serve EV adopters and further promotes EV adoption rate.

Moreover, the increasing residential and commercial EV charging load is significantly impacting power systems especially distribution systems. Especially, a DC fast charging station requires high uncertain EV charging power, which challenge the system operations. If a charging station is unable to be finely control, EV charging load will accelerate the transmission devices like the transformer aging. This effect limits the station design capacity. One option to resolve this problem is to upgrade the distribution system. However, this method neglects the transmission devices’ potential to deal with these flexible peak loads which only last several minutes. By using optimal control strategies, these capacity potentials can be fully used.

Charging station can also act as an EV aggregator which aggregates EV charging demands and represents EV owners to bid power in energy market. Using appropriate bidding and control strategies, the station operator can arbitrage in an energy market, significantly reduce EV charging cost, and maximize the station’s profit. In addition, if a station is installed with batteries or the connected EVs are used as distributed energy storage by vehicle-to-grid technique, an EV aggregator is able to participate in an ancillary service market by providing frequency regulation capacity and responding real-time regulation signals. From power system operator’s perspective, integrating
large-scale EV charging stations which provides frequency regulation services will increase the integration scale of uncertain renewable energy.

1.2 Dissertation Scope and Contributions

This dissertation addresses two EV charging network related problems: designing charging station network optimal layout and Creating EV charging station control strategies. Chapter 2 focuses on charging station network design problem. Chapter 3 and 4 resolve the problem of the optimal schedule for EV charging, strategies to purchase and sell electricity, and how to provide frequency regulation services.

1.2.1 Charging Network Optimal Layout Design

An issue that limits consumer adoption of EVs is “range anxiety”. “Range anxiety” refers to drivers’ concerns with being stranded with a discharged EV battery and the associated delays to their journeys due to long recharging times. “Range-anxiety” is one of the most important concerns for those who plan to choose EVs, which can be significantly relieved by a well-developed charging station network. Chéron and Zins (1997), Eberle and von Helmolt (2010), Franke and Krems (2013a) find that many consumers are unwilling to purchase EVs if a ‘safety net’ of charging stations that covers their anticipated driving needs is not available.

Chapter 2 introduces a stochastic flow capturing station location model (SFCLM) to plan the location of DC fast charging stations. Different from previous flow-based station location models introduced by Hodgson (1990) and extended by Kuby and Lim (2005), Kim and Kuby (2012) and Wang and Lin (2009), the proposed SFCLM extends the flow-based model and overcomes several issues in flow-based models. First, EV charging station location optimization is complicated by uncertainties in
where future EV charging demands will be geographically located within the study region. The proposed SFCLM overcomes this future uncertainty by bootstrapping EV driving patterns from tour-record data. Second, the previous flow-based models are unable to deal with large-scale vehicle flow uncertainty which should be considered when planning EV charging infrastructure. This issue is resolved by incorporating a simulation-based stochastic optimization method to solve SFCLMs. Finally, the robustness of EV charging network in uncertain EV-flow context is shown.

We also apply an SFCLM in a case study in Central Ohio with different charging deviation distance assumptions. We find that the optimal charging station locations of choosing smaller number of stations is optimal when choosing a larger number of stations. We also find that if EV drivers would like to deviate 10 minutes from their intended routes 80% EVs can be served by a relatively small network.

1.2.2 Charging Station Control Strategies

Clement-Nyns et al. (2010) find that uncontrolled EV charging (i.e., EVs charging without any coordination) can result in power losses and voltage deviations on the local distribution network. Razeghi et al. (2014) study the impacts of charging 10 EVs in an uncontrolled fashion on a residential transformer. They demonstrate that the resulting charging loads can result in catastrophic failure of the transformer and conclude that management of charging loads is critical for prolonging transformer life. Weiller (2011) estimates the impacts of EV recharging using standard wall outlet on electric loads in the U.S. The results of this analysis show that residential, workplace, and retail shopping center loads can be increased by 74%.
Chapters 3 and 4 introduce centralized EV charging station control models which use stochastic optimization models to schedule EV charging load within flexible charging windows, create electricity purchasing and selling strategies, and optimize responses to frequency regulation signals. The optimal strategy minimizes the station operation cost based on forecasting energy market states (e.g., electricity prices, ancillary service prices, frequency regulation signals, etc.) and simulating charging station state (e.g., EV arrival times, PV generator’s outputs, etc.). Compared with the typical previous work of Pantoš (2012) and Momber et al. (2015) which use stochastic program with EV-usage uncertainty, the proposed models in these two chapters account for renewable generation and energy market uncertainty besides EV-usage uncertainty. Moreover, the methods propose simulation-based algorithms instead of heuristic scenario reduction techniques which guarantee high-quality control strategies with stochasticities.

The proposed methods are applied on a DC fast charging station location in central Ohio. Several configureations of DC fast charging station are simulated for one year. The proposed optimal control strategies can significantly reduce transmission system overloads and prolong the station tranformers’ lives compared with other heuristic methods.
Chapter 2: A Stochastic Flow-Capturing Model to Optimize the Location of Fast-Charging Stations with Uncertain Electric Vehicle Flows

2.1 Introduction

An issue that limits consumer adoption of EVs is ‘range anxiety.’ Range anxiety refers to drivers’ concerns with being stranded with a discharged EV battery and the associated delays to their journeys due to long recharging times. Chéron and Zins (1997), Eberle and von Helmolt (2010), Franke and Krems (2013a) find that many consumers are unwilling to purchase EVs if a ‘safety net’ of charging stations that covers their anticipated driving needs is not available.

The EV charging infrastructure in many parts of the world is not sufficiently well developed to alleviate range anxiety, calling for added infrastructure development. An important issue in developing such an infrastructure is the type of charging technology used. EV charging technologies can be broadly broken into two categories. The first includes what are known as Level-1 and -2 charging technologies. These technologies deliver low-power charging (typically less than 6.6 kW), meaning that EV recharging times are long. For instance, a Nissan Leaf with 30 kWh battery could take as many as 15 hours to fully recharge using Level-1 technology. Thus, Level-1 and -2 technologies
are amenable to use for charging overnight or at a destination where an EV is parked for a prolonged period of time (e.g., a workplace in the case of a commuter vehicle).

The other category includes DC fast-charging technologies, which deliver high-power charging. These technologies typically deliver between 50 kW and 200 kW of charging power and could fully recharge a Nissan Leaf in as little as seven minutes. Given the substantively shorter charging times, fast-charging technologies could be used for charging *en route* from origin to destination, in addition to charging at a vehicle destination. Given these fundamental differences in the use of slow- and fast-charging technologies, different techniques are appropriate to optimize the location of these two charging station types.

EV charging station location belongs to the broader class of refueling-infrastructure problems. The existing literature on refueling infrastructure problems can be categorized into two techniques: set-covering and flow-capturing. These two methods rely on models with a similar underlying mathematical structure. The difference between the methods, in the context of refueling-infrastructure-location problems, is how one models refueling demand. In set-covering models, refueling requirements are treated as being demands at specific spatial locations. The objective of such models can include maximizing demands that are covered or minimizing refueling distances. Because of these properties, the set-covering method is mostly applied to home-based refueling-infrastructure siting. The set-covering method is used by Frade et al. (2011) and Xi et al. (2013). Frade et al. (2011) introduce a maximal-covering model to locate Level-1 and -2 EV charging stations in Lisbon, Portugal. Their model forecast charging demand from each region and maximizes the amount of EV charging demand that is satisfied by the network of stations based on a maximum charging-distance.
tolerance. Different from Frade et al. (2011), Xi et al. (2013) develop a simulation-optimization model to locate Level-1 and -2 public charging stations and consider EV charging times. One novelty of the work of Xi et al. (2013) is that their model determines the number of chargers in each station and simulates the EV charging and queuing process. A recently applied set-covering based median method to plan refueling infrastructure is the c-means clustering approach, which is proposed by Shi and Zheng (2014). In this method, EV charging demands are represented by coordinates in a two-dimensional plane. The demands are then clustered using the fuzzy c-means method and the refueling stations are assumed to be placed at the cluster centers.

Flow-capturing models assume that the refueling demands are from vehicle flows and the refuelings occur *en route*. Hodgson (1990) is one of the earliest examples of a flow-capturing location model which maximizes vehicle flows served by the network. Kuby and Lim (2005) extend this work by incorporating a vehicle-range parameter and introduce additional binary variables to indicate the combination of location candidates. This model is further extended by Kim and Kuby (2012) and includes shortest-time based deviation. Tan and Lin (2014) extend the flow-capturing location model by considering uncertain vehicle flows. Different from previous work which maximize vehicle flow served by the refueling station network, Wang and Lin (2009) propose a flow-capturing model to minimize the total cost of locating alternative-fuel stations. Their model constrains the facility locations to ensure that they cover the alternative-fuel vehicle flows along their shortest paths from origin to destination. Li et al. (2016) proposed a dynamic multi-period multi-path refueling-location model to optimize an EV charging network. Their model focuses on EV use for intercity
trips and minimizes the cost of installing new stations and relocating existing stations. He et al. (2015) introduce a bi-level model to locate EV charging stations while minimizing social cost under budget constraints. Nicholas et al. (2004) and Lin et al. (2008) introduce a flow-capturing based median model to optimize the refueling station using convenience level. Dong et al. (2016) propose a flow-capturing based median model to design EV fast charging infrastructure’s layout. This method is based on shared nearest neighbor clustering algorithm for station allocation and an optimization model for station sizing. One difficulty in the use of flow-capturing models is that they are normally formulated as computationally complex mixed integer programming models. He et al. (2015) address this computational issue using a genetic algorithm while Li et al. (2016) use a heuristic-based genetic algorithm.

As noted above, fast charging infrastructure is much more amenable to use en route from origin to destination than slower Level-1 and -2 technologies. Because the charging demands occur as vehicles are en route, flow-capturing methods are more appropriate for fast-charging station location planning than set-covering. This is because set-covering methods assume that charging demands occur at fixed locations. Moreover, Kuby et al. (2013) find that 88.2% drivers of alternative-fuel vehicles refuel their cars en route. In this work, we seek an optimization based-approach to station location planning. This is because refueling infrastructure is a capital-intensive and long-lived investment. These properties mean that charging stations must be placed judiciously to ensure that their upfront costs are recovered over their lifetimes.

EV charging station location optimization is complicated by uncertainties in where future EV charging demands will be geographically located within the study region. One approach to overcoming this issue, which we take, is to bootstrap EV driving
patterns from tour-record data. Regional planning commissions or similar organizations typically have access to or maintain such datasets, which are used for economic and infrastructure forecasting and planning. An important limitation of tour-record data is that they do not typically specify all of the information needed for EV modeling. For instance, tour records often report the origin and destination of a vehicle trip, but not the route taken. We use a shortest-path model to translate tour records into vehicle paths for purposes of EV charging demand modeling.

Another important limitation of existing flow-capturing models is that they do not adequately capture uncertainty in where EV charging demands appear in the study region. Although Tan and Lin (2014) use a stochastic flow-capturing model in their analysis, they assume a small scenario tree with on the order of six EV-flow scenarios. Our solution method dynamically determines the scenario tree size to provide a desired tolerance on the error in estimating the problem’s objective function value. As a result, we solve problems with on the order of 1000 scenarios. Our method also allows us to provide statistical bounds on the objective function value. As such, our model and solution approach can be said to provide charging station locations that are more robust to uncertainties in EV-charging demand.

The remainder of this chapter is organized as follows. In Section 2.2 we detail the approach taken to modeling fast EV charging station location and the input data that our method uses. An important consideration in developing our modeling approach is that it can be used with datasets that regional planning commissions and similar organizations typically have access to. This design feature allows our model to be used widely. We model EV flows by bootstrapping from tour-record data and use a shortest-path model to translate tour records into complete EV trips. EV charging
station locations are optimized using a two-stage stochastic flow-capturing location model (SFCLM). The first stage of the problem is when the charging stations are placed and the second stage determines which of the EV flows, which are unknown in the first stage, are ‘caught’ by the stations built. The first-stage objective is to maximize the expected number of EVs that can be caught by the stations that are built. Because the first-stage objective function cannot be directly computed, we use a sample-average approximation (SAA) method and an averaged two-replication procedure (A2RP) to solve the SFCLM and provide statistical bounds on the quality of the solution found.

Section 2.3 introduces the case study, based on Central Ohio, to which we apply the SFCLM. All of the inputs used to build the SFCLM are obtained from the Mid-Ohio Regional Planning Commission (MORPC). Because we develop our modeling approach around data available from a regional planning commission, it can be widely used beyond the case study presented here.

Section 2.4 summarizes our case study results. We find that the optimal station locations are fairly robust in the sense that as more stations are built, the optimal locations of the starting stations do not change. We demonstrate the benefit of the charging-station network in terms of how many EVs are able to complete their daily trips with and without access to the public charging stations. We also conduct simulations to estimate how many chargers would need to be installed at each charging station to allow their use without EVs having to queue. We finally compare the SFCLM to a deterministic version of the model. We show that if there are a limited number of charging stations to be built, the SFCLM outperforms the deterministic model in terms of the expected number of EVs caught and should be used in place of a
deterministic model. As the number of stations to be built increases, the SFCLM and the deterministic model choose very similar station locations. Thus, the deterministic model could be used as the SFCLM provides little incremental value and greater computational cost. Section 2.5 concludes this chapter.

2.2 Modeling and Solution Approach

Our model determines the location of a fixed number of fast EV charging stations within a given planning region to maximize the expected volume of EVs that can be served by the station network. As noted above, an important consideration in developing this model is that it use data that are readily available from regional planning commissions or similar organizations that maintain transportation-infrastructure and infrastructure-usage data. Our model accounts for uncertainty in EV flows, given that the charging infrastructure is being built in anticipation of future EV adoption. Moreover, our work considers limited EV range and drivers' behavior in deviating from their intended shortest paths from origin to destination to use a charging station.

Section 2.2.1 details the necessary inputs to our SFCLM and gives its formulation. We then describe in Section 2.2.2 the data sets that are readily available from most regional planning commissions. We also detail the assumptions and data processing done to convert these data into the necessary inputs to our SFCLM. Section 2.2.3 summarizes the challenges in solving the SFCLM and the method that we employ to obtain a near-optimal solution and statistical bounds on its quality.
2.2.1 Stochastic Flow-Capturing Location Model Formulation and Required Data

The key inputs to an SFCLM are: (i) EV flow volumes on different activity chain (ACH) on the road network, (ii) a set of candidate locations where charging stations can be built, and (iii) an assumption regarding how EV drivers would use charging stations built. An ACH is a chain of the trips constituting a daily tour. Thus, an ACH is similar to the definition of trip/activity chain, as introduced in Kang and Recker (2009). However, our ACH does not include timestamps, unlike a trip/activity chain.

As noted before, we formulate this problem as a two-stage stochastic integer linear program. The locations of a fixed number of charging stations are determined in the first stage. This is done with the objective of maximizing the expected number of EVs that can be ‘captured’ by the charging station network. ‘Capturing’ an EV means that its driver would be willing to deviate from his or her original intended route to one of the charging stations built, if the vehicle battery’s state of charge (SoC) is sufficiently low. The second stage determines which EVs can be captured by the charging stations built. The scenarios modeled in the second stage represent different possible volumes of vehicle flows along each ACH. The scenarios are intended to capture uncertainty in where EV charging demands may appear when making charging station-placement decisions.

To formulate the SFCLM, we first introduce the sets and parameters:

- $Q$: set of ACH indices on the road network;

- $K$: set of all candidate charging station locations on the road network;
$N_q$: set of candidate charging station locations that are capable of capturing EVs traveling on ACH $q$;

$p$: maximum number of stations to be built in the region; and

$\tilde{f}_q$: number of EVs traveling on ACH $q$.

Note that $\tilde{f}_q$ is a second-stage parameter, which is random in the first stage. We let $\tilde{f} = (\tilde{f}_1, \tilde{f}_2, \cdots, \tilde{f}_{|Q|})$ denote a vector of all of the $\tilde{f}_q$’s. We also define the decision variables:

$x_k$: binary variable that equals 1 if a station is placed at location $k$ and equals 0 otherwise; and

$y_q$: binary variable that equals 1 if EVs on ACH $q$ are ‘captured’ by a charging station and equals 0 otherwise.

The SFCLM is then formulated as:

\[
\begin{align*}
\max_x & \quad \mathbb{E} \left[ g(x, \tilde{f}) \right] \\
\text{s.t.} & \quad \sum_{k \in K} x_k \leq p; \quad (2.2) \\
& \quad x_k \in \{0, 1\}, \quad \forall k \in K; \quad (2.3)
\end{align*}
\]

where:

\[
\begin{align*}
g(x, \tilde{f}) &= \max_y \sum_{q \in Q} \tilde{f}_q y_q \quad (2.4) \\
\text{s.t.} & \quad \sum_{k \in N_q} x_k \geq y_q, \quad \forall q \in Q; \quad (2.5) \\
& \quad y_q \in \{0, 1\}, \quad \forall q \in Q. \quad (2.6)
\end{align*}
\]
The first-stage objective function (2.1) maximizes the expected volume of EVs that can be captured by the charging station network. This objective is implicitly defined in terms of the optimal objective function value of the second-stage problem, which is given by (2.4)–(2.6). The first-stage problem includes constraint set (2.2), which limits the number of stations built, and integrality constraint set (2.3).

The second-stage objective function (2.4) maximizes the volume of EVs that can be caught by the charging station network under a particular scenario of EV flows. The different EV flow-scenarios are defined by the second-stage parameter, \( \tilde{f}_q \), which is random in the first stage. As noted before, \( g(x, \tilde{f}) \) is defined as the optimal second-stage objective value and defines the first-stage objective function. The second-stage problem includes constraint set (2.5), which captures the spatial relationship between the candidate charging station locations and EV flows. This constraint set says that EVs traveling along ACH \( q \) can only be captured if at least one charging station in the set \( N_q \) is built. Constraint set (2.6) imposes integrality on the second-stage EV flow-capture variables.

The SFCLM has relatively complete recourse. This means that for every feasible first-stage solution there exists at least one feasible second-stage solution. Relatively complete recourse is trivially easy to show, because \( y = 0 \) is always feasible in the second-stage problem. This result implies that \( g(x, \tilde{f}) > -\infty \) for any realization of \( \tilde{f} \). Because \( g(x, \cdot) \) is measurable, we have that:

\[
\mathbb{E}\left[g(x, \tilde{f})\right] > -\infty,
\]

ensuring that the first-stage objective function is well defined.
2.2.2 Input Data and Data Processing

Regional planning commissions typically have two datasets that are valuable inputs to our SFCLM. The first is a graph model of the road network, which captures all of the major intersections and road segments. In some cases, vehicle speed data for road segments are also available. The second is vehicle use data, in the form of tour records. Tour-record data provide a detailed accounting of typical daily vehicle usage patterns. Tour-record data normally specify the origin and destination of, mode of transportation (e.g., privately owned vehicle, public transportation, or ridesharing) used in, and departure and return time of each trip. Trips taken by privately owned vehicles, which are the ones we focus on in our analysis, are associated with vehicle identifiers. Combining all of the trips assigned to a particular vehicle identifier allows the full sequence of trips taken with that vehicle during a typical day to be determined.

We model EV trips by bootstrapping from the tour-record data. This is done by randomly sampling which vehicles in the tour-record data are EVs and which are not. In our analysis, we apply the assumed penetration rate of EVs uniformly across the vehicles in the tour-record data. More sophisticated sampling methods can be used, however. For instance, Xi et al. (2013) randomly sample EVs from tour-record data, but account for the effects of socioeconomic and demographic factors on EV adoption. They do this by adapting a regression model to determine the relative propensity of people living in different areas of their study region to adopt EVs. They then apply a geographically varying penetration rate to different areas of the study region to sample which vehicles are EVs. Alternate approaches could use spatial regression models to capture direct spatial effects on clustering of EV ownership.
Once the vehicles in the tour records are sampled to determine which subset are EVs, $\tilde{f}$, the EV flows along different ACHs, can be computed. An important limitation of tour-record data, however, is that they do not specify the route taken by a vehicle to get from origin to destination. We associate driving routes to each vehicle trip in the tour-record data by employing the following assumption that each driver takes the shortest path from origin to destination.

**Assumption 1** (Shortest-Path). *Each vehicle drives along a shortest path for each trip in a tour.*

Assumption 1 allows us to determine the route of each EV traveling on an ACH by applying a shortest-path model to the graph representation of the road network. If vehicle speed data for the road segments are available, the shortest-path problem can be formulated to provide a minimum-time path. Otherwise, if road speed data are not available, the shortest-path problem can be formulated to provide a minimum-distance path.

In addition to modeling the path that an EV takes from its origin to its destination, we must determine EV driver’s range-anxiety, a threshold SoC, below which an EV driver tends to recharge his or her vehicle. Moreover, we must also determine what charging stations each EV can be captured by on each trip. We do so by appealing to the following additional range-anxiety and charging deviation assumption.

**Assumption 2** (Range-Anxiety Threshold). *An EV driver will seek to recharge his or her vehicle if its SoC falls below $\tau\%$.*

Assumption 2 suggests that only long-distance EV tours are considered if we assume that each EV begins its daily tour with a fully charged battery. That is, the
model only includes EVs with a sufficiently long daily tour for the battery SoC to fall below $\tau \%$ in the course of the day. For the other EVs, their daily tours are too short for their drivers to seek to use a charging station.

**Assumption 3 (En-route Recharging).** _Only EVs with daily tours require a maximum of one battery recharging per day are considered._

Assumption 3 is based on two facts. First, it is unnecessary to account for EVs that recharge more than once per day for daily commute. The National Household Travel Survey reports that the longest tract-average commute distance is 75.05 miles, which is within the range given by a single daily EV recharging. MORPC tour-record data for Central Ohio show that only 0.008% of vehicles have daily tours that exceed the range offered with one daily recharging.

Second, it is intractable to solve a problem in which EV recharges more than once if the candidate station-location pool is sufficiently large. A typical method to modeling multiple recharging is to use flow-refueling location model (FRLM) introduced by Kuby and Lim (2005). The FRLM introduces additional binary variables to model EV range. The number of binary variables exponentially increases in the number of candidates charging station locations. Thus, for example, our case study, which considers 222 station candidate locations would require approximately one million binary variables to represent candidate station combinations. Such a problem would be intractable in a stochastic optimization framework. Interested readers are referred to the work of Kuby and Lim (2005) for further details on the FRLM.

**Assumption 4 (Charging Deviation).** _An EV can be captured by a charging station on a trip if it passes within a fixed $r$ km radius of the charging station while on its path from its origin to destination._
Assumption 4 says that an EV driver is willing to use an EV charging station that is a fixed maximum distance ‘out of the way’ from his or her planned path from origin to destination. Otherwise, if a charging station is too far out of the way, the driver would not use it due to the inherent inconvenience of doing so. Kim and Kuby (2012) calculate the shortest time-based deviation distance and use an exponentially decayed capturing rate. We opt to simply use a fixed maximum distance and analyze the effects of driver deviation tolerance. This is because the time complexity of calculating a shortest-time-based deviation distance is of exponential order with respect to the number of ACHs and the candidate pool. Li and Huang (2014) also find that the computational effort of solving a multipath refueling location model exactly is exponentially increasing. A capturing circle with a fixed maximum distance is accurate enough to estimate deviation distance, especially in a metropolitan area.

For notational convenience, we define $L(q)$ as the length of ACH $q$ and $R$ as the EV range. We also define $L_{O_k}(q)$ as the shortest-time distance between $q$’s origin and the first intersection of the route and candidate location $k$’s “capturing” area. Combining Assumptions 1-4 with a road-network graph, Algorithm 1 determines $N_q$, the set of candidate charging station locations that can capture EVs traveling between different ACHs. Step 5 ensures that candidate $k$ can serve $q$ if the driver has “range anxiety” because of low SoC. Step 11 guarantees that a long tour $q$ is able to complete with en-route recharging at candidate station $k$.

### 2.2.3 Solution Technique

The SFCLM requires the second-stage scenarios, which are unknown when making first-stage charging station-placement decisions, representing possible EV flows. The
Algorithm 1 Create $N_q$

1: for $q \in Q$ do
2: \hspace{0.5cm} $N_q \leftarrow \varnothing$
3: \hspace{0.5cm} if $\mathcal{L}(q) \leq \mathcal{R}$ then
4: \hspace{1cm} for $k \in K$ do
5: \hspace{1.5cm} if $\mathcal{L}_{O_k}(q)/\mathcal{L}(q) \in [\tau/100, 1]$ then
6: \hspace{2cm} $N_q \leftarrow N_q \cup \{k\}$
7: \hspace{1cm} end if
8: \hspace{1cm} end for
9: \hspace{0.5cm} else if $\mathcal{R} < \mathcal{L}(q) \leq 2\mathcal{R}$ then
10: \hspace{1cm} for $k \in K$ do
11: \hspace{1.5cm} if $\mathcal{L}_{O_k}(q) \in [\max\{\tau\mathcal{L}(q)/100, \mathcal{L}(q) - \mathcal{R}\}, \mathcal{R}]$ then
12: \hspace{2cm} $N_q \leftarrow N_q \cup \{k\}$
13: \hspace{1cm} end if
14: \hspace{1cm} end for
15: \hspace{0.5cm} end if
16: end for

first-stage objective function is also implicitly defined by the second-stage scenarios.

As noted above, the input data used to model second-stage EV flows are generated by bootstrapping from the tour-record data. This amounts to generating numerous scenarios, in which different vehicles represented in the tour-record data are EVs. The number of such scenarios is exponential in the size of the tour-record data, however. Thus, the SFCLM would be intractable if every possible second-stage scenario is included in the model. For instance, the tour-record data for the Central Ohio case study that we introduce in Section 2.3 has about 1.3 million personal vehicles. This would yield approximately $9.9 \times 10^{291338}$ second-stage scenarios if all possible EV/non-EV combinations are modeled.

We propose an SAA method to deal with this model intractability. Broadly speaking, an SAA method rewrites the original SFCLM and approximates the first-stage
objective function by using a small subset of randomly generated second-stage problems. The second-stage problems are generated by random sampling and the sample size is dynamically determined to achieve a desired tolerance on the error in estimating the first-stage objective function.

To formally define the SAA, we first define $f_1, f_2, \ldots, f_M$ as $M$ random samples of the random vector, $\tilde{f}$. Using these $M$ realizations, the sample average function is defined as:

$$\bar{g}_M(x) = \frac{1}{M} \sum_{m=1}^{M} g(x, f^m).$$

The corresponding SAA problem is then defined as:

$$\max_{x \in X} \bar{g}_M(x), \tag{2.7}$$

where $X$ is the first-stage feasible region, which is defined by constraint sets (2.2) and (2.3). Note, also, that $X$ is invariant to the random samples, $f_1, f_2, \ldots, f_M$, that implicitly define the objective of the SAA problem, because of relatively complete recourse.

