12-2020

Micro Grid Control Optimization with Load and Solar Prediction

Shaju Saha
Utah State University

MICRO GRID CONTROL OPTIMIZATION WITH LOAD AND SOLAR PREDICTION

by

Shaju Saha

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

of

MASTER OF SCIENCE

in

Computer Science

Approved:

Nicholas Flann, Ph.D. Major Professor	Regan Zane, Ph.D. Committee Member
John Edwards, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Interim Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2020

ABSTRACT

Micro Grid Control Optimization with Load and Solar Prediction

by

Shaju Saha, Master of Science

Utah State University, 2020

Major Professor: Nicholas Flann, Ph.D.
Department: Computer Science

Technology advances in solar energy production and battery storage have made microgrids feasible in many situations. Given the large investment required, it is important to control the microgrid in an efficient manner to minimize operational costs or maximize profit. In this work we explore the application of model-based machine learning to predict both solar production and load demands to produce an optimal charge/discharge schedule for the microgrid battery. To calculate the costs accurately a model of battery cost is applied as a function of holding charge and rate of discharge/charge. Studies where performed to assess how the accuracy of prediction and the influence of time of day pricing effect profit/loss and charging schedule. Additionally, to aid during the design phase, multiple microgrid configurations of battery and solar capacity were evaluated using the optimal scheduler.This thesis also investigates using load shifting to support more customers when resources are limited. The results show that the operational cost of a microgrid configuration is significantly influenced by the time-of-day pricing and utility buy-back policies. For instance, given the same load and solar power profile, utility policies in California will produce twice the profit compared to policies in Texas.

(63 pages)

PUBLIC ABSTRACT

Micro Grid Control Optimization with Load and Solar Prediction

Shaju Saha

Using renewable energy can save money and keep the environment cleaner. Installing a solar PV system is a one-time cost but it can generate energy for a lifetime. Solar PV does not generate carbon emissions while producing power. This thesis evaluates the value of being able to make accurate predictions in the use of solar energy. It uses predicted solar power and load for a system and a battery to store the energy for future use and calculates the operating cost or profit in several designed conditions. Various factors like a different place, tuning the capacity of sources, changing buy/sell schedule are considered to verify the results. Combining real battery cost makes this work more reliable from the existing system. The prediction error also considered while testing the results.

To all the close people....

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

EV	Electric Vehicle
RL	Reinforcement Learning
VIA	Value Iteration Algorithm

CHAPTER 1

INTRODUCTION

Environmental concerns and the improving performance and reduced cost of electric vehicles (EV) are creating increasing EV sales. However, an increase in EVs may have detrimental effects on power system performance due to increased demand for electrical energy. International commitments to reduce carbon dioxide (CO_2) emissions- the most common and pervasive greenhouse gas- has fuelled efforts to decarbonize the traditional transport sector. Nowadays electric vehicles (EV) have become a viable alternative to the conventional fossil-fuelled vehicles. With the gradual increases in EV adoption, the power load profile in distribution networks is prone to significant change. There is a need to increase EV charging infrastructure help to increase clean energy adoption, reduce carbon emission, to alleviate peak charging loads, and increase the convenience for EV users.

1.1 Background

To maintain the growth in EV adoption, more charging stations are needed throughout the transportation network. Currently, most of the big cities and along most interstate freeways have charging stations to provide service for EVs but there are few EV charging stations in remote areas. The following figure 1.1 shows the monthly electric car sale amounts in the years 2017, 2018, and 2019.

Most of the existing EV charging stations are buying energy from the national grid and then selling to the EV user. This is a naive approach to support the rapidly growing community. To improve the quality of this process and increase the use of clean energy, some of the stations use solar power for charging. This has two benefits. Firstly, it reduces the carbon emission for growing electricity demand, and secondly if the grid goes down these stations can still provide the some service for EVs.

Overall, the purchase price of EVs is still higher than ICE (internal combustion engine)

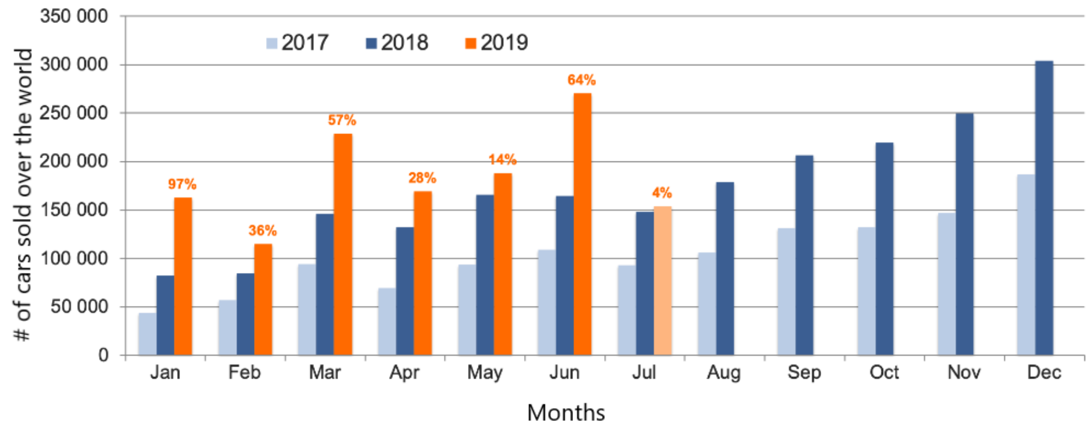

Fig. 1.1: Number of car sold in last three years

cars, but considering full costs of ownership EVs are becoming comparable or cheaper than ICE vehicles. The following figure 1.2 shows the average price of EVs and diesel cars through the period starting from July 2018 to June 2019 [1].The report illustrates that the past year's price dropped from \$64,300 to \$55,600, 13.4% specifically. But it is still high compared to the average price of other vehicles. [2]According to a study in 2018 by the University of Michigan's Transportation Research Institute, the operational cost of EV is less than half of the cost of gas-powered cars. This study mentioned the exact amount cost for EV is \$485 per year where the cost of the gas-powered car is \$1,117.

1.2 Related Work

Most existing literature has designed their charging scheduling model based on the assumption that future EV arrivals and electricity prices are known to charging stations when pricing and scheduling decisions are made.

In recent years, numerous day-ahead scheduling approaches have been proposed for this problem [3, 4]. For instance, in order to handle the uncertainty in electricity price, [3] developed a robust optimization approach for residential EV charging scheduling. Similarity, [5] proposed an information-gap-decision based approach to deal with the uncertainty in electricity price and optimize day-ahead scheduling of EV fleet. In [6, 7], EV fleet was formulated as a probabilistic virtual battery model, and scenario-based robust approaches

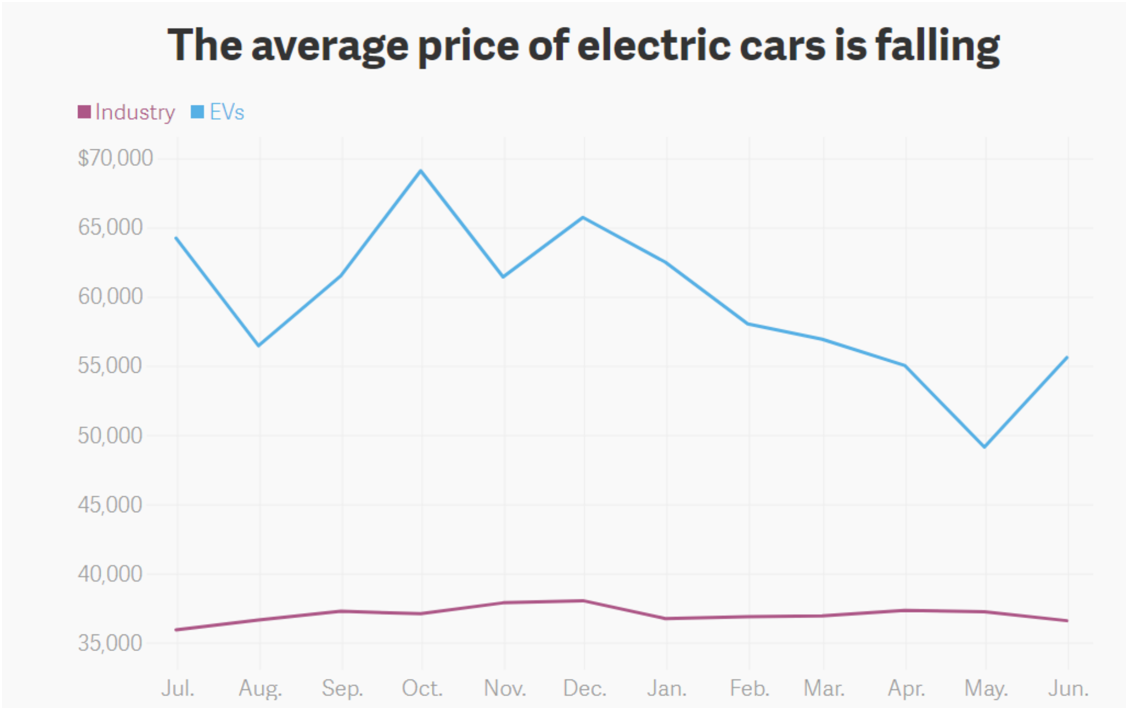

Fig. 1.2: Average price of EV and total

were proposed to deal with the uncertainty of the EV users' commuting behavior and the balancing requests.

[8] studied the day-ahead scheduling of battery swapping stations where the uncertainty of the battery demand and the electricity price was modeled by inventory robust optimization and multi-band robust optimization, respectively. Due to the existence of randomness in traffic conditions, users' commuting behavior, and pricing process of the utility, EV arrival and departure time, EV energy consumption, and electricity prices are dynamic and time-varying. Therefore, efficiently managing EV charging/discharging to reduce the cost becomes challenging. Real-time scheduling strategies that can respond to dynamic charging demand and time-varying electricity prices have attracted a lot of attention recently. For example, [9] developed a strategy to coordinate multiple EVs charging in a parking station in response to real-time curtailment requests. [10] offered a formulation for the coordinated charging problem which considered the plug-in and plug-off frequency.

Recently, model-free approaches that do not need any system model information has

achieved great success in complex decision-making application [11]. This success has inspired the development of model-free approaches for smart grid applications [12,13]. Compared to the model-based approach, the advantage of the model-free approach is that it can learn a good control policy based on reinforcement learning (RL) and does not rely on any knowledge of the system [13]. Neural networks are universal approximators [14] and have been widely used for RL [15,16]. In recent years, deep neural networks have achieved promising results in learning a complex mapping from high-dimensional data. By utilizing deep neural networks, deep RL has obtained significant success in many complex decision-making applications. For instance, a deep Q-network has achieved a level comparable to that of a professional human in the Atari 2600 [11] in 2015. Long Short-Term Memory (LSTM) are a type of recurrent neural networks which are useful when dealing with sequential data because of feedback connections. It has already shown promising results in predicting sequential information because of its capability to exploit long term dependencies among the different sequences. However, to the best of our knowledge, the application of LSTM in RL structure for an optimal policy in real time EV charging/discharging problems have not been reported in literature.

In this work, a hybrid approach is applied, where the EV charging/discharging scheduling problem is formulated as a value-iteration algorithm from the charging station's perspective. The objective is to find an optimal charging/discharging policy to take full advantage of the predicted real-time demand while fulfilling a user's driving demand. The schedule determination is computed using value iteration, a model-based approach, but with the demand and supply predicted using model-free learning approaches. The approach also uses the anticipated electricity prices in a specific time slot and the battery State Of Charge (SOC) in that time slot as inputs to compute real-time charging/discharging polices.

The rest of the thesis is organized as follows. First the problem investigated is clearly defined in Chapter Two. Then the methods used, both the model based and model free are described in Chapter Three. Chapter Four gives the results obtained. Chapter Five provides a summary of the work performed, conclusions drawn and considers the potential

for future work.

CHAPTER 2

Problem Definition

This thesis formulates the real-time EV charging/discharging scheduling problem from the charging station's perspective. At time slot t, this work observes the system state s_t which includes the information about the remaining charge in the charging station's battery, the anticipated 48-hour electricity prices and predictions of load and solar energy. Based on this information, the optimizer will choose the charging/discharging action a_t. This action represents the amount of energy that the station battery will charge or discharge during this time interval. After executing this action, the optimizer can observe the new system state s_{t+1} and choose the new charging/discharging action a_{t+1} for time step $t + 1$. Thus to summarize the problem can define as

"*Given a fixed battery capacity in a charging station, predicted time-series load and real-time energy price, find an optimal sequence of charging or discharging actions for the EV station that either sell electricity to the customer or to buy electricity from grid such that profit of charging station is maximized.* ".

Additionally, given the optimal scheduler, the problem of determining an optimal battery configuration is also determined by an enumeration of possible battery sizes, which can be deployed in a charging station in order to observe the optimal profit for a specific charging station.

2.1 System Design

The system design for this research is combined with some subsections. Which is equally important to validate our learning idea. For the general system, everyone used one source of energy supply for this problem specifically the system used solar power and electricity from the national grid as a supply. The system considers a spot for the load side for this well-structured problem it is the total load for specific EV charging stations. To

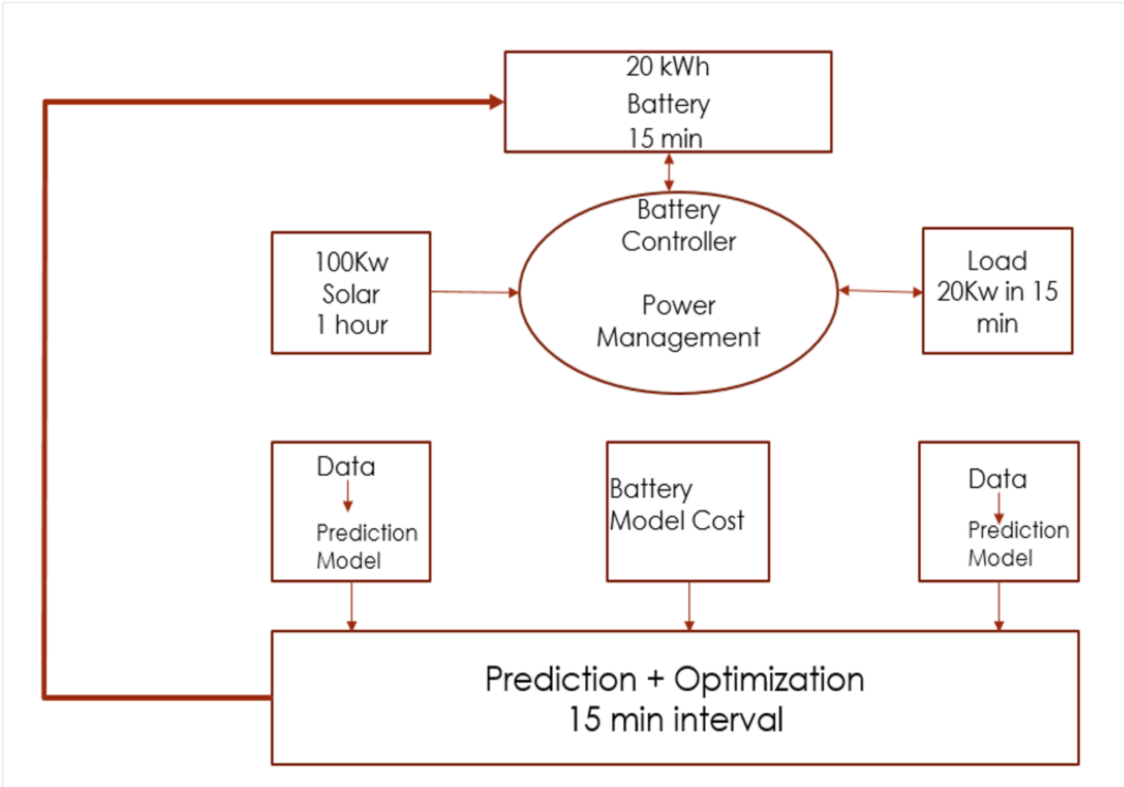

Fig. 2.1: System model for EV charging station.

make a balance between supply and demand this design introduces a storage section in the entire system as a battery.