Algorithm 2 gives a high-level overview of the steps taken to solve the SAA problem. The SAA method begins with two inputs—an initial sample size and a desired relative tolerance. The method then randomly generates samples and solves the resulting SAA problem. If the relative optimality gap of the solution to the SAA problem is sufficiently small, then the algorithm terminates. Otherwise, $M$ is increased and the algorithm begins with a new set of random samples. We now detail the L-shaped method used to solve the SAA method and our approach to assessing solution quality in Steps 3 and 4 of the algorithm. We then discuss the convergence properties of Algorithm 2.
Algorithm 2 SAA Method

1: **input:** initial sample size $M$, relative tolerance $d$
2: generate i.i.d. random samples, $f^1, f^2, \ldots, f^M$
3: solve SAA problem (2.7) using L-shaped method with branch-and-cut
4: assess quality of SAA problem solution
5: if relative optimality gap $\leq d$ then
6: stop
7: else
8: increase $M$ and return to Step 2
9: end if

Solution of SAA Problem

We solve the SAA problem in Step 3 of Algorithm 2 using a branch-and-cut method. Moreover, an L-shaped method is used to solve subproblems at each node of the branching tree. To explain our solution method, we rewrite the SAA problem as:

\[
\max_{\theta, x} \frac{1}{M} \sum_{m=1}^{M} \theta^m \quad (2.8)
\]
\[
\text{s.t. } \sum_{k \in K} x_k \leq p; \quad (2.9)
\]
\[
x_k \in \{0, 1\}, \quad \forall k \in K. \quad (2.10)
\]

This problem has a new set of decision variables, $\theta^1, \theta^2, \ldots, \theta^M$. The variable $\theta^m$ is defined as the optimal second-stage objective function value in scenario $m$. Put another way, $\theta^m$ represents $g(x, f^m)$ in the SAA problem. The scenario-$m$ second-stage problem is formulated as:

\[
g(x, f^m) = \max_y \sum_{q \in Q} f_q^m y_q^m \quad (2.11)
\]
\[
\text{s.t. } \sum_{k \in N_q} x_k \geq y_q^m, \quad \forall q \in Q; \quad (2.12)
\]
\[
y_q^m \in \{0, 1\}, \quad \forall q \in Q. \quad (2.13)
\]
We explicitly place $m$ superscripts on the $y$ variables in this problem, to highlight the fact that the $y$'s could take different values under different scenarios. It is straightforward to show that there is an integral optimal solution to the linear relaxation of second-stage problem (2.11)–(2.13). This can be shown using the total unimodularity of the recourse matrix corresponding to constraint set (2.12). Indeed, because $f^m \geq 0$, it is straightforward to explicitly derive an optimal integral solution to the second-stage problem. For each $q \in Q$, if $x_k = 1$ for at least one $k \in N_q$ then $y^m_q = 1$ is optimal. Otherwise, if $x_k = 0$ for all $k \in N_q$, then only $y^m_q = 0$ is feasible.

It is also straightforward to show that the second-stage problem is separable, because there are no constraints linking the $y$'s with one another. Following from these two properties, problem (2.11)–(2.13) can be rewritten as $|Q|$ linear programs, one corresponding to each ACH, $q$. The scenario-$m$ second-stage linear program corresponding to ACH $q$ is:

$$\max_{y^m_q} \quad f^m_q y^m_q$$

s.t.

$$\sum_{k \in N_q} x_k \geq y^m_q, \quad (\pi^m_q);$$

$$y^m_q \leq 1, \quad (\mu^m_q);$$

$$y^m_q \geq 0.$$

We let $\pi^m_q$ and $\mu^m_q$ denote dual variables associated with constraints (2.15) and (2.16), respectively.

We can derive Benders’s-type optimality cuts, which define the values of the $\theta$’s in the SAA problem in terms of dual information obtained from the second-stage
subproblems. These cuts are defined as:

\[ \theta^m \leq \sum_{q \in Q} \left( \mu^m_{q} + \pi^m_{q} \cdot \sum_{k \in N_q} x_k \right), \quad (2.18) \]

where \( \pi^m_{q} \) and \( \mu^m_{q} \) are optimal values for the dual variables corresponding to problem (2.14)–(2.17).

These optimality cuts are iteratively added to SAA problem, giving what we refer to as the L-shaped SAA problem. Our overall solution approach is as follows. The incumbent L-shaped SAA problem is solved to find values for the \( x \)'s and \( \theta \)'s. The values of the \( x \)'s are then input to the second-stage subproblems and the resulting objective function values are compared to the \( \theta \)'s found in the SAA problem. If the \( \theta \)'s overestimate the true second-stage objective function values, then optimality cuts are added and the new L-shaped SAA problem is re-solved. Otherwise, the algorithm terminates.

To give the formulation of the L-shaped SAA problem, we define \( J_1, J_2, \ldots, J_M \) as the number of optimality cuts for scenarios 1, 2, \ldots, \( M \) that have been added to the SAA problem. We also define \( \pi^m_{q,j} \) and \( \mu^m_{q,j} \) as the optimal dual variable values used to generate the \( j \)th optimality cut for scenario \( m \). The L-shaped SAA problem with \( J \) optimality cuts is then formulated as:

\[
\max_{\theta, x} \quad \frac{1}{M} \sum_{m=1}^{M} \theta^m \\
\text{s.t.} \quad \sum_{k \in K} x_k \leq p; \quad (2.20) \\
x_k \in \{0, 1\}, \quad \forall k \in K; \quad (2.21) \\
\theta^m \leq \sum_{q \in Q} \left( \mu^m_{q,j} + \pi^m_{q,j} \cdot \sum_{k \in N_q} x_k \right), \quad \forall m = 1, \ldots, M, j = 1, \ldots, J_m. \quad (2.22)
\]
L-shaped SAA problem (2.19)–(2.22) is a mixed-integer program, which we solve using the branch-and-cut-based method outlined in Algorithm 3. In Step 1 we initialize the L-shaped SAA problem to have no optimality cuts and then iteratively add any optimality cuts needed in the remaining steps of the algorithm. Step 3 solves the incumbent L-shaped SAA problem with the optimality cuts found thus far. We then, in Steps 4 through 6, iteratively consider each second-stage subproblem and determine if \( \hat{\theta}_m \) overestimates the true second-stage objective value that \( \hat{x} \) gives. If so, a new optimality cut is generated in Steps 7 through 9 and added to the L-shaped SAA problem. This iterative procedure is repeated until the \( \hat{\theta} \)'s provide a proper estimate of the second-stage objective function value, in the sense that the optimality gap is less than some fixed tolerance, \( \rho \) (Step 12).

**Algorithm 3** Branch-and-Cut Solution Method for L-Shaped SAA Problem

1: \( J_m \leftarrow 0, \forall m = 1, \ldots, M \)
2: repeat
3: \( (\hat{\theta}, \hat{x}) \leftarrow \arg \max_{\theta, x} \) (2.19) s.t. (2.20)–(2.22)
4: for \( m = 1, \ldots, M \) do
5: \( y^m_q \leftarrow \arg \max_g \) (2.14) s.t. (2.15)–(2.16) with \( f^m_q \) in (2.14) \( \triangleright \) \( \pi^m_q \) and \( \mu^m_q \) denote corresponding duals
6: if \( f^m_q y^m_q < \hat{\theta}_m \) then \( \triangleright \) add a new scenario-\( m \) optimality cut
7: \( J_m \leftarrow J_m + 1 \)
8: \( \pi^m_q, J_m \leftarrow \pi^m_q \)
9: \( \mu^m_q, J_m \leftarrow \mu^m_q \)
10: end if
11: end for
12: until \( \frac{1}{M} \sum_{m=1}^{M} \hat{\theta}_m - \frac{1}{M} \sum_{m=1}^{M} g(\hat{x}, f^m) < \rho \)
Assessing Quality of SAA Solution

In Step 4 of Algorithm 2 we assess the quality of the SAA solution that is found in Step 3. This is done by appealing to the results of Kleywegt et al. (2002), Bayraksan and Morton (2006). Kleywegt et al. (2002) show the asymptotic properties of the difference between the objective function value of the SAA problem and the true problem. More specifically, let $g^*$ denote the optimal objective function value of the true SFCLM:

$$g^* = \max_x \mathbb{E} \left[ g(x, \tilde{f}) \right]$$

s.t. \( \sum_{k \in K} x_k \leq p; \)

\( x_k \in \{0, 1\}, \quad \forall k \in K; \)

where $g(x, \tilde{f})$ is defined by (2.4)–(2.6). Also let:

$$X^* = \left\{ x \in X \left| \mathbb{E} \left[ g(x, \tilde{f}) \right] = g^* \right. \right\},$$

be the set of $x$’s that are optimal in the true SFCLM, where $X$ is (as before) defined by constraint sets (2.2) and (2.3). Also define $\hat{g}^M$ as the optimal objective function value of the SAA problem with $M$ random samples of the stochastic parameter:

$$\hat{g}^M = \max_{x \in X} \bar{g}_M(x).$$

Kleywegt et al. (2002) prove that for $M$ sufficiently large we have:

$$\sqrt{M}(g^* - \hat{g}^M) \overset{d}{\to} \min_{x \in X^*} Z(x),$$

where $Z(x)$ are normally distributed random variables with mean zero and covariance equal to the covariance of $g(x, \tilde{f})$. This asymptotic property implies that the random
variable, \((g^* - \hat{g}^M)\), is not normally distributed if \(|X^*| > 1\) (i.e., if there are multiple optimal solutions to the true SFCLM).

We can further show, by the concavity of \(g(x, \tilde{f})\) (cf. the proof of Proposition 1 below, which shows that \(g(x, \tilde{f})\) is a concave piecewise-linear function) and from Jensen’s inequality, that \(\hat{g}^M\) is a biased estimator of \(g^*\). More specifically, we have:

\[
\mathbb{E}[\hat{g}^M] \geq \mathbb{E}[\hat{g}^{M+1}] \geq g^*.
\]

A frequent approach to estimating the optimality gap is to employ a batch-mean method. This method provides an optimality gap estimator that is asymptotically normally distributed.

Bayraksan and Morton (2006) suggest an A2RP, which is a batch-mean-based method, to estimate the optimality gap. We define the optimality gap as:

\[
\eta_{x^*_M} = \mathbb{E}[\hat{g}^M] - \mathbb{E}[g(x^*_M, \tilde{f})],
\]

where \(x^*_M\) is the optimal solution obtained in Step 3 of Algorithm 2 (i.e., the final solution obtained in Algorithm 3). A2RP provides valid confidence intervals on the optimality gap of problems that satisfy the following three assumptions.

**Assumption 5.** \(g(\cdot, \tilde{f})\) is continuous on \(X\) with probability 1.

**Assumption 6.** \(\mathbb{E}\left[\sup_{x \in X} (g(x, \tilde{f}))^2\right] < +\infty\).

**Assumption 7.** \(X \neq \emptyset\) and is compact.

We show in the following proposition that the SFCLM satisfies all three of these assumptions.

**Proposition 1.** The SFCLM satisfies Assumptions 5 through 7.
Proof. Assumption 6 follows immediately from analyzing the second-stage problem of the SFCLM, which is given by (2.4)–(2.6). More specifically, note that \( g(x, \tilde{f}) \leq \sum_{q \in Q} \tilde{f}_q \) and that \( \sum_{q \in Q} \tilde{f}_q \) is bounded above by the number of EVs in the study region, which is assumed to be finite. Thus, \( (g(x, \tilde{f}))^2 \) is guaranteed to be finite.

To show Assumption 7 is satisfied, note that \( X \neq \emptyset \) as long as \( p \geq 0 \). Moreover, so long as \( p \) is finite, \( X \) is bounded. We also note that \( X \) is closed because it is defined as an integer lattice. Thus, \( X \) is compact.

Finally, to show that Assumption 5 is satisfied, without loss of generality, we consider the scenario-\( m \) second-stage problem. First, we write the dual of the problem’s linear relaxation and from strong duality know that:

\[
g(x, f^m) = \min_{\pi, \mu} \sum_{q \in Q} \left( \mu^m_q + \pi^m_q \cdot \sum_{k \in N_q} x_k \right)
\]

subject to:

\[
\pi^m_q + \mu^m_q \leq f_q^m, \quad \forall q \in Q; \tag{2.23}
\]

\[
\pi^m_q, \mu^m_q \geq 0, \quad \forall q \in Q. \tag{2.24}
\]

We let \( \Pi \) denote the feasible region of the dual, which is defined by constraint sets (2.23) and (2.24). We know that \( \Pi \) is bounded by inspecting constraint sets (2.23) and (2.24). For notational convenience, define:

\[
\hat{\Pi} = \left\{ (\pi^{m,1}, \mu^{m,1}), (\pi^{m,2}, \mu^{m,2}), \ldots, (\pi^{m,\Lambda}, \mu^{m,\Lambda}) \right\},
\]

as the set of extreme points of polyhedron \( \Pi \). We know that \( \Lambda < +\infty \). Thus, we can rewrite the function \( g(x, f^m) \) as:

\[
g(x, f^m) = \min_{\lambda \in \{1, \ldots, \Lambda\}} \sum_{q \in Q} \left( \mu^{m,\lambda}_q + \pi^{m,\lambda}_q \cdot \sum_{k \in N_q} x_k \right),
\]

which is a piecewise-linear concave function on \( \text{Conv}(X) \). Thus, \( g(\cdot, \tilde{f}) \) is continuous on \( X \). \( \square \)
The A2RP works by comparing $x^*_M$ to an optimal solution that would be obtained from solving an SAA problem with a different sample of the random vector. We outline the A2RP here and refer interested readers to the work of Bayraksan and Morton (2006) for full details. We begin by defining first- and second-moment estimators of $\eta_{x^*_M}$. This is done by defining two new random samples, $\hat{f}^1_1, \hat{f}^2_1, \ldots, \hat{f}^{L/2}_1$ and $\hat{f}^1_2, \hat{f}^2_2, \ldots, \hat{f}^{L/2}_2$, each of size $L/2$. The first-moment estimators are then defined as:

$$G^i_{L/2}(x^*_M) = \max_{x \in X} \left\{ \frac{2}{L} \sum_{l=1}^{L/2} g(x, \hat{f}^i_l) \right\} - \frac{2}{L} \sum_{l=1}^{L/2} g(x^*_M, \hat{f}^i_l), \forall i = 1, 2.\right\}$$

Also define:

$$x^{i*}_{L/2} = \arg \max_{x \in X} \frac{2}{L} \sum_{l=1}^{L/2} g(x, \hat{f}^i_l),$$

as a maximizer of the first term defining $G^i_{L/2}(x^*_M)$. For notational convenience we define:

$$\hat{g}_{L/2}(x) = \frac{2}{L} \sum_{l=1}^{L/2} g(x, \hat{f}^i_l).$$

Thus, we can write:

$$G^i_{L/2}(x^*_M) = \hat{g}_{L/2}(x^{i*}_{L/2}) - \hat{g}_{L/2}(x^*_M), \forall i = 1, 2.$$ 

The second-moment estimators of the optimality gap are given by:

$$s^2_{L/2}(x^{i*}_{L/2}) = \frac{2}{L - 2} \sum_{l=1}^{L/2} \left[ (g(x^{i*}_{L/2}, \hat{f}^i_l) - g(x^*_M, \hat{f}^i_l)) - (\hat{g}_{L/2}(x^{i*}_{L/2}) - \hat{g}_{L/2}(x^*_M)) \right]^2.$$

Next, compute the estimates:

$$\bar{G}_{L/2}(x^*_M) = \frac{1}{2} \left[ G^1_{L/2}(x^*_M) + G^2_{L/2}(x^*_M) \right],$$

and:

$$\bar{s}^2_{L/2} = \frac{1}{2} \left[ s^2_{L/2}(x^{1*}_{L/2}) + s^2_{L/2}(x^{2*}_{L/2}) \right].$$
Bayraksan and Morton (2006) show that for any $\alpha \in (0, 1)$:

$$\left[0, \bar{G}_{L/2}(x^*_M) + \frac{t_{L-1, \alpha} \bar{s}_{L/2}}{\sqrt{L}}\right],$$

where $t_{L-1, \alpha}$ is a $t$-statistic with $L - 1$ degrees of freedom is a valid one-sided $(1 - \alpha)$ confidence interval on $\eta_{x^*_M}$. We state, but do not prove (interested readers are referred to their work for complete details), their formal result showing this.

**Theorem 1.** Suppose Assumptions 5 through 7 hold, that $x^*_M \in X$, and that the samples, $\hat{f}_1^1, \hat{f}_1^2, \ldots, \hat{f}_{L/2}^1$ and $\hat{f}_2^1, \hat{f}_2^2, \ldots, \hat{f}_{L/2}^2$, are independent and identically distributed from the same distribution as $\tilde{f}$. Then $\forall \alpha \in (0, 1)$ we have that:

$$\liminf_{L \to +\infty} \text{Prob}\left\{\eta_{x^*_M} \leq \bar{G}_{L/2}(x^*_M) + \frac{z_{\alpha} \bar{s}_{L/2}}{\sqrt{L}}\right\} \geq 1 - \alpha,$$

where $z_{\alpha}$ is a $z$-score.

**Convergence Properties of SAA Method**

An important question in implementing Algorithm 2 is whether it is guaranteed to converge and, if so, the rate at which it does. We examine this in two parts. First, we examine whether Algorithm 3, which is used to solve the SAA problem in Step 3 of Algorithm 2 converges. We then examine whether we are guaranteed to find a solution with a desired optimality gap in Step 5 of Algorithm 2.

Algorithm 3 is guaranteed to converge because, at its heart, it is a branch-and-cut algorithm. The branching tree of the SAA problem will have at most $2^p$ nodes. Moreover, there are a finite number of optimality cuts that can be added to the L-shaped SAA problem, which is bounded by the number of extreme points of the second-stage subproblems (cf. the proof of Proposition 1). Nevertheless, Algorithm 3 has exponential time complexity. One method to speed up the algorithm, which we
employ, is to add cuts as necessary whenever a solution of the incumbent L-shaped SAA problem is found. That is to say, we do not have to find an integer-optimal solution to the incumbent L-shaped SAA problem in Step 3 of Algorithm 3. Rather, as solutions (either integer-feasible solutions or solutions to the linear relaxation) are found in the branching tree, optimality cuts can be added to the L-shaped SAA problem. This can significantly improve computational performance.

As for convergence of Algorithm 2 in Step 5, we rely on the results of Kleywegt et al. (2002), which can be used to show that Algorithm 2 converges and its convergence rate. We begin by defining:

\[ \hat{X}_M^\epsilon = \{ x \in X \mid |\bar{g}_M(x) - \hat{g}^M| \leq \epsilon \} , \]

as the \( \epsilon \)-optimal solution set of SAA problem (2.7). We similarly define:

\[ X^\epsilon = \{ x \in X \mid |\bar{g}_M(x) - g^*| \leq \epsilon \} , \]

as the \( \epsilon \)-optimal solution set of true SFCLM. Kleywegt et al. (2002) show the following result, which we state but do not prove (interested readers are referred to their work for the details of the result).

**Theorem 2.** The following two results hold:

1. \( \hat{g}^M \to g^* \) with probability 1 as \( M \to +\infty \); and

2. the event \( \{ \hat{X}_M^\epsilon \subset X^\epsilon \} \) occurs with probability 1 for any \( \epsilon > 0 \) so long as \( M \) is sufficiently large.
This theorem implies that so long as we allow for a strictly positive optimality gap and allow $M$ to increase sufficiently, Algorithm 2 is guaranteed to converge. Kleywegt et al. (2002) also have a result showing that the Algorithm has an exponential convergence rate, so long as the following assumption is satisfied.

**Assumption 8.** For every $x_1 \in X$ and $x_2 \in X \setminus X^\epsilon$, the moment-generating function of the random variable, $g(x_1, \tilde{f}) - g(x_2, \tilde{f})$, is finite-valued in a neighborhood of 0.

We show that the SFCLM satisfies this assumption in the following proposition.

**Proposition 2.** The SFCLM satisfies Assumption 8.

*Proof.* We have that:

$$-\infty < 0 \leq g(\cdot, \tilde{f}) \leq \sum_{q \in Q} \tilde{f}_q < +\infty.$$  

The lower bound comes about because the lowest feasible second-stage objective value is attained by setting all of the $y$’s equal to zero. The upper bound come about because in a practical setting, there are a finite number of EVs. Thus, $g(x_1, \tilde{f}) - g(x_2, \tilde{f})$ is bounded and the moment-generating function is finite-valued. \hfill \Box

We now state, but do not prove, the convergence rate result of Kleywegt et al. (2002).

**Theorem 3.** Suppose Assumption 8 holds. Then for all $\epsilon \geq \delta \geq 0$ there exists a $\gamma(\delta, \epsilon) > 0$ such that:

$$\text{Prob} \left\{ \hat{X}_M^\epsilon \not\subseteq X^\epsilon \right\} \leq |X \setminus X^\epsilon| e^{-M\gamma(\delta, \epsilon)}.$$  

34
2.3 Case Study

We apply the modeling approach outlined in Section 2.2 to a case study based on the 12-county Central-Ohio region. This region covers roughly 6000 km$^2$. According to data obtained from MORPC, this region had approximately 1.3 million light-duty vehicles as of the end of 2010.

More specifically, MORPC provides three datasets that we rely on: (i) geographic information system (GIS) data on the Central-Ohio region, including the road network; (ii) tour-record data specifying the use of the 1.3 million light-duty vehicles in Central Ohio on a typical workday; and (iii) average-vehicle-speed data for the road segments of the road network. Importantly, regional planning commissions or similar organization typically have these data for regions in the United States and elsewhere. This means that the approach that we take to construct a case study and apply the SFCLM can be used in other areas beyond the case study presented here.

The GIS data contain three important data types for constructing our case study. One is the boundaries of 1805 traffic analysis zones (TAZs). MORPC models vehicle travel within the tour-record data as occurring between these TAZs. The second is GIS data for the Central-Ohio road network. This consists of roughly 20000 road segments and their lengths, and 26000 intersections between the road segments. Finally, MORPC provides data on the location of the 222 major commercial retail locations in the Central-Ohio region.

In our case study, we consider these 222 retailers as candidate locations for EV charging stations. We select these as candidate locations for three reasons. First, commercial retail locations are likely to have the physical space and existing electrical infrastructure to accommodate fast EV charging stations. A single fast charger may
draw 200 kW of power. Thus, a station with multiple chargers may add more than 1 MW of load to the distribution system, which could not be easily accommodated by a residential circuit. Second, retail locations are likely to have a commercial rationale for EV charging station development. Recharging an EV, even with a fast charging technology, may take longer than refueling a gasoline vehicle. Because an EV driver may be parked at a charging station for several minutes while making retail purchases, this would provide a convenient opportunity for vehicle recharging. Indeed, He et al. (2013) present retailers offering EV charging as a potential business model for developing fast EV charging infrastructure. Co-locating a charging station with a retail shopping location may also provide an incentive for an EV driver to use a charging station that is out of the way from his or her intended route from origin to destination. A third rationale is that the retail locations in Central Ohio are relatively evenly placed within the study region. Thus, the candidate locations provide adequate coverage of where EVs may be driven in the future.

The tour-record data provide modeled data on the use of the 1.3 million light-duty vehicles in Central Ohio during a typical workday. The data are generated using the multi-step tour-based approach described by Sadeghi-Barzani et al. (2014). The data include about 2.4 million daily tours, of which about 2.1 million involve the use of a personal vehicle, made by 688000 households. Each tour is associated with a specific vehicle in the set of vehicles modeled. Each tour includes multiple trips. Each trip is characterized by the vehicle’s starting and ending TAZs, the departure and return time from the starting and ending TAZs, and an indicator of whether the vehicle makes a sub-tour en route in each trip. If there is a sub-tour, the tour record indicates the sub-tour destination TAZ.
For instance, a vehicle may begin and end a tour at TAZ $i$ and be used for commuting to TAZ $j$. The tour record for this particular vehicle would indicate a departure trip from TAZs $i$ to $j$ at a departure time and a return trip from TAZ $j$ to $i$ at a return time. Moreover, if the vehicle stops at a retail shopping location in TAZ $k$ en route from TAZ $j$ to $i$ in the evening, there would be a subtour in the tour record data. Combining the trips and subtour would indicate that the particular vehicle’s tour is $(i \rightarrow j \rightarrow k \rightarrow i)$ with timestamps. An ACH is a vehicle trip chain without timestamps (e.g., $(i \rightarrow j \rightarrow k \rightarrow i)$). Although the tour-record data specify the sequence points that a vehicle travels to, they do not specify the paths taken between the points.

Finally, MORPC has average-vehicle-speed data for a typical workday for all of the road segments in the road network dataset. MORPC reports road speeds for four time periods: 5:00 am to 9:00 am, 10:00 am to 2:00 pm, 3:00 pm to 7:00 pm, and 8:00 pm to 12:00 am.

Figure 2.1 summarizes the steps taken to use the MORPC data to produce our case study. We begin by randomly sampling from among the 1.1 million light-duty vehicles in the tour-record data to determine which are EVs and which are conventional gasoline vehicles. Our case study assumes a 3% EV penetration rate, meaning that we randomly determine with probability 0.03 which of the modeled vehicles in the tour-record data are EVs and which are not. Note that each random sample ultimately corresponds to a realization of the exogenous random variable, $\tilde{f}$, in the SFCLM. Moreover, when we solve the SAA problem or use the A2RP we generate $f^m$’s or $f^t$’s by randomly sampling which vehicles in the tour-record data are EVs.
Next, for each realization of which vehicles are EVs and not, we rely on Assumption 1 to determine the path taken by each EV to get from its origin to destination on each trip. This is done by solving a shortest-path problem. More specifically, we solve a shortest-time-path problem, using the average-vehicle-speed data for each road segment in the road network. MORPC provides vehicle-speed data for four time periods. Thus, for each EV tour we use road-speed data corresponding to the time block during which the trip takes place. The shortest-path problem provides the route taken by each EV for each of its individual vehicle tours.
Then, we use Algorithm 1 and rely on Assumptions 2-4 to determine which EVs could be caught by a charging station at each of the 222 candidate locations. In doing so, we need to determine recharging deviation distance tolerance $r$ and “range-anxiety” threshold $\tau\%$. Kuby et al. (2013) investigate refueling behavior of early alternative-fuel vehicle (AFV) adopters, which shows that the percentage of AFV drivers who refueled with less than $3/8$ left in their tanks is 83.4%. They also find that the median refueling deviation tolerance distance of AFV drivers is between 0.8 and 1.3 miles and more than 75% drivers deviate more than 9 minutes (i.e., about 4 miles with 45 mph average speed).

Based on these findings, we use $\tau\% = 40\%$ as EV driver range-anxiety threshold and use $r = 1.6$ and 6.4 km as station capturing radius. Thus, for each EV trip in a tour, we determine its intended path from the shortest-time-path model. If this intended path comes within $r = 1.6$ or 6.4 km of a particular candidate location and the EV SoC satisfies a criterion (i.e., $[0.38, 1]$ for $L(q) < R$; $\max\{0.38, \frac{L(q)}{R} - 1\}, 1]$ for $R < L(q) \leq 2R$), then that candidate location could ‘capture’ that EV.

### 2.4 Case Study Results

Our case study results are presented in this section. We begin by summarizing the results using our SFCLM. We first report the optimal locations of charging stations and expected number of EVs that can be caught as a function of the number of charging stations built. We also report bounds on solution quality given by the SFCLM. We next present results illustrating the benefit of the public charging station network, in terms of the increase in the number of simulated EVs that can complete their typical daily tours with the charging stations present. We also provide insights
into the number of chargers that must be deployed at the charging stations selected by
the SFCLM for EVs to be able to use the stations. We finally contrast the results of
our SFCLM with a deterministic variant of the model, in which EV flows are assumed
to be known and equal to a sample mean of $\tilde{f}$. Contrasting the results of the two
models estimates the value of the stochastic solution given by the SFCLM. Note that
we use recharging capturing radius $r = 1.6$ km unless any another value is specified.

2.4.1 Stochastic Flow-Catching Location Model Results

We examine the optimal placement of charging stations as a function of the total
number of charging stations built. This is done by solving the SFCLM and varying
the value of $p$. Figure 2.2 shows the optimal locations of six, 11, 21, and 31 charging
stations. The figure shows two important properties of the optimal locations of the
charging stations. First, the optimal charging station locations are fairly robust to
the total number of stations built. That is to say, each of the six locations where a
charging station is built with $p = 6$ is close to an optimal charging station location
with higher values of $p$. This is also true of the stations built with $p = 11$ and $p = 21$.
This finding that the optimal station locations are roughly nested implies that if the
region has limited resources to build an initial charging station network, it can build
those stations without having to be concerned about future station deployments. If
additional stations are built later on, the locations of the starting stations would still
be close to optimal as the network is expanded.

The second property shown in Figure 2.2 concerns the locations of the stations
themselves. The Central-Ohio region is centered around the city of Columbus, which
is relatively (compared to the surrounding region) densely populated. Columbus
is surrounded by a beltway (Interstate 270). The region outside of this beltway is sparsely populated (i.e., it consists of suburban and rural communities). There are also three smaller cities outside of the beltway—Marysville, Delaware, and Newark. If the network is limited to having a small number of charging stations, the stations are located on the periphery of the beltway. This provides good coverage of many EVs, which tend to use the beltway for at least a portion of one trip daily. With \( p = 6 \) there is only one charging station built in the center of the beltway (near the central business district of Columbus).

As the number of charging stations built increases, most of the additional stations are still built on the periphery and inside of the beltway. This is because such locations still provide more coverage of EVs than locations outside of the beltway. As the number of stations built increases further, some are placed in the three smaller cities outside of the beltway.
Figure 2.3 shows the expected number of EVs that can be caught as a function of $p$. With a 3% EV-penetration level, there would be approximately 40000 EVs (in expectation) in the study region. The figure shows that if $p = 222$ (i.e., a charging station is built at every candidate location considered), a maximum of 20000 EVs are caught in expectation. Thus, approximately half of the EVs are not driven within a 1-mile radius of any of the candidate locations. Assumptions 1 and 4 are likely conservative in estimating how many EVs can be caught by the stations built. This is because drivers may adjust their destinations and driving patterns more than our assumptions account for if fast charging stations are built at retail shopping locations.

![Figure 2.3: Expected number of EVs caught as a function of $p$.](image)

The vertical axis on the right-hand side of Figure 2.3 normalizes the expected number of EVs that can be caught as a percentage of the maximum number that can be caught with 222 charging stations built. This axis shows that a relatively small number of charging stations can catch a majority of the EVs. For instance,
26 charging stations, which account for 12% of the candidate locations, could capture more than 80% of the maximum number of EVs caught by all 222 stations. Thirty-six stations would need to be built to be able to capture 90% of the maximum number that can be caught by all 222 stations.

Figure 2.4 shows a 95% confidence interval on the relative optimality gap of the SFCLM, obtained from the A2RP. This optimality gap confidence interval is defined as:

$$\frac{G_{L/2}(x^*_M) + (z_{\alpha}s_{L/2})/\sqrt{L}}{\hat{g}^M} - \hat{g}^M.$$ 

All of the cases (i.e., with different values of $p$) are solved using at most $M \leq 1000$ random samples to generate the SAA problem and $L/2 = 500$ random samples for each of the two replications used in the A2RP. Figure 2.4 shows that these sample sizes are sufficient for obtaining a high-quality solution with an optimality gap of less than 0.0045 (0.45%). Solutions with tighter optimality gaps would require larger samples to generate the SAA problem. This would involve a commensurate increase in computational efforts. Our computational simulations are conducted on the Oakley Cluster at the Ohio Supercomputer Center (cf. Center (1987) for details of the Oakley Cluster) with a CPU-clock limit of 48 hours. Given that the SFCLM is a planning problem that can be solved long in advance of when charging stations are built, a 48-hour solution time limit is reasonable for such a problem.

2.4.2 Effect of Recharging Deviation Tolerance

We examine the effect of the charging-deviation radius, $r$, on the optimal charging station layout in two ways. First, Figure 2.5 shows the optimal location of $p = 6$ charging stations with a radius of $r = 6.4$. Three of the charging stations are in
the same location, two in very slightly different locations, and one in a very different location when compared to the $r = 1.6$ case. This sixth station is moved from far east of the city center to a location just outside the beltway. This is because the increased charging-deviation radius allows a station at the new location to serve EVs on the beltway and on two other major arterial routes. Moreover, this station can now serve vehicles from the airport and around a major shopping center in east Columbus. This is not possible with $r = 1.6$.

Figure 2.6 shows the effect of the optimal charging-station network given by the SFCLM assuming $p = 6$ and $r = 1.6$ km, if drivers were a higher charging-deviation radius than assumed (i.e., if $r$ is actually greater than 1.6km). The blue solid line in the figure shows the expected number of EVs that would not be able to complete their daily tours, as a function of the true value of $r$ (i.e., if $r$ is not actually equal to 1.6km). With the assumed 3% EV-penetration level, there are about 39000 EVs

Figure 2.4: 95% confidence interval on relative optimality gap of solution given by SFCLM as a function of $p$. 

![Graph showing 95% confidence interval on relative optimality gap as a function of number of charging stations.](image-url)
in expectation. Of these, an expected 14000 have driving patterns that exceed the assume 80 mile range. If six charging stations are built, and expected 10000 of these 14000 EVs can complete their daily trips if the true value of \( r \) is 1.6km. As the true value of \( r \) increases, more of these EVs are able to complete their daily trips, with all but 5\% of the EVs unable to complete their trips with \( r = 1.6 \) km.

The red dashed line shows the expected number of EVs with daily trips longer than their assumed 80 mile range that can complete their daily trips with the charging stations. This is given as a percentage of the expected number of EVs with daily trips longer than the vehicle range. The figure shows that many EVs can complete their daily trips with only six charging stations, if drivers are willing to deviate from their intended path from origin to destination.
Figure 2.6: Expected number of EVs with typical daily driving tours longer than 86 miles, and percentage of those EVs that can complete their tours with a six-station EV-charging network if drivers can deviate at most a given distance.

### 2.4.3 Benefit of Public Charging Network to EV Owners

From the perspective of EV owners, an EV-charging network can be evaluated using at least two criteria. One is whether the stations are in locations that allow them to be used by EV owners. Objective function (2.1) of our SFCLM maximizes the expected number of prospective EV owners that can use at least one charging station without having to drive more than 1 mile out of the intended path from origin to destination.

A second criterion is what benefit the charging-station network provides in allowing EV owners to complete their daily tours. The expected number of EVs that can complete their daily tours can differ from the value of objective function (2.1). This is because some EVs that can use a charging station do not drive far enough in a single day to deplete the vehicle battery. Conversely, some EVs may have very long
daily tours, which they cannot complete solely using public charging stations. Both types of EVs may still derive benefits from the charging-station network in terms of the second criterion. For instance, an EV with a very long daily tour may not be able to complete its tour using only the public charging stations chosen by the SFCLM. However, it may be able to complete its tour using the public charging stations and other charging infrastructure (e.g., at a workplace).

We examine this second criterion by conducting a further simulation using the optimized charging-station locations determined by the SFCLM. Specifically, we fix the locations of the charging stations determined by the SFCLM and randomly generate 365 samples of \( \tilde{f} \) (assuming the same 3% EV penetration), by bootstrapping from the tour record data. For each random sample of \( \tilde{f} \), we assume that each EV begins the day with a fully charged battery (e.g., using an at-home charger). We next simulate the SOC of each EV battery during the course of its daily tour, assuming that any EV that passes within the 1-mile capture radius of one of the charging stations stops at the station and fully recharges its battery.

We use this simulation to estimate the expected number of EVs that can successfully complete their daily tours (i.e., the SOC of the EV battery does not reach zero before completing the tour) during the course of a typical year. We contrast this with the number of EVs that would be able to successfully complete their daily tours without any public charging stations (this is done by repeating the same simulation with no public charging stations).

Figures 2.7 and 2.8 summarize the simulation results for a six-station EV-charging network. First, the solid line in Figure 2.7 shows the expected number of EVs that have daily driving tours that are longer than a given EV range. Our assumed 3%
EV-penetration level corresponds to about 40,000 EVs in expectation. Thus, an EV with a 70-mile driving range does not provide sufficient energy for a quarter of drivers to complete a typical daily tour without recharging midday. The stars in Figure 2.7 show the fraction of EVs that have a typical daily tour longer than a given EV range (i.e., the fraction of the solid line in the figure) that can complete their tours if a six-station EV-charging network is built. The figure shows that a six-station network allows more than 60% of EVs that would otherwise not be able to complete their daily tours (without the charging stations) to do so.

Figure 2.7: Expected number of EVs with typical daily driving tours longer than a given EV range, and percentage of those EVs that can complete their tours with a six-station EV-charging network.

Figure 2.8 summarizes the expected number of EVs that are able to complete their typical daily tours with and without a six-station EV-charging network, as a function of the EVs’ range. The figure shows that for EVs with ranges of less than 110 miles (which includes the 2016 Nissan Leaf, 2016 BMW i3, and 2017 Mitsubishi
i-MiEV), a six-station EV-charging network has significant benefits in allowing more drivers to be able to complete their typical driving activities using an EV.

The estimates of the impacts of the charging-station network shown in Figures 2.7 and 2.8 are likely conservative. This is because our simulation assumes that EV owners will only deviate at most 1 mile from the intended path between origin and destination to use a public charging station. In practice, EV owners may plan their driving tours based on the location of public charging stations (e.g., shopping at a retailer that has a charging station available). Moreover, drivers that choose to purchase an EV will do so based on their driving patterns. Our simulations indicate that 4 EVs (in expectation) have typical daily tours that are longer than 210 miles. Drivers with such driving patterns are less likely to purchase EVs.
2.4.4 Charging Station Sizing

Another question in developing public EV charging stations is the number of chargers that should be deployed at each station. Xi et al. (2013) study this question in the context of deploying Level-1 and -2 charging infrastructure by simulating a queuing model of EVs waiting to use an available charger upon arrival at a charging station. We take a similar approach here, building off of the simulation conducted in Section 2.4.3 to estimate the number of EVs wanting to use each charging station in the network throughout the course of a typical day.

Specifically, we fix the locations of the charging stations determined by the SFCLM and use the 365 samples of $\bar{f}$ randomly generated for the simulation conducted in Section 2.4.3. For each random sample, we assume that each EV begins the day with a fully charged battery and simulate the SOC of each EV battery during the course of its daily tour. We assume that EV drivers use an SOC-threshold policy to determine their use of a public charging station. If the SOC of a vehicle battery drops below a predefined threshold, the driver uses the first charging station that is within 1 mile of his or her planned route and fully recharges the vehicle battery. The simulation determines the arrival times of EVs to charging stations. The charging duration is determined based on an assumed power capacity of the charger and the SOC of the vehicle battery upon arrival to the charging station.

Figure 2.9 summarizes the results of this simulation with six charging stations that have 100 kW chargers, EVs that have 85-mile driving ranges, and drivers with 20% SOC thresholds. This means that an EV driver is assumed to use a public charging station if and only if the SOC of his or her vehicle battery drops below 20% and a charging station is available within 1 mile of the planned route. Figure 2.9 shows the
cumulative distribution of the number of vehicles that is charging at any given time. For instance, Station 1 is unoccupied 60% of the time and fewer than one EV charging 80% of the time. The results of this simulation can be used to determine the number of chargers to deploy, based on a desired service-related criterion. As an illustrative example, the horizontal solid line in Figure 2.9 is at the 95th percentile. The figure shows that deploying one charger in Stations 3 and 5, two in Station 6, three in Station 1, four in Station 2, and five in Station 4 will result in a 95% probability that no EVs will have to queue to use a charger.

![Cumulative distribution of EVs in six stations](image)

**Figure 2.9**: Cumulative distribution of the number of EVs in each of six stations with 100-kW chargers assuming 85-mile EV driving ranges and 20% SOC thresholds.

Figure 2.10 summarizes the results of the simulation assuming 200-kW chargers for the same six stations. The cumulative distributions in Figure 2.9 first-order stochastically dominate those in Figure 2.10. This dominance is because of the higher power
capacity in the latter case. EVs have shorter charging times with 200-kW chargers, meaning that there will always be fewer EVs simultaneously needing to use a charger.

Figure 2.10: Cumulative distribution of the number of EVs in each of six stations with 200-kW chargers assuming 85-mile EV driving ranges and 20% SOC thresholds.

2.4.5 Comparison of Stochastic and Deterministic Flow-Catching Location Models

An important question in developing our SFCLM is the benefit that it provides relative to a deterministic version of the model. This is because the SFCLM requires the use of Monte Carlo-based simulation and an L-shaped algorithm to find a near-optimal solution and determine its quality. If a deterministic model can provide a solution of similar quality with less work involved, the value of a SFCLM would be minimal and the deterministic model should be used instead.
We compare our SFCLM to a deterministic flow-capturing location model (DF-CLM) in two ways. First, we examine the optimal locations chosen by the SFCLM and DFCLM to place charging stations. We then examine how the charging station locations given by the two models compare in terms of the expected number of EV that can be caught.

To formulate the DFCLM, we define $\bar{f}_q = \mathbb{E}[\tilde{f}_q]$ as the expected value of $\tilde{f}_q$. The DFCLM is then formulated as:

$$\max_{x,y} \sum_{q \in Q} \bar{f}_q y_q$$

subject to

$$\sum_{k \in K} x_k \leq p; \hspace{1cm} (2.26)$$

$$\sum_{k \in N_q} x_k \geq y_q, \hspace{1cm} \forall q \in Q; \hspace{1cm} (2.27)$$

$$x_k \in \{0, 1\}, \hspace{1cm} \forall k \in K; \hspace{1cm} (2.28)$$

$$y_q \in \{0, 1\}, \hspace{1cm} \forall q \in Q; \hspace{1cm} (2.29)$$

where $Q, K, N_q, p, x_k$ and $y_q$ retain the same definitions introduced in Section 2.2.1.

The DFCLM has a very similar structure to the SFCLM. Objective function (2.25) maximizes the number of EVs caught by the network of charging stations built, based on expected EV flows on each ACH. The constraints of the DFCLM correspond to the analogous constraints of the SFCLM—constraint (2.26) limits the number of stations built, constraints (2.27) define the EV flows caught in terms of the stations built, and constraints (2.28) and (2.29) impose integrality.

As noted in Section 2.2.3, there are an exponentially large number of EV-flow scenarios. Thus, it is intractable to compute $\bar{f}$ exactly. For this reason, we use a sample average of $\bar{f}$ in defining objective function (2.25). More specifically, we use
the same $M$ random samples of $\tilde{f}$ used to define the SAA problem in solving the SFCLM to compute the sample average of $\tilde{f}$.

Figures 2.11 and 2.12 show the optimal locations of six and 31 charging stations, respectively, chosen when the SFCLM and DFCLM are applied to the same underlying central-Ohio case study data. Comparing the optimal station locations reveals some interesting properties of the station locations. First, the DFCLM results in the same general station-location strategy as the SFCLM. If there is a limited number of charging stations to be built (e.g., with $p = 6$), the DFCLM chooses to locate them around the periphery of the beltway. This is because the beltway sees high expected use by EVs. As the number of stations to be built increases, the DFCLM places a greater number of stations inside the beltway and in the three smaller cities outside of Columbus.

Figure 2.11: Optimal charging station locations from SFCLM and DFCLM with $p = 6$. 
Figure 2.12: Optimal charging station locations from SFCLM and DFCLM with $p = 31$.

Although the DFCLM follows the same general strategy as the SFCLM, the two models nevertheless place stations at different locations. This is most notable in the $p = 6$ case shown in Figure 2.11. The DFCLM places one of the six stations in the small town of Marysville (northwest of Columbus), whereas the SFCLM only places a station in this town with $p \geq 21$ (cf. Figure 2.2). With $p = 31$ the locations chosen by the DFCLM and SFCLM are quite similar. Although they are not identical, each of the 31 station locations chosen by the DFCLM is in very close proximity to a station location chosen by the SFCLM.

We study the effects of building charging stations at these different locations by comparing the expected number of EVs caught by the charging station networks selected by the SFCLM and DFCLM. Specifically, we fix the station locations determined by each model and simulate the number of EVs that can be caught under 1000 randomly generated samples of $\tilde{f}$. Figure 2.13 shows the expected number of
EVs that can be caught when up to 31 charging stations are built. For greater numbers of charging stations the locations selected by the SFCLM and DFCLM result in virtually the same expected number of EVs able to be caught. This is consistent with what is shown in Figure 2.12—if 31 stations are being placed the SFCLM and DFCLM chose locations that are in extremely close proximity to one another, resulting in virtually the same expected number of EVs able to be caught. This is true for greater numbers of charging stations as well, as there is a fraction of a percent difference in the expected number of EVs that can be caught by the two station networks.

![Figure 2.13: Expected number of EVs caught with charging station locations chosen by SFCLM and DFCLM as a function of p.](image)

The two lines in Figure 2.13 show the expected number of EVs that can be caught using the station locations determined by the SFCLM and DFCLM. The dots show the percentage difference between the expected number of EVs that can be caught by the
two sets of station locations and is a measure of the value of the stochastic solution given by the SFCLM. The important finding illustrated in Figure 2.13 is that if a limited number of charging stations is to be built, the added effort of using a SFCLM is worthwhile. This is because the station locations chosen by the DFCLM underperform relative to those chosen by the SFCLM by 2-22%. If there are limited resources for building a network of charging stations, the SFCLM provides a more robust set of station locations. As the available resources for building charging stations increases, the SFCLM has limited incremental value compared to the DFCLM. This is because stations are being added to provide coverage of areas that with relatively low EV traffic volume and the gains from these added stations are marginal.

2.5 Conclusions and Future Work

This chapter introduces a SFCLM to optimize the locations of a limited number of EV fast charging stations within a given study region. We formulate this problem as a two-stage stochastic integer program, in which the station locations are fixed in the first stage and EV flows between different ACHs are determined in the second stage. An important feature of our model is that it uses as inputs data that are typically available from regional planning commissions or similar organizations. We apply an L-shaped SAA algorithm to decompose and solve the SFCLM and apply an A2RP to estimate the optimality gap of the final solution. A benefit of the A2RP is that it allows us to dynamically determine the number of scenarios modeled in the SAA problem to achieve a desired solution quality. Our SFCLM and solution method can be contrasted with other stochastic charging station-location models that use a very limited scenario tree size and do not provide an assessment of solution quality.
We apply our model to a case study based on Central Ohio. This is a comprehensive case study with 1.3 million vehicles, 20000 road segments, $9.9 \times 10^{391338}$ second-stage scenarios, and 222 candidate charging station locations. Our SAA method is able to determine near-optimal station locations using less than 48 hours of CPU time, which is reasonable for a planning problem of this type. We demonstrate that the station locations are robust to the number of stations ultimately built. We further demonstrate the benefit of the public charging stations in allowing for greater use of EVs using only a small number of charging stations with a limited number of chargers. We also demonstrate the benefit of the SFCLM by contrasting it with a DFCLM. We find that if the number of stations to be built is limited, the SFCLM has value in maximizing the expected number of EVs that can be caught. As the number of stations to be built increases, the DFCLM could be used instead as the station networks determined by the two models become similar.
3.1 Introduction

Electric vehicles hold great promise to improve the energy efficiency and environmental impacts of transportation. However, widespread EV use brings uncertain impacts to electric power systems, especially at the distribution level. Clement-Nyns et al. (2010) find that uncontrolled EV charging (i.e., EVs charging without any coordination) can result in power losses and voltage deviations on the local distribution network. Razeghi et al. (2014) study the impacts of charging 10 EVs in an uncontrolled fashion on a residential transformer. They demonstrate that the resulting charging loads can result in catastrophic failure of the transformer and conclude that management of charging loads is critical for prolonging transformer life. Weiller (2011) estimates the impacts of EV recharging using standard wall outlet on electric loads in the United States. The results of this analysis show that residential, workplace, and retail shopping center loads can be increased by 74%.

One way to accommodate these impacts of widespread EV use is to upgrade distribution-system infrastructure, including transformers. This solution would see the distribution system sized to accommodate the anticipated peak load. Indeed,
because distribution-infrastructure investments are typically long-lived (e.g., distribution transformers often have a design life of 20.5 years), the system would be sized based on anticipated future peak loads. This is an inefficient solution, however, because the peak load on many distribution circuits may only be reached a few hours each year. This means that the distribution system would have excess unused capacity the overwhelming majority of the time. Uncontrolled EV charging may exacerbate this inefficiency, because uncontrolled EV charging tends to give the distribution load profile more extreme peaks.

An alternate solution is control or management of EV charging loads. The basic premise of charging control is that an EV may be connected to a charging station for a longer duration of time than is required to fully recharge its battery. If so, the charging demand could be shifted within this window of time. By properly managing such flexible EV loads, the peaks in the distribution load profile can be reduced, reducing the need for expensive transformer upgrades. Moreover, controlling EV charging can increase the load factor of the distribution system, meaning that the distribution infrastructure is used more efficiently.

The literature typically focuses on two forms of EV-charging control: centralized and decentralized. Decentralized control is a price-signal based method to coordinate EV charging. This method usually requires individual EVs and an EV aggregator (EVA) to communicate their demands for charging energy and the availability of energy in an iterative before reaching an equilibrium. In this context, an equilibrium is a set of charging loads that are optimal from the perspective of the EVs (in terms of satisfying their demands for charging energy) and the EVA (in terms of being able to feasibly serve the loads). Ma et al. (2013) introduce a price-based decentralized
control scheme. In their proposal, EVs communicate their charging demands iteratively based on a set pricing scheme. They demonstrate that the iterative scheme can reach an equilibrium under mild conditions. Wu et al. (2012) propose a decentralized scheme which uses a price policy that encourages individual EVs to provide frequency regulation. Bayram et al. (2013) propose an admission-control mechanism which applies congestion pricing to mitigate station-level overloads and guarantee quality of service among EVs. Xi and Sioshansi (2014) introduce a decentralized pricing scheme that conveys both price and quantity from the power system operator to an EVA or individual EVs.

Conversely, centralized control relies on a single entity to manage EV charging. Thus, it relies on EV owners letting someone else determine when their vehicles are recharged. Hu et al. (2014) introduce a linear programming model that determines an optimal EV-charging schedule to minimize charging cost to an EVA while preventing distribution grid congestion. Rotering and Ilic (2011) propose a dynamic programming model to control and optimize accumulated EV charging demand. Sundström and Binding (2012) develop a quadratic program, which minimizes the operation cost of an EVA while imposing distribution-network constraints.

More recent works pay increasing attention to uncertainty in when EVs may arrive at a charging station and their charging demand upon arrival. Pantoš (2012) proposes a stochastic optimization model with uncertainty in EV-usage patterns to create strategies for an individual EV that wants to participate in energy and ancillary service markets by providing charging flexibility. Momber et al. (2015) introduce
an EV-charging control model for a risk-averse EVA. They use conditional value-at-risk as the risk metric in the objective function. Their work focuses on modeling uncertainty in EV-usage patterns and energy prices.

The use of stochastic optimization techniques typically raises computational challenges, because of the immense number of scenarios needed to capture all of the uncertainties modeled. Pantos (2012), Momber et al. (2015) deal with this issue through scenario reduction, wherein a large starting set of scenarios is reduced to a smaller set that is meant to represent the range of possible sample paths in the starting scenarios. This use of scenario reduction raises two important and related questions, however. The first is whether scenario reduction guarantees a high-quality solution. The second is how to choose an appropriate starting sample size and reduced sample size that gives a desired solution quality with the least amount of computational effort. This second issue is especially important if a charging-control model is to be used for actual real-time control of EV-charging loads.

In this paper we introduce a centralized control model that concentrates on high-power fast-charging stations. The aim of the model is to optimize EV-charging loads within a fixed window of time after each EV arrives at the charging station to minimize costs. The costs modeled include a penalty, based on the associated accelerated aging, of operating the distribution transformer above its rated capacity. The core methodology of our approach is a two-stage stochastic linear optimization problem. We model uncertainties in EV arrival times and charging demands upon arrival. The model also captures uncertainties in energy prices, non-EV load (which shares the distribution transformer), and energy from distributed energy resources (DERs), such as a photovoltaic (PV) solar panel. To obtain a high-quality solution within
the minimum amount of time, we use sample-average approximation (SAA) and an L-shaped method to solve the problem. We also apply the averaged two-replication procedure (A2RP) of Bayraksan and Morton (2011) to control solution quality and the sample sizes in the SAA.

The remainder of this paper is organized as follows. Section 3.2 gives an explicit formulation of our station-control model. This model schedules EV charging using a two-stage stochastic optimization framework. It also allows EV charging to be cooptimized with the use of DERs, such as PV and a battery energy storage system (BESS). Section 3.3 outlines our proposed solution technique, which uses a Monte Carlo-based SAA method to solve the station-control problem and an A2RP to control solution quality and sample sizes. We then demonstrate the use of the model using a case study based on Central Ohio. Section 3.4 provides detailed case-study data. Case study results are summarized and discussed in Section 3.5. This includes a comparison of our proposed station-control model to simple heuristics in terms of operation costs, transformer loading and aging, and how much EV charging energy is unserved. On all of these metrics, our model outperforms the heuristics. We also conduct some sensitivity analyses, in which the design of the charging station or some parameters relating to EV use are varied. This analysis shows the sensitivity of the resulting load profile to different possible EV-usage cases.

The contributions of this work are developing a charging station-control model that captures unique characteristics of fast charging stations that are not considered in other works in the literature. This includes the ability to model and optimize the use of DERs and capturing uncertainties around EV usage, non-EV load, energy prices, and supply from the DERs. Moreover, we use a Monte Carlo-based SAA technique to
efficiently solve the resulting two-stage stochastic program, without having to rely on scenario reduction. Finally, our use of an A2RP allows us to dynamically determine the size of the randomly sampled scenario tree to provide a solution satisfying a desired quality level at minimal computational cost.

### 3.2 Electric Vehicle Charging Station-Control Model

This section details our EV charging station-control model. The model assumes that the charging station is electrically co-located with a building, with which it shares a distribution-level transformer. This transformer must serve the electrical load of the building, as well as EV-charging loads. The station can also have DERs in the form of a PV panel and a BESS.