The following figure 2.1 shows the configuration of the actual system though for experiment purpose this research tune some of the system equipment. For calculating the actual cost of operation, this work introduces a battery model that provides the actual cost of operating the battery based on the age of the battery, how much energy holds into the battery, and how speedily the battery is charged. Finally, all this information added to a reinforcement learning agent to select the best action sequences marked as an optimizer in figure 2.1.

2.1.1 Solar Prediction

To ensure better performance the system needs to introduce solar power as clean energy. For analyzing the result this work varying the prediction hours from 6 to 48 hours. As it

is hard to predict for a long period this work tries historical data analysis and forecasted numerical weather data as a feature vector in different Machine learning algorithms. Picking a suitable solar panel capacity which makes it more useful for an existing station this work tries different solar power capacity. Mostly this research problem focused on 20 to 100 kWh solar capacity. Different time steps also tried to maximize the accuracy of the forecast starting from 15 minutes to 6 hours.

2.1.2 Load Prediction

In this research, the method defined load as the demand for an EV charging station in a time period. The unit used for load calculation was KWh. Historical data analysis is an effective approach found by analyzing prior work to predict the load. For this work collected past data from Salt Lake City EV charging stations. For the prediction, the system machine learning algorithms described in Section**??**. To make it integrate load prediction for this project, the prediction period was varied from 6 to 48 hours. For generalization with solar, the prediction period was changed to 6 hours.

2.1.3 Optimization

The most significant portion of this study is the optimization of operational costs to maximize possible profit. The optimizer accounts for solar power, load, and a grid pricing policy as input and produces future action sequences as output. The optimizer was evaluated on different experiment situations to demonstrate correct functionality. Profit or loss was calculated based on the amount of electricity this system needs to buy from the grid or sell to the grid to service its loads.

2.2 Design the Research Experiment

This research designed experiments to validate the correctness and and then to understand the optimal schedules produced. In the following subcategories, we briefly mention all the research experiments.

2.2.1 Calibrate the source, load, and battery size

One benefit of the optimization approach is to assist in the design of a microgrid configuration based load and solar data. Results are presented that assist in the identification of optimal battery and solar sizes based on predictions of load and solar energy, while taking into account time-of-day pricing.

2.2.2 Use different day-of-pricing

Optimizer was designed to operate independent of a specific area or nation. Every nation has a different pricing policy, and in the united states, every utility has its different pricing policy. The kind of pricing policy significantly influences the resultant profit and the charge/discharge action sequence. For example,in Utah, the power supply company divides the day into three different pricing policies called peak hour, mid-peak hour and off-peak hour. The peak hour electricity price is higher than the other two and the mid-peak hour price is higher than the off-peak. Southern California also has the same three pricing policy but the hours of the for peak and off-peak are different so the best-picked action order also should be changed.

2.2.3 Battery modeling

For accurate calculation of operating costs, this work considered a realistic model of battery cost incurred during operation. It calculated two types of battery costs to make the cost more realistic. First, how fast system charges or discharges the battery. From past research, it has been shown that fast charge or discharge of the battery causes a decrease the total lifetime. Research shows that holding the electricity in a battery is also a cause of damage. According to the expert to operate the battery between 20 to 80% is best. Greater than or less than the limit can hamper the battery health.

2.2.4 Complex buy/sell arrangement

The research also considered the microgrid system operating in a different situations such as when no external power source is available or has failed. The work examined four

different conditions to operate the system. The first one is the normal scenario the system can buy and sell on the grid. The second one is it can buy from the grid but can not sell it to the grid. Third, is it cannot buy from the grid but can sell electricity to the grid. Lastly, if it was in a remote location that it didn't have any connection with the grid.

2.2.5 Actual versus predicted profit

This work used solar prediction and load prediction as an input of this system. Prediction always has a certain amount of error which is called prediction error. In this research design, this work evaluated the impact on action order if the system knows specifically what will happen future. For this analysis, it initially determined the action and total profit with 48 hours predicted solar and load. After that, it used the data where the first six hours of solar and load are real and the next 42 hours were are predicted. Continuing this process until all the data was known, not predicted. In this way, the system was evaluated to determine the impact of prediction accuracy has on the prediction of operating profit.

2.2.6 Tracking versus Non-tracking

The research in the field of solar panels is growing every day and new technologies are being developed to generate more solar power. Tracking the sun is one of the new technologies that can produce more energy compared to fixed panels. In tracking, amount of solar power collection increases in the morning and the evening. From the data analysis, this research found this is the time when the load is also high. An empirical study was conducted to evaluate the potential benefits of tracking vs. fixed panel mounts under various scenarios.

2.2.7 Load Shifting

According to the load pattern, there maybe a high demand for power for electric vehicles outside of sunny hours of the day. If there is no grid available, this results in charging stations not being able to provide power before sunrise or after sunset. On the other hand, during the middle of the day with an abundance of solar power, there maybe a loss of power

due to a lack of efficient storage. Using large-capacity batteries to store said excess power could result in a waste of the battery capacity during the rest of the day when storage is not required. Therefore to address this problem, this thesis looks at ways to shift the load demands to solar power abundant parts of the day.

CHAPTER 3

Method

The method used to optimize the EV station is a value iteration algorithm (VIA), [17] which is an example of dynamic programming. The VIA will calculate the optimal reward for each state the charging station can enter and then choose an action that will advance the state to the next, most beneficial state. The VIA will do this by using the energy generated by the solar panels, the current load on the station, and the current state of the battery, for all possible states.

At the end of a given time period, the battery percentage can be in any of its predefined states, for example, a battery with 20KWh capacity could be broken into ten, 2,000 watt-hour bins. Using the final state of the battery, the VIA calculates the value of the energy stored in the battery at that time. Then the VIA calculates the value of being in the state previous until each time step for the period is filled out for every possible state.

By calculating an intermediate reward based on the purchase or sale of electricity plus the wear on the battery, the learning algorithm will be able to calculate the ideal times to purchase electricity from the main power grid. It will also be able to decide when to sell the stored energy back to the main power grid, and when to use the energy stored in the battery to accommodate the load on the EV charging station.

3.1 Reinforcement learning

Reinforcement learning is a powerful and yet universal method that learns the best actions to take in all states of a domain. To identify the optimal operating policy, the applicant need only specify the set of actions possible in each state, the resulting new state, and the benefit or loss of the action, called the immediate reward.

Before considering the EV charging station, consider a generic domain of well-defined states and actions. The system will assume that the state is observable and that a prob-

abilistic model exists of the domain that returns the next state given an action. Under these simplified assumptions, the problem can be represented as a value function learning problem and solved using a form of dynamic programming.

Given S_t, the state at t, then $S_{t+1} = A(S_t, a)$, where a is the action and Do executes the action, producing the next state. For a given state S_t, the set of actions available are $A(S_t)$.

Then let define $V(S_t)$ as the total reward for the domain being in state S_t. Optimal operation will always take actions such that the states occupied maximize total reward. Then $V(S_t)$ may be defined as:

$$V(S_t) = \max_{a \in A(S_t)} R(S_t, a) + \gamma V(Do(S_t, a)) \tag{3.1}$$

Where $R(S_t, a)$ is the immediate reward of being in state S_t and taking action a, and $0 < \gamma \leq 1.0$ is the discount rate.

The optimal policy always selects the action that maximizes future total reward. Then the optimal policy is defined:

$$\Pi(S_t) = \arg \max_{a \in A(S_t)} R_t(S_t, a) + \gamma V(Do(S_t, a))$$

Reinforcement learning is a method that solves for V and Π, given the model. In this case, the model is a specification of a representation of S, and a definition of $A()$, $Do()$ and $R()$. Additionally, there may be some states for which there are no actions available and the total reward is known, these states are referred to as terminal with $t = T$.

V is complied into a look-up table, mapping the current state to a number. Consider this table of all possible states, each with a V value. The value iteration method first assigns values to terminal states, then repeatedly applies Equation 3.1 as an update rule until all the values in the table are at a near fixed point. The order of the updates could be made in a specific way, such as from the terminal states backwards, or asynchronously based on states that are experienced during operation.

3.1.1 Value-Iteration for RL

Environment: Physical world in which the agent operates.

Action (A): All the possible moves that the agent can take.

State (S): Current situation returned by the environment.

Reward (R): An immediate return send back from the environment to evaluate the last action.

Policy(π): Policy is the strategy that the agent employs to determine next action based on the current state. Thus this policy function returns an action given a current environment state.

$$\pi(S) : S \rightarrow A \tag{3.2}$$

State transition model ($p(s_{t+1}|s_t, a_t)$): State transition model defines how the agent enters into a new state s_{t+1} from it's current state s_t having taken an action a_t. Reward Model ($p(r_{t+1}|s_t, a_t)$: Reward model describes the real number (*termed as reward*) that the agent receives from the environment after performing an action and entering to the next state. Discounting Factor (γ): It controls the importance of future rewards Value Function (V_s^π): The value function represents how good is a state for an agent to be in. It is expressed as expected total discounted reward agent get when starting from a state s and reach to the terminal state after vising all immediate states following a fixed policy π. Thus, for a given policy π to select actions, the corresponding value function is given by

$$V_s^\pi = E[\sum_{i=1}^{T} \gamma^{i-1} r_i | S_t = s] \;\; \forall s \in S \tag{3.3}$$

Among all possible value-functions (*under different polices*), there exist an optimal value function (*optimal policy*) that has higher value than other functions for all states and thus, it is denoted by

$$V_s^* = max_\pi V_s^\pi \;\; \forall s \in S \tag{3.4}$$

The optimal policy π^* is the policy that corresponds to optimal value function. So,

$$\pi^* = argmax_\pi V_s^\pi \quad \forall s \in S \tag{3.5}$$

3.1.2 State

This algorithm's state has consisted of battery charge condition, hours in the day, load in that specific time, solar energy in that specific time.

Battery Charge Condition

The battery condition in the EV stations is divided into 100 slots, each slot corresponding to the remaining charge in percentage in the battery. For this problem, designed policy such that battery charges always remains at a certain limit to maximize the battery lifetime. The system maintains that battery charge of both $< 20\%$ and $> 80\%$ were injurious to the health of the battery.

Hours in the Day

As the system is maximizing the savings at the end of the day, that's why the day is described in 24 discreet spaces, each space is considered as one hour.

Demand in a specific time-slot

For every hour the system is predicting the excepted load for the charging station, Which is considered as a load in our proposed system.

Solar energy in that specific time

For every hour the research is predicting the excepted solar power in the charging station location, Which is considered as a source in our proposed system.

3.1.3 Action

action is consisted of how much charging or discharging happens in the battery.

Charging Quantity

During taking action RL agents can get charged whatever it needs.

Discharging Quantity

But during discharging RL agent discharges just according to the demand of the coming EVs.

3.1.4 Reward

In the RL method reward is an essential way to evaluate the best action sequences. In this designed model this work used two types of rewards to calculate the exact cost of the system.
Reward for buy/sell ($R_{buy/sell}$): Reward for taking action buy or sell.
Reward for battery modeling ($R_{battery}$): Total cost in the battery if the system took this action.

Reward for buy/sell

This part of the reward is only considered the taken buy and sell action. Originally it considered the price of the electricity and the amount of electricity sold or bought from the grid at that time step. The only exception was in the experiment where the system can not buy or sell from the grid. The following equations give the entire idea of this reward. If action represented buy amount system makes this amount negative.
$R_{buy/sell}$: Reward for buy/sell.
$ElectricityPrice(t)$: Price of electricity at time t.
$buy/sellamount$: The amount of electricity buys or sells at that timestep.

$$R_{buy/sell} = ElectricityPrice(t) * buy/sellamount \tag{3.6}$$

Special edition of this reward was for the situation can not buy or can not sell. *action*:

The action was taken at that timestep.

$$R_{buy/sell} = abs(action) \tag{3.7}$$

Reward for battery modeling

Calculating the battery cost was an important part to make the system more realistic. From prior research work, there are two factors involved in estimating battery cost. First a holding cost that is a function of the duration and amount of energy stored. Second, a charging cost that is a function of the power transfer, the number of cycles and the operating range.

Holding cost was added if the battery went more than 80% of total battery capacity. Prior work has shown that operating the battery from 20% to 80% state of charge was the best for the battery. For this model, system design strictly fixed that battery can not go less than 20% of its total capacity.

$R_{holding}$: Reward for holding in battery.

$BatteryPenalty$: Battery penalty rate for more than 80%.

$BatteryPercent(t)$: Battery level at this time step t.

$BatteryCapacity$: Total battery capacity.

$$R_{holding} = (BatteryPenalty * BatteryPercent - 80 * BatteryPenalty)/BatteryCapacity \tag{3.8}$$

The charging cost was calculated on the charging speed in that timestep. Higher rates of charge and discharge are the principle reason for decreased battery life in the long run. It is exponentially related to battery life.

$R_{charging}$: Reward for charging/discharging speed in battery.

$BatteryCost$: New battery price.

$BatteryEfficiency$: Battery efficiency when battery is new.

$ChargingSpeed$: How fast the battery is charged/discharged.

$BatteryLifePanalty$: The battery life loss if battery charge/discharge in 1C mode.

$$Price = (BatteryCost/(BatteryEfficiency * BatteryCapacity)) * BatteryLifePanalty \tag{3.9}$$

$$R_{charging} = (price/e(1)) * e(ChargingSpeed) \tag{3.10}$$

$$R_{battery} = R_{charging} + R_{holding} \tag{3.11}$$

The proposed system total reward was calculated from the above rewards. The system reward equation is mentioned in the following equation.

$$R_{total} = R_{buy/sell} - R_{battery} \tag{3.12}$$

CHAPTER 4

Results

4.1 Overview

In this research, the overall system integrates solar energy prediction, electrical load prediction, and day-of-pricing. Once data is collected and cleaned, the system is ready to run the designed experiments mentioned in chapter three. In these studies, additional design decisions are made such as what type of battery to use, and the time period and specific locations to study. Data preprocessing aligned the load and solar datasets with their corresponding timestamps.

4.2 Environment

All the experiment was run on the windows 10 enterprise machine with Intel Core i7-3770 with 3.40GHz and 16 GB RAM. For all the experiments Python 3.6.3 was used. Some Python packages were used to perform the experiments designed for this research. For the timestamp aligned DateTime is applied. For the reinforcement learning implementation, the environment used the gym pack. Determining the sun position for tracking studies used Astral. In addition, the program used some popular tools for calculation, visualization and data accusation which are NumPy, matplotlib, pandas, etc.

4.3 Experiment Results

To assist in interpretation of the results and to enable easy comparison among alternative experimental conditions, this research used a uniform format for the results. The result figures consist of five different graphs, aligned vertically on the same time scale, see Figure 4.1. The top graph among the five represents the input to the whole system. Input is the predicted solar power in orange and represented as a positive value. Demand is the

predicted load for that timestamps in gray color, shown as a negative value. Both represent in the bar chart with the KW unit on the y-axis since power is measured.

The second graph in Figure 4.1 represents the state of charge of the battery for each time step for two days in blue. The unit used here is KWh for the y-axis since energy is measured.

The third graph in Figure 4.1 describes the policy determined by the optimization algorithm as a charge/discharge sequences from the battery for the maximum profit. Maroon means charge the battery at a power value and Sea green denotes the discharge power. The fourth graph in Figure 4.1 shows the buy and sell electricity power to/from the grid or fail/waste measure of power if off-grid. The colors used to represent the buy amount is salmon and the sell amount is lime. The final (bottom) graph keeps track of the actual cost or profit which can be made with the situation.

4.3.1 Results for single situation of source, load, and battery size

According to this experiment design, this work first fixed an ideal scenario and generated the resulting graph. After that, the system considers a higher load, solar power, and battery size. In the ideal case, the system fixed the battery size 20KWh, the maximum load was 40KW and the maximum solar power was 30KW.

According to the resulting graph for ideal design, illustrated in Figure 4.1, optimizer determines an optimal policy based on the time of day electricity price. It uses the stored electricity or solar power to fulfill the demand or sell back to the grid when the day of electricity is highest. On the other hand, it buys the electricity from the grid to support the load or store in the battery. The optimized cost for this ideal case is close to 24 dollars for February 1 and 2. For this experiment, in this work keep the smallest timestamp as fifteen minutes long.

Effect of an increase the battery capacity

In this case, the system increased the battery size to 50KWh and keep the remaining parameters the same. By examining the graph, the optimal policy changed in two areas

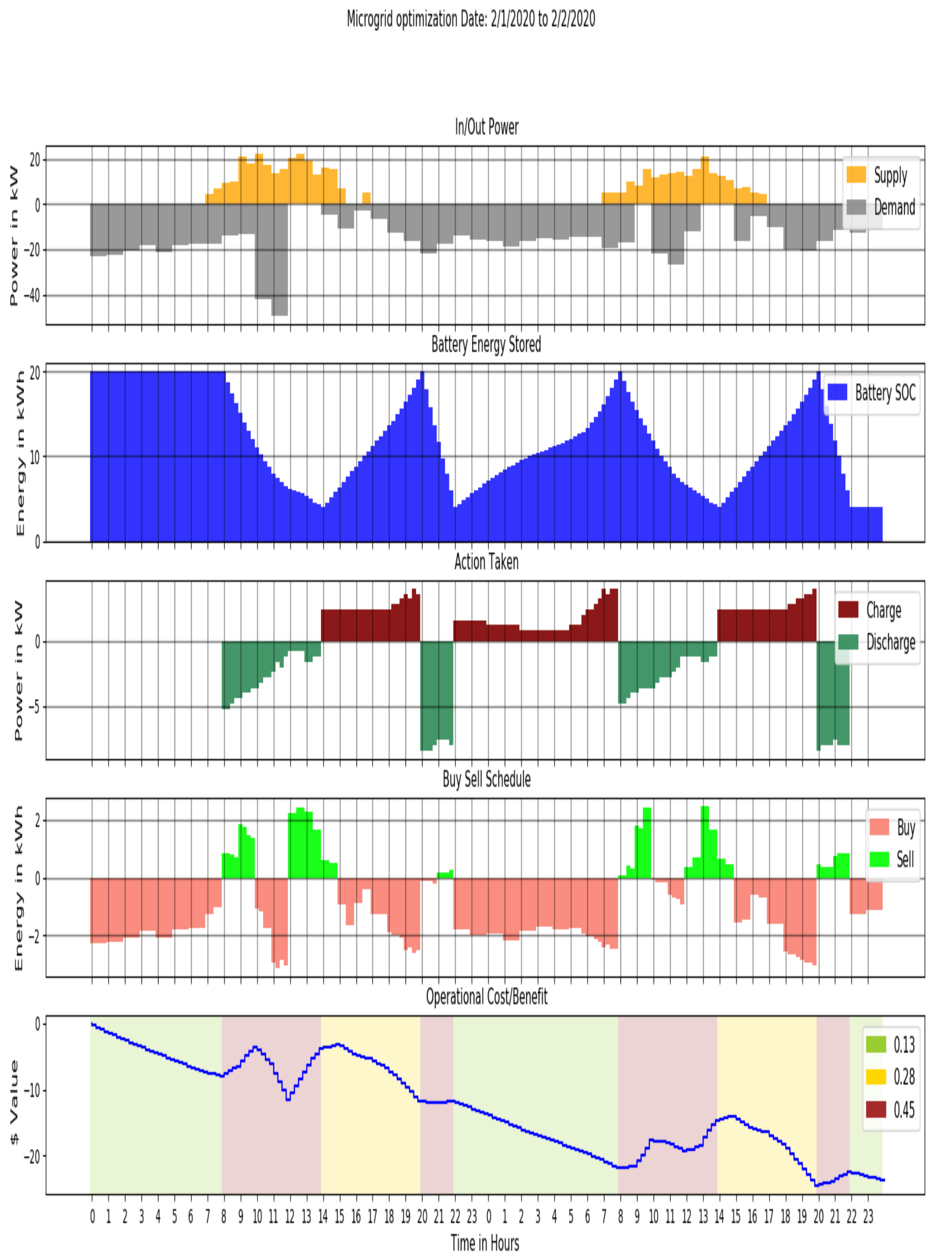

Fig. 4.1: Time cost optimization for battery size 20KWh, solar capacity 30KW and maximum load 40KW from 02/01/2020 to 02/02/2020.

compared to the earlier study with a battery size of 20Kwh. First, now the optimizer buys more electricity from the grid in the off-pick hour and then sells it back in peak hours. Second, it takes more constant action sequences to fulfill the goal of this optimizer. The following Figure 4.2 display the entire graph. Overall in two days, the system can profit or make 5 dollars from this optimizer and also support the existing demand.

Increase the solar capacity

In this experiment, the method only changes the maximum solar power capacity from 40KW to 100KW. Analyzing the output of this configuration from the Figure 4.3, the system found that most of the peak time is aligned with the time when it had solar power available. For the increased solar power, it did not need to buy the electricity when the price is high to support the loads. As mention that most of the solar generation was in peek hours, the optimizer can sell it at a higher price and make more money. In this case, the system can support the entire load of the system and income around 38 dollars per two days as profit.

Increase the load capacity

This experiment considers a future with more EVs added in the system and load increased by a constant factor. Here, the maximum load was increased from 40KW to 100KW. The results shown in 4.4 show the opposite result compared to increased solar above. Optimizer needs to buy more energy from the supplier and the costs are increased to support the required load. To support this total load optimizer required around 150 dollars.

4.3.2 Results for use different day-of-pricing

Day-of-Pricing indicates the pricing policy from the power generation company. Based on the use of electricity in different hours of the day, the power generation company sets the price of electricity. In this experiment, the method used two separate states' pricing policies to understand the taken action sequences' validity. One of the pricing policies from Southern

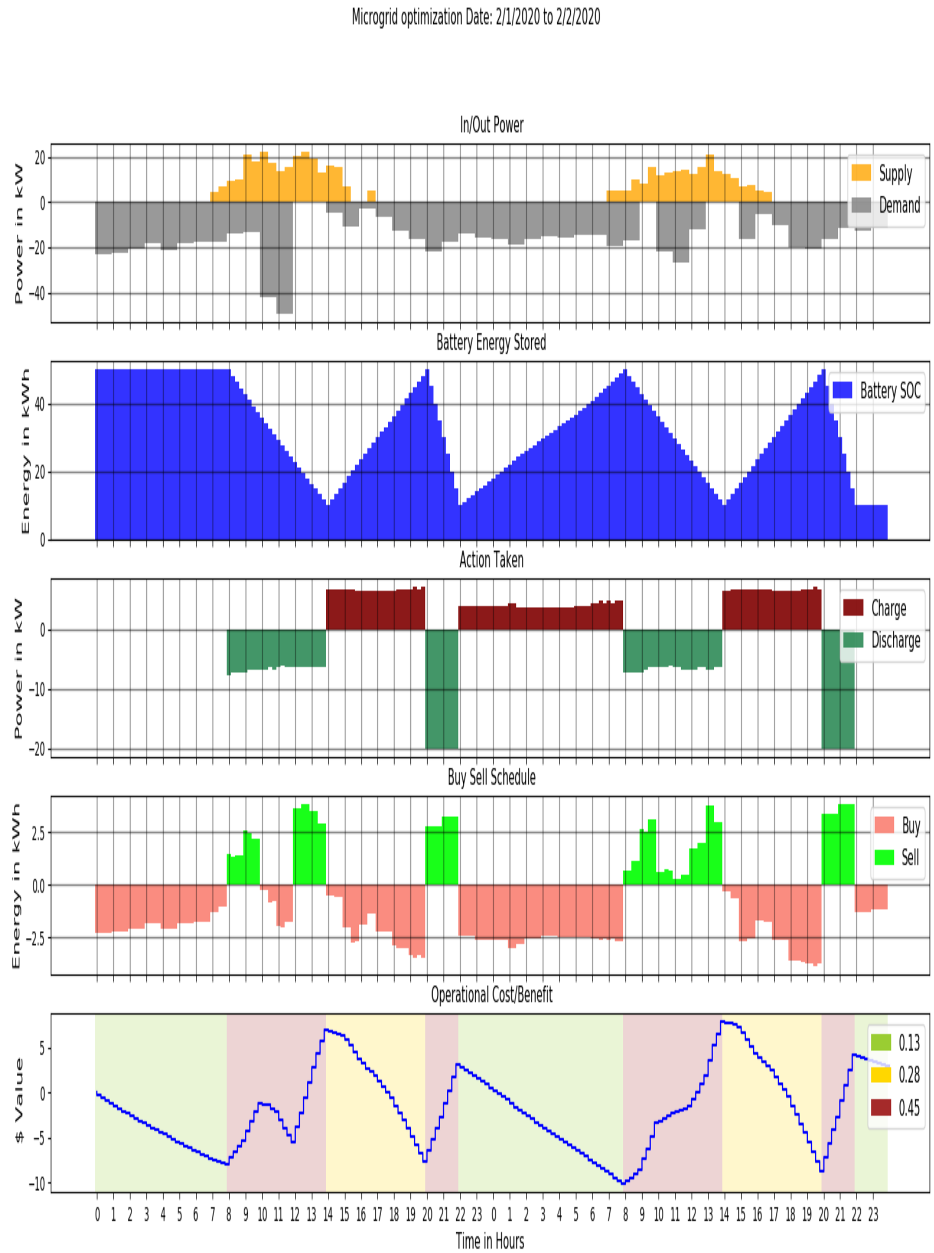

Fig. 4.2: Time cost optimization for battery size 50KWh, solar capacity 30KW and maximum load 40KW from 02/01/2020 to 02/02/2020.

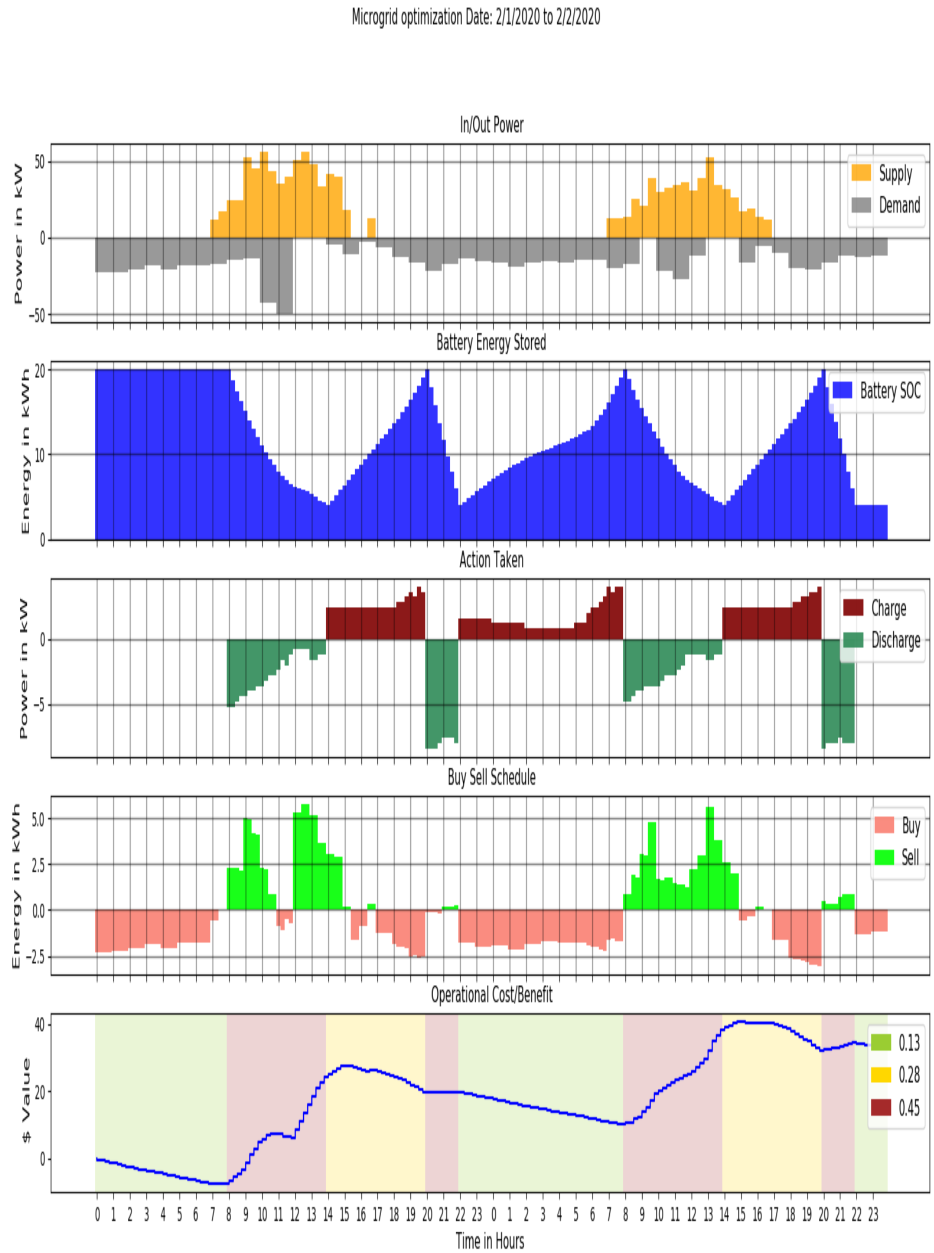

Fig. 4.3: Time cost optimization for battery size 20KWh, solar capacity 100KW and maximum load 40KW from 02/01/2020 to 02/02/2020.