There are two types of EVs that arrive to the charging station. The first, which we term inflexible EVs, must have their charging demands met immediately. The others, which we term flexible EVs, have a fixed window of time within which to be recharged. The station-control model determines how to schedule flexible charging loads and the use of the DERs to meet station load at minimum cost. Our model formulation assumes a one-minute time scale at which control decisions are modeled and made. A one-minute time interval is reasonable for the types of dynamics (e.g., EV arrivals and departures and PV output) modeled.

The station-control model is formulated as a two-stage stochastic linear optimization problem, which minimizes expected charging station-operation costs over a fixed time horizon. The first stage represents the present time unit (i.e., the next minute), for which here-and-now EV-charging and DER-operational decisions are made. The second stage represents the remainder of the model horizon. The model assumes
that the system state (e.g., energy prices, EV and non-EV loads, and PV output) are known for stage one, but that these are uncertain for stage two. A scenario tree represents possible realizations of these random variables in subsequent minutes of the model’s time horizon.

Our model determines here-and-now decisions for the next minute only. Our proposed use of this model is in a rolling-horizon fashion. That is to say, the model is used starting from minute $j$ to determine minute-$j$ station-control decisions. The model then rolls forward one minute, updates the stage-one system state and stage-two scenario forecasts, and is re-solved to determine minute-$(j + 1)$ decisions. This process repeats to control the charging station. We use the notational convention throughout our formulation that any random parameter with the subscript $t$ is known with certainty at time $t$ (and thereafter) but may be uncertain before time $t$. We use the subscript $\omega$ to represent different second-stage scenarios. We also let the subscript $j$ denote the current minute (i.e., random parameters and decision variables with a $j$ subscript represent known stage-one parameters and here-and-now decisions).

### 3.2.1 Model Parameters

We begin by defining the following model parameters.

- $\Omega$: Set of second-stage scenarios
- $T$: Optimization horizon [min]
- $W$: Charging window for flexible EVs [min]
- $H$: EV-charger nameplate capacity [kW]
- $R$: Nameplate transformer capacity [kW]
- $R^C(\cdot)$: Penalty for operating transformer above nameplate capacity [$$/min]
- $S^E_{max}$: Maximum state of charge (SOC) of the BESS [kW-min]
- $S^E_{min}$: Minimum SOC of the BESS [kW-min]
- $S^D$: Charging and discharging capacity of the BESS [kW]
- $\mu^C$: Charging efficiency of the BESS
- $\mu^d$: Discharging efficiency of the BESS
Our model has a $T$-minute optimization horizon. The flexible EVs are assumed to have a $W$-minute window of time, within which they must be recharged. We assume that all flexible EVs have the same charging duration for notational brevity. We could, alternately, subdivide the flexible EVs based on the window of time within which they must be recharged. We do not do this, as it would further complicate the model notation.

Each EV is assumed to connect to a $\bar{H}$ kW charger. MW and MWh are typically used as units to measure power and energy, respectively, in electricity systems. We use kW and kW-min throughout our formulation, however. This is because the amount of power transacted in a charging station is on the order of kW and our model has a one-minute temporal granularity.

The charging station is connected to the electric power system through a distribution transformer, which has a nameplate capacity of $\bar{H}$ kW. As is common utility practice, the transformer can be operated above this nameplate capacity. Doing so has the effect of reducing its usable life. We capture this effect of overloading the transformer through $R^{C}(\cdot)$, which gives the per-minute cost of overloading the transformer as a function of the amount that the transformer is overloaded (based on the expected loss of transformer life and its replacement cost).

The BESS is represented by its capacity and efficiency in our model. The BESS is assumed to have maximum and minimum SOCs, $\bar{S}^{E,+}$ and $\bar{S}^{E,-}$, respectively. The maximum SOC represents the amount of energy that can be safely stored in the BESS without overloading it. The minimum SOC, which may be nonzero, is modeled because some storage technologies can suffer extreme cycle-life loss if the SOC gets too low. The BESS also has a power capacity, $\bar{S}^{P}$, which limits the rate at which it can
be charged or discharged. We also assume that energy can be lost when charging or discharging the BESS. The parameters, $\mu^c$ and $\mu^d$, represent these efficiencies. They also implicitly define the cost of storing energy in the BESS (through the energy lost).

### 3.2.2 State Parameters

We now define two sets of state parameters. The first, which we term deterministic state parameters, provide all of the system state information up to the current time (i.e., minute $j$). These parameters include exogenous state information (e.g., the minute-$j$ energy price and amount of non-EV and EV-charging loads), as well as decisions that have already been made (e.g., energy that has already been recharged into flexible EVs that arrived before minute $j$). The second, which we term stochastic state parameters, provide possible realizations of the system state in the future (i.e., after minute $j$). Because these parameters represent possible future states of the system (e.g., EV arrivals, PV output, energy prices, and non-EV loads), they are indexed by the second-stage scenarios, $\omega$.

#### Deterministic State Parameters

- $L_j$: Minute-$j$ inflexible energy demand [kW]
- $F_t$: Total charging demand of flexible EVs that arrived to the charging station in minute $t$, where $t \in \{j-W+1, \ldots, j\}$ [kW-min]
- $N_t$: Total number of flexible EVs that arrived to the charging station in minute $t$, where $t \in \{j-W+1, \ldots, j\}$
- $\bar{f}_t$: Total charging demand of flexible EVs that arrived to the charging station in minute $t$, where $t \in \{j-W+1, \ldots, j\}$, that has been satisfied as of the beginning of minute $j$ [kW-min]
- $V_j$: Minute-$j$ PV output [kW]
- $p_j$: Minute-$j$ electricity price [$$/kW-min]
- $x_j$: Beginning minute-$j$ SOC of the BESS [kW-min]
The charging station faces two types of demands. The first are inflexible demands. These include non-EV demands (i.e., the load of the commercial building sharing the transformer with the charging station) and the charging loads of inflexible EVs. Both of these loads are included in $L_j$.

The second are the demands of flexible EVs. Flexible-EV demands are represented by two parameters. $F_t$ represents the total charging demand of flexible EVs that arrived to the charging station in minute $t$ and $N_t$ represents the total number of EVs that arrived in minute $t$. We assume that the $F_t$ kW-min of charging demand are uniformly spread among the $N_t$ EVs. This assumption is made for notational convenience. We could further subdivide the EVs based on their battery SOC upon arrival to the charging station. We let $\bar{f}_t$ denote the total cumulative charging demand of flexible EVs that arrived in minute $t$ that has been satisfied as of the beginning of minute $j$. This is a deterministic state parameter that represents previously made charging-control decisions.

The parameters $F_t$, $N_t$, and $\bar{f}_t$ are defined for values of $t \in \{j - W + 1, \ldots, j\}$. This is because of the assumed $W$-minute charging window. Any EVs that arrived prior to minute $(j - W + 1)$ would no longer be in the station as of minute $j$.

We also define $\xi_j = (L_j, F_t, N_t, \bar{f}_t, V_j, p_j, x_j)$, where $t \in \{j - W + 1, \ldots, j\}$, as the deterministic state-parameter vector.

**Stochastic State Parameters**

- $L_{t,\omega}$: Minute-$t$ inflexible energy demand in scenario $\omega$ [kW]
- $F_{t,\omega}$: Total charging demand of flexible EVs that arrive to the charging station in minute $t$ of scenario $\omega$ [kW-min]
- $N_{t,\omega}$: Total number of flexible EVs that arrive to the charging station in minute $t$ of scenario $\omega$
- $V_{t,\omega}$: Minute-$t$ PV output in scenario $\omega$ [kW]
- $p_{t,\omega}$: Minute-$t$ electricity price in scenario $\omega$ [$/kW-min]$
The stochastic state parameters are all analogous to the deterministic state parameters, except that they are indexed by scenario. This is because the stochastic state parameters represent the second-stage system state, which is unknown when making minute-$j$ here-and-now decisions. We define $\psi_{t,\omega} = (L_{t,\omega}, F_{t,\omega}, N_{t,\omega}, V_{t,\omega}, p_{t,\omega})$ as a vector of scenario-$\omega$ minute-$t$ stochastic state parameters. We also define $\psi_\omega = (\psi_{j+1,\omega}, \ldots, \psi_{j+T,\omega}) \in \Psi$ as the scenario-$\omega$ sample path of stochastic state parameters.

### 3.2.3 Decision Variables

We define two sets of decision variables: those corresponding to first- and second-stage decisions.

#### First-Stage Decision Variables

- $e^s_j$: Minute-$j$ power sales to the power system [kW]
- $e^b_j$: Minute-$j$ power purchases from the power system [kW]
- $e^c_j$: Minute-$j$ power charged into the BESS [kW]
- $e^d_j$: Minute-$j$ power discharged from the BESS [kW]
- $x_{j+1}$: Beginning minute-$(j + 1)$ SOC of BESS [kW-min]
- $f_{j,\tau}$: Total kW provided in minute $j$ to recharge flexible EVs that arrived to the charging station in minute $\tau$, where $\tau \in \{j - W + 1, \ldots, j\}$ [kW]
- $f_j$: Total kW provided in minute $j$ to recharge flexible EVs [kW]
- $v_j$: Minute-$j$ transformer overload [kW]

The first-stage decision variables pertain to the minute-$j$ here-and-now decisions. The first pair of variables, $e^s_j$ and $e^b_j$, represent energy transacted between the charging station and the power system. Depending on local demand and energy production from the DERs, the charging station may purchase energy from or sell energy to the power system (in net) at a given point in time. These transactions are settled at the energy price, $p_j$. 
The variables, $e^c_j$ and $e^d_j$, represent BESS charging and discharging decisions and $x_{j+1}$ represents the resulting beginning minute-$(j + 1)$ SOC. We let $f_{j,\tau}$ denote the amount of charging load provided to flexible EVs that arrived to the charging station in minute $\tau$ and $f_j$ the sum of the $f_{j,\tau}$’s.

We also define $a_j = (e^s_j, e^b_j, e^c_j, e^d_j, x_{j+1}, f_{j,\tau}, f_j, v_j)$, where $\tau \in \{j - W + 1, \ldots, j\}$ as the first-stage decision-variable vector.

### Second-Stage Decision Variables

- $e^s_{t,\omega}$: Minute-$t$ power sales to the power system in scenario $\omega$ [kW]
- $e^b_{t,\omega}$: Minute-$t$ power purchases from the power system in scenario $\omega$ [kW]
- $e^c_{t,\omega}$: Minute-$t$ power charged into the BESS in scenario $\omega$ [kW]
- $e^d_{t,\omega}$: Minute-$t$ power discharged from the BESS in scenario $\omega$ [kW]
- $x_{t+1,\omega}$: Beginning minute-$(t + 1)$ SOC of the BESS in scenario $\omega$ [kW-min]
- $f_{t,\tau,\omega}$: Total kW provided in minute $t$ of scenario $\omega$ to recharge flexible EVs that arrived to the charging station in minute $\tau$ [kW]
- $f_{t,\omega}$: Total kW provided in minute-$t$ of scenario $\omega$ to recharge flexible EVs [kW]
- $v_{t,\omega}$: Minute-$t$ transformer overload in scenario $\omega$ [kW]

The second-stage decision variables are analogous to the first-stage decision variables, except that they represent second-stage recourse decisions. As such, the second-stage variables are all indexed by the second-stage scenarios, $\omega$. We define $u_{t,\omega} = (e^s_{t,\omega}, e^b_{t,\omega}, e^c_{t,\omega}, e^d_{t,\omega}, x_{t+1,\omega}, f_{t,\tau,\omega}, f_{t,\omega}, v_{t,\omega})$ as the scenario-$\omega$/minute-$t$ second-stage decision-variable vector. We also define $u_{\omega} = (u_{j+1,\omega}, \ldots, u_{j+T,\omega})$ as the scenario-$\omega$ second-stage decision-variable vector.

### 3.2.4 Constraints

We define two types of problem constraints: first- and second-stage constraints. In the first stage, we first have a power-balance constraint:

$$L_j + f_j + e^c_j = V_j + e^d_j + e^b_j - e^s_j.$$  (3.1)
This constraint ensures that power consumed within the charging station for serving inflexible or flexible loads or charging the BESS equals the sum of power produced by the PV panel, discharged from the BESS, or transacted with the power system.

We next have constraints:

\[ v_j + \bar{R} \geq e_j^b - e_j^s; \]  

(3.2)

and:

\[ -v_j - \bar{R} \leq e_j^b - e_j^s; \]  

(3.3)
imposing the transformer’s capacity. These constraints restrict net power transactions with the power system, \((e_j^b - e_j^s)\), based on the transformer capacity, \(\bar{R}\). The constraint allows the transformer to operate above its capacity, in which case \(v_j\) must take on a positive value.

We next have an energy-balance constraint defining the evolution of the BESS SOC:

\[ x_{j+1} = x_j + \mu^c \cdot e_j^c - e_j^d / \mu_d; \]  

(3.4)
bounds on the BESS SOC:

\[ \bar{S}^{E-} \leq x_{j+1} \leq \bar{S}^{E+}; \]  

(3.5)

and power limits on charging and discharging:

\[ e_j^c, e_j^d \leq \bar{S}^p. \]  

(3.6)

The next set of constraints involve flexible-EV recharging. We first have constraints:

\[ f_{j, \tau} \leq \bar{H} \cdot N_{\tau}; \forall \tau = j - W + 1, \ldots, j; \]  

(3.7)
which restrict the total amount that flexible EVs that arrived in minute \( \tau \) can be recharged based on the charger capacity and the total number of EVs. We next have a constraint defining the total amount that flexible EVs are recharged in minute \( j \):

\[
f_j = \sum_{\tau=j-W+1}^{j} f_{j,\tau}. \tag{3.8}
\]

We also have sets of constraints that ensure that none of the flexible EVs are over- or undercharged. The first:

\[
f_{j,\tau} \leq F_\tau - \bar{f}_\tau; \forall \tau = j - W + 1, \ldots, j; \tag{3.9}
\]

restricts the amount recharged in EVs that arrived in minute \( \tau \) to be no greater than the remaining unfulfilled charging demand as of the beginning of minute \( j \). The second:

\[
F_\tau - \bar{f}_\tau - f_{j,\tau} \leq \bar{H} \cdot N_\tau \cdot (W - j + \tau - 1); \forall \tau = j - W + 1, \ldots, j; \tag{3.10}
\]

ensures that the remaining unfulfilled charging demand as of the end of minute \( j \) of EVs that arrived in minute \( \tau \) could be feasible met if they are recharged at the charger’s maximum capacity for the remainder of the charging window.

We finally have non-negativity constraints:

\[
e^a_j, e^b_j, e^c_j, e^d_j, x_{j+1}, f_j, v_j \geq 0; \tag{3.11}
\]

and:

\[
f_{j,\tau} \geq 0, \forall \tau = j - W + 1, \ldots, j. \tag{3.12}
\]

The second-stage constraints have a very similar structure, in that they impose the same sets of restrictions. The major difference between the first- and second-stage constraints is that the second-stage constraints are scenario-dependent. We define
the scenario-ω second-stage scenarios here. We first have a set of power-balance constraints:

\[ L_{t,\omega} + f_{t,\omega} + e_{t,\omega}^c = V_{t,\omega} + e_{t,\omega}^d + e_{t,\omega}^b - e_{t,\omega}^s; \forall t = j + 1, \ldots, j + T. \]  

(3.13)

We next have transformer-capacity constraints:

\[ v_{t,\omega} + \bar{R} \geq e_{t,\omega}^b - e_{t,\omega}^s; \forall t = j + 1, \ldots, j + T. \]  

(3.14)

and:

\[ -v_{t,\omega} - \bar{R} \leq e_{t,\omega}^b - e_{t,\omega}^s; \forall t = j + 1, \ldots, j + T. \]  

(3.15)

Next are constraints for the BESS. The constraint sets:

\[ x_{j+2,\omega} = x_{j+1,\omega} + \mu^c \cdot e_{j+1,\omega}^c - e_{j+1,\omega}^d / \mu^d; \]  

(3.16)

and:

\[ x_{t+1,\omega} = x_{t,\omega} + \mu^c \cdot e_{t,\omega}^c - e_{t,\omega}^d / \mu^d; \forall t = j + 2, \ldots, j + W; \]  

(3.17)

define the evolution of the BESS SOC in each minute. Constraint sets:

\[ \bar{S}E,- \leq x_{t,\omega} \leq \bar{S}E,+; \forall t = j + 2, \ldots, j + W + 1; \]  

(3.18)

and:

\[ e_{t,\omega}^c, e_{t,\omega}^d \leq \bar{S}P; \forall t = j + 2, \ldots, j + W; \]  

(3.19)

impose the SOC and charging and discharging power limits of the BESS, respectively.

We next have constraints related to flexible-EV charging. The first set imposes the chargers’ capacities:

\[ f_{t,\tau,\omega} \leq \bar{H} \cdot N_{t,\omega}; \forall t = j + 1, \ldots, j + T; \tau = t - W + 1, \ldots, t; \]  

(3.20)
and the second defines total flexible-EV charging in each minute of each scenario:

\[ f_{t,\omega} = \sum_{\tau=t-W+1}^{t} f_{t,\tau,\omega}; \forall t = j + 1, \ldots, j + T. \] (3.21)

We finally need additional sets of constraints to ensure that all of the flexible EVs are recharged. First, we ensure that all vehicles that have already arrived as of minute \( j \) will be fully recharged within the charging window under all scenarios:

\[ \sum_{t=j+1}^{\tau+W} f_{t,\tau,\omega} + f_{j,\tau} = F_{\tau} - \bar{f}_{\tau}; \forall \tau = j - W + 1, \ldots, j. \] (3.22)

Next, we ensure that all flexible EVs that arrive after minute \( j \) and that have a charging window that ends before the \( T \)-minute model horizon are fully recharged:

\[ \sum_{t=\tau}^{\tau+W} f_{t,\tau,\omega} = F_{\tau}; \forall \tau = j + 1, \ldots, j + T - W. \] (3.23)

We finally ensure that the remaining EVs (i.e., those that will arrive after minute \( j \) but have a charging window that ends after the \( T \)-minute model horizon) are not overcharged but have been charged sufficiently to be able to recharge before their charging windows expire. These constraint sets are written as:

\[ \sum_{t=\tau}^{j+T} f_{t,\tau,\omega} \leq F_{\tau}; \forall \tau = j + T - W + 1, \ldots, j + T; \] (3.24)

and:

\[ F_{\tau} - \sum_{t=\tau}^{j+T} f_{t,\tau,\omega} \leq \bar{H} \cdot N_{\tau,\omega} \cdot (W - j - T + \tau); \forall \tau = j + T - W + 1, \ldots, j + T. \] (3.25)

We finally have non-negativity constraints:

\[ e_{t,\omega}^a, e_{t,\omega}^b, e_{t,\omega}^c, e_{t,\omega}^d, x_{t+1,\omega}, f_{t,\omega}, v_{t,\omega} \geq 0; \forall t = j + 1, \ldots, j + T; \] (3.26)

and:

\[ f_{t,\tau,\omega} \geq 0; \forall t = j + 1, \ldots, j + T; \tau = t - W + 1, \ldots, t. \] (3.27)
For notional convenience, we also define the first-stage feasible region as:

\[ A_j = \{ a_j | (3.1)-(3.12) \} . \]

We similarly define the second-stage feasible region on sample path \( \omega \) as:

\[ U_\omega = \{ u_\omega | (3.13)-(3.27) \} . \]

We finally note that (3.10) ensures that this problem has relatively complete recourse. These constraints ensure that all of the EVs that are in the charging station as of minute \( j \) are charged sufficiently so that they could be fully recharged within the \( W \)-minute charging window.

### 3.2.5 Objective Function

To define the objective function, we begin by defining the minute-\( j \) charging station-operation cost as:

\[ c_j(a_j, \xi_j) = p_j \cdot (e_j^b - e_j^s) + R^C(v_j). \]  

(3.28)

The first term in (3.28) gives the cost of energy transacted with the power system, while the second term represents the cost of overloading the distribution-level transformer. We similarly define the second-stage cost in minute \( t \) of scenario \( \omega \) as:

\[ c_{t,\omega}(u_{t,\omega}, \psi_{t,\omega}) = p_{t,\omega} \cdot (e_{t,\omega}^b - e_{t,\omega}^s) + R^C(v_{t,\omega}). \]

The two terms in this function have the same interpretations as those in (3.28). The charging station-control problem is then formulated as:

\[ \min_{a_j \in A_j} \{ g(a_j) = c_j(a_j, \xi_j) + \mathbb{E} [r(a_j, \psi_\omega)] \} , \]

(3.29)
where we define:

\[ r(a_j, \psi_\omega) = \min_{u_\omega \in U_\omega} \sum_{t=j+1}^{j+T} c_{t,\omega}(u_{t,\omega}, \psi_{t,\omega}). \]  

(3.30)

Objective function (3.29) minimizes the sum of stage-one cost, which is the first term, and expected second-stage cost, which is the second term. The second-stage is cost is defined by (3.30) as the sum of costs accumulated in each of minutes \((j + 1)\) through \((j + T)\).

### 3.3 Solution Technique

First-stage objective function (3.29) is implicitly defined by the second-stage problems. As such, all of the second-stage problems must be solved to obtain an optimal solution to the station-control problem. The second-stage problems are defined by forecasting second-stage scenarios through the end of the \(T\)-minute model horizon. These second-stage scenarios are defined by different realizations of the second-stage parameter vector, \(\psi_\omega\). Some of the second-stage parameters are continuous random variables, implying that there are an infinite number of second-stage parameter vectors and, as such, an infinite number of second-stage problems.

These properties of the station-control problem make it computationally challenging. Our proposed use of the station-control problem is in a rolling-horizon fashion, whereby it is re-solved every minute. Thus, the station-control problem must be solved quickly. To accomplish this, we apply a SAA approach to quickly obtain near-optimal solutions within seconds. We also apply a sequential sampling procedure (SSP) proposed by Bayraksan and Morton (2011) to estimate and control the quality of the solutions obtained from the SAA.
We proceed in this section by first defining the SAA problem and then detailing the algorithm used to solve it. We next discuss the gap and variance estimators used to assess the quality of a solution given by the SAA problem and determine the stopping criterion for our solution algorithm. We finally discuss the convergence properties of our solution method.

### 3.3.1 Sample Average Approximation Problem

An SAA problem replaces the infinite number of second-stage problems that define the true station-control problem with a finite random sample of second-stage scenarios. We solve the SAA problem in an iterative fashion, in which the number of second-stage scenarios sampled is increased until the solution to the SAA problem satisfies a desired termination criterion.

To formulate the SAA problem, we first define $\psi_1, \ldots, \psi_{m_\kappa}$ as $m_\kappa$ independent and identically distributed (i.i.d.) samples of the second-stage sample paths (i.e., samples of $\psi_\omega$). The $\kappa$ subscript on $m_\kappa$ denotes the iteration number of the algorithm employed to solve the SAA problem (cf. Algorithm 4). $\psi_1, \ldots, \psi_{m_\kappa}$ are randomly generated by a Monte Carlo simulation. The method used to generate the sequence of $m_\kappa$’s is further discussed in Section 3.3.4 (cf. Equation (3.39) in particular) and all of the $m_\kappa$’s are integral. The SAA problem is formulated as:

$$
\min_{a_j \in A_j} \left\{ g^{m_\kappa}(a_j) = \frac{1}{m_\kappa} \sum_{m=1}^{m_\kappa} [c_j(a_j, \xi_j) + r(a_j, \psi_m)] \right\}; \quad (3.31)
$$

where we define:

$$
r(a_j, \psi_m) = \min_{u_m \in U_m} \sum_{t=j+1}^{j+T} c_{t,m}(u_{t,m}, \psi_{t,m}). \quad (3.32)
$$

Algorithm 4 provides a high-level overview of the method used to solve the SAA problem. Step 1 begins by choosing sample-size sequences, $\{m_\kappa\}$ and $\{l_\kappa\}$, and a
termination criterion, Υ (cf. Sections 3.3.3 and 3.3.4). The sample-size sequence, \( \{l_\kappa\} \), is used to estimate the optimality gap of the solution to the SAA problem solved in each iteration. Steps 3–8 are the main iterative loop. In Step 4 of iteration \( \kappa \), the \( m_\kappa \) i.i.d. random samples are generated to define an SAA problem. This problem is solved using an L-shaped method (cf. Section 3.3.2) in Step 5. In Step 6, \( l_\kappa \) i.i.d. samples of the random variables are generated to estimate the optimality gap of the solution to the SAA problem found in Step 5. This gap estimator is used in the termination criterion (cf., Sections 3.3.3 and 3.3.4). The iterative process repeats, with the sample size of the SAA problems increasing, until stopping criterion \( \Upsilon \), is satisfied.

**Algorithm 4 SAA Method**

1: **input:** sample-size sequence \( \{m_\kappa\}, \{l_\kappa\} \), stopping criterion \( \Upsilon \)  
2: \( \kappa \leftarrow 0 \)  
3: **repeat**  
4: generate \( m_\kappa \) i.i.d. random samples, \( \psi_1, \ldots, \psi_{m_\kappa} \) of second-stage sample path  
5: solve SAA problem (3.31) using L-shaped method  
6: generate \( l_\kappa \) i.i.d. random samples and optimality gap estimator  
7: \( \kappa \leftarrow \kappa + 1 \)  
8: **until** \( \Upsilon \) is satisfied  

The following subsections provide further details of the L-shaped method (cf. Step 5) and the selection of the sample-size sequences and stopping criterion (cf. Step 1).

### 3.3.2 Solution of SAA method

We use an L-shaped method to solve the SAA problem in Step 5 of Algorithm 4. This method works by doing a Benders-type decomposition of (3.31). We define the
master problem:

\[
\min_{a_j \in A_j, \theta} \frac{1}{m} \sum_{m=1}^{m_k} \left[ c_j(a_j, \xi_j) + \theta_m \right]; \quad (3.33)
\]

where \( \theta_1, \ldots, \theta_m \) is a set of new variables. The variable, \( \theta_m \), is an estimate of the second-stage objective value under sample path \( \psi_m \). Thus, \( \theta_m \) is estimating \( r(a_j, \psi_m) \).

Problem (3.33) relaxes the optimality of the second-stage decisions, which is required to solve (3.31). Thus, any optimal solution of (3.33) should be checked to verify whether it satisfies or violates second-stage optimality. If one of the \( \theta_m \)'s violates this optimality, \( i.e., \) if:

\[
\theta_m < r(a_j, \psi_m);
\]

for some \( m \), then an optimality cut should be added to (3.33). These optimality cuts are obtained from the optimal solution of the dual of (3.32). More specifically, let \( \pi_m \) and \( \Pi_m \) denote the variable vector and feasible region, respectively, of the dual of the sample path-\( m \) second-stage problem. The dual of (3.32) can then be written as:

\[
\max_{\pi_m \in \Pi_m} \phi(\pi_m, a_j, \psi_m). \quad (3.34)
\]

Note that \( \phi(\pi_m, a_j, \psi_m) \) is the optimal objective-function value of the sample path-\( m \) second-stage problem with first-stage decision-variable values, \( a_j \). Because of strong duality, \( r(a_j, \psi_m) \) is equal to the optimal objective-function value of (3.34). Thus, the separated optimality cuts are:

\[
\theta_m \geq \phi(\pi_m^*, a_j, \psi_m); \quad (3.35)
\]

where \( \pi_m^* \) is the optimal dual-variable vector of (3.34).

Benders’s-type decomposition schemes typically also require adding feasibility cuts to the master problem. This is because master problem (3.33) relaxes feasibility of the
second-stage problems (i.e., first-stage decisions are made without explicit consideration of second-stage constraints). Feasibility cuts are not necessary for our problem, however, because (3.10) guarantees that Problem (3.29) has relatively complete recourse. Because Constraints (3.10) are imposed in master problem (3.33) as well (it is minimizing over the same first-stage feasible set, $A_j$), feasibility cuts are not necessary.

The L-shaped method iteratively adds optimality cuts to master problem (3.33). To more explicitly specify this iterative technique, we first define the L-shaped master SAA problem. To do this, we let $K_1, \ldots, K_{m^\kappa}$ denote the number of optimality cuts that have already been added to the problem for each of sample paths, $\psi_1, \ldots, \psi_{m^\kappa}$, respectively. We also let $\pi_m^k$ denote the optimal dual vector, which is obtained from (3.34), to generate the $k$th optimality cut for sample path $m$. The L-shaped master SAA problem is then given by:

$$\min_{a_j, \theta} \frac{1}{m^\kappa} \sum_{m=1}^{m^\kappa} [c_j(a_j, \xi_j) + \theta^m]$$

s.t. $a_j \in A_j$;

$$\theta^m \geq \phi(\pi_m^k, a_j, \psi_m); \quad \forall m = 1, \ldots, m^\kappa; k = 1, \ldots, K_m.$$  \hspace{1cm} (3.37)

For notational convenience, we define the feasible region of the L-shaped master problem as:

$$\tilde{A}_j = \{(a_j, \theta) \in A_j \times \mathbb{R}^{m^\kappa} | (3.37)\}.$$ 

Algorithm 5 outlines the L-shaped method. The set of optimality cuts is first initialized to be empty in Step 1. In Step 3, the incumbent master problem is solved to obtain the solution, $(\hat{a}_j, \hat{\theta})$. We then iterate through each second-stage sample path and solve the cut-separation problem in Step 5. If, in Step 6, we find that the value
of $\hat{\theta}_m$ underestimates the true second-stage objective value, $\phi(\hat{\pi}_m, \hat{a}_j, \psi_m)$, then a new sample path-$m$ optimality cut is added in Steps 7 and 8. This iterative procedure repeats until $\hat{\theta}$ does not underestimate the true second-stage objective value of any sample path.

**Algorithm 5** L-Shaped method for SAA Problem

1: $K_m \leftarrow 0, \forall m = 1, \ldots, m_{\kappa}$
2: repeat
3: $(\hat{a}_j, \hat{\theta}) \leftarrow \arg \min_{(a_j, \theta) \in \hat{A}_j} (3.36)$
4: for $m = 1, \ldots, m_{\kappa}$ do
5: $\hat{\pi}_m \leftarrow \arg \max_{\pi_m \in \Pi_m} (3.34)$
6: if $\phi(\hat{\pi}_m, \hat{a}_j, \psi_m) > \hat{\theta}_m$ then $\triangleright$ add a new sample path-$m$ optimality cut
7: $K_m \leftarrow K_m + 1$
8: $\pi_{Km} \leftarrow \hat{\pi}_m$
9: end if
10: end for
11: until $\frac{1}{m_{\kappa}} \sum_{m=1}^{m_{\kappa}} \hat{\theta}_m \geq \frac{1}{m_{\kappa}} \sum_{m=1}^{m_{\kappa}} r(\hat{a}_j, \psi_m)$

### 3.3.3 Gap and Variance Estimators

The termination criterion, $\Upsilon$, in Algorithm 4 requires us to assess the quality of the solution found in Step 5. We do this by applying SSP with A2RP. More specifically, SSP gives gap and variance estimators using the optimality-gap estimators provided by the A2RP. We first introduce the gap and variance estimates provided by the A2RP.

A2RP is a batch-mean-based method suggested by Bayraksan and Morton (2006). We begin by defining $g^*$ to be the optimal objective value of the original problem:

$$g^* = \min_{a_j \in A_j} g(a_j);$$
where \( g(a_j) \) is defined in (3.29). We also let:

\[
A^* = \{ a_j \in A_j | g(a_j) = g^* \};
\]
declare the set of optimal solutions of (3.29). We similarly define \( \hat{g}^{m\kappa} \) as the optimal objective-function value of SAA problem (3.31) with \( m\kappa \) randomly generated sample paths:

\[
\hat{g}^{m\kappa} = \min_{a_j \in A_j} g^{m\kappa}(a_j);
\]
and \( \hat{A}^{m\kappa} \) as the corresponding optimal-solution set:

\[
\hat{A}^{m\kappa} = \{ a_j \in A_j | g^{m\kappa}(a_j) = \hat{g}^{m\kappa} \}.
\]
The optimality gap and gap variance are defined, respectively, as:

\[
\mu(a_j) = g(a_j) - g^*;
\]
and:

\[
\sigma^2(a_j) = \text{Var} ( r(a_j, \psi_\omega) - r(a_j^*, \psi_\omega) ) ;
\]
where \( a_j^* \in A^* \).

A2RP compares an optimal solution obtained from solving the SAA problem with the \( m\kappa \) random sample paths in Step 5 of Algorithm 4, to an optimal solution obtained from solving an SAA problem with a different set of stochastic parameters. We only provide an outline of the A2RP method here and refer interested readers to the work of Bayraksan and Morton (2006) for further details.

We begin by defining two new randomly generated sample paths, \( \dot{\psi}^1_1, ..., \dot{\psi}^{l\kappa/2}_1 \) and \( \dot{\psi}^1_2, ..., \dot{\psi}^{l\kappa/2}_2 \), each of size \( l\kappa/2 \). The optimality-gap estimators are then defined as:

\[
G^i_{l\kappa/2}(a^{m\kappa}_j) = g^{l\kappa/2,i}(a^{m\kappa}_j) - \min_{a_j \in A_j} g^{l\kappa/2,i}(a_j) ; \forall i = 1, 2;
\]

82
where \( a_j^{m,\kappa} \) is the optimal solution from Step 5 of Algorithm 4 (i.e., \( a_j^{m,\kappa} \in \hat{A}^{m,\kappa} \)) and \( g^{l_\kappa/2,i}(\cdot) \) is defined by (3.31) with the parameter \( m_\kappa \) on the right-hand side replaced with \( l_\kappa \). Note that:

\[
g^{l_\kappa/2,i}(a_j) = c_j(a_j, \xi_j) + \frac{2}{l_\kappa} \sum_{l=1}^{l_\kappa/2} r(a_j, \psi_i^l).
\]

We also define:

\[
a_j^{l_\kappa/2,i} = \arg \min_{a_j \in A_j} g^{l_\kappa/2,i}(a_j);
\]

as the minimizer of the second term defining \( G_{l_\kappa/2}(a_j^{m,\kappa}) \). Thus, we can write:

\[
G_{l_\kappa/2}(a_j^{m,\kappa}) = g^{l_\kappa/2,i}(a_j^{m,\kappa}) - g^{l_\kappa/2,i}(a_j^{l_\kappa/2,i}); \forall i = 1, 2.
\]

The variance estimators of the optimality gaps are defined as:

\[
s_{l_\kappa/2}^2(a_j^{l_\kappa/2,i}) = \frac{2}{l_\kappa - 2} \sum_{l=1}^{l_\kappa/2} \left[ \left( c_j(a_j^{m,\kappa}, \xi_j) + r(a_j^{m,\kappa}, \psi_i^l) \right) - \left( c_j(a_j^{l_\kappa/2,i}, \xi_j) + r(a_j^{l_\kappa/2,i}, \psi_i^l) \right) \right]^2;
\]

for all \( i = 1 \) and 2. The mean and variance estimates are calculated as:

\[
\bar{G}_{l_\kappa/2}(a_j^{m,\kappa}) = \frac{1}{2} \left[ G_{l_\kappa/2}^1(a_j^{m,\kappa}) + G_{l_\kappa/2}^2(a_j^{m,\kappa}) \right];
\]

and:

\[
\bar{s}_{l_\kappa}^2(a_j^{m,\kappa}) = \frac{1}{2} \left[ s_{l_\kappa/2}^2(a_j^{l_\kappa/2,i}) + s_{l_\kappa/2}^2(a_j^{l_\kappa/2,i}) \right].
\]

### 3.3.4 Stopping Criterion and Sample Sizes

We use an SSP stopping criterion proposed by Bayraksan and Morton (2011). We denote the stopping criterion as \( \Upsilon \). \( \Upsilon \) should ensure that the algorithm provides a
high-quality solution in a finite number of iterations. The stopping criterion, \( \Upsilon \), is defined as:

\[
\tilde{G}_{l,2}(a_j^{m,\kappa}) \leq h' \bar{s}_{l,\kappa}(a_j^{m,\kappa}) + \epsilon'; \tag{3.38}
\]

where \( h' > 0 \) is a relative tolerance and \( \epsilon' > 0 \) ensures finite stopping. We define the first iteration that satisfies \( \Upsilon \) as:

\[
\Gamma = \inf_{\kappa \geq 1} \{ \kappa | \tilde{G}_{l,2}(a_j^{m,\kappa}) \leq h' \bar{s}_{l,\kappa}(a_j^{m,\kappa}) + \epsilon' \}.
\]

Stopping criterion (3.38) requires that the optimality gap, \( \tilde{G}_{l,2}(a_j^{m,\kappa}) \), be no more than a fraction, \( h' \), of its standard deviation, \( \bar{s}_{l,\kappa}(a_j^{m,\kappa}) \). Criterion (3.38) is satisfied with probability 1 if \( l_\kappa < \infty \), \( m_\kappa < \infty \), and \( \epsilon' > 0 \). In fact, \( \epsilon' \) limits the order of magnitude of the optimality gap when the algorithm terminates. For instance, an algorithm terminates if \( \tilde{G}_{l,2}(a_j^{m,\kappa}) = 10^{-8} \) and \( \epsilon' = 10^{-7} \) no matter what value \( \bar{s}_{l,\kappa}(a_j^{m,\kappa}) \) takes.

Thus, \( \epsilon' \) should be appropriately small to avoid interfering with Algorithm 4 stopping when the optimality gap and the corresponding standard deviation are large. For further details of finite stopping with positive \( \epsilon' \), we refer interested readers to the work of Bayraksan and Morton (2011).

Stopping criterion (3.38) is defined in terms of parameters \( h' \) and \( \epsilon' \). However, the confidence interval on the optimality gap, once the stopping criterion is satisfied, is actually given in terms of:

\[
h \cdot \bar{s}_{l,\kappa}(a_j^{m,\kappa}) + \epsilon;
\]

where \( h > h' \) and \( \epsilon > \epsilon' \). The reason that the confidence interval is inflated (\textit{i.e.}, is given in terms of \( h \) and \( \epsilon \) as opposed to \( h' \) and \( \epsilon' \)) is because of the sequential nature of the sampling method employed. We choose the sample-size sequence, \( \{m_\kappa\} \), such
that:

\[ m_\kappa \geq \left( \frac{1}{h - h'} \right)^2 (\eta_q + 2q[\log(\kappa)]^2); \]  

(3.39)

where:

\[ \eta_q = \max \left\{ 2\log \left( \frac{\sum_{i=1}^{+\infty} \frac{\ell - q \log \iota}{\sqrt{2\pi \alpha}}}{1} \right), 1 \right\}; \]

(1 - \alpha) is the desired confidence level of the confidence interval obtained, with \( \alpha \in (0, 1) \), and \( q > 0 \) is a parameter that governs the sample sizes. The value of \( q \) is chosen to minimize the computational effort, which is defined by the total number of samples required in Algorithm 4 (i.e., \( \sum_{\kappa=1}^{\Gamma} m_\kappa \)). Further details on the choice of \( q \) are provided by Bayraksan and Morton (2011).

Algorithm 6 outlines the procedure employed to determine the stopping criterion and the sample size sequence. For notational convenience, we define \( \Delta h = h - h' \) and \( \Delta \epsilon = \epsilon - \epsilon' \). The algorithm takes two inputs in Step 1—an initial sample size, \( m_0 \), and a desired number of iterations, \( \Gamma_0 \), for Algorithm 4. These inputs are used to determine the \( \Upsilon \) and \( \{m_\kappa\} \). \( \Gamma_0 \) represents a tradeoff between the amount of computational work involved in solving the station-control problem and the tightness of the confidence interval on the optimality gap obtained. A higher value of \( \Gamma_0 \) tightens the confidence interval but also increases:

\[ \sum_{\kappa=1}^{\Gamma_0} m_\kappa; \]

the number and size of the SAA problems that must be solved. Step 4 calculates \( \eta_q \).

In practice, the:

\[ \sum_{i=1}^{+\infty} \frac{\ell - q \log \iota}{\sqrt{2\pi \alpha}}; \]  

(3.40)

term is approximated by replacing \( +\infty \) in the upper limit of the sum with a large but finite number. Note that (3.40) converges to a finite value because \( q > 0 \).
Algorithm 6 Stopping Criterion and Sample Sizes Initialization

1: \textbf{input:} $m_0$ and $\Gamma_0$ \hspace{1cm} $\triangleright m_0$: initial sample size, $\Gamma_0$: desired iterations
2: $l_0 \leftarrow 2 \cdot \lceil m_0/2 \rceil$
3: choose $q$ based on $\Gamma_0$ \hspace{1cm} $\triangleright$ cf. Bayraksan and Morton (2011)
4: $\eta_q \leftarrow \max\{2 \log(\sum_{i=1}^{+\infty} e^{-q \log i}/\sqrt{2\pi i}), 1\}$
5: $\Delta h \leftarrow \sqrt{\eta_q m_0}$ based on (3.39) with $\kappa = 1$
6: \textbf{for} $\varrho = 1, \ldots, \rho$ \textbf{do}
7: solve (3.31) with a $m_0$ second-stage sample paths using Algorithm 5 \hspace{1cm} $\triangleright$ let $a_j^{m_0^*}$ denote solution
8: solve A2RP problems to obtain $\bar{G}_{l_n/2}(a_j^{m_0^*})$ and $s^2(a_j^{m_0^*})$
9: \textbf{end for}
10: $\bar{G} \leftarrow \frac{1}{\rho} \sum_{\varrho=1}^{\rho} \bar{G}_{l_n/2}(a_j^{m_0^*})$
11: $\bar{f}^2 \leftarrow \frac{1}{\rho} \sum_{\varrho=1}^{\rho} \bar{s}^2_{l_n}(a_j^{m_0^*})$
12: $h' \leftarrow \beta \cdot (\bar{G}/\bar{f})$
13: $h \leftarrow h' + \Delta h$
14: $\epsilon' \leftarrow 1 \times 10^{-7}$
15: $\epsilon \leftarrow 2 \times 10^{-7}$
16: generate $\{m_\kappa\}$ using (3.39)
17: $l_\kappa \leftarrow 2 \cdot \lceil m_\kappa/2 \rceil, \forall \kappa$
18: \textbf{output:} $h', h, \epsilon', \epsilon, \{m_\kappa\}$, and $\{l_\kappa\}$

In Steps 6–9 we generate and solve $\rho$ SAA problems with sample sizes of $m_0$. The results of these initial SAA problems are used to generate a set of gap and variance estimates, which are averaged in Steps 10 and 11. These average gap and variance estimates are used in Step 12 to determine $h'$. The reasoning behind choosing $h'$ in this way is that the initial $\rho$ SAA problems give us a sense of the ratio between the optimality gap and its variance for near-optimal solutions. $\bar{G}$ and $\bar{f}$ are random samples of these two values, thus we use $\bar{G}$ and $\bar{f}$ to determine $h'$. Indeed, $h'$ specifies the instance stochasticity and $h' \propto \bar{G}/\bar{f}$. The reason for solving multiple SAA problems is to avoid a biased set of samples and to estimate the average uncertainty using a set of near-optimal solutions. Values of $\rho$ between 2 and 5 are common and we use $\rho = 2$. 

86
\( \beta \) is a user-specified positive relative stopping constant and is usually fixed between 0.9 and 1. We use \( \beta = 1 \). In this setting, the tightness of stopping criterion (3.38) is decided by the variance of the objective function around the optimal solution. For instance, (3.38) is tight if the variance is small. Otherwise, a relatively loose stopping criterion is given.

The values of \( \epsilon' \) and \( \epsilon \) are fixed in Steps 14 and 15. In practice, these two parameters can be set to any reasonable positive numbers, based on observed behavior of the optimality and variance of the problem being solved. Finally, the sample-size sequences, \( \{m_\kappa\} \) and \( \{l_\kappa\} \), is generated in Steps 16 and 17, using (3.39). We generate the \( m_\kappa \)'s as:

\[
m_\kappa \leftarrow \left\lceil \left( \frac{1}{h - h'} \right)^2 \left( \eta_q + 2q[\log(\kappa)]^2 \right) \right\rceil.
\]

A consequence of using the sample sizes and stopping criterion found in Algorithm 6 is that:

\[
[0, h \cdot \bar{s}_{l_\kappa}(a_{j^*}^{m_{\tau^*}}) + \epsilon];
\]

is a valid \((1 - \alpha)\) confidence interval on the optimality gap, so long as \( \Delta h \) has order of magnitude less than \( \Delta \epsilon / M \), where \( M \) is the supremum of the optimality gap's variance. Further details regarding the asymptotic validity of the stopping criterion are given in Section 3.3.6.

### 3.3.5 Convergence Properties of SAA

An important question in applying Algorithm 4 is whether there are any guarantees that SAA problem (3.31) converges to the true station-control problem given by (3.29). Convergence of the SAA problem relies on the results of Shapiro et al. (2014). Before answering this question, we first place an additional assumption on...
the stochastic properties of the station-control problem and show four additional properties of the original and SAA problems. These are then used to appeal to a theorem shown by Shapiro et al. (2014), which provides our convergence result.

**Assumption 9.** $p_{t,\omega}$ has a finite moment-generating function (MGF). Moreover, random vectors $V_{t,\omega}$, $L_{t,\omega}$, $N_{t,\omega}$ and $F_{t,\omega}$ are all bounded.

The assumption on the MGF of $p_{t,\omega}$ says that there exists a $\gamma_0 > 0$ such that for all $\gamma \in (-\gamma_0, \gamma_0)$ we have that $M(\gamma) = \mathbb{E}[e^{\gamma p_{t,\omega}}] < +\infty$, where $M(\gamma)$ is the finite MGF of $p_{t,\omega}$.

We now show the additional properties of the original and SAA problems through the following lemma and propositions.

**Lemma 1.** Suppose that there exist $N$ independent random variables, $X_1, X_2, \ldots, X_N$, for some $N \in \mathbb{Z}^+$ and that $X_i$ has finite MGF $\forall i = 1, \ldots, N$. Define

$$Z = \min \{X_1, X_2, \ldots, X_N\}.$$ 

Then, the random variable, $Z$ has a finite MGF.

**Proof.** Because each $X_i$ has a finite MGF, this means that if we define:

$$M_i(\gamma) = \mathbb{E}[e^{\gamma X_i}] = \int_0^{+\infty} \text{Prob} \{e^{\gamma X_i} \geq y\} \, dy; \forall i = 1, \ldots, N;$$

then there exist values $\gamma^i_0 > 0, \forall i = 1, \ldots, N$ such that:

$$M_i(\gamma) < +\infty, \forall i = 1, \ldots, N; \gamma \in (-\gamma^i_0, \gamma^i_0).$$

Define:

$$M_Z(\gamma) = \mathbb{E}[e^{\gamma Z}] = \int_0^{+\infty} \text{Prob} \{e^{\gamma Z} \geq y\} \, dy.$$
It is trivial to show that $M_Z(0) = 1$. Fix a value of $\gamma_0 = \min\{\gamma_0^1, \ldots, \gamma_0^N\}$. For any $\gamma \in (0, \gamma_0)$ we have that:

$$M_Z(\gamma) = \int_0^{+\infty} \text{Prob}\left\{ e^{\gamma Z} \geq y \right\} dy = \int_0^{+\infty} \text{Prob}\left\{ e^{\gamma X_k} \geq y \right\} dy$$  \hspace{1cm} (3.41)

$$= \int_0^{+\infty} \prod_{i=1}^{N} \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} dy$$  \hspace{1cm} (3.42)

$$\leq \int_0^{+\infty} \text{Prob}\left\{ e^{\gamma X_k} \geq y \right\} dy$$  \hspace{1cm} (3.43)

$$= M_k(\gamma)$$

$$< +\infty.$$

Equation (3.41) follows because we have that $e^{\gamma \cdot x}$ monotonically increases with respect to $x$ for $\gamma > 0$. As such:

$$\left\{ e^{\gamma \min\{X_1, \ldots, X_N\}} \geq y \right\} \equiv \left\{ e^{\gamma X_k} \geq y; \forall i = 1, \ldots, N \right\}.$$  

Equation (3.42) then follows because of the assumed independence of the $X_i$’s. We define the $k$ in (3.43) as:

$$k \in \arg \max_i \left\{ \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} \right\}.$$  

Inequality (3.43) follows because:

$$0 \leq \text{Prob}\left\{ e^{\gamma X_i} \leq y \right\} \leq \text{Prob}\left\{ e^{\gamma X_k} \leq y \right\} \leq 1; \forall i = 1, \ldots, N.$$  

We next examine the case of any $\gamma \in (-\gamma_0, 0)$, which has that:

$$M_Z(\gamma) = \int_0^{+\infty} \text{Prob}\left\{ e^{\gamma Z} \geq y \right\} dy = \int_0^{+\infty} \text{Prob}\left\{ \max\{e^{\gamma X_1}, \ldots, e^{\gamma X_N}\} \geq y \right\} dy$$  \hspace{1cm} (3.44)
\[
\int_0^{+\infty} \left( 1 - \text{Prob}\left\{ \max\{e^{\gamma X_1}, \ldots, e^{\gamma X_N}\} < y \right\} \right) dy \\
= \int_0^{+\infty} \left( 1 - \prod_{i=1}^{N} \text{Prob}\left\{ e^{\gamma X_i} < y \right\} \right) dy \\
= \int_0^{+\infty} \left( 1 - \prod_{i=1}^{N} \left( 1 - \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} \right) \right) dy \\
\leq \sum_{j=1}^{J} \int_{\chi_j} \left( 1 - \left( 1 - \text{Prob}\left\{ e^{\gamma X_{\nu_j}} \geq y \right\} \right)^N \right) dy \\
\leq \sum_{j=1}^{J} \int_{\chi_j} \left( 1 - \left( 1 - N \cdot \text{Prob}\left\{ e^{\gamma X_{\nu_j}} \geq y \right\} \right) \right) dy \\
= N \cdot \sum_{j=1}^{J} \int_{\chi_j} \text{Prob}\left\{ e^{\gamma X_{\nu_j}} \geq y \right\} dy \\
= N \cdot \sum_{i=1}^{N} \int_{Y_i} \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} dy \\
\leq N \cdot \sum_{i=1}^{N} \int_0^{+\infty} \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} dy \\
= N \cdot \sum_{i=1}^{N} M_i(\gamma) \\
< +\infty.
\]

Equation (3.44) holds because \(e^{\gamma x}\) monotonically decreases with respect to \(x\) for \(\gamma < 0\) and (3.45) follows because of the independence of the \(X_i\)'s. The intervals, \(\chi_j\), in (3.46) are defined as \(\chi_j = [\delta_j^-, \delta_j^+]\) and partition the interval \([0, +\infty)\) such that \(|\mathcal{V}_j| = 1\), where:

\[
\mathcal{V}_j = \arg\min_i \left\{ 1 - \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} \mid y \in \chi_j \right\}.
\]

That is, there is a single random variable, \(X_i\), which dominates all of the others in the sense that:

\[
1 - \text{Prob}\left\{ e^{\gamma X_i} \geq y \right\} < 1 - \text{Prob}\left\{ e^{\gamma X_j} \geq y \right\}.
\]
for all \( j \neq i \) and \( y \in \chi_j \). The breakpoints of the intervals (i.e., the \( \delta_j^- \)'s and \( \delta_j^+ \)'s) are crossing points at which the dominant random variable changes. Thus, we have that 
\[
\bigcup_{j=1}^J \chi_j = [0, +\infty) \text{ and } \chi_j \cap \chi_k = \emptyset, \forall j \neq k.
\]

Each interval, \( \chi_j \), has a corresponding index of the \( X_i \)'s, \( \nu_j \), associated with it. We have that \( \nu_j \in \mathcal{V}_j \). Thus, \( X_{\nu_j} \) is the random variable that minimizes the probability, 
\[
1 - \text{Prob}\{ e^{X_i} \geq y \}, \text{ on the interval } \chi_j.
\]
We also define \( Y_i = \bigcup_{\nu_j \sim i} \chi_{\nu_j} \). Because \( |\mathcal{V}_j| = 1 \), \( Y_i \cap Y_k = \emptyset, \forall i \neq k \). We also note that \( Y_i \subset [0, +\infty) \).

Inequality (3.46) follows from these definitions of \( \chi_j \) and \( \nu_j \). Inequality (3.47) then follows by applying Bernoulli’s inequality to (3.46). Thus, we have that the random variable, \( Z \), has a finite MGF.

**Proposition 3.** If Assumption 9 is satisfied, then the random variable, \( r(a_j, \psi_\omega) \), has a finite MGF.

**Proof.** We begin by rewriting (3.30) as:
\[
r(a_j, \psi_\omega) = \min_{u_\omega \in U_\omega} \sum_{t=\nu_j}^{j+T} p_{t,\omega} \cdot (e_{t,\omega}^c - e_{t,\omega}^d + f_{t,\omega} + L_{t,\omega} - V_{t,\omega}) + C(v_{t,\omega}).
\]

For notational convenience, we define:
\[
\tilde{r}(u_\omega, a_j, \psi_\omega) = \sum_{t=\nu_j}^{j+T} p_{t,\omega} \cdot (e_{t,\omega}^c - e_{t,\omega}^d + f_{t,\omega} + L_{t,\omega} - V_{t,\omega}) + C(v_{t,\omega}).
\]

We define the random vector:
\[
\zeta = (V_{j+1,\omega}, \ldots, V_{j+T,\omega}, L_{j+1,\omega}, \ldots, L_{j+T,\omega}, N_{j+1,\omega}, \ldots, N_{j+T,\omega}, F_{j+1,\omega}, \ldots, F_{j+T,\omega}).
\]

For all \( a_j \in A_j \) and \( \hat{u}_\omega \in U_\omega \), the conditional random variable, \( \{ h(\hat{u}_\omega, a_j, \psi_\omega) | \zeta \} \), has a finite MGF. This is because (3.10) guarantees relatively complete recourse and because of Assumption 9. Thus, we have that the conditional random variable:
\[
\{ r(a_j, \psi_\omega) | \zeta \} = \min_{\lambda \in \Lambda(a_j, \psi_\omega)} \{ \tilde{r}(u_\omega^\lambda, a_j, \psi_\omega) | \zeta \};
\]

91
where $\Lambda(a_j, \psi_\omega)$ is the set of extreme-points of feasible region $U_\omega$, given $a_j$ and $\psi_\omega$. Note that $|\Lambda(a_j, \psi_\omega)|$ is finite. Based on Lemma 1, the conditional random variable $\{r(a_j, \psi_\omega)|\zeta\}$ has a finite MGF. Assumption 9 ensures that $\zeta$ is bounded. Therefore, $r(a_j, \psi_\omega)$ has a finite MGF.