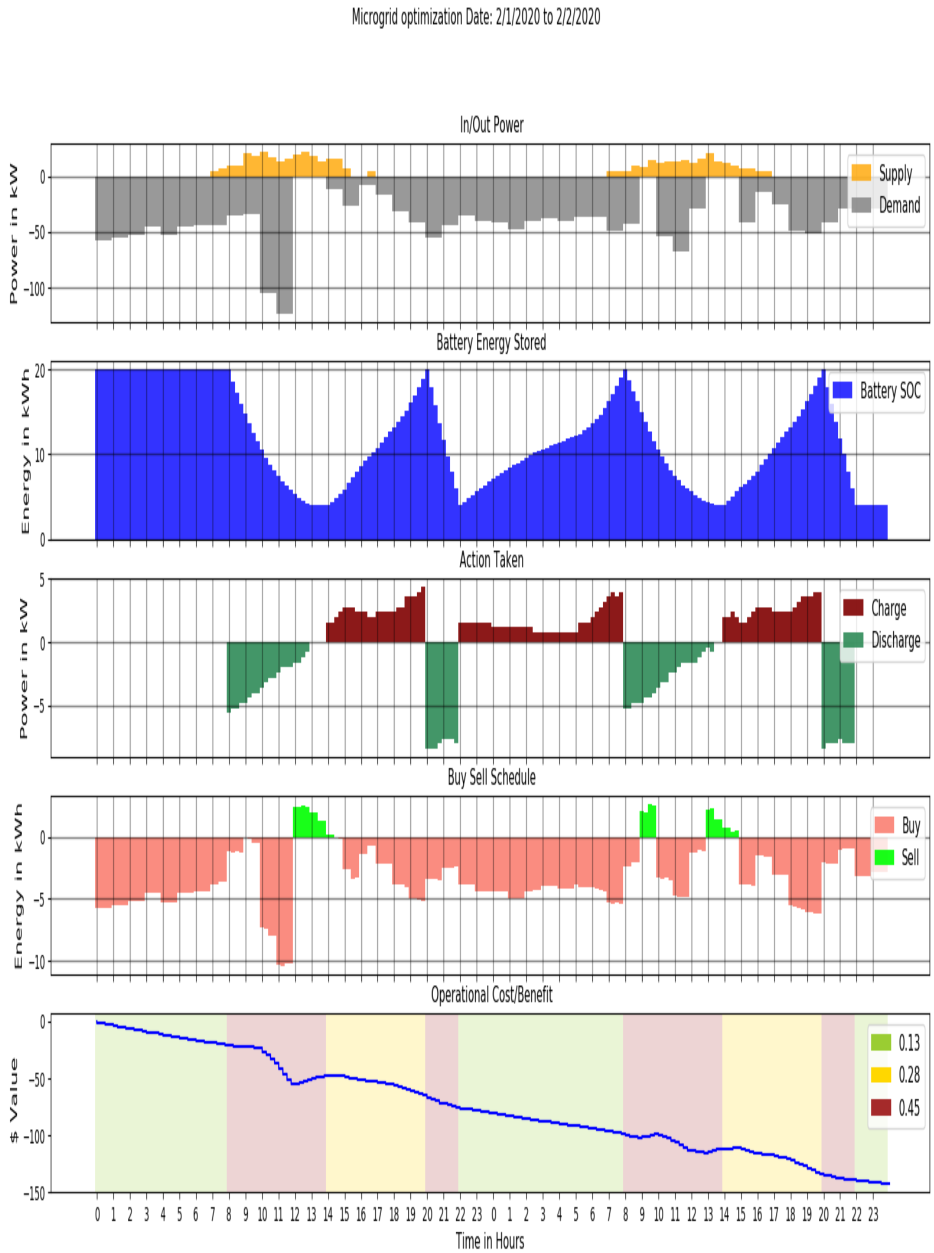

Fig. 4.4: Time cost optimization for battery size 20KWh, solar capacity 30KW and maximum load 100KW from 02/01/2020 to 02/02/2020.

California [18] and the other is from Texas [19]. For both cases, the optimizer assumes that the battery size is 20KWh, maximum solar capacity is 50KW and the maximum load is also 50KW.

Southern California

In table-5.1 and 5.2 shows the pricing policies of the southern California area for winter and summer respectively. Costs divide the day into three parts of pricing called off-peak, mid-peak and peak hours. The price of the peak hour electricity is highest than mid-peak and the price of off-peak is the lowest.

Table 4.1: Southern California pricing policy for Winters

Category	Times	Price in dollar
Peak	8 a.m. to 2 p.m. and 8 p.m. to 10 p.m.	0.45
Mid-Peak	2 p.m. to 8 p.m.	0.28
Off-peak	10 p.m. to 8 a.m. and Weekends and Holidays	0.13

Table 4.2: Southern California pricing policy for summer

Category	Times	Price in dollar
Peak	2 p.m. to 8 p.m.	0.45
Mid-Peak	8 a.m. to 2 p.m. and 8 p.m. to 10 p.m.	0.28
Off-peak	10 p.m. to 8 a.m. and Weekends and Holidays	0.13

According to the result in Figure 4.5 which is the pricing policy of southern California, the total cost of support the whole two days demand is almost 30 dollars. Before the beginning of the peak hours, the optimizer keeps the battery full either buying from the national grid or utilizing solar power. Taking this action is justifiable because the system can sell the energy to the grid with a high price or support the load without buying from the grid. The optimizer keeps the battery level as low as possible at the starting of the off-peak hour. As the optimizer has the knowledge of pricing differences between hours and the predictions of load and solar production, it attempts to make more money using this information.

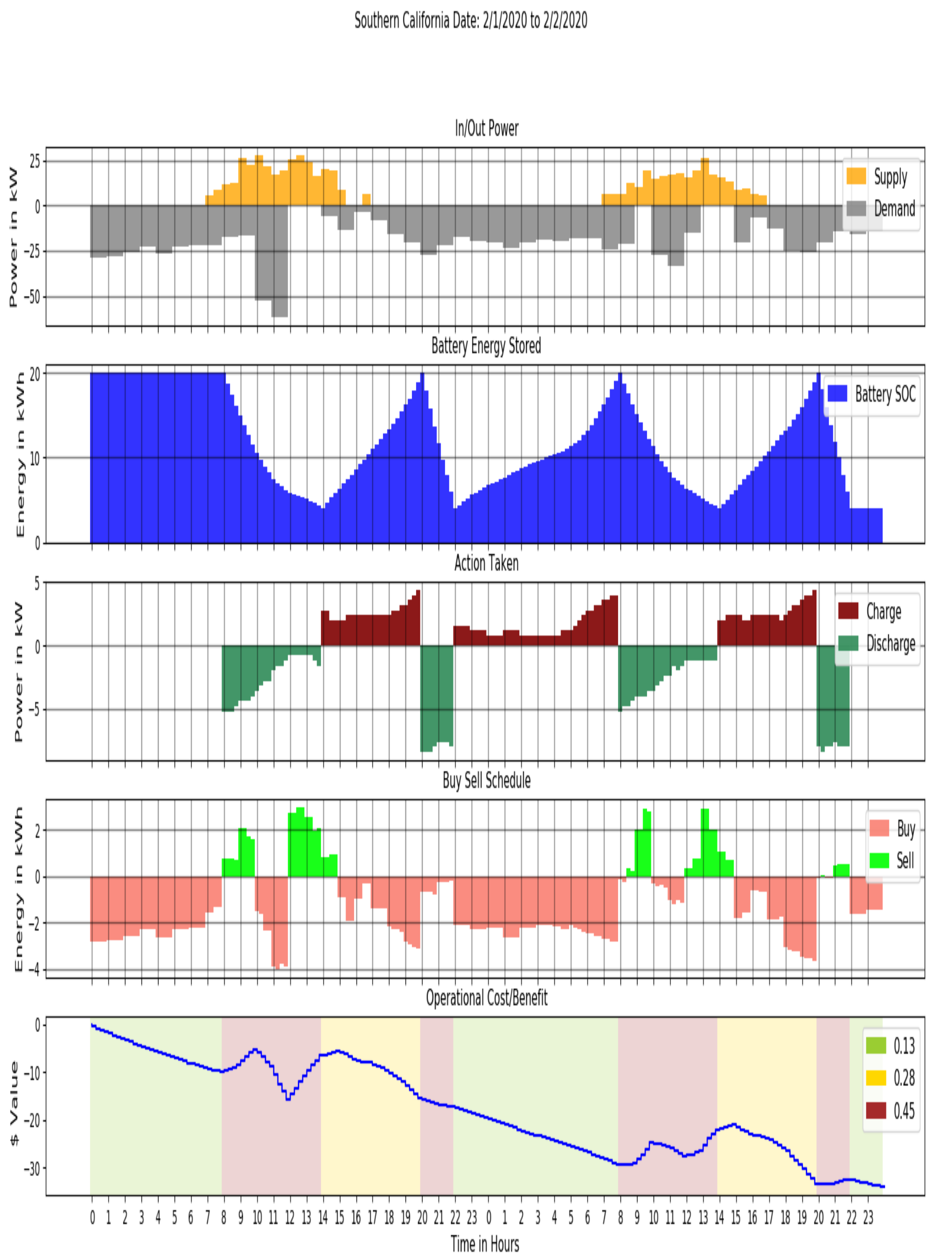

Fig. 4.5: Time cost optimization for Southern California from 02/01/2020 to 02/02/2020

Texas

The pricing policy of Texas is quite different than southern California. The price of the unit electricity is less than in California and the difference between peak hours and off-peak hours also less. The following tables 4.3 and 4.4 represents the winter and summer price respectively. The reason to choose Texas is it also divided a day into three different pricing policies.

Table 4.3: Texas pricing policy for Winters

Category	Times	Price in dollar
Peak	8 a.m. to 12 a.m. and 6 p.m. to 8 p.m.	0.132
Mid-Peak	12 a.m. to 6 p.m.	0.094
Off-peak	8 p.m. to 8 a.m. and Weekends and Holidays	0.065

Table 4.4: Texas pricing policy for summer

Category	Times	Price in dollar
Peak	12 a.m. to 6 p.m.	0.132
Mid-Peak	8 a.m. to 12 a.m. and 6 p.m. to 8 p.m.	0.094
Off-peak	8 p.m. to 8 a.m. and Weekends and Holidays	0.065

The resulting Figure 4.6 describes the result with the Texas pricing policy. In contrast to California, it requires lower money to support the same load in Texas. As the price is lower overall, the optimizer does not fill the battery before the peak hours. Also, it does not sell the app the energy from the battery before the off-peak hours. As the system takes into account costs of battery operation (using the battery model described in Section **??** below) it blocks frequent charge and discharge actions and the holding of energy in the battery since these result in battery degradation. Total cost for support the load is almost half from California at 16 dollars.

4.3.3 Results for Battery modeling

Battery Modeling is one of the innovative additions in this research. Battery modeling takes into account how operations on the battery influence long term costs to the user which

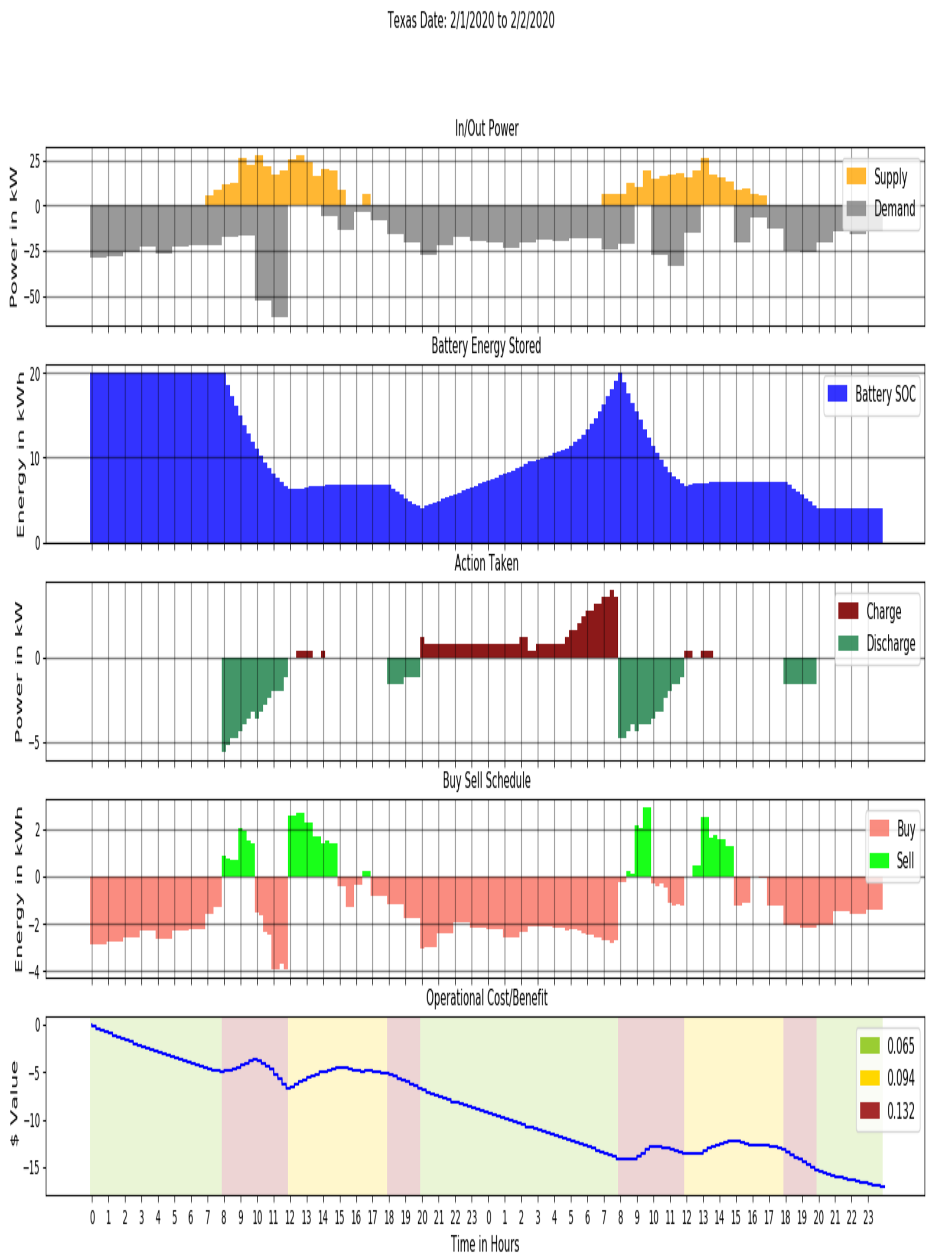

Fig. 4.6: Time cost optimization for Texas from 02/01/2020 to 02/02/2020

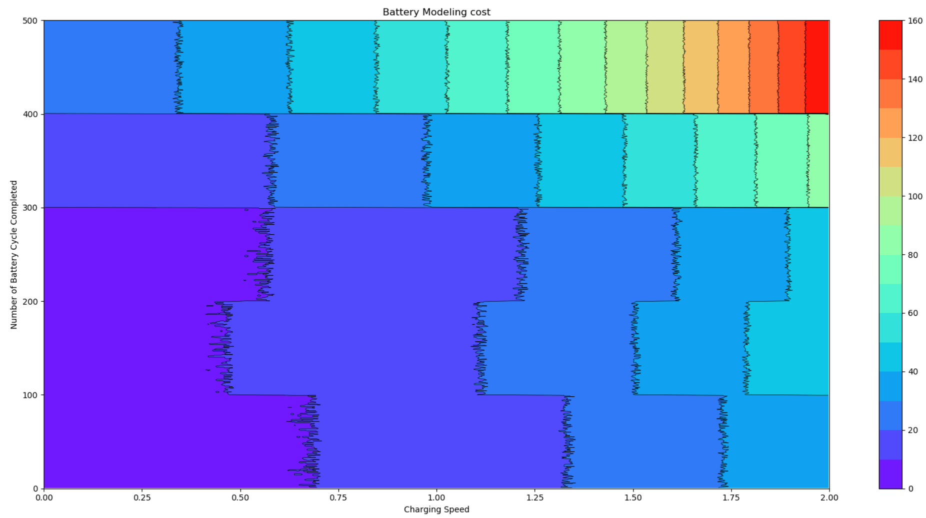

Fig. 4.7: Estimated cost for battery based on battery cycle and charging speed.

are ignored in the ideal case. This makes the research work more accurate in assessing the total operational cost of the microgrid. As described in Section 3.1.4, the battery modeling categorizes this cost into two parts: a holding cost and a charging cost. For this experiment, four different scenarios were explored to fully understand the impacts of these two kinds of battery costs. The first experiment is without any battery cost, the second adds only the holding cost, the third applies only the charging cost and last is with a full battery model, which is a combination of both.