**Proposition 4.** Suppose that Assumption 9 is satisfied. Then the following two properties hold for Problem (3.29):

a. for every $a_j \in A_j$, $r(a_j, \psi_\omega)$ is finite for almost every $\psi_\omega \in \Psi$; and

b. the function $g(a_j)$ is well defined and finite-valued for all $a_j \in A_j$.

**Proof.** To show Property a, we note that (3.10) guarantees that the problem has relatively complete recourse. This guarantees the existence of a feasible second-stage solution for all $a_j \in A_j$ and for almost every $\psi_\omega \in \Psi$ (i.e., the event $\{r(a_j, \psi_\omega) = +\infty\}$ has measure zero $\forall a_j \in A_j$). Moreover, because $U_\omega$ is bounded for all $\psi_\omega \in \Psi$, we have that $r(a_j, \psi_\omega) > -\infty$ for almost every $\psi_\omega \in \Psi$ and for every $a_j$.

To show Property b, we first note that it is easy to show that $r(a_j, \psi_\omega)$ is piecewise-linear convex in $a_j$ for almost every $\psi_\omega \in \Psi$. Thus, $\mathbb{E}[r(a_j, \psi_\omega)]$ is well defined. Because of Proposition 3, the first moment of $r(a_j, \psi_\omega)$ is finite. We further have that $c_j(a_j, \xi_j)$ is well defined and finite. Thus, $g(a_j)$ is well defined and finite. □

Proposition 4 implies that both $g(a_j)$ and $g^M(a_j)$ are convex and continuous on $A_j$.

**Proposition 5.** The objective function of SAA problem (3.31), $g^M(a_j)$, uniformly converges to the objective function of real problem (3.29), $g(a_j)$, with probability 1 (w.p.1) on $A_j$, as $M \to +\infty$. 92
Proof. We know from Proposition 3, that the second moment of \( r(a_j, \psi_\omega) \) is finite. Because \( c_j(a_j, \xi_j) \) is also finite, we have that \( \text{Var} (c_j(a_j, \xi_j) + r(a_j, \psi_\omega)) < +\infty \). For notational convenience, we define:

\[
\varsigma^2(a_j) = \text{Var} (c_j(a_j, \xi_j) + r(a_j, \psi_\omega)).
\]

Due to the law of large numbers, \( g^M(a_j) \) pointwise converges to \( g(a_j) \) w.p.1. By Chebyshev’s inequality, for every fixed \( a_j \in A_j \) and every \( \varepsilon > 0 \) there exists an \( M(a_j, \varepsilon) \in \mathbb{Z}^+ \) such that for all \( m > M(a_j, \varepsilon) \) we have that:

\[
\text{Prob} \{|g^m(a_j) - g(a_j)| < \varepsilon\} \geq 1 - \frac{\varsigma^2(a_j)}{m \varepsilon}.
\]

Define:

\[
\varsigma_{\text{max}}^2 = \sup_{a_j \in A_j} \varsigma^2(a_j);
\]

and:

\[
M(\varepsilon) = \max_{a_j \in A_j} M(a_j, \varepsilon);
\]

for any \( \varepsilon > 0 \). Thus, for all \( a_j \in A_j \) and for all \( \varepsilon > 0 \), we have that:

\[
\text{Prob} \{|g^m(a_j) - g(a_j)| < \varepsilon\} \geq 1 - \frac{\varsigma_{\text{max}}^2}{m \varepsilon};
\]

if \( m > M(\varepsilon) \). Therefore, \( g^M(a_j) \) converges to \( g(a_j) \) uniformly on \( A_j \) w.p.1 as \( M \to +\infty \).  

Using the results of Propositions 4 and 5, we can appeal to the following theorem shown by Shapiro et al. (2014), which gives the convergence of an SAA. We only state this result and refer interested readers to the work of Shapiro et al. (2014) for the proof and further details.

**Theorem 4.** Suppose that the following four properties hold:

93
a. $A^* \neq \emptyset$;

b. $g(a_j)$ is finite valued and continuous on $A_j$;

c. $g^{m_k}(a_j)$ converges to $g(a_j)$ w.p.1, as $m_k \to +\infty$; and

d. $\hat{A}^{m_k} \neq \emptyset$ w.p.1 for $m_k$ sufficiently large.

Then, $g^{m_k} \to g^*$ and $\mathbb{D}(\hat{A}^{m_k}, A^*) \to 0$ w.p.1 as $m_k \to +\infty$, where:

$$\mathbb{D}(\hat{A}^{m_k}, A^*) := \max_{a_j \in \hat{A}^{m_k}} \text{dist}(a_j, A^*).$$

Proposition 4 implies Property b. We can get Property c from Proposition 5. Properties a and d are satisfied because of feasibility of the first stage and the relatively complete recourse property, which comes from Constraint (3.10).

3.3.6 Asymptotic Validity and Finite Termination

The two other important questions about Algorithm 4 are whether it provides an asymptotically valid confidence interval on the optimality gap and whether the algorithm terminates, by satisfying termination criterion $\Upsilon$, in a finite number of iterations. We answer the two questions by relying on the results of Bayraksan and Morton (2011).

To answer these questions, we first show the following proposition.

**Proposition 6.** If Assumption 9 is satisfied, $r(a_j, \psi_\omega)$ is Lipschitz continuous on $A_j$ w.p.1.

**Proof.** To show that $r(a_j, \psi_\omega)$ is Lipschitz continuous, we need to show that there exists a Lipschitz constant, $\mathcal{N}(\psi)$, such that:

$$|r(a_j^1, \psi_\omega) - r(a_j^2, \psi_\omega)| \leq \mathcal{N}(\psi)\|a_j^1 - a_j^2\|;$$
w.p.1 for all $a^1_j, a^2_j \in A_j$.

From Proposition 4 we know that $r(a_j, \psi_\omega)$ is finite for almost every $\psi_\omega \in \Psi$. We further know that $r(a_j, \psi_\omega)$ is piecewise-linear convex on $A_j$. Thus, $r(a_j, \psi_\omega)$ is subdifferentiable and has finite subgradients for all $a_j \in A_j$ and almost every $\psi_\omega \in \Psi$. Thus, there must exist a finite Lipschitz constant, $N(\psi)$, such that the Lipschitz condition is satisfied for almost every $\psi_\omega \in \Psi$. This implies that the Lipschitz constant has a finite MGF. Therefore, $r(a_j, \psi_\omega)$ is Lipschitz continuous w.p.1.

We can now use Proposition 6 to appeal to the following theorem proven by Bayraksan and Morton (2011), which gives the asymptotic validity of the confidence interval and finite termination of the SAA algorithm. We only state this theorem and refer interested readers to the work of Bayraksan and Morton (2011) for further details and the proof.

**Theorem 5.** Suppose that the following five properties hold:

a. $g^{m_\kappa} \to g^*$ and $\mathbb{D}(\hat{A}^{m_\kappa}, A^*) \to 0$ w.p.1 as $m_\kappa \to +\infty$;

b. $G_{l_\kappa}(a_j) \geq D_{l_\kappa}(a_j)$, w.p.1 for all $a_j \in A_j$ and $l_\kappa \geq 1$, where:

$$G_{l_\kappa}(a_j) = g^{l_\kappa}(a_j) - g^{l_\kappa}(a_j^*);$$

and:

$$D_{l_\kappa}(a_j) = g^{l_\kappa}(a_j) - g^{l_\kappa}(a_j^*);$$

where $g^{l_\kappa}(\cdot)$ is defined by (3.31) with the parameter $m_\kappa$ on the right-hand side replaced with $l_\kappa$, $a_j^* \in A^*$, and $a_j^{l_\kappa*} \in \hat{A}^{l_\kappa}$, where:

$$\hat{A}^{l_\kappa} = \left\{ a_j \in A_j \left| g^{l_\kappa}(a_j) = \min_{a_j \in A_j} g^{l_\kappa}(a_j) \right. \right\};$$

95
c. \( \liminf_{\kappa \to \infty} s_{\kappa}^2(a_j) \geq \sigma^2(a_j) \), w.p.1 for all \( a_j \in A_j \);

d. \( \sqrt{l_\kappa} (D_\kappa(a_j) - \mu(a_j)) \xrightarrow{p} N(0, \sigma^2(a_j)) \) as \( l_\kappa \to +\infty \) for all \( a_j \in A_j \), where \( \xrightarrow{p} \) means convergence in probability; and

e. \( r(a_j, \psi_\omega) \) is Lipschitz continuous on \( A_j \).

If the sample sizes used in Step 4 of Algorithm 4 satisfy (3.39) and the algorithm terminates by satisfying (3.38), then for any \( \epsilon > \epsilon' > 0 \), \( q > 0 \), and \( \alpha \in (0, 1) \) we have that:

\[
\lim_{h \to h'} \inf \Pr \{ \mu(a_j^{\text{mr}*}) \leq h\bar{s}(a_j^{\text{mr}*}) + \epsilon \} \geq 1 - \alpha;
\]

where \( a_j^{\text{mr}*} \) is the final solution found when Algorithm 4 terminates. Moreover, for any \( h > h' > 0 \) we have that:

\[
\Pr \{ \Gamma < +\infty \} = 1.
\]

Our original and SAA problems satisfy the requirements of Theorem 5. Theorem 4 implies Property a. We get Property b from the fact that \( g^\kappa(a_j^{*}) \leq g^\kappa(a_j^*) \). Bayraksan and Morton (2006) show that Property c holds if \( r(a_j, \psi_\omega) \) is continuous with respect to \( a_j \in A_j \) w.p.1, \( \mathbb{E} \left[ \sup_{a_j \in A_j} r^2(a_j, \psi_\omega) \right] < +\infty \) and \( A_j \) is non-empty and compact. Property d is satisfied because of the central limit theorem for i.i.d. random variables. Finally, Proposition 6 gives Property e.

3.4 Case Study

The station-control model and solution technique introduced in Sections 3.2 and 3.3 are demonstrated using a case study. The case study is used to study the optimal scheduling of EV charging loads and the resulting cost of operating the charging
station. These cost calculations can be used to estimate the impact of building a public EV-charging station from energy- and infrastructure-cost perspectives. We also contrast the charging profiles and resulting cost given by our proposed model to three simple heuristic charging techniques. This is done to demonstrate the benefits of the proposed model in scheduling flexible EV-charging loads more intelligently than simple heuristics.

The case study is based on the Central-Ohio region. It assumes that a public EV-charging station is built in the parking area of a major retail shopping center in northeastern Columbus, Ohio. The shopping center is connected to the electricity-distribution system through an $R = 500$ kVA transformer. This transformer, which is shared by the buildings in the retail shopping center, is also used by the EV-charging station. Thus, the transformer must serve the total EV and non-EV loads. Station operations are modeled over a representative year, using electricity-price, weather, and vehicle-usage data for the year 2013.

A number of the case study parameters (e.g., the non-EV load, electricity prices, and transformer characteristics) are fixed. We examine the effects of other parameters through a fractional factorial experiment. Specifically, we examine the effects of the EV charger power capacity, whether distributed PV and BESS are installed, the EV penetration level, and the charging window through this experimental design.

We detail the datasets used and how these data are processed in constructing our case study in Sections 3.4.1–3.4.7. Section 3.4.8 describes our fractional factorial experimental design.
3.4.1 Electric Vehicle Usage

EV-usage data, which are implicitly needed to determine EV arrival times and the charging demands of EVs upon their arrivals, are modeled using vehicle-driving data provided by the Mid-Ohio Regional Planning Commission (MORPC). More specifically, our work relies on two MORPC datasets: (1) tour-record data specifying the use of the approximately 1.1 million light-duty vehicles in the Central-Ohio region and (2) geographic information system data for the region, which include average vehicle-speed data for all of the road segments in the road network.

The tour records provide modeled data on the use of the 1.1 million light-duty vehicles in Central Ohio during a typical day. The data are generated using the multi-step tour-based approach discussed by Sadeghi-Barzani et al. (2014). Each tour is associated with a specific vehicle in the modeled dataset and information on subtours (e.g., stopping at a retail shopping location *en route* while commuting from work to home) are included as well. The dataset also provides the starting and ending times and locations of each tour and subtour.

The MORPC data are used to simulate EV-usage patterns by following a four-step process. The first step is to determine what subset of the 1.1 million light-duty vehicles in the MORPC dataset are EVs by bootstrapping from the tour-record data. This is done by assuming an EV penetration level (as discussed in Section 3.4.8, we consider cases in which the EV-penetration level is 3% and 8% of the light-duty vehicle fleet of Central Ohio) and randomly selecting a corresponding number of vehicles in the tour-record data to be EVs.

The second step is to determine the path taken by each randomly selected EV as it completes each driving trip during the day. This is done by assuming that each
EV owner takes the shortest-time path on each vehicle trip. Using this assumption and the road-network and average vehicle-speed data provided by MORPC, the EV routes are determined by solving a shortest-time-path problem.

Once the vehicle paths are determined, the third step is to determine the SOC of each EV battery as it travels along its path. This is done using an assumed battery capacity of 24 kWh and a driving efficiency of 3.73 kWh/km. These values correspond to reported characteristics of a Nissan Leaf.

The fourth step is to determine which EVs use the modeled charging station. In doing so, we make two additional assumptions. The first is that an EV driver only uses the charging station if the vehicle comes within 1.6 km (1.0 mile) of the charging station while traveling along the intended shortest path from origin to destination. The second is that an EV driver will only use the charging station if the EV battery’s SOC is below some set threshold value, which we take to be 50%. This threshold value is meant to capture the perceived inconvenience of using a public charging station. An EV driver only uses one if the SOC of the EV battery is sufficiently low.

The arrival time of an EV that does use the charging station is determined based on its departure time to begin the tour that brings it within the 1.6 km radius of the station (which is reported in the tour-record data) and the driving time to arrive to the station (which is determined from the shortest-time-path problem). Its charging demand is determined based on its modeled battery SOC upon arrival at the station (assuming that the EV driver wishes its battery fully recharged). We further assume that 80% of EVs are flexible, in the sense that the charging station can schedule their charging within a $W$-minute window of time. The others are assumed to be inflexible and begin charging immediately at the $H$ kW rated charger capacity upon arrival to
the station. The loads of such EVs are modeled by including them in the $L_j$ and $L_{t,\omega}$ state parameters. EVs are randomly assigned to be either flexible or inflexible.

Numerical simulations show that the temporal distribution of the number of vehicles that come within the 1.6 km capture radius of the charging station is bimodal. 18.6% of the vehicles that come within the capture radius in a typical day arrive between 7:00 am and 9:00 am and another 19.0% arrive between 4:00 pm and 7:00 pm, which are typical commuting times. However, very few of the EVs that come within the capture radius during the morning commute have a sufficiently low SOC to use the charging station. Less than 0.5% of the EVs that come within the capture radius between 7:00 am and 9:00 am have a battery SOC below the 50% threshold, meaning that the overwhelming majority of them would not use the charging station in the morning. This is because we assume that each EV begins each day with a fully charged battery (from overnight at-home vehicle recharging) and the relatively short cumulative driving distance that each EV has gone in the morning. Conversely, 5% of EVs that come within the capture radius between 4:00 pm and 7:00 pm have a battery SOC below the 50% threshold, meaning greater use of the charging station during the evening commute. Averaging across all of the hours, about 3% of EVs that come within the capture radius during the course of a typical day have a battery SOC below the 50% threshold. With an EV penetration level of 3% of the total light-duty vehicle fleet in Central Ohio, the charging station has 17 expected EV arrivals per day with an SOC threshold of 50%. This number increases to 45 expected daily EV arrivals with an EV penetration level of 8%.
3.4.2 Non-Electric Vehicle Load

We model the non-EV load using data from the year 2013 provided by American Electric Power Ohio (AEP Ohio), the distribution utility company that operates in Central Ohio. The data consist of anonymized load profiles on a number of 500 kVA commercial transformers, recorded at a 15-minute time resolution. We fit an order-16 autoregressive model (i.e., an AR(16) model) to generate forecasts of non-EV load for the scenarios used in the station-control model.

3.4.3 Photovoltaic Generation

We model the real-time output of the distributed PV generator using historical weather data from the year 2013 for the city of Columbus. These weather data are input to version 5 of the PVWatts model, the use of which is detailed by Dobos (2014). PVWatts uses weather data to simulate the output of a hypothetical PV generator. We obtain historical solar-insolation data from the National Solar Radiation Database (NSRDB). As discussed by Sengupta et al. (2014), the NSRDB models ground-level solar insolation using satellite-image data. Other weather data needed by PVWatts (i.e., wind speed and temperature) are obtained from the National Oceanic and Atmospheric Administration. We run PVWatts with a one-minute temporal resolution, to obtain simulated one-minute PV-output data.

We fit an AR(6) model to the PV-output data simulated by the PVWatts model. This model is used to generate forecasts of PV availability for the scenarios used in the station-control model. Figure 3.1 shows modeled PV availability from a 200 kW panel on 5 December, 2013. The solid curve shows actual output available from the panel between 14:54 and 15:51. The dotted line and the pairs of squares shows the
mean and three-standard deviation band around this mean of the forecast of PV availability as of 15:11. The line with the circles shows one sample path of forecasts as of 15:11.

Figure 3.1: Modeled PV Availability From a 200 kW Panel on 5 December, 2013: Actual and Forecast Mean, Three-Standard Deviation Band, and Scenario-1 Forecast as of 15:11

### 3.4.4 Electricity Price

Real-time electricity prices are forecasted using an AR(8) model that is fit to hourly historical real-time prices from 2013. Specifically, we use real-time prices for the AEP zone of the PJM Interconnection market, which covers the AEP Ohio service territory. Our model assumes that the real-time price is constant over each one-hour period, which is consistent with wholesale pricing practice in PJM. As is common practice in wholesale electricity markets, our model does allow for negative energy prices.
3.4.5 Battery Energy Storage System Characteristics

We assume that the BESS has upper and lower bounds on its SOC of $S_{E,+} = 4200$ kW-min (70 kWh) and $S_{E,-} = 1200$ kWh (20 kWh), respectively, and a power capacity of $S_P = 100$ kW. We also assume charging and discharging efficiencies of $\mu^c = \mu^d = 0.99$. These values are typical for a lithium-ion BESS, as reported by Xi and Sioshansi (2016).

3.4.6 Transformer Characteristics

Our case study assumes that the charging station is connected to the power system through a 500 kVA transformer, which it shares with the retail shopping center where the station is located. Our model allows the transformer to operate above its 500 kVA design capacity. However, doing so places additional strain on the transformer, causing it to age and deteriorate more rapidly.

We capture this accelerated aging using the model introduced by Gong et al. (2011). This model simulates the effects of operating a transformer above its design capacity on the temperature of the windings in the transformer, which is the component that is typically prone to failure from such overloading. More specifically, we use the model to estimate the service life of the transformer if it is operated at its design capacity of 500 kVA. We then determine what effect operating the transformer above the 500 kVA design capacity has on reducing its service life. This reduction in service life is combined with an assumed $10000$ replacement cost for a transformer, to derive the following piecewise-linear convex penalty function:

$$R^C(v_j) = \begin{cases} 
1.16v_t, & \text{if } v_t \in [0, 0.4\bar{R}); \\
42.65(v_t - 0.4\bar{R}) + 232.70, & \text{if } v_t \in [0.4\bar{R}, 0.6\bar{R}); \\
764.62(v_t - 0.6\bar{R}) + 4497.35, & \text{if } v_t \in [0.6\bar{R}, 0.8\bar{R}); \\
12309.73(v_t - 0.8\bar{R}) + 80959.50, & \text{if } v_t \in [0.8\bar{R}, +\infty); 
\end{cases}$$

103
for operating the transformer above its design capacity.

### 3.4.7 Temporal Resolution

Our model uses a one-minute temporal granularity in scheduling EV charging and making BESS-charging and -discharging decisions. The model has a $T = 60$ minute optimization horizon. This is a reasonable model horizon, because each EV can be fully recharged within at most 30 minutes, based on the assumed capacity of the EV batteries and the power capacity of the EV chargers (i.e., 30 minutes is how long it would take for a fully depleted EV battery to be fully recharged). As a result, the primary control decision, which is how much EV-charging load to schedule in the current minute as opposed to deferring for later, can only be delayed by at most 30 minutes into the future. As such, the state of the charging station more than 60 minutes in the future is expected to have a relatively insignificant effect on the current-minute scheduling decision.

### 3.4.8 Fractional Factorial Experimental Design

The operations of a public EV charging station depends on a number of factors that are either uncertain or need to be determined by the station designer. We examine the impacts of five such factors—$H$ (the EV charger power capacity), whether distributed PV or BESS are installed in the charging station, the EV penetration level (as a percentage of the light-duty vehicle fleet in Central Ohio), and $W$ (the charging window for flexible EVs)—through a fractional factorial experiment. We assume that each of these five factors have two possible values. Table 3.1 lists the eight cases, with different combinations of factors, examined. Note that a full factorial experiment would require the evaluation of $2^5 = 32$ cases. We opt to conduct the
fractional design summarized in Table 3.1, which assumes only first- and second-order interactions amongst the factors. We conduct this fractional experiment because a full factorial experiment would be computationally expensive.

Table 3.1: Fractional Factorial Experimental Design

<table>
<thead>
<tr>
<th>Case</th>
<th>$\bar{H}$ [kW]</th>
<th>PV</th>
<th>BESS</th>
<th>EV Penetration</th>
<th>$W$ [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>no PV</td>
<td>no BESS</td>
<td>3%</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>no PV</td>
<td>BESS</td>
<td>3%</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>200 kW</td>
<td>no BESS</td>
<td>3%</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>200 kW</td>
<td>BESS</td>
<td>3%</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>120</td>
<td>no PV</td>
<td>no BESS</td>
<td>8%</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>no PV</td>
<td>BESS</td>
<td>8%</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>200 kW</td>
<td>no BESS</td>
<td>8%</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>120</td>
<td>200 kW</td>
<td>BESS</td>
<td>8%</td>
<td>40</td>
</tr>
</tbody>
</table>

We can examine the effects of different factors by averaging the cases corresponding to that factor. For instance, we can estimate the operating cost with a BESS installed in the charging station by averaging the operating costs given in Cases 2, 4, 6, and 8. We can then contrast this with the operating cost without a BESS installed, which is given by averaging across Cases 1, 3, 5, and 7. This averaging allows us to determine the impacts on operating cost of having a BESS installed, without other factors confounding the results.

3.5 Case-Study Results

We use the case study outlined in Section 3.4 to analyze the EV-charging station in two ways. First, in Section 3.5.1 we demonstrate the use of the station-control model over a one-year period assuming a fixed station design that consists
of a 70 kWh/100 kW BESS, a 200 kW distributed PV system, $\bar{H} = 120$ kW EV chargers, an EV-penetration level of 8%, and a $W = 40$ minute charging window for flexible EVs. We also compare the charging profiles given by our station-control model to three simple heuristics, demonstrating the benefits of our proposed model. This analysis is done using the fractional factorial experiment, which is outlined in Section 3.4.8, to examine cases with EV penetration levels of 3% and 8% and with 30- and 40-minute charging windows. Next, in Section 3.5.2 we examine the sensitivity of charging-station operations to having the distributed BESS and PV system and the power capacity of the EV chargers. This analysis is also conducted using the fractional factorial experiment.

### 3.5.1 Station-Control Model Results

We first examine the charging profiles and resulting transformer-load profiles given by the station-control model under a fixed station-design setting. This setting assumes that the charging station consists of a 70 kWh/100 kW BESS, 200 kW distributed PV system, and $\bar{H} = 120$ kW EV chargers. We further assume an EV-penetration level of 8% and a $W = 40$ minute charging window for flexible EVs. Figure 3.2 shows the total load on the transformer over a 2.5-hour-long period of time if EV charging loads are scheduled using the proposed station-control model. It also shows the transformer load if EV charging is instead scheduled using a simple first-come, first-served (FCFS) heuristic. This FCFS heuristic assumes that each EV immediately begins charging at the $\bar{H} = 120$ kW charger capacity until it is fully recharged (i.e., it assumes that the flexibility in scheduling EV charging loads is not used).
Figure 3.2: Transformer-Load Profile From Station-Control Model and Using First-Come, First-Served Heuristic

Figure 3.2a shows that under a FCFS heuristic, the transformer is operated above its 500 kVA design capacity whereas the station-control model schedules EV charging loads to keep the load below this level. Over the course of the entire day shown in Figure 3.2, the FCFS heuristic results in the transformer operating above the 500 kVA design capacity for 54 minutes, as opposed to never operating above this capacity with the proposed station-control model.
Figure 3.2b shows that the distributed PV and BESS play limited roles in helping to manage EV-charging loads. As discussed in Section 3.4.1, most of the EV arrivals to the charging station occur during the evening commute home, when PV output is beginning to taper off as the sun is setting. As a result, the primary benefit of the distributed PV system is to help reduce the use of grid-purchased energy to serve the non-EV load midday.

Moreover, EV loads are mostly managed by shifting them within the W-minute charging window. The BESS is only used in a limited fashion when there is insufficient flexibility to keep the transformer load below its 500 kVA capacity. The BESS is only discharged in 17 minutes of the entire day shown in Figure 3.2. During these 17 minutes it is discharged, on average, at 70 kW, which is far short of its $S^P = 100$ kW capacity. The reason for this is that shifting EV-charging loads is a cost-free means of keeping the transformer load below 500 kVA. Using the BESS imposes a small cost of about 2% (relative to the wholesale energy price), due to the loss of energy when charging and discharging the BESS. As a result, the model prefers shifting EV-charging loads and only uses BESS when there is insufficient flexibility to keep the load below 500 kVA. Otherwise, if there is no BESS capacity available (or in cases without a BESS installed) and there is insufficient flexibility availability to shift EV-charging loads, the model is forced to load the transformer above its 500 kVA capacity. Instances in which this occurs are illustrated in Sections 3.5.1 and 3.5.2.

The station-control model and the solution method proposed in Section 3.3 are implemented using CPLEX version 12.6.1. When run on a computer with a 3.0 GHz Intel Pentium G3220 Processor and 16 GB of memory, each instance of the one-minute station-control models takes an average of 39.44 seconds of CPU time to run.
Effect of First-Come, First-Served Policy on Transformer Aging

Figure 3.2 shows that one of the most severe impacts of relying on a simple heuristic, such as FCFS, is that the transformer can be easily overloaded, especially in the afternoon when many EVs simultaneously arrive to the station. Figure 3.3 shows the effects on the expected operational life of the 500 kVA transformer of relying on a FCFS heuristic. It shows the expected lifetime of the transformer under different EV penetration levels and with different threshold SOC levels at which EV owners use the public charging station.