In Figure 4.7 it shows the cost calculation concerning charging speed based on [20]. To determine this cost, the number of charge-discharge cycles completed is recorded. In the figure, the y-axis represents the battery completed cycle and the x-axis shows the charging speed in amps.

Without Battery Modeling

For this experiment, the system fixed the battery capacity of 20KWh, load maximum at 60KW, and solar power capacity 50KW with a southern California pricing policy. Without the battery modeling, the optimizer always took the naive step and constantly filled the battery into 100% to support the load and buy cheep energy from the grid. At the

transaction time step when it moves from peak hour to off-peak hours, the optimizer took the action to sell the entire stored energy and make the highest profit when selling price was highest. This is the only time it took the action from the battery which is shown in Figure 4.8. As the system is not considered the charging speed in the reward function it took this high charging rate action. The total cost to support the load was around $30.

Only using holding cost

Now for this experiment design, battery holding cost was added. Figure 4.9 shows the best action sequence can cost less to operate. After adding the holding cost, the optimizer shifts from the prior policy to fill the battery the whole time. The battery storage level is maintained between 20% to 80% except for the transaction time from peak to off-peak. At that transaction time, it fills the battery and uses it in the peak hours to minimize operating costs.

Only using charging cost

The research added a charging cost which takes into account wear on the battery due to high charge or discharge rates. Figure 4.7 showed the cost charging depends on the speed and number of charge-discharge cycles. For this experiment, the system set the battery cycle to zero. Analyzing the output of the results in Figure 4.10, the optimizer avoided a high charging speed. Now it chooses a smaller charging or discharging rate to fill the battery before the peak hours and did the opposite at the beginning of off-peak hours. The entire cost of this scenario is around 32 dollars for two days.

With total battery modeling

From the former research, it is clear that the total battery cost is best modeled taking into account both battery holding costs and the charging speed costs. So, in this research, the optimizer considered both operating costs as a component of the reward function. Following the result in Figure 4.11, the optimizer now bypasses the high charging speed and also decides to hold less electricity in the battery. When the optimizer considered an action

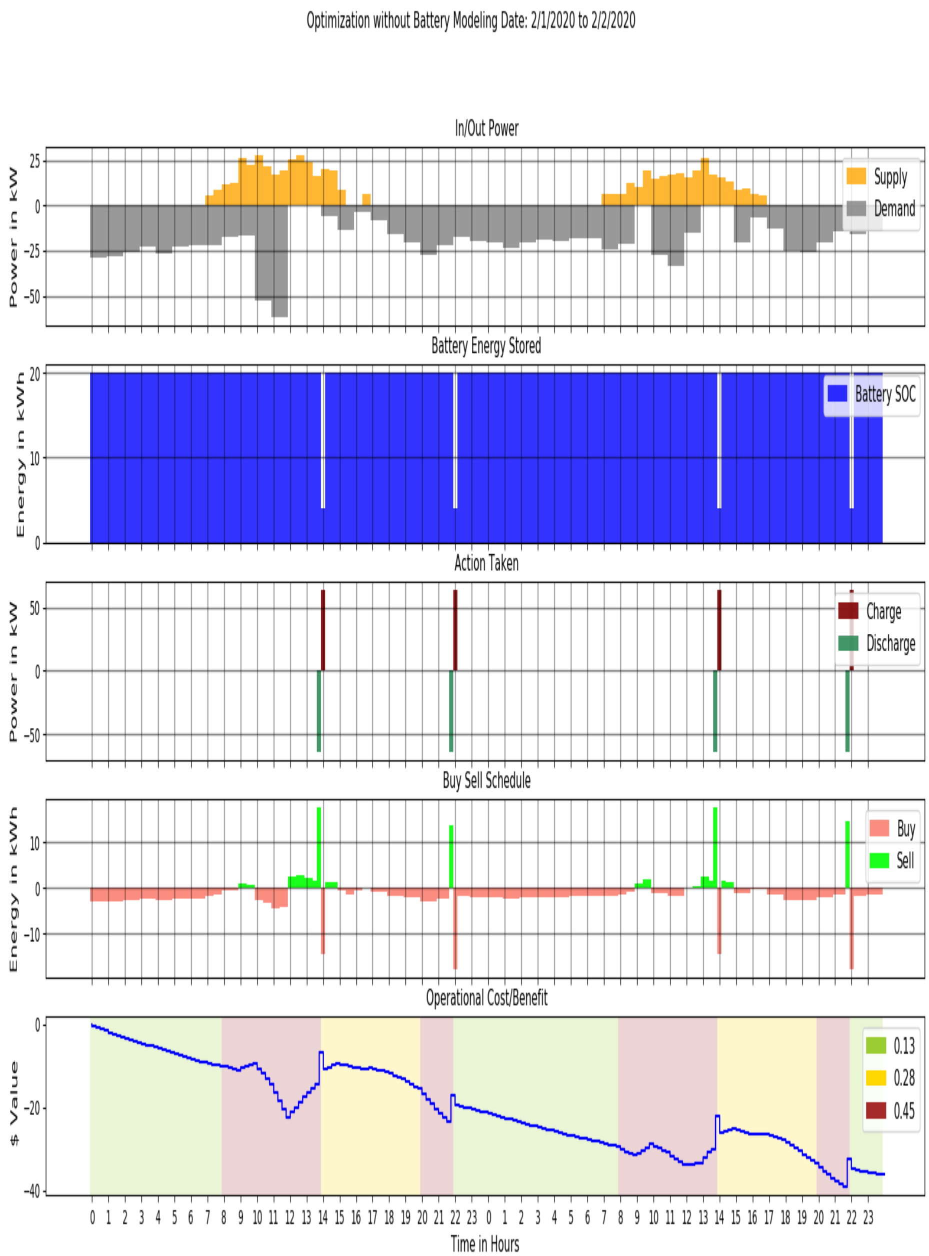

Fig. 4.8: Optimizer output without battery modeling.

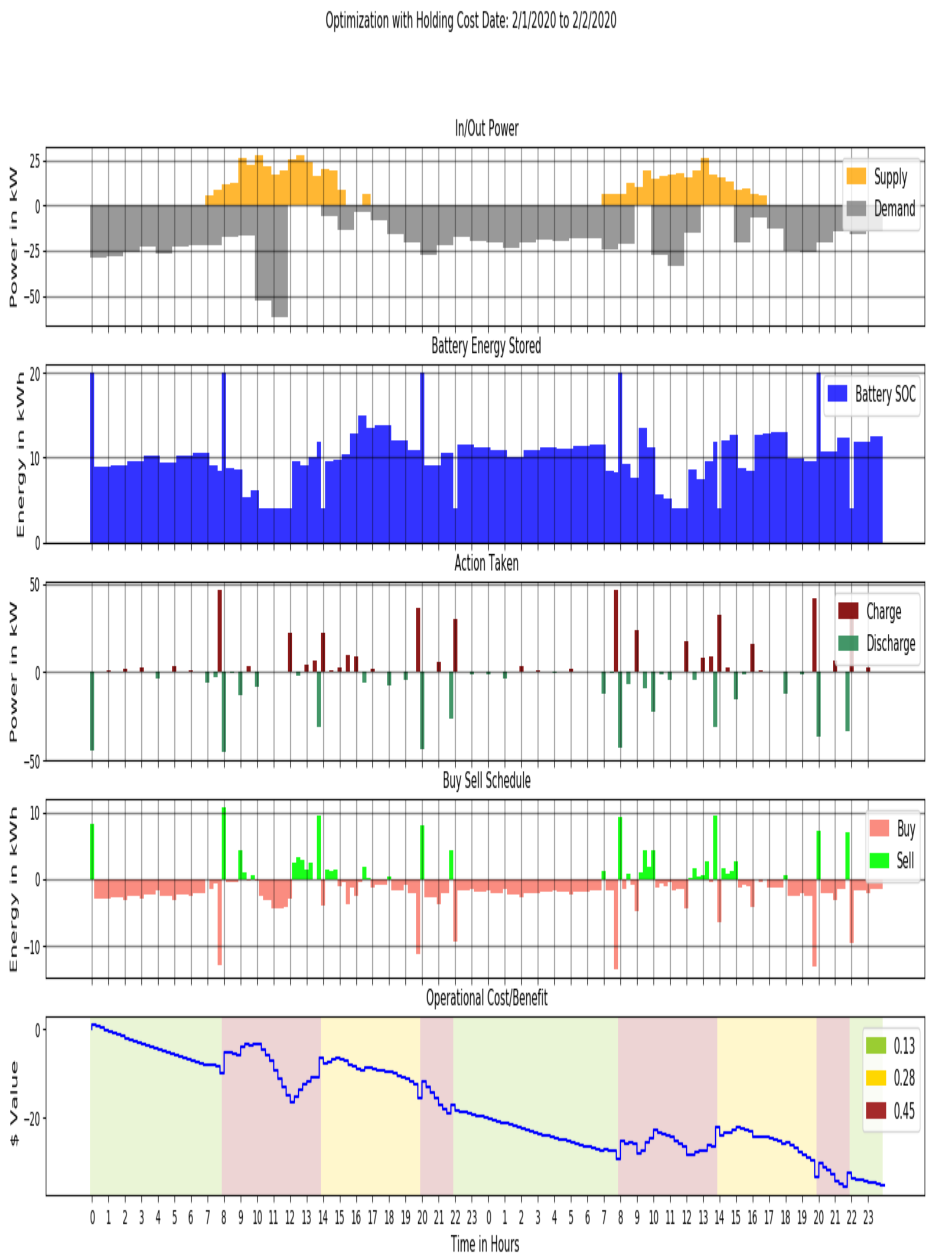

Fig. 4.9: Optimizer output with holding cost.

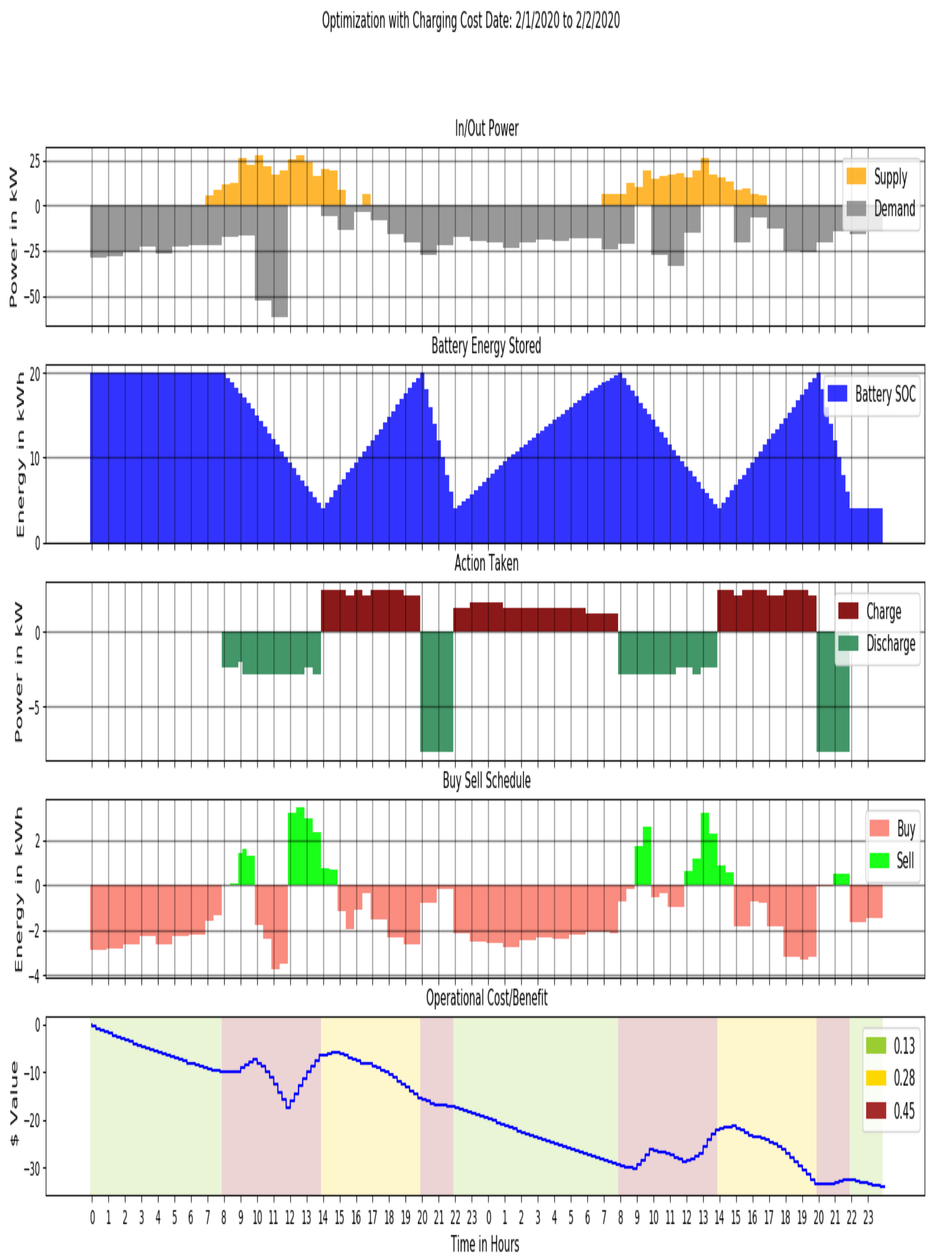

Fig. 4.10: Optimizer output with charging cost.

for discharging it takes large charge steps and after a safe level of the battery, it takes small steps to discharge. On the other hand, when charging it initially chooses small steps and after some time, it chose the larger steps to charge. It is clear evidence of considering the holding cost compare to the prior results. The system still exploits the short period of peak hours to discharge the battery and make more money.