We estimate the expected lifetime of the transformer by randomly simulating 1000 replications of EV-usage, PV generation, and non-EV load data over a one-year period. Each of the 1000 resulting load profiles are then simulated in the transformer aging model of Gong et al. (2011), by assuming that the load profiles repeat year after year, to determine the lifetime of the transformer under that load profile. Thus,
we do not account for load growth over the lifetime of the transformer, which would accelerate transformer aging relative to what is reported in Figure 3.3.

Distribution transformers normally have a planned operational life of 20.5 years. The figure shows that in many instances, the FCFS heuristic does not reduce the expected life of the transformer below this level. However, in cases in which the SOC threshold is below 40% and the EV-penetration level is above 4%, the expected life of the transformer can be reduced, on average, by up to 32% relative to the 20.5-year design life. Moreover, some of the random replications used to simulate transformer aging result in load profiles with extreme peaks that reduce the expected life of the transformer to 0.057 years (21 days). Although the likelihood of such a load profile is small, this result nevertheless points out the importance of managing EV-charging loads from the perspective of reducing the cost of having to replace distribution infrastructure.

**Comparison of Station-Control Model and Heuristic Scheduling Techniques**

FCFS is a relatively simple heuristic in which the flexibility of EV charging is not exploited in any way, because EV-charging loads are served immediately without any scheduling. Thus, we examine two other heuristic charging-control techniques.

The first, which we refer to as Constrained FCFS, has the same basic premise as the FCFS heuristic. When an EV arrives to the station, it immediately begins charging at the $\bar{H}$ kW capacity of the charger, so long as doing so does not overload the transformer. If charging the newly arrived EV would overload the transformer, the newly arrived EV instead queues and waits for one of the EVs that is currently charging to finish charging, at which point the newly arrived EV begins charging. If another EV arrives while an earlier arrival is queuing, this newly arrived EV also queues. The
Constrained FCFS heuristic is guaranteed not to overload the transformer, because each EV only charges so long as there is transformer capacity available. However, the Constrained FCFS heuristic may result in an EV not being able to fully recharge within the $W$-minute window of time.

We refer to the other heuristic as Uniform Charging. Upon arrival to the charging station, each EV charges at $(F/W)$ kW, where $F$ is the total charging demand of the EV upon arrival to the station measured in kW-min. Thus, the Uniform Charging heuristic has each EV spread its charging demand uniformly over the $W$-minute charging window. By design, the Uniform Charging heuristic is guaranteed to fully recharge each EV. However, it is possible that the transformer can be overloaded if too many EVs arrive within a short span of time.

Figure 3.4 shows the load-duration curves of the 1000 hours of the year with the highest load on the transformer when EV charging is scheduled using our proposed station-control model and the three heuristics. The four plots show the load-duration curve for four different settings, in which the charging window is either 30 or 40 minutes and the EV penetration level is 3% or 8% of the light-duty vehicle fleet of Central Ohio. The figures are generated by conducting the fractional factorial design discussed in Section 3.4.8, and each figure shows the corresponding average load-duration curve. For instance, Figure 3.4a, which corresponds to having a 3% EV penetration, shows the average load-duration curve among Cases 1–4 in Table 3.1.

As expected, the figure shows that the proposed model significantly reduces the number of hours in which the transformer is operated above its 500 kVA capacity. The only heuristic that outperforms the station-control model on this metric is the Constrained FCFS heuristic, which is designed to keep the transformer load below
Figure 3.4: Load-Duration Curve of 1000 Hours of the Year With Highest Transformer Loads Under Station-Control Model and Three Heuristics

500 kVA at all times. Tables 3.2 and 3.3 summarize the performance of the station-control model and the three heuristics on two key metrics.

Table 3.2 lists the average (using the fractional factorial experimental design) number of hours in which the transformer is operated above its 500 kVA capacity in the different cases examined and the peak transformer load. Values for the Constrained FCFS heuristic are not listed, because, by definition, this heuristic never overloads the transformer. The table shows that the station-control model outperforms the
heuristics in terms of the number of hours in which the transformer is overloaded. However, in some of the cases the station-control model puts a higher peak load on the transformer. This is because there is an inherent tradeoff when there are many EVs needing to have their charging loads satisfied. The Uniform Charging heuristic spreads these loads more evenly, which can result in a lower peak load but the transformer being overloaded in more hours. The station-control model, conversely, opts
to concentrate these loads in a fewer number of hours, which can give a higher peak 
but fewer overloaded hours.

Unlike the other heuristics and station-control model, the Constrained FCFS 
heuristic can result in EVs not being fully recharged. Table 3.3 summarizes the 
average (using the fractional factorial experiment) number of EVs that are not fully 
recharged over the course of the year and the total unserved EV charging load if 
the Constrained FCFS heuristic is employed. With our assumed EV efficiency of 
3.73 kWh/km, the Constrained FCFS results in the equivalent of between 1518 km 
and 12779 km of EV charging demand not being satisfied over the course of the year.

3.5.2 Parameter Sensitivity Analysis

We now examine the effects of changing a number of the model parameters in 
our case study, using the fractional factorial experimental design discussed in Sec-
tion 3.4.8. Specifically, we examine the effect of removing the BESS and distributed 
PV generator and the effect of reducing the power capacity of the EV chargers. This 
analysis is done by simulating the station-control model over a one-year period under 
the different cases listed in Table 3.1, and averaging the resulting costs, load profiles, 
and other key metrics across the cases corresponding to each parameter examined.

Battery Energy Storage System

Figure 3.5 shows the average load-duration curves of the 1000 hours of the year 
with the highest load on the transformer with and without the distributed BESS 
installed in the charging station. The figure shows that the BESS improves the 
loading on the transformer by allowing some relief of the transformer during high-
load periods. The BESS reduces the average duration of time that the transformer
is overloaded during the year from 36.6 hours without the BESS to 0.4 hours with the BESS. The BESS also reduces the average peak load on the transformer from 766.7 kW to 691.4 kW. This comes with an increase in the average number of hours that the transformer is operating close to its capacity, however, because the BESS must be charged to provide this relief. Without the BESS the transformer is operated at between 400 kVA and 500 kVA for an average of about 786.4 hours as opposed to 834.8 hours with the BESS.

Figure 3.5: Average Load-Duration Curve of 1000 Hours of the Year With Highest Transformer Loads With and Without Battery Energy Storage System

The BESS also increases total average energy consumption of the distribution circuit (i.e., including EV and non-EV loads) over the year by about 1 MWh. The increased load is due to energy losses when storing and discharging energy into and from the BESS. However, the BESS gives an average annual reduction in operation costs of about 72%, from $720000 without the BESS to $200000 with the BESS. These cost savings are entirely from reduced transformer aging, which are very slightly offset by the increased energy consumption.
Photovoltaic Generator

Figure 3.5 shows the average load-duration curves of the 1000 hours of the year with the highest load on the transformer with and without the distributed PV generator. While the PV generator does improve the load profile, its impact is more limited compared to the BESS. This is because PV generation tends to taper off during the afternoon and evening commute hours, when most EVs arrive at the charging station. More specifically, on average about 75.3% of daily EV-charging demand arrives to the charging station after 5:00 pm. However, the PV generator only produces 0.5% of its total output after 5:00 pm.

![Figure 3.6: Average Load-Duration Curve of 1000 Hours of the Year With Highest Transformer Loads With and Without Photovoltaic System](image)

Nevertheless, the PV generator does reduce the average load on the transformer by 12% from 316.7 kW to 279.6 kW. It has a more limited effect on the peak load, reducing it from 745.0 kW to 713.1 kW. The PV generator also reduces the average duration of time that the transformer is overloaded by 42% from 23.4 hours to
13.6 hours. However, the PV generator has a greater load-reduction effect during relatively low-load period than during peak-load periods. During hours that the total EV and non-EV load is between 0 kW and 500 kW, the PV generator reduces transformer loading by an average of 12.0%. On the other hand, the PV generator only gives an average of a 6.8% reduction in transformer loading during hours that the total EV and non-EV load is greater than 500 kW. There is, nevertheless, a 58% reduction in average annual station-operation costs from $650000 without the PV generator to $270000 with the PV generator.

**Electric Vehicle Charger Power Capacity**

Figure 3.5 shows the average load-duration curves of the 1000 hours of the year with the highest load on the transformer with 120 kW and 50 kW EV chargers. 50 kW chargers reduce transformer loading by simply reducing the maximum amount of power that the EVs can draw. In net, the duration of time that the transformer is overloaded during the year is reduced on average by 43% from 23.6 hours to 13.4 hours with 50 kW chargers. There is also a 15% average reduction in the peak load from 786.4 kW to 671.7 kW.

**3.6 Conclusion**

This chapter presents a model to schedule EV charging. Our proposed use of the model is to manage EV-charging loads at a public station that has fast chargers installed. Fast chargers can result in extreme peaks in the load on a radial distribution line serving the station. One way to accommodate such a load profile is to increase the capacity of the distribution infrastructure, including distributed transformers. This is likely an inefficient solution, however, because the peaks in the load profile only
Figure 3.7: Average Load-Duration Curve of 1000 Hours of the Year With Highest Transformer Loads With 120 kW and 50 kW Electric Vehicle Chargers

occur during a limited window of time when EVs are commuting home from work. Our model manages these loads by using the flexibility in recharging EVs while they are parked. Our model also allows for the use of distributed resources, such as PV and BESS, to help manage the load.

We also propose a solution technique, which employs SAA and Benders’s decomposition, to efficiently solve the resulting two-stage station-control problem. We employ an SSP to estimate and control the quality of the solutions obtained from the SAA. Our numerical experiments show that when implemented on a standard computer, the proposed technique is able to give solutions that have an optimality-gap-to-standard-deviation ratio of 0.9 within an average of 39.44 seconds. This means that the model could be employed in the control system of a public charging station.

We demonstrate the use of the model using a case study based on Central Ohio. We show the benefit of the proposed model in serving EV-charging loads while minimizing the impacts of the resulting load profile on the distribution transformer. We
also show the benefits of the proposed model compared to using simple heuristics, such as Constrained FCFS and Uniform Charging. Our results suggest that without a control model, such as the one that we propose, a public EV fast charging station would likely require costly transformer upgrades.

We also conduct a fractional factorial experiment to determine the impacts of distributed BESS and PV on managing charging-station loads. While these technologies have some benefits, they are relatively small compared to the benefits obtained from directly managing and scheduling EV-charging loads. Thus, our results suggest that if EVs have flexibility in having their charging loads scheduled, this flexibility provides an effective and cost-free means of reducing the distribution-impacts of EV charging. One question that our work does not address is what the optimal mix of transformer upgrades and distributed PV and BESS are to minimize the infrastructure and energy cost of operating the charging station. This is an area of future research that our model could be employed to address.
Chapter 4: Stochastic Control Strategies for EV Charging
Station Control and Frequency Regulation

4.1 Introduction

Ancillary services are necessary to maintain power system balance between electricity supply and demand which ensures system security and reliability. Mismatches between the supply and demand leads to a frequency deviation. A large deviation from nominal frequency is unacceptable for power quality. In practice, frequency control is one of the most important ancillary services. Rebours et al. (2007b,a) discuss that frequency regulation, spinning reserve, and non-spinning reserve are three most common services used for frequency control. Frequency regulation is used to fine-tune the frequency by matching the supply and demand in the time scale from several seconds to one minute. Due to the quick response feature, the frequency regulation service has the highest price of all ancillary services.

Energy storage, microgrids and plug-in electric vehicles are good sources to participate in the regulation market, provide regulation capacity, and respond to frequency regulation signals. Arsie et al. (2009) and Drury et al. (2011) assess the value of energy storage for providing ancillary services to the power system. Kempton and
Tomić (2005a,b) and Dallinger et al. (2011) analyze the potential of PEVs participating in the frequency regulation market and the effectiveness of EVs providing regulation services. Bessa and Matos (2012) summarize the roles of an EV aggregator (e.g., public charging station operator, microgrid with EV charging overnight, etc.) and indicate that one of its important goals is to provide ancillary services to power system operators.

In this chapter, we introduce a stochastic station control method for EV aggregators. This method is used to create optimized operation strategies for a DC fast charging station equipped with a battery energy storage system and a distributed photovoltaic generator. The operational strategies include rescheduling EV charging load, managing the battery energy storage system and photovoltaic generator, arbitraging in the real-time energy market, and participating in the frequency regulation market. The proposed station control model includes several uncertainties such as uncertain EV arrivals, random regulation signals and stochastic PV output. We model this problem as a stochastic optimization problem which is able to use the solution techniques introduced in Chapter 3 to obtain a high-quality solution. A case study on a potential DC fast charging station in Central Ohio shows that the station control model is able to provide accurate frequency regulation signal responses. It also shows that the strategies are also able to protect the distribution system from overheating by mitigating flexible EV charging load within a window.

4.2 Control Model

We model a public DC fast charging station with distributed PV panels. The station also has a battery energy storage system installed. We assume that the EV
charging load and the other non-EV load from the neighboring buildings connect to the power grid through the same transformer. That is, the load on the transformer consists of non-EV load (e.g., electricity used for building lighting, air-conditioning, etc.) and EV-charging load. We also assume that this charging station is able to participate in the ancillary service market and provides frequency regulation to energy market. The regulation frequency service is provided in the form of providing an hourly amount of regulation capacity to the power system operator at the beginning of each hour. This amount keep the capacity being constant during the hour.

We formulate this problem as a 2-stage stochastic programming model. It minimizes the expected station operation cost within a optimization horizon. The first stage represents the current time step; and the model determines the current optimal station operation strategy. The second stage represents the charging station operations from the next time step to the end of the optimization horizon.

The station operation strategy consist of five types of decisions: the amount of power to purchase from the grid; how much energy to charge to and discharge from BESS; the amount of frequency regulation capacity to sell; the real-time response to the regulation signal; and the charging schedule of EVs. The first stage of the model is based on realized market information, such as electricity and frequency regulation prices, and station information, such as the number and charging demand of EVs, PV panel output, and the storage level.

Each scenario of the second stage represents a sample path based on forecasting market and station states in subsequent time steps. For notation convenience, we assume that the current time step is minute \( j \). The variables and parameters at and before minute \( j \) are known; the others after minute \( j \) are unrealized. For those
unknown parameters and variables, we use the superscript $\omega$ to indicate that they correspond to scenario $\omega$. We indicate past known variables with a hat, meaning that the variable has already been decided and is fixed.

### 4.2.1 Model Parameters

- $\Omega$: Set of second-stage scenarios
- $T$: Optimization horizon [min]
- $W$: Charging window for flexible EVs [min]
- $H$: EV-charger nameplate capacity [kW]
- $R$: Nameplate transformer capacity [kW]
- $R^C(\cdot)$: Penalty for operating transformer above nameplate capacity [$/\text{min}]
- $S^{E,+}$: Maximum SOC of the BESS [kW-min]
- $S^{E,-}$: Minimum SOC of the BESS [kW-min]
- $S^P$: Charging and discharging capacity of the BESS [kW]
- $\mu^c$: Storage charging efficiency
- $\mu^d$: Storage discharging efficiency
- $\phi$: Frequency regulation service unsatisfied penalty constant

The maximum and minimum storage levels (i.e., $S^{E,+}$ and $S^{E,-}$) are used to protect lithium-ion batteries from extreme cycle-life degradation. $\mu^c$ and $\mu^d$ are the storage charging and discharging efficiency, respectively. They also implicitly represent a cost of storing energy in BESS. $\phi$ is used to penalize the unfulfilled frequency regulation demand that the station has committed itself to provide. This parameter satisfies:

$$
\phi > \max \left\{ \frac{p_t^c}{p_t} \left( \frac{1}{\delta_j^i} - 1 \right) \bigg| \forall \delta_j^i > 0, \forall p_t > 0, \forall i = u, d \right\};
$$

where $p_t$, $p_t^c$ and $\delta_j^i$ are defined in Section 4.2.2.

We assume that flexible EVs have a charging window of $W$ minute. The charging window is usually wider than the time span which is required to fully charge an empty EV battery. This assumption guarantees problem feasibility of rescheduling and fully
fulfilling EV charging demands. \( T \) defines the optimization horizon within which the expected cost is minimized.

### 4.2.2 State Parameters

#### Deterministic State Parameters

\( L_j: \) Minute-\( j \) inflexible energy demand [kW]

\( F_t: \) Total charging demand of flexible EVs that arrived to the charging station in minute \( t \), where \( t \in \{ j - W + 1, \ldots , j \} \) [kW-min]

\( N_t: \) Total number of flexible EVs that arrived to the charging station in minute \( t \), where \( t \in \{ j - W + 1, \ldots , j \} \) [kW-min]

\( \bar{f}_t: \) Total charging demand of flexible EVs that arrived to the charging station in minute \( t \), where \( t \in \{ j - W + 1, \ldots , j \} \), that has been satisfied as of the beginning of minute \( j \) [kW-min]

\( V_j: \) Minute-\( j \) PV output [kW]

\( p_j: \) Minute-\( j \) electricity price [$\$/kW-min]

\( p_j^C: \) Minute \( j \) frequency regulation capacity price [$kW-min]

\( x_j: \) Beginning of minute \( j \) SOC of the BESS [kW-min]

\( \delta^u_j: \) Minite \( j \) regulation-up ratio of the regulation-up amount to capacity

\( \delta^d_j: \) Minite \( j \) regulation-down ratio of the regulation-down amount to capacity

\( \bar{f}_t \) is realized and fulfilled flexible EV charging demand. We assume that station operator receives frequency regulation signal, \( \delta^u_j \) and \( \delta^d_j \), at the beginning of each time unit; the signals keep constant in the time unit. The regulation signal indicates the proportion of the amount of required regulation power to the committed regulation capacity. We also define a state-parameter vector, \( \xi_j = (L_j, F_t, N_t, \bar{f}_t, V_j, p_j, p_j^C, x_j, \delta^u_j, \delta^d_j) \).

#### Stochastic State Parameters

\( L_{t,\omega}: \) Minute-\( t \) inflexible energy demand in scenario \( \omega \) [kW]

\( F_{t,\omega}: \) Total charging demand of flexible EVs that arrived to the charging station in minute \( t \) of scenario \( \omega \) [kW-min]

\( N_{t,\omega}: \) Total number of flexible EVs that arrived to the charging station in minute \( t \) of scenario \( \omega \) [kW-min]

\( V_{t,\omega}: \) Minute-\( t \) PV output in scenario \( \omega \) [kW]

\( p_{t,\omega}: \) Minute-\( t \) electricity price in scenario \( \omega \) [$\$/kW-min]

\( p_{C,t,\omega}: \) Minute-\( t \) frequency regulation capacity price in scenario \( \omega \) [$kW-min]
\( \delta_{u,\omega}^t \): Minute-\( t \) regulation-up ratio of the regulation-up amount to capacity in scenario \( \omega \)

\( \delta_{d,\omega}^t \): Minute-\( t \) regulation-down ratio of the regulation-down amount to capacity in scenario \( \omega \)

The stochastic state parameters are generated using forecasting models. These parameters define each second-stage scenario, \( \omega \). We use an EV transportation simulation to generate \( F_{t,\omega} \) and \( N_{t,\omega} \). We define \( \psi_{t,\omega} = (L_{t,\omega}, F_{t,\omega}, N_{t,\omega}, p_{t,\omega}, p^C_{t,\omega}, \delta_{u,\omega}^t, \delta_{d,\omega}^t) \) as the vector of stochastic state parameters in minute \( t \) in scenario \( \omega \), and \( \psi_\omega = (\psi_{j+1,\omega}, \ldots, \psi_{j+T,\omega}) \in \Psi \) is a parameter vector from minute \( j + 1 \) to the problem end of horizon.

### 4.2.3 Decision Variables

#### First-stage Decision Variables

- \( e_{s,j} \): Minute-\( j \) power sales to the power system [kW]
- \( e_{b,j} \): Minute-\( j \) power purchases from the power system [kW]
- \( e_{c,j} \): Minute-\( j \) power charged into the BESS [kW]
- \( e_{d,j} \): Minute-\( j \) power discharged from the BESS [kW]
- \( e_{r,j} \): Minute-\( j \) power capacity for regulation [kW]
- \( e_{up,j} \): Minute-\( j \) served regulation-up power [kW]
- \( e_{down,j} \): Minute-\( j \) served regulation-down power [kW]
- \( s_{up,j} \): Minute-\( j \) unserved regulation-up power [kW]
- \( s_{down,j} \): Minute-\( j \) unserved regulation-down power [kW]
- \( x_{j+1} \): Beginning minute-\((j + 1)\) SOC of the BESS [kW-min]
- \( f_{t,\tau} \): Total kW provided in minute \( j \) to recharge flexible EVs that arrived to the charging station in minute \( \tau \), where \( \tau \in \{j - W + 1, \ldots, j\} \) [kW]
- \( f_j \): Total kW provided in minute \( j \) to recharge flexible EVs [kW]
- \( v_j \): Minute-\( j \) transformer overload [kW]

We define \( a_j = (e_{s,j}^j, e_{b,j}^j, e_{c,j}^j, e_{d,j}^j, e_{r,j}^j, e_{up,j}^j, e_{down,j}^j, s_{up,j}^j, s_{down,j}^j, x_{j+1}, f_j, f_j, v_j) \) as the first-stage decision vector of minute \( j \), the entries of which consists all first-stage decision variables above.
Second-stage Decision Variables

\( e_{s,t,\omega} \): Minute-\( j \) power sales to the power system in scenario \( \omega \) [kW]
\( e_{b,t,\omega} \): Minute-\( j \) power purchases from the power system in scenario \( \omega \) [kW]
\( e_{c,t,\omega} \): Minute-\( j \) power charged into the BESS in scenario \( \omega \) [kW]
\( e_{d,t,\omega} \): Minute-\( j \) power discharged from the BESS in scenario \( \omega \) [kW]
\( e_{r,t,\omega} \): Minute-\( j \) power capacity for regulation in scenario \( \omega \) [kW]
\( e_{up,t,\omega} \): Minute-\( j \) served regulation-up power in scenario \( \omega \) [kW]
\( e_{down,t,\omega} \): Minute-\( j \) served regulation-down power in scenario \( \omega \) [kW]
\( s_{up,t,\omega} \): Minute-\( j \) unserved regulation-up power in scenario \( \omega \) [kW]
\( s_{down,t,\omega} \): Minute-\( j \) unserved regulation-down power in scenario \( \omega \) [kW]
\( x_{t+1,\omega} \): Beginning minute-(\( j + 1 \)) SOC of the BESS in scenario \( \omega \) [kW-min]
\( f_{t,\tau,\omega} \): Total kW provided in minute \( j \) to recharge flexible EVs that arrived to the charging station in minute \( \tau \), where \( \tau \in \{ j - W + 1, \ldots, j \} \) in scenario \( \omega \) [kW]
\( f_{t,\omega} \): Total kW provided in minute \( j \) to recharge flexible EVs in scenario \( \omega \) [kW]
\( v_{t,\omega} \): Minute-\( j \) transformer overload in scenario \( \omega \) [kW]

The second stage decision variables define the optimal operation strategies if scenario \( \omega \) is realized. Note that scenario \( \omega \) is a sample path, which is defined by the state parameters from minute \( j + 1 \) to \( j + T - 1 \) (i.e. \( j + 1 \leq t \leq j + T - 1 \)). We define \( u_{t,\omega} = (e_{s,t,\omega}, e_{b,t,\omega}, e_{c,t,\omega}, e_{d,t,\omega}, e_{r,t,\omega}, e_{up,t,\omega}, e_{down,t,\omega}, s_{up,t,\omega}, s_{down,t,\omega}, x_{t+1,\omega}, f_{t,\tau,\omega}, f_{t,\omega}, v_{t,\omega}) \) as the decision vector of minute \( t \) and scenario \( \omega \) and define \( a_\omega \) as the second stage decision vector on sample path \( \omega \). We also define \( u_\omega = (u_{j+1,\omega}, \ldots, u_{j+T,\omega}) \) as the scenario-\( \omega \) second-stage decision-variable vector.

### 4.2.4 Constraints

For each time step, total energy supplied by the system must equal total energy consumption. For the current minute, this balance constraint is

\[
e_{b,j} + e_{d,j} + e_{down,j} + V_j = e_{s,j} + e_{c,j} + e_{up,j} + f_j + L_j,
\]

Note that this constraint ensures the power balance of all operations: EV charging, BESS charging and discharging, the PV output, and serving frequency regulation.
demand. The net power transaction, \( (e^b_j - e^s_j) \), is restricted by transformer capacity constraints:

\[
v_j + \bar{R} \geq e^b_j - e^s_j, \tag{4.2}
\]

and

\[
-v_j - \bar{R} \leq e^b_j - e^s_j. \tag{4.3}
\]

Note that the constraints allow overloading the transformer. The BESS state evolution is defined by the BESS state-transition constraint:

\[
x_{j+1} = \hat{x}_j + \mu c e^c_j - e^d_j / \mu_d. \tag{4.4}
\]

The BESS SOC and power charging and discharging is bounded by:

\[
\bar{S}_E, - \leq x_{j+1} \leq \bar{S}_E, +, \tag{4.5}
\]

and

\[
e^c_j, e^d_j \leq \bar{S}_P. \tag{4.6}
\]

The next set of constraints is related to EV recharging. The first set:

\[
f_{j, \tau} \leq \bar{H} \cdot N^\tau_j; \forall \tau = j - W + 1, \ldots, j, \tag{4.7}
\]

restricts the total amount that flexible EVs that arrived in minute \( \tau \) can be recharged based on the charger capacity and the total number of EVs. The total amount of EV recharging power in minute \( j \) is defined as:

\[
f_j = \sum_{\tau=j-W+1}^j f_{j, \tau}. \tag{4.8}
\]

The following set of constraints ensure that none of the EVs is over- or undercharged. The first:

\[
f_{j, \tau} \leq F_{\tau} - \bar{f}_\tau; \forall \tau = I(j), \ldots, j; \tag{4.9}
\]

127
restricts the amount of energy recharged in EVs that arrived in minute \( \tau \) to be no greater than the remaining unfulfilled charging demand. The second:

\[
F_{\tau} - \bar{f}_{\tau} - f_{j, \tau} \leq \bar{H} \cdot N_{\tau} \cdot (W - j + \tau - 1); \forall \tau = I(j), \ldots, j. \tag{4.10}
\]

ensures that their remaining unfulfilled charging demand at the end of minute \( j \) could be feasible met if they are recharged at the maximum capacity for the remainder of the charging window.