Enumeration of alternative battery and solar configurations

It is clear that when a microgrid is connected to a grid, the buy-sell price policy of the grid has a significant impact on the operational profit and loss. The following Figure 4.12 helps understand how following the optimizer action sequence leads to a linear increase in profit based on the battery and solar capacity. If the solar capacity is increased, then we can store extra solar power and sell it when the price is high. Similarly, if the battery capacity is increased, then we can buy electricity when the price is low and sell them when the price is high. In both of these cases, the energy is saved at low price and sold at high price, leading to a linear increase in profit when following the optimizer's actions sequence.

On increasing the solar power by 5kW, a profit of 1.6 dollars per day is obtained. On increasing the battery capacity by 5kWh, a profit of 0.25 dollars per day is obtained.

4.3.4 Results for Complex buy/sell arrangement

All the previous results have assumed that the microgrid may buy and sell electricity at any time. However, power companies have been changing their policies in buying back energy from individual producers, either preventing it, or providing very little financial benefit. Additionally, the profitability of microgrids that are off grid is of great interest for deployment at more remote locations. The following studies consider alternative scenarios where electricity may or may not be purchased and when electricity may or may not be sold.

Can buy and Can sell

This is the normal case of an electric charging station situated in a city with unlimited

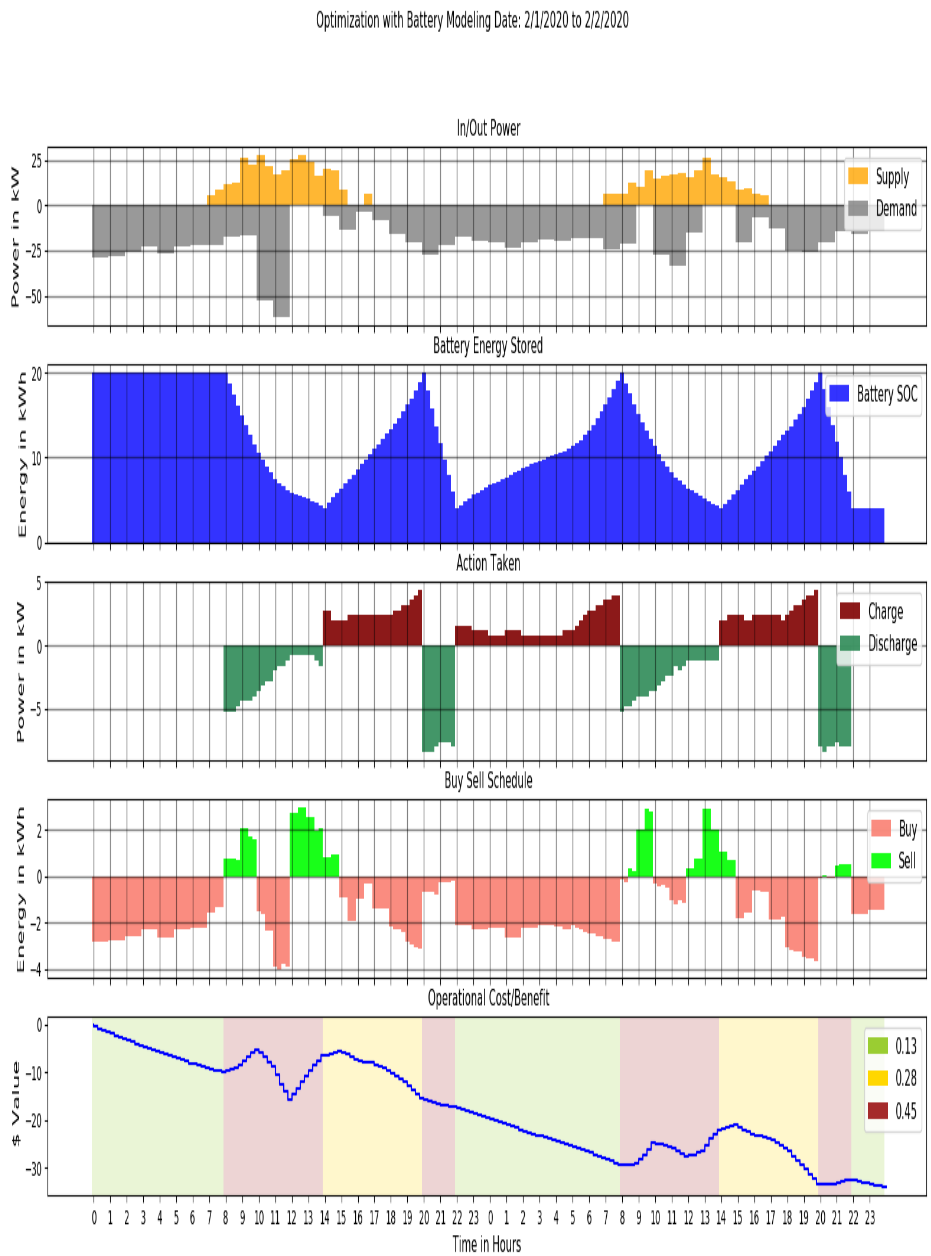

Fig. 4.11: Optimizer output with Battery Modeling.

Fig. 4.12: Profit calculation tuning solar capacity and battery size.

electricity and the likelihood of interrupting electricity connection is close to zero. The optimizer can assume that the EV station can buy electricity anytime when needed and also sell to the grid at its convenience. The following Figure 4.13 shows the selected action sequences and the fourth graph shows the buying and selling amount of electricity for every time period. The operation cost for this designed situation was around 34 dollars.

Cannot buy and Cannot sell

In this experiment, the microgrid is located at a remote place without a grid connection and where it cannot buy extra electricity to support the customer and cannot sell the extra stored energy for more profit. As it was not buying any electricity, the optimizer did not spend any money but it can earn money by serving customers. In this scenario, the reward policy was modified to provide service to as many customers as possible. Even after these modifications, the load demands surpassed the power capacity and it was found that the optimizer failed to provide service for three to four hours every two days as shown in Figure 4.14.

4.3.5 Results for Actual versus predicted profit

Optimizer took two days of solar power prediction and two days of load prediction as an input and based its profit loss calculations on these values. However, all predictions have errors, so it is important to understand the effect of this prediction error will have on the performance of the optimizer. For this experiment, the optimizer operates under different combinations of predicted and actual values and compares the resulting profit and loss calculations. First it uses the purely predicted values. Next, it selects the case six hours ahead, which means it had actual data for the first six hours and for the next 42 hours it used prediction data. It generates actions for 12 hours ahead, 18 hours ahead to 48 hours ahead. Comparing the profit among the results of these experiments helps to understand what is the effect of actual values of inputs and predicted inputs.

Figures 4.15 and 4.16 are the representation of predicted input and actual input respectively. For comparison, the system uses the same configuration of batteries and load

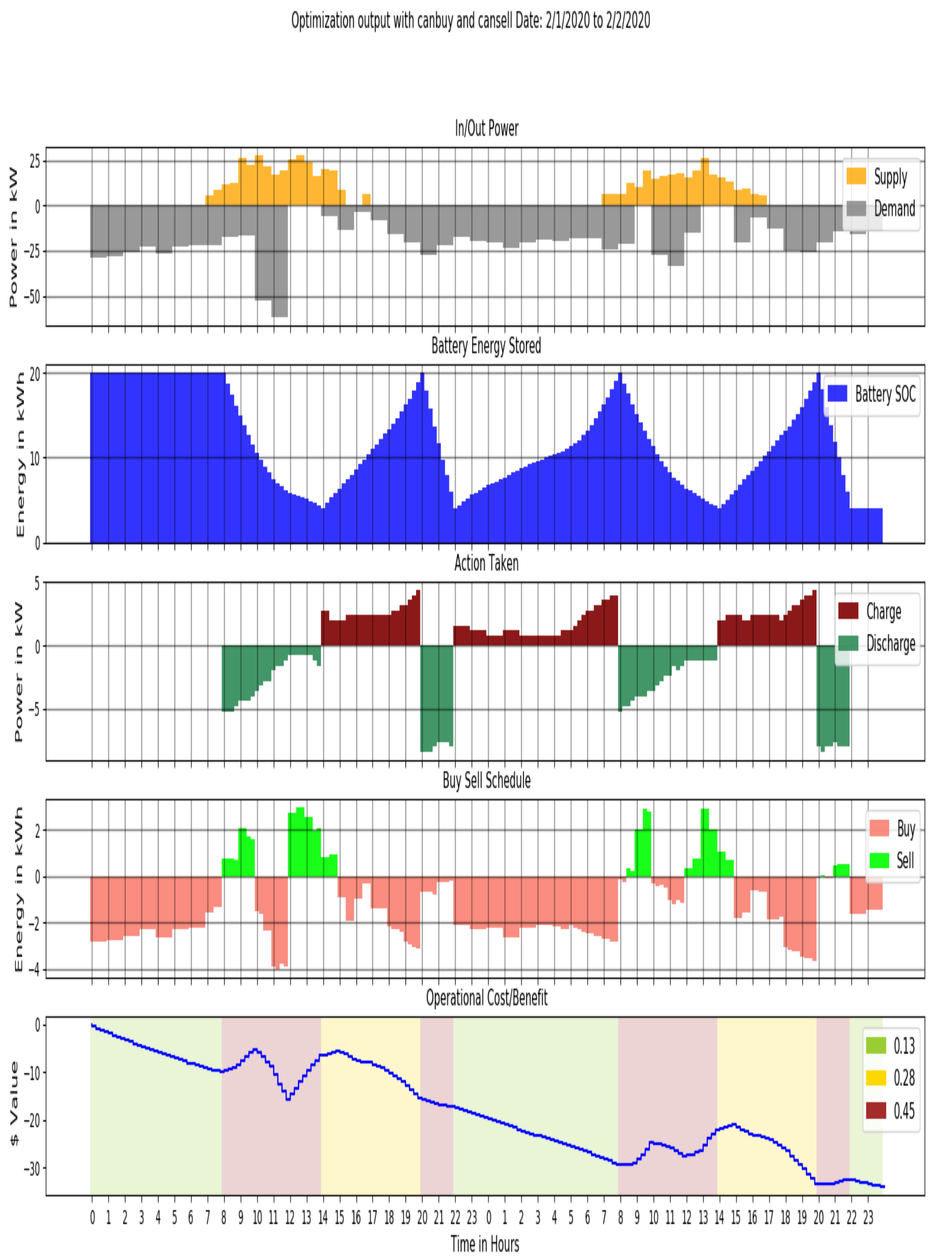

Fig. 4.13: Optimizer output in can buy and can sell situation.

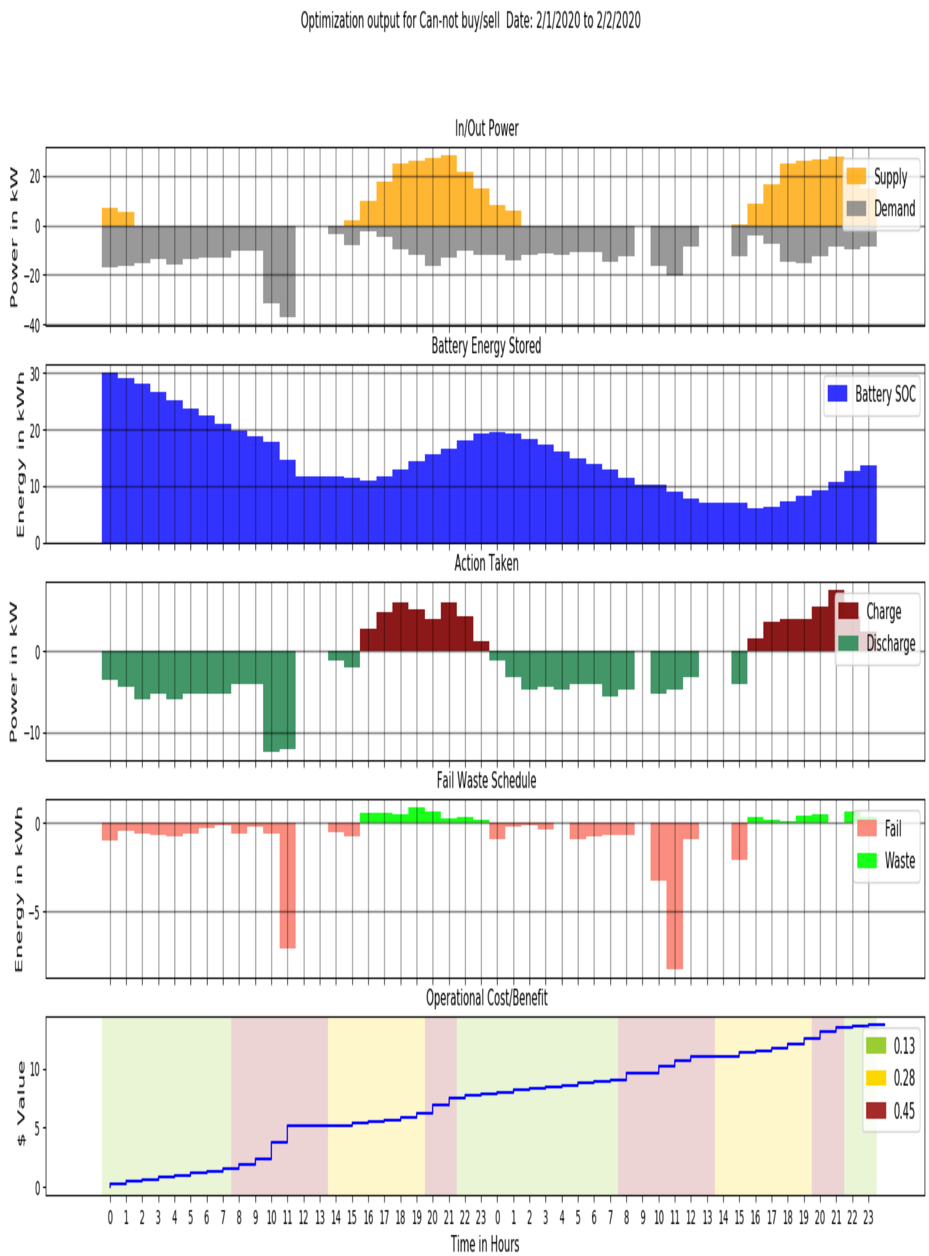

Fig. 4.14: Optimizer output in can-not buy and can-not sell situation.

for both inputs. The optimizer identified an increase of 15 dollars in profit if knows it the actual input. But surprisingly, the optimizer kept the battery in the same states in both experiments for the most part. The actual solar power values are higher than the ones predicted. This leads to the optimizer either wasting power or providing support to extra customers if possible.

Previous studies have focused on a representative two day period to understand the behavior of the optimizer under different scenarios. To obtain a more general understanding, experiments were run from 02/01/2020 to 02/25/2020 and took the average in profit and loss over all the runs. Figure 4.17 shows a line chart of the profit for the first seven days from February and the average profit graph for 25 days. The Y-axis represents how much money the optimizer can make and the X-axis shows how many hours ahead out input is. In conclusion, the actual solar power being higher than the predicted solar power, leads to a minimal change in our policy than vice-versa where it would require a much larger change to the policy to adjust to the lack of power.

4.3.6 Results for tracking versus Non-tracking

Two kinds of solar systems are used in the real world mostly. One is non-tracking and another is tracking. Non-tracking solar panels are positioned and oriented in a fixed direction. On the other hand, tracking solar panel automatically change the position of the panel to follow the sun for the entire day. Tracking solar generates more extra power during morning and evening, a time when loads can be high. In the situation when the optimizer cannot buy or sell in a grid, battery size performed an important role in supporting the load when tracking was not employed. Results show that the same battery size can support more people in an EV charging station when tracking is deployed.

Non-Tracking

In this experiment, the battery capacity was 30KWh, maximum solar power and load were 50KW. The result in the figure 4.18 shows that in 14-time steps out of 48, the optimizer failed to support the load significantly. Here the optimizer only sells electricity to EVs and

00 0Hours Ahead.png

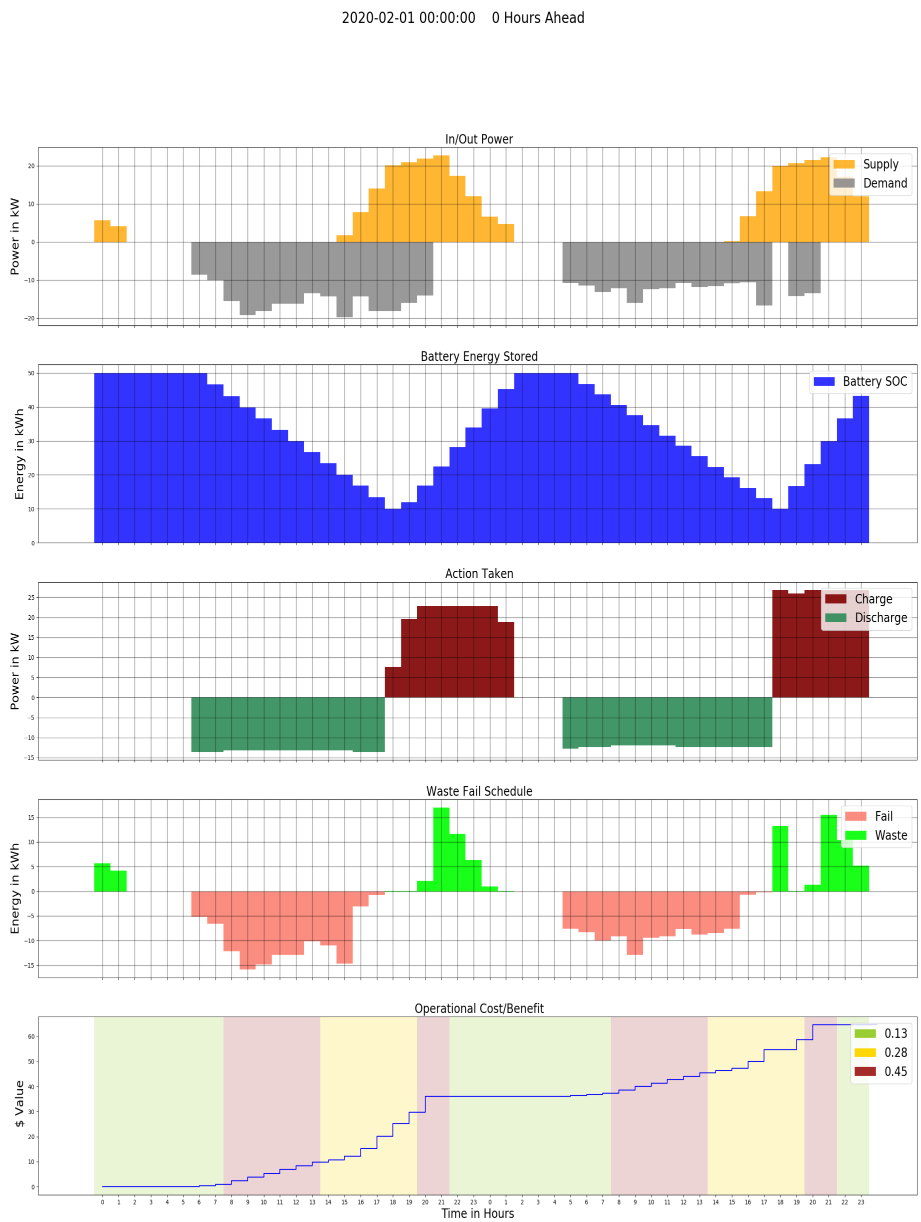

Fig. 4.15: Profit calculation For predicted inputs.

00 48Hours Ahead.png

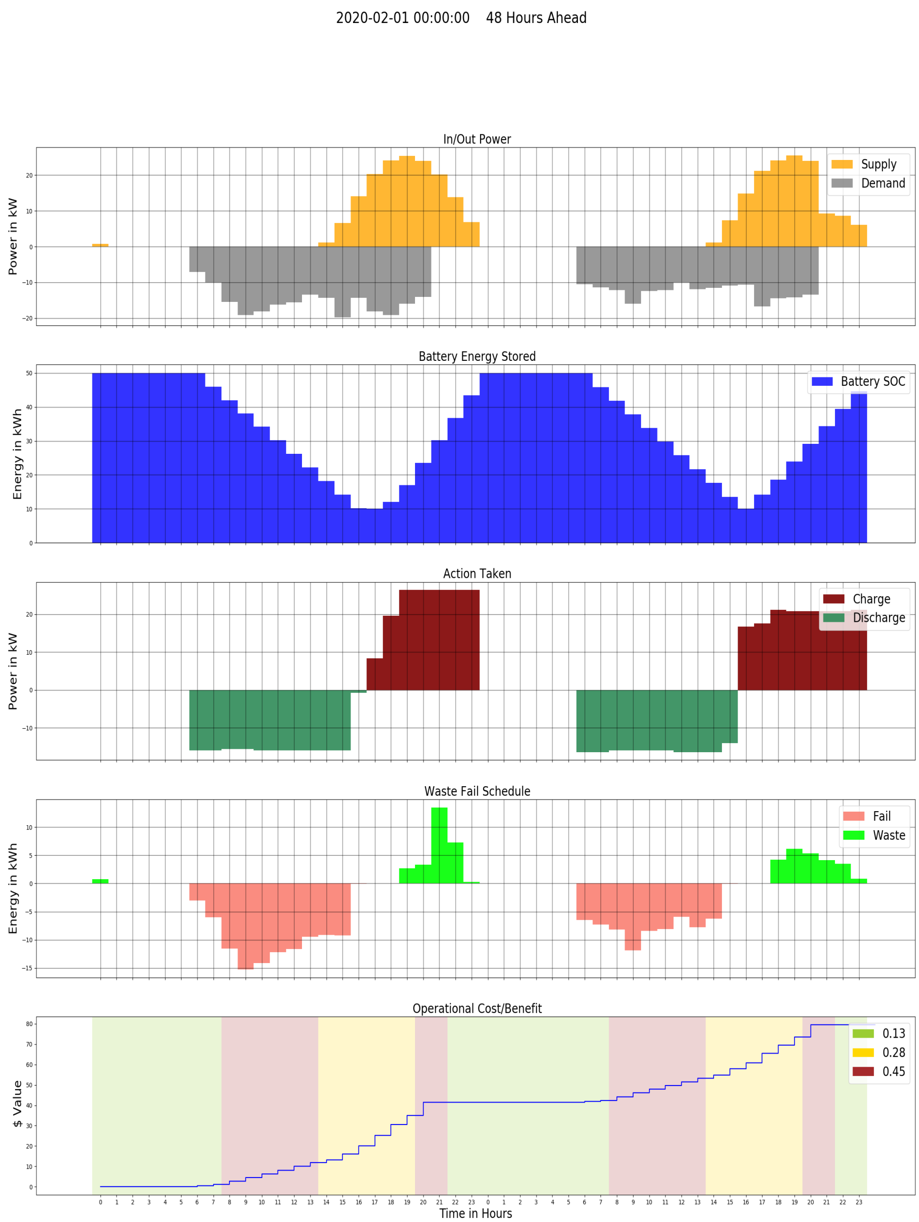

Fig. 4.16: Profit calculation For actual inputs.

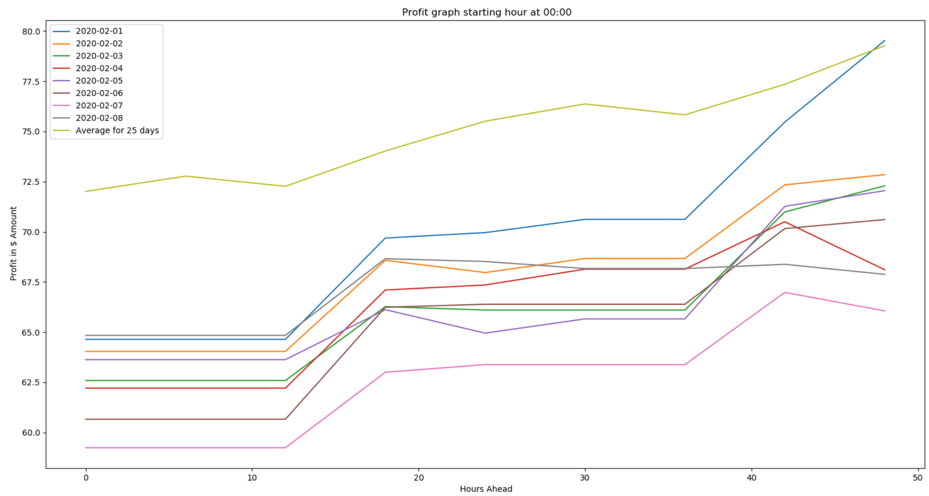

Fig. 4.17: Average profit calculation.

it does not buy from the grid. The profit is calculated based on how much energy it supports to user multiplied by price. In total, it failed to support 107 KWh electricity to the user.

Tracking

In the tracking system, the system generally generates more energy than non-tracking, which was helpful when the EV station in a remote place. In the early morning and evening, tracking collects more energy that a fixed configuration. The optimizer failed 3 times out of 48 times to serve the customers, which is much lower than when it uses non-tracking solar panels. Only 33 KWh was short from the complete client fulfillment state and optimizer made around 14 dollars from users.

4.3.7 Results for Load shifting

From our experiments when performing with tracking, there was a lack of power supply for consumers during the start and end of the day, and excess of power during the middle of the day. This problem can be solved by easily increased battery capacity but the existing stations cannot make this change overnight. Therefore it is important to look for other solutions such as shifting the load to hours of the day where there is an abundance of

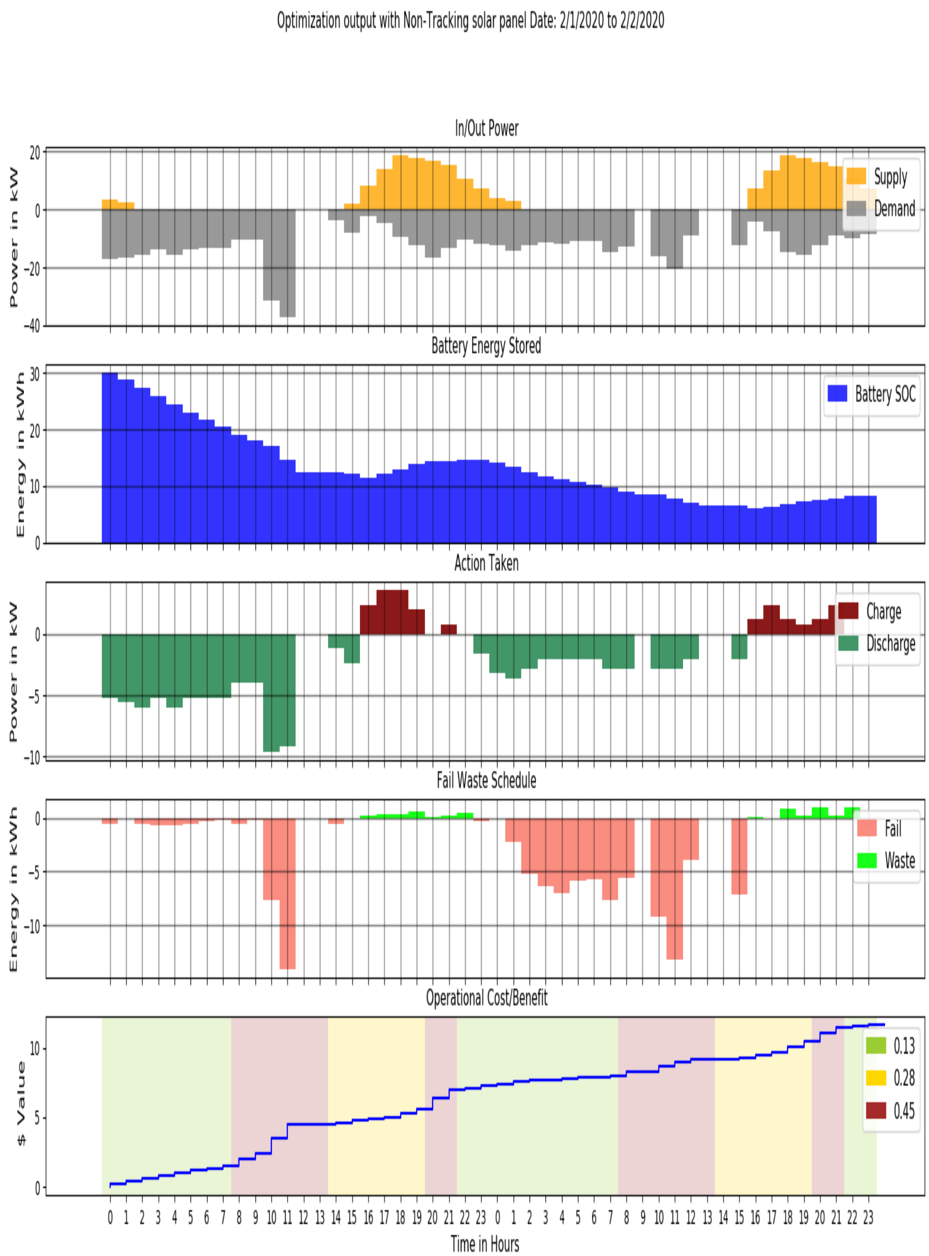

Fig. 4.18: Profit calculation For non-tracking solar panel.