We have another set of constraints that define the mechanism of providing frequency regulations. We have a regulation-capacity state-transition constraint:

\[
e_r^j = e_r^{j-1}; \text{ if } j - \lfloor \frac{j}{60} \rfloor \neq 0, \tag{4.11}
\]

which ensures that the regulation capacity remains the same over an one-hour period. If minute \( j \) is the regulation-capacity bidding time step, this constraint is unnecessary. Otherwise, regulation capacity of minute \( j \) has already been decided in the previous time step. We allow that some of regulation energy demand is unfulfilled:

\[
s_{up}^j = \max\{\delta_u^r e_r^j - e_{up}^j, 0\}, \tag{4.12}
\]

and

\[
s_{down}^j = \max\{\delta_d^r e_r^j - e_{down}^j, 0\}. \tag{4.13}
\]

If the station is unable to provide the required amount of regulation demand, it only serves part of the demand. In this case, the station must to pay a penalty for the unfulfilled part. In PJM, unfulfilled regulation leads to cutting the station’s performance score, which reduces per-unit regulation capacity profit in the future.

The fulfilled regulation demands, \( e_{up}^j \) and \( e_{down}^j \), are upperbounded:

\[
e_{up}^j \leq \delta_u^r e_r^j; \tag{4.14}
\]
and:

\[ e_j^{\text{down}} \leq \delta_j e_j^r, \quad (4.15) \]

by regulation energy demand. We also have an upper bound the regulation capacity:

\[ e_j^r \leq \frac{\bar{S}_P + \bar{R}}{\delta_j}, \quad (4.16) \]

which ensures \( e_j^r \) is below the system regulating ability. Note that \( \delta_j \) is a constant related to the situation that the transformer is overloaded by providing regulation demand. If \( \delta_j \) equals 1, the station is not allowed overloading the transformer to provide regulation demands. If \( \delta_j \) is greater than one, the overload is allowed. In the case that \( \delta_j > 1 \), the station can get the advantage from the difference between the committed regulation capacity and the real served amount.

We finally have non-negativity constraints:

\[ e_j^s, e_j^b, e_j^c, e_j^d, e_j^{\text{up}}, e_j^{\text{down}}, f_j, v_j \geq 0; \quad (4.17) \]

and:

\[ f_{j,\tau} \geq 0; \forall \tau = j - W + 1, \ldots, j. \quad (4.18) \]

The second-stage constraints have a similar structure and put the same set of restrictions. Compared with the first-stage the second stage constraints are defined per scenario. For scenario \( \omega \), we define the constraints as follows. The power balance constraints are defined as:

\[ e_{t,\omega}^b + e_{t,\omega}^d + e_{t,\omega}^{\text{down}} + V_{t,\omega} = e_{t,\omega}^s + e_{t,\omega}^c + e_{t,\omega}^{\text{up}} + f_{t,\omega} + L_{t,\omega}, \forall t = j + 1, \ldots, T(j). \quad (4.19) \]

The transformer capacity constraints are given as:

\[ v_{t,\omega} + \bar{R} \geq e_{t,\omega}^b - e_{t,\omega}^s; \forall t = j + 1, \ldots, j + T, \quad (4.20) \]
and

\[-v_{t,\omega} - \bar{R} \leq e^b_{t,\omega} - e^a_{t,\omega}; \forall t = j + 1, \ldots, j + T. \]  \hspace{1cm} (4.21)

BESS state-transition follows:

\[x_{j+2,\omega} = x_{j+1} + \mu e^e_{j+1,\omega} - e^d_{j+1,\omega}/\mu_d \]  \hspace{1cm} (4.22)

and

\[x_{t+1,\omega} = x_{t,\omega} + \mu e^e_{t,\omega} - e^d_{t,\omega}/\mu_d; \forall t = j + 1, \ldots, j + T. \]  \hspace{1cm} (4.23)

The BESS SoC is bounded by

\[S^E_{E,\omega} \leq x_{t,\omega} \leq S^E_{E,\omega}^{+}; \forall t = j + 2, \ldots, j + T + 1, \]  \hspace{1cm} (4.24)

and BESS charging and discharging have limits:

\[e^e_{t,\omega}, e^d_{t,\omega} \leq \bar{S}^P; \forall t = j + 1, \ldots, j + T. \]  \hspace{1cm} (4.25)

The following set of constraints is related with flexible EVs charging. The charger’s capacity constraints are defined as:

\[f_{t,\tau,\omega} \leq \bar{H} \cdot N^{r}_{t,\omega}; \forall t = j + 1, \ldots, j + T, \forall \tau = I(t), \ldots, t. \]  \hspace{1cm} (4.26)

EV recharging power in each time unit is:

\[f_{t,\omega} = \sum_{\tau=t-W+1}^{t} f_{t,\tau,\omega}; \forall t = j + 1, \ldots, j + T. \]  \hspace{1cm} (4.27)

We have additional sets of constraints to ensure that flexible EVs are recharged. First, we ensure that all vehicles that have already arrived as of minute \( j \) will be fully recharged within the charging window under all scenarios:

\[\sum_{t=j+1}^{\tau+W} f_{t,\tau,\omega} + \tilde{f}_{j,\tau} = F_{\tau} - \bar{f}_{\tau}; \forall \tau = j - W + 1, \ldots, j. \]  \hspace{1cm} (4.28)
Next, we ensure that all flexible EVs that arrive after minute $j$ and that have a charging window that ends before the $T$-minute model horizon are fully recharged:

$$
\sum_{t=\tau}^{\tau+W} f_{t,\tau,\omega} = F_\tau; \forall \tau = j+1, \ldots, j+T-W.
$$

(4.29)

We finally ensure that the remaining EVs are not overcharged but have been charged sufficiently to be able to recharge before their charging window expire. These constraint are written as:

$$
\sum_{t=\tau}^{j+T} f_{t,\tau,\omega} \leq F_\tau; \forall \tau = j + T - W + q, \ldots, j + T; 
$$

(4.30)

and

$$
F_\tau - \sum_{t=\tau}^{j+T} f_{t,\tau,\omega} \leq H \cdot N_{\tau,\omega} \cdot (W - j - T + \tau); \forall \tau = j + T - W + 1, \ldots, j + T. 
$$

(4.31)

We also have a set of constraints to meet the rules of providing frequency regulation. First, we have constraints:

$$
e^{r}_{t,\omega} = e^{r}_{t,\omega}; \forall t = j + 2, \ldots, T_1 - 1, T_1 + 1, \ldots, T + j, 
$$

(4.32)

which define the regulation capacity between two bids. $T_1 \in \{ t \mid t \neq \lfloor \frac{t}{60} \rfloor \}$ is the frequency regulation capacity bidding time unit. We also have the constraints:

$$
s^{up}_{t,\omega} = \max \{ \delta^{u}_{t,\omega} e^{r}_{t,\omega} - e^{up}_{t,\omega}, 0 \}, 
$$

(4.33)

and

$$
s^{down}_{t,\omega} = \max \{ \delta^{d}_{t,\omega} e^{r}_{t,\omega} - e^{down}_{t,\omega}, 0 \}. 
$$

(4.34)

which define the unfulfilled regulation demand. The fulfilled regulation demands, $e^{up}_{t,\omega}$ and $e^{down}_{t,\omega}$, are bounded above:

$$
e^{up}_{t,\omega} \leq \delta^{u}_{t,\omega} e^{r}_{t,\omega}; \forall t = j + 1, \ldots, j + T; 
$$

(4.35)
and:
\[ e_{t,\omega}^{\text{down}} \leq \delta_{t,\omega}^{d} e_{t,\omega}^{r}; \forall t = j + 1, \ldots, j + T. \] (4.36)

The regulation capacity is also bounded above:
\[ e_{t,\omega}^{r} \leq \frac{\bar{S}^{p} + \bar{R}}{\delta_{j}}; \forall t = j + 1, \ldots, j + T. \] (4.37)

We finally have non-negativity constraints:
\[ e_{t,\omega}^{s}, e_{t,\omega}^{b}, e_{t,\omega}^{c}, e_{t,\omega}^{d}, e_{t,\omega}^{up}, e_{t,\omega}^{down}, f_{t,\omega}, v_{t,\omega} \geq 0; \forall t = j + 1, \ldots, j + T; \] (4.38)

and:
\[ f_{t,\omega} \geq 0; \forall t = j + 1, \ldots, j + T; \tau = t - W + 1, \ldots, t. \] (4.39)

### 4.2.5 Objective Functions

We define charging-station operation cost in scenario \( \omega \) s:
\[
c_{t,\omega}(u_{t,\omega}, \psi_{\omega}) = \sum_{t=j+1}^{j+T} \left\{ p_{t,\omega}(e_{t,\omega}^{b} - e_{t,\omega}^{s}) + R^{C}(v_{t,\omega}) + p_{t,\omega}^{C} e_{t,\omega}^{r} + p_{t,\omega}(1 + \phi)(s_{t,\omega}^{up} + s_{t,\omega}^{down}) \right\},
\]
where \( R^{C}(v_{t,\omega}) \) is a piecewise linear convex function and the last term is the penalty cost on unfulfilled regulation energy. The minute-\( j \) cost function is:
\[
c_{j}(a_{j}, \xi_{j}) = p_{j}(e_{j}^{b} - e_{j}^{s}) + R^{C}(v_{j}) + p_{j}^{C} e_{j}^{r} + p_{j}(1 + \phi)(s_{j}^{up} + s_{j}^{down}).
\]

We also define the first-stage feasible region:
\[
A_{j} = \{a_{j} \mid (4.1) - (4.18)\};
\]
and the second-stage scenario-\( \omega \) feasible region:
\[
U(\psi_{\omega}) = \{u_{\omega} \mid (4.19) - (4.39)\}.
\]
We define the station operation problem as:

$$
\min_{a_j \in A_j} \quad c_t(a_j, \xi_j) + \mathbb{E}[f(a_j, \psi_\omega)],
$$

(4.40)

where the first term is the minute-\textit{j} station operation cost using operation strategy, \(a_j\). The second term is the expected operation cost of the future time periods in the optimization horizon:

$$
f(a_j, \psi_\omega) = \min_{u_\omega \in U(\psi_\omega)} \sum_{t=j+1}^{T(j)} c_t(u_{t,\omega}, \psi_{t,\omega}).
$$

(4.41)

### 4.3 Solution Method

This problem has similar properties with the model in Chapter 3. The model has relatively complete recourse. We also assume the tail of each second-stage parameter’s distribution is exponentially bounded. The solution methods that is introduced in section 3.3 can also be applied to solve this problem.

### 4.4 Case Study

The station-control model with frequency regulation service introduced in section 4.2 is demonstrated using a case study. This case study is also based on the same study conducted in Chapter 3. The charging station locates in a parking lot in northeastern Columbus, Ohio. This location is one of the optimal DC fast charging stations found by solving the SFCLM in Chapter 2. The electricity-distribution system of the shopping center is connected to the power system through an \(\bar{R} = 500\) kVA transformer, which is shared by the buildings and the parking lot. Station operations are simulated over a representative year, using electricity market information (\textit{e.g.},

133
electricity price, frequency regulation service price, and regulation signal), weather, and vehicle-usage data for the year 2013.

We fix a number of parameters (e.g., electricity price, frequency regulation service price, regulation signal, PV output, and transformer characteristics) and examine the effects of the BESS, symmetry of the regulation signal, and whether regulation service is provided on the performance and revenue of the charging station adopting the proposed control model.

Electric vehicle-usage and EV interarrival process simulation is conducted based on tour-record data provided by MORPC. EV travel data are introduced in Section 3.5 and EV arrival process simulation is detailed in Section 3.4.1. Transformer characteristics and the overload-penalty function, $R^C(\cdot)$, are introduced in Section 3.4.6. The overload-penalty function, $R^C(\cdot)$, is calculated based on transformer-aging simulations on a 500 kVA transformer with a 20-year design life. The model uses a one-minute temporal granularity in scheduling EV charging, making BESS-charging and -discharging decisions, and optimizing responses to regulation signals. The model has $T = 60$ minute optimization horizon as proposed in Section 3.4.7. The other details of the datasets used and how these data are processed in constructing our case study are introduced in Sections 4.4.1 to 4.4.5.

### 4.4.1 Non-Electric Vehicle Load

We model the non-EV load using data from the year 2013 provided by American Electric Power Ohio (AEP Ohio), which is the distribution utility company in Central Ohio. The data is consist of anonymized load profiles on a number of 500 kVA commercial transformers, recorded at a 15-minute time resolution. Each transformer
supports one specific building. Each building has the same daily pattern of energy consumption. It is because one building has the same electric applicants and users whose behaviors of using applicants can be forecasted. We fit an ARIMA $(2, 1, 0) \times (0, 1, 1)_{96}$ model to capture the pattern of building loads as shown in Figure 4.1. This figure compares the realized non-EV load with ARIMA-forecasting load. This ARIMA model is able to capture non-EV load daily patterns. The model is also used to generate each 15-minute non-EV load for each scenario.

![Figure 4.1: Comparison of Realized Non-EV Load and ARIMA Forecasting Non-EV Load](image)

Figure 4.1: Comparison of Realized Non-EV Load and ARIMA Forecasting Non-EV Load
4.4.2 Frequency Regulation Signal

The regulation signal is modeled in the terms of regulation ratio, which is defined as the fraction of the absolute regulation energy to the reserved regulation power capacity. The pattern of the market regulation power capacity can be forecasted. We use hour-averaged frequency regulation capacity history data from PJM to fit an ARIMA $(1, 1, 1) \times (0, 1, 1)_24$. Figure 4.2 compares the real regulation capacity to the ARIMA forecasting. The absolute regulation energy is random and unpredictable. Hence, we fit 24 Gaussian distributions for each hour in a day and randomly generate 1-minute regulation energy using truncated Gaussian distributions. The regulation energy are bounded by the regulation capacity and zero.

4.4.3 Photovoltaic Generation

The real-time PV generation profile is created by the PVWatts model, with historical weather data from the year 2013 for the city of Columbus as inputs. PVWatts uses weather data to calculate a hypothetical PV generator given installation characteristics. We run PVWatts with the same data as introduced in Section 3.4. This model is used to generate forecasts of PV availability for the scenarios used in the station-control model and get a one-minute temporal resolution generation.

The output of PV generator is determined by three types of solar irradiation: direct, diffusion and global solar irradiations. The three types of irradiations are significantly related to longitude, latitude, elevation and weather. On a one-hour scale, the noise of PV generation is averaged over that hour. On a finer scale (i.e., 1 minute), PV generation is able to forecast.
Therefore, we fit an ARIMA \((5, 2, 0) \times (2, 1, 0)_{24}\) to forecast one-hour average PV generation. We compare the real-time PV generation and the ARIMA forecasting output as shown in Figure 4.3. The output of each minute is mainly decided by solar zenith angle and the weather. We also find that, by adjusting one-minute output by hourly average PV generations, the noise follows a Laplace distribution with mean zero. Therefore, we use a Laplace distribution to generate the PV output deviation from one-hour mean.
4.4.4 Electricity and Regulation Capacity Price

We fit 24 Lognormal distributions to model electricity prices in each hour of one day. The same approach is also used to model regulation-capacity prices. In a given problem, we first identify which hour in a day the current problem belongs to. Then, we use the corresponding Lognormal distribution to randomly generate the electricity price profile and regulation capacity price profile with one-hour resolution. The hourly electricity prices and regulation capacity market clearing prices are obtained from
PJM Interconnection market for the AEP zone, which covers the AEP Ohio service territory. Note that the negative prices are not allowed in our model.

4.4.5 Battery Energy Storage System Characteristics

We assume that the BESS has upper and lower bounds on its SOC of \( S_{E,+} = 4200 \text{ kW-min}(70 \text{ kWh}) \) and \( S_{E,-} = 1200 \text{ kW-min}(20 \text{ kWh}) \), respectively, and a power capacity of \( S_P = 100 \text{ kW} \). We also assume charging and discharging efficiencies of \( \mu^c = \mu^d = 0.99 \). These values are typical for a lithium-ion BESS Xi and Sioshansi (2016).

4.4.6 Experiment Cases

The performance of the operational strategies from this charging station-control model is affected by a number of factors. In this case study, we examine the effects of \( H \) (the EV charger power capacity), symmetry of regulation signal (i.e., the probability of receiving regulation-up signal (RRUP)), whether BESS is installed, and whether frequency regulation service is provided. Table 4.6 lists the six cases that we examined. Note that we assume that the EV penetration level (as a percentage of the light-duty vehicle fleet in Central Ohio) is 8%; \( W \) (the charging window for flexible EVs) is 40 minutes; and the station is installed with a 200 kW PV generator.

Cases 2 through 6 are obtained by modifying one factor in Case 1. We can estimate the effects of the modified factor on operations by comparing each case to Case 1.

4.5 Case Study Results

We use the case study described in Section 4.4 to analyze the operational strategies of the control model. First, in Section 4.5.1, we examine the operation strategies in
### Table 4.6: Test Cases

<table>
<thead>
<tr>
<th>Case</th>
<th>( \bar{H} ) [kW]</th>
<th>BESS</th>
<th>RRUP [%]</th>
<th>Regulation Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>BESS</td>
<td>50%</td>
<td>service</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>no BESS</td>
<td>50%</td>
<td>service</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>BESS</td>
<td>55%</td>
<td>service</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>BESS</td>
<td>45%</td>
<td>service</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>BESS</td>
<td>50%</td>
<td>no service</td>
</tr>
<tr>
<td>6</td>
<td>200</td>
<td>BESS</td>
<td>50%</td>
<td>service</td>
</tr>
</tbody>
</table>

Case 1, in which BESS is installed, the station provides regulation service; the EV charger capacity, \( \bar{H} \), is 100 kW; and RRUP is 50%. Second, in Section 4.5.2, we analyze the effects of the EV-charger capacity, the value of RRUP, whether BESS is installed, and whether provides regulation service. We use the experimental cases listed in Table 4.6 to conduct this analysis.

#### 4.5.1 Station-Control Model Results

We first examine the strategies of providing frequency regulation services and fulfillment of the committed services. This charging station is installed with a 200 kW distributed PV generator and BESS system. We assume the EV peentration level in Central Ohio is 8% and the charging window, \( W \), is 40 minutes. We further make an assumption that 50% of frequency-regulation signals require regulation up (\( i.e., \) to generate more or consume less electricity). The others are assumed to require regulation down. Figure 4.4 shows the frequency regulation capacity and fulfillment profiles and BESS SOC profile over 8 hours of a typical day. Comparing Figures 4.4a and 4.4b, shows that the amount of regulation capacity is highly related with the difference between regulation capacity and electricity prices: higher regulation capacity
is provided if the difference is larger. In hour 19-20, the regulation capacity price is $90 per MW higher than the electricity price. During this period, the control model offers 375 kW regulation capacity. This large capacity offer results in the station not fully fulfill the regulation signals in some time steps and a penalty being charged.

Figure 4.4: Frequency Regulation Capacity and Fulfillment and BESS State of Charge Profiles
Because we assume that the energy market only allow the symmetric frequency regulation capacity (i.e., capacity of regulation-up equals regulation-down) and RRUP is 50%, the load on the transformer in some time steps is negative, meaning that the station is providing power to the grid in net. In Case 1, positive load is on the transformer in 50.09% of a day using the strategies from the control model; the transformer overload time is 66-minutes; and the maximum load is 705.6 kW. Figure 4.5 illustrates the transformer positive load duration. It shows that the transformer works at its capacity for 18% of its positive load period. The transformer always operates at its capacity in negative load period.

![Transformer Positive Load Duration of Case 1](image)

Table 4.7 shows the frequency regulation quality provided by the station using the station-control strategies, which is closely related with whether the regulation service is accepted by the energy market planner. For example, PJM calculates each frequency regulation provider’s performance score. It only accepts regulation providers whose historical or test performance score is better than a given standard.
The performance score is closely related to the response accuracy. Table 4.7 shows that the regulation signal can be fully fulfilled in 95.5% time; and on average 85.23% of regulation demands can be fulfilled during periods that demand is not fully fulfilled. This shows that the control strategies satisfy accuracy requirements on frequency regulation.

Table 4.7: Control Model Frequency Regulation Performance in Case 1

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Regulation-up</th>
<th>Regulation-down</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfulfilled Time Daily [min]</td>
<td>64.25</td>
<td>64.25</td>
<td>0</td>
</tr>
<tr>
<td>Unfulfilled Amount Daily [kWh]</td>
<td>65.81</td>
<td>65.81</td>
<td>0</td>
</tr>
<tr>
<td>Fully Fulfilled Time Percentage</td>
<td>95.5 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Fulfilled Percentage</td>
<td>85.23 %</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.5.2 Sensitivity Analysis

We compare each of Cases 2 through 5 with Case 1 and analyze the effects of the 5 parameters in the station-control model. Table 4.8 compares the average daily operation cost, transformer overload time and positive load time of the six cases. Table 4.9 shows the frequency regulation service quality of these cases. By comparing the daily average positive load time and RRUPs in each case, we find that the time of transformer with positive load is close correlated to RRUPs. This means that the transformer has outflow of power with a high probability if the signal is regulation-up and vice versa. Moreover, the regulating-down demands, as shown in Table 4.9, can always be fully fulfilled by the charging station in this case study. This is because the charging station is an electricity consumer, which requires power to support inflexible
### Table 4.8: Daily Operation Cost, Transformer Overload Time and Positive Load Time Sensitivity Analysis

<table>
<thead>
<tr>
<th>Case</th>
<th>Positive Load Time per Day [hr]</th>
<th>Overload Time per Day [hr]</th>
<th>Operation Cost per Day [$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.02</td>
<td>0.433</td>
<td>473.89</td>
</tr>
<tr>
<td>2</td>
<td>12.04</td>
<td>1.283</td>
<td>670.07</td>
</tr>
<tr>
<td>3</td>
<td>13.08</td>
<td>0.2</td>
<td>334.74</td>
</tr>
<tr>
<td>4</td>
<td>10.81</td>
<td>0.3</td>
<td>277.96</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>0.417</td>
<td>166.33</td>
</tr>
<tr>
<td>6</td>
<td>11.84</td>
<td>0.417</td>
<td>1364.87</td>
</tr>
</tbody>
</table>

Hence, this type of charging station tend to fulfill regulation-down more accurately (i.e., using more power). This observation is also shown by comparing Cases 3 and 4 in which RRUPs are 55% and 45%, respectively. Cases with higher probility of

### Table 4.9: Quality of Frequency Regulation Service

<table>
<thead>
<tr>
<th>Case</th>
<th>Unfulfilled Time per Day [min]</th>
<th>Unfulfilled Amount per Day [kWh]</th>
<th>Fully Fulfilled Time [%]</th>
<th>Fulfilled Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64.25</td>
<td>65.81</td>
<td>95.54 %</td>
<td>99.34 %</td>
</tr>
<tr>
<td>2</td>
<td>70.5</td>
<td>54.67</td>
<td>95.10 %</td>
<td>99.22 %</td>
</tr>
<tr>
<td>3</td>
<td>59.97</td>
<td>59.67</td>
<td>95.85 %</td>
<td>99.39 %</td>
</tr>
<tr>
<td>4</td>
<td>59.57</td>
<td>66.58</td>
<td>95.37 %</td>
<td>99.41 %</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>6</td>
<td>52.33</td>
<td>43.27</td>
<td>96.37 %</td>
<td>99.58 %</td>
</tr>
</tbody>
</table>
regulation down have smaller station operation costs, as well as better frequency fulfillment accuracy.

Moreover, the BESS system, by comparing Cases 1 and 2, is able to significantly reduce the transformer overload time and the daily operation cost and improve the service quality of frequency regulation. Even though charging station operations with regulation service is not good as the station without this service comparing Cases 1 and 5, the regulation service is also beneficial to the station operator if the regulation signal is asymmetric and has more regulation-down demands. In addition, the daily operation costs shown in Table 4.8 are calculated by our control model, which puts a high penalty (i.e., $p_t(1 + \phi)$) on not fully fulfilled regulation requirement to ensure high-quality regulation service.

Last, we can conclude from Case 6 that 200 kW EV charger capacity improves the regulation response accuracy but increases the peak load by 12% from 705 kW to 787 kW, which also increases operation cost in which transformer overload is piecewise convex penalized.

4.6 Conclusion

In this chapter, we propose a stochastic station control method for DC fast charging station installed energy storage and distributed renewable generator which allow the charging station to participate in the ancillary service market to provide frequency regulation services. This method uses a Monte Carlo-based sample-average approximation to sequentially solve and assess a series of sample stochastic optimization problems and their solution qualities until a high quality near-optimal solution is formed.
We apply this method to a potential DC fast charging station candidate location in Central Ohio and examine the load features of the control strategies from the proposed model. We exam the station operation in six different configurations and simulate each of them for one year. We demonstrate that the proposed method is able to fully respond to frequency regulation more than 95% of the time. The average response accuracy is greater than 99%, showing that the proposed stochastic station control method can provide high-quality regulation services to the ancillary service market.
Chapter 5: Conclusion and Future Works

The market share of electric vehicle keeps growing, while public charging demands are increasing. This rising demand is challenging social planners in how to wisely invest in infrastructure to serve EVs. At the same time, EV charging loads are growing to be a nonnegligible effect on power system and distribution system operation.

In this dissertation, a stochastic flow capturing station location model with stochastic EV flows is first introduced, which is modeled as a flow-capturing based refueling location model. This model aims to help transportation planner wisely setup public DC fast charging station network in an environment in which EV adoption and distribution rapidly growing and changing. This model addresses a problem about which locations are the best to refuel the maximum amount of potential EVs with long-distance commutes *en route* and how many chargers each station should be installed with. An EV-flow simulation model is introduced to simulate electric vehicle flows in a transportation network based on light-duty vehicle tour-record data. This simulation model. It is also used in charging station control models.

Second, a stochastic charging station control model is introduced. It creates station-control strategies for a DC fast-charging station. This model is able to get potentials of the existing distribution system and relieve the existing distribution constraints by mitigating flexible EV charging load. It is also able to provide frequency...
regulation strategies for the station participating in the ancillary service market. This station control model aims to provide station operators an optimization scheme that optimally manages EV refueling, battery energy storage charging and discharging, and frequency regulation operations. This model can protect the distribution devices from overloading without upgrading the distribution system.

The proposed recharging station location model and the station control method still have some limits which may be solved in the future. The proposed flow-capturing station location model is based on bootstrapping EVs from light duty vehicles which is acceptable but is not accurate enough. If the EV penetration level is able to forecast based on demographic and geographic data, EV-flow patterns will be more accurate; and the proposed model can create better charging station layout. In addition, the current EV flow simulation is based on a dataset describing alternative-vehicle drivers’ behaviors. As the amount of electric vehicles increases, it is possible to survey EV drivers’ refueling behaviors. This survey will contribute to a more sophisticated EV-flow simulation; and it will be a concrete basis of EV infrastructure planning-related research.

For the charging-station control models, we have created control strategies with one-minute time steps. In practice, photovoltaic generation are changing real-time. Frequency regulation signals are received every 2 seconds in PJM. In order to respond to these stochasticities in real-time, an optimal control function is necessary to respond to these continuous variations. It is an interesting question how to use the stochastic station control models to create optimal control functions for a given time span. It fits the real world practices. However, even though the easiest affine control functions is used, a station-control model will become a nontrivial nonlinear
and nonconvex optimization problem. This is a real challenge for creating realtime optimal control functions.
References