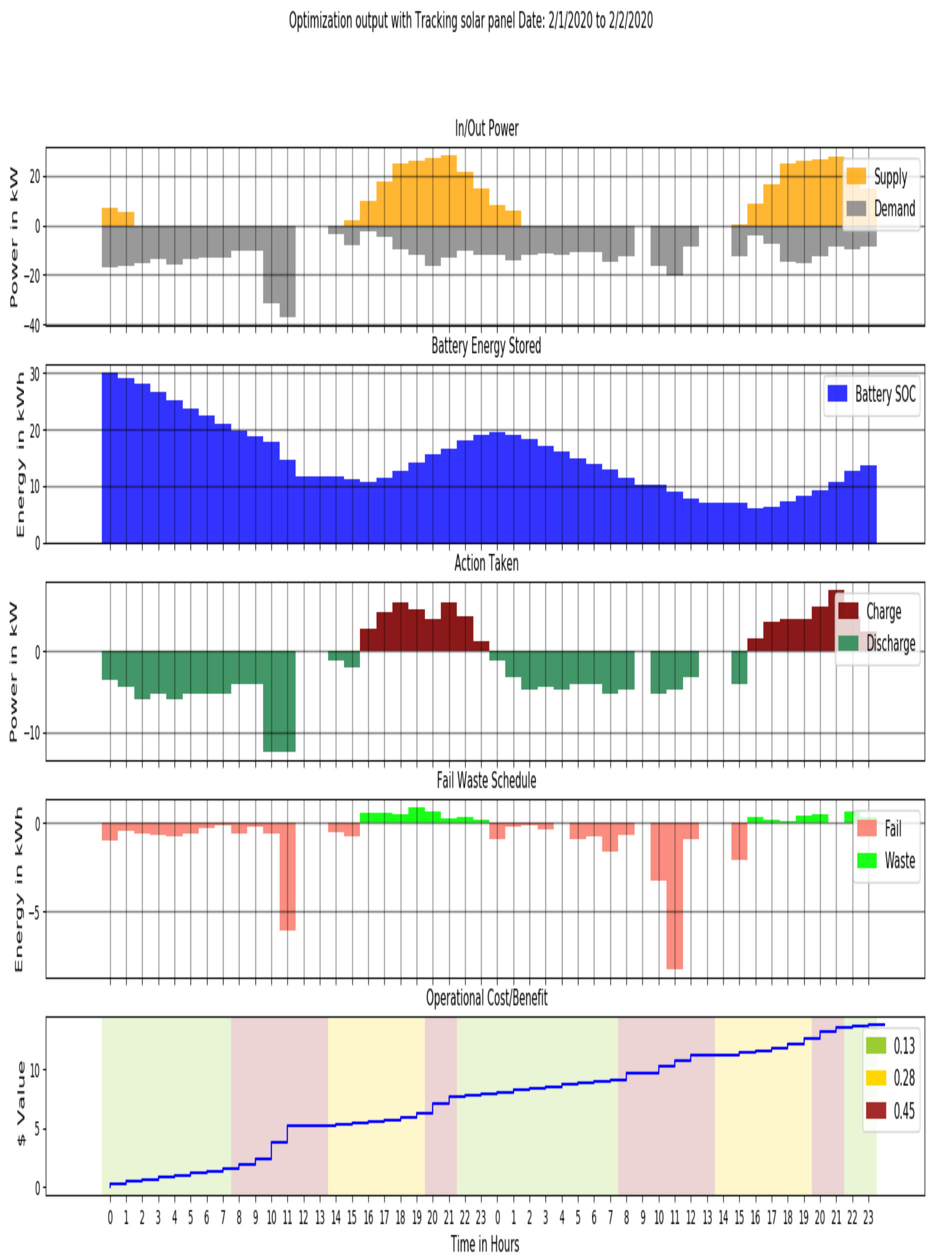

Fig. 4.19: Profit calculation For tracking solar panel.

power.

An experiment has been performed regarding load shifting, where the environment was kept static. The static factors in this experiment were, the battery capacity was 40KWh, maximum solar power and load were 50KW.

As shown in Figure 4.20, during 7-9 and 19-20 hrs of the day, the station failed to support customers due to a lack of stored power. On the other hand, during 11-16 hours, there was a wastage of power due to lack of storage. On experimentation with load shifting, as seen in Figure 4.21, there was no deficit of power during 7-9 and 19-20 hrs. Similarly, there was less wastage of power during 12-15 since the load was shifted to these plentiful hours.

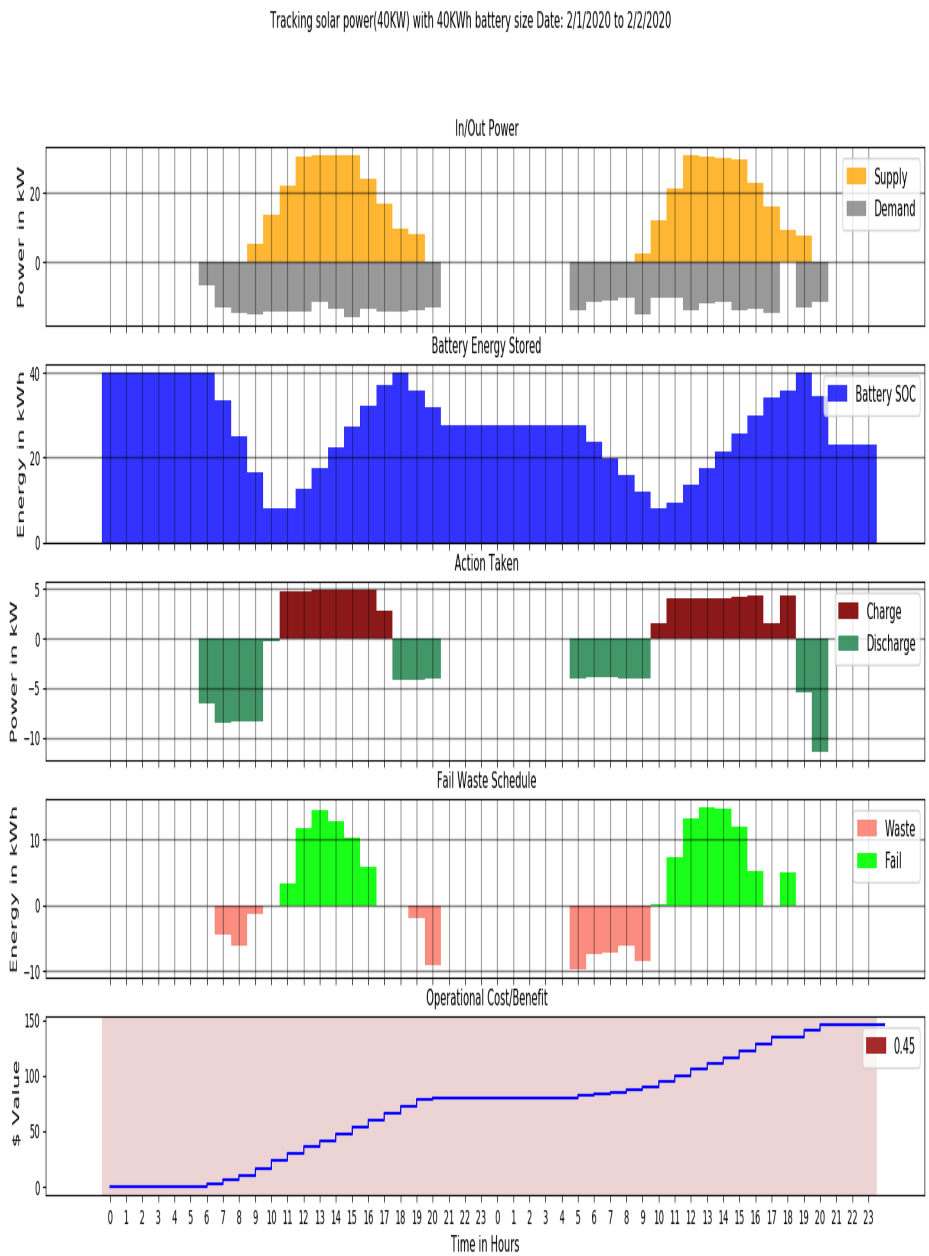

Fig. 4.20: Micro grid Optimization Before Load Shifting.

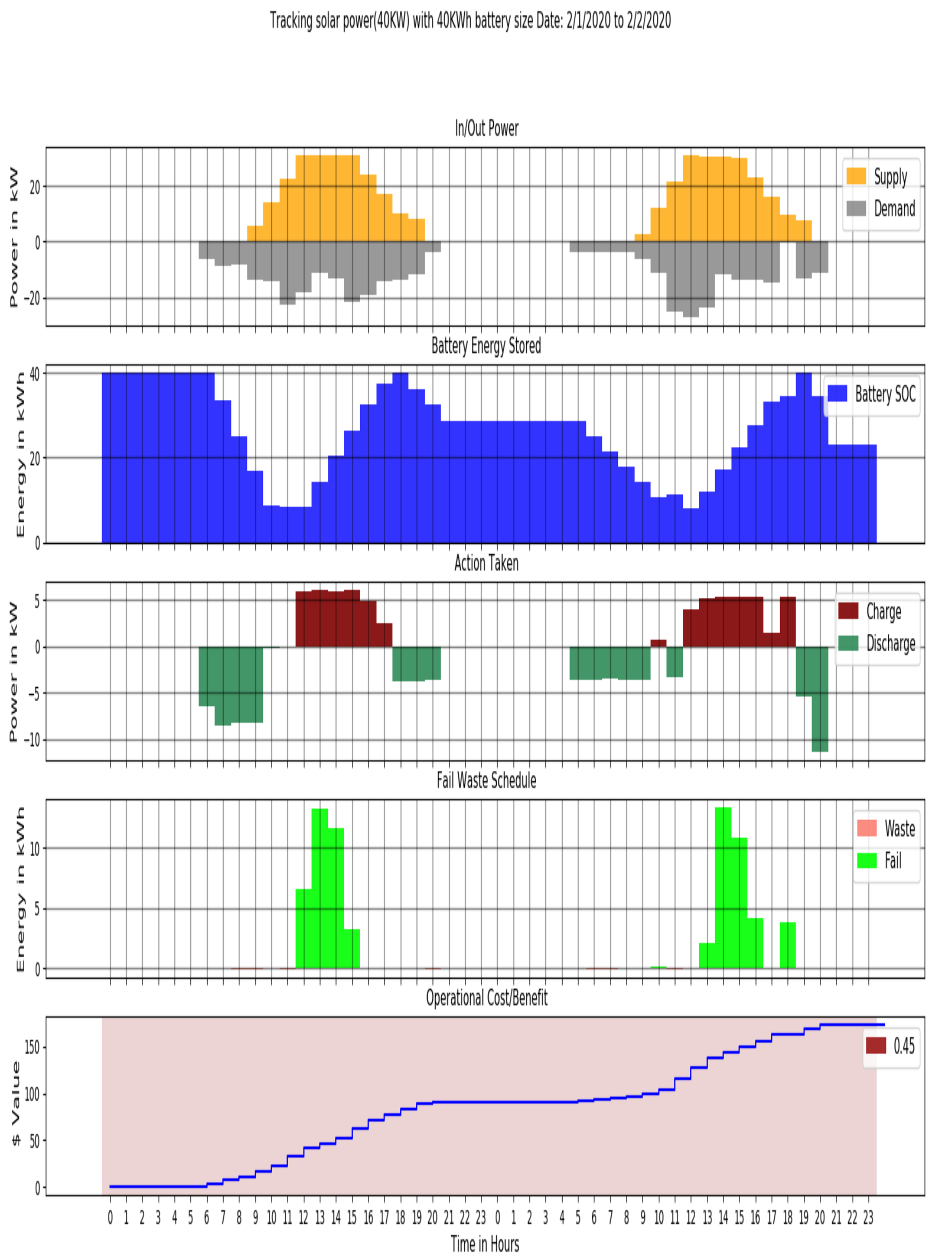

Fig. 4.21: Micro grid Optimization After Load Shifting.

CHAPTER 5

Conclusion and Future work

5.1 Conclusion

This thesis works on designing an energy management system, with the goal of energy efficiency and cost optimization of a microgrid. These optimizations cany help in making microgrid and energy management systems more popular for commercial use.

It is difficult to find a single action sequence for all scenarios with respect to microgrid or energy management systems. Therefore, this thesis explores the robustness of the RL agent by simulating different scenarios. In all of the designed experiments, the results produce optimal action sequences. This thesis also takes the novel approach of including a battery modelling policy in the reward structure, thereby also extending battery life. Additionally, load shifting helped redistribute power demands to the times of the day with abundance of solar power so that it is possible to provide service with lower battery capacity.

5.2 Future work

Our current approach predict actions 48 hrs ahead without a significant overhead in computational power since the time complexity of the algorithm is to the square of the number of states. Optimizing the algorithm to have lower time complexity will be beneficial. The load shifting approach assumes that it is viable to shift consumer demands at a particular time of the day. Realistically this might be be difficult to do, therefore further research is required to find realistic incentives such as price modelling to achieve this.

REFERENCES

[1] mjcoren, "The average price of electric cars is falling," *https://theatlas.com/charts/1Vd8XXNOQ*, 07 2019.

[2] M. Sivak and B. Schoettle, "Relative costs of driving electric and gasoline vehicles in the individual u.s. states," 2018.

[3] M. A. Ortega-Vazquez, "Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty," *IET Generation, Transmission & Distribution*, vol. 8, no. 6, pp. 1007–1016, 2014.

[4] O. Erdinc, N. G. Paterakis, T. D. Mendes, A. G. Bakirtzis, and J. P. Catalão, "Smart household operation considering bi-directional ev and ess utilization by real-time pricing-based dr," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1281–1291, 2014.

[5] J. Zhao, C. Wan, Z. Xu, and J. Wang, "Risk-based day-ahead scheduling of electric vehicle aggregator using information gap decision theory," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1609–1618, 2015.

[6] M. G. Vaya and G. Andersson, "Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets under uncertainty," *IEEE Transactions on Power Systems*, vol. 30, pp. 2375–2385, 2015.

[7] M. G. Vayá and G. Andersson, "Self scheduling of plug-in electric vehicle aggregator to provide balancing services for wind power," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 886–899, 2015.

[8] M. R. Sarker, H. Pandžić, and M. A. Ortega-Vazquez, "Optimal operation and services scheduling for an electric vehicle battery swapping station," *IEEE transactions on power systems*, vol. 30, no. 2, pp. 901–910, 2014.

[9] L. Yao, W. H. Lim, and T. S. Tsai, "A real-time charging scheme for demand response in electric vehicle parking station," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 52–62, 2016.

[10] G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. L. Lewis, "Scalable real-time electric vehicles charging with discrete charging rates," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2211–2220, 2015.

[11] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[12] Z. Wen, D. O'Neill, and H. Maei, "Optimal demand response using device-based reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2312–2324, 2015.

[13] F. Ruelens, B. J. Claessens, S. Vandael, B. De Schutter, R. Babuška, and R. Belmans, "Residential demand response of thermostatically controlled loads using batch reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2149–2159, 2016.

[14] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 5, no. 4, pp. 455–455, 1992.

[15] P. J. Werbos, W. Miller, and R. Sutton, "A menu of designs for reinforcement learning over time," *Neural networks for control*, pp. 67–95, 1990.

[16] H. He, Z. Ni, and J. Fu, "A three-network architecture for on-line learning and optimization based on adaptive dynamic programming," *Neurocomputing*, vol. 78, no. 1, pp. 3–13, 2012.

[17] R. S. Sutton and A. G. Barto, "Reinforcement learning:an introduction," *The MIT Press,Cambridge, Massachusetts, London, England*, 01 2015.

[18] S. by Green Convergence, "What is time of use? tou energy rates explained," *https://www.greenconvergence.com/blog/2014/september/what-is-time-of-use-tou-energy-rates-explaine*, 09 2014.

[19] xcelenergy, "Texas time of use rate," *https://www.xcelenergy.com/staticfiles/xe-responsive/Marketing/TX-Time-of-use-rate-FAQ.pdf*, 01 2018.

[20] S. Choi and H. Lim, "Factors that affect cycle-life and possible degradation mechanisms of a li-ion cell based on licoo2," *Journal of Power Sources*, vol. 111, pp. 130–136, 09 2002.