



**Optimal Policies for the Management of a
Plug-In Hybrid Electric Vehicle Swap Station**

THESIS

MARCH 2015

Rebecca S. Widrick, Second Lieutenant, USAF

AFIT-ENS-MS-15-M-135

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENS-MS-15-M-135

OPTIMAL POLICIES FOR THE MANAGEMENT OF A
PLUG-IN HYBRID ELECTRIC VEHICLE SWAP STATION

THESIS

Presented to the Faculty
Department of Operational Sciences
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Rebecca S. Widrick, B.S. Mathematics
Second Lieutenant, USAF

MARCH 2015

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENS-MS-15-M-135

OPTIMAL POLICIES FOR THE MANAGEMENT OF A
PLUG-IN HYBRID ELECTRIC VEHICLE SWAP STATION

THESIS

Rebecca S. Widrick, B.S. Mathematics
Second Lieutenant, USAF

Committee Membership:

Dr. Sarah G. Nurre
Chair

Lt Col Matthew J. Robbins, PhD
Member

Abstract

Optimizing operations at plug-in hybrid electric vehicle (PHEV) battery swap stations is internally motivated by the movement to make transportation cleaner and more efficient. A PHEV swap station allows PHEV owners to quickly exchange their depleted PHEV battery for a fully charged battery. The PHEV-Swap Station Management Problem (PHEV-SSMP) is introduced, which models battery charging and discharging operations at a PHEV swap station facing nonstationary, stochastic demand for battery swaps, nonstationary prices for charging depleted batteries, and nonstationary prices for discharging fully charged batteries. Discharging through vehicle-to-grid is beneficial for aiding power load balancing. The objective of the PHEV-SSMP is to determine the optimal policy for charging and discharging batteries that maximizes expected total profit over a fixed time horizon. The PHEV-SSMP is formulated as a finite-horizon, discrete-time Markov decision problem and an optimal policy is found using dynamic programming. Structural properties are derived, to include sufficiency conditions that ensure the existence of a monotone optimal policy. A computational experiment is developed using realistic demand and electricity pricing data. The optimal policy is compared to two benchmark policies which are easily implementable by PHEV swap station managers. Two designed experiments are conducted to obtain policy insights regarding the management of PHEV swap stations. These insights include the minimum battery level in relationship to PHEVs in a local area, the incentive necessary to discharge, and the viability of PHEV swap stations under many conditions.

I dedicate this thesis to my husband.

Acknowledgements

It is with immense gratitude that I acknowledge the support and guidance from my advisor, Dr. Sarah Nurre. I would also like to thank Dr. JD Robbins for his invaluable insight to our project and for his continued support of my research.

Rebecca S. Widrick

Table of Contents

	Page
Abstract	iv
Dedication	v
Acknowledgements	vi
List of Figures	viii
List of Tables	ix
I. Introduction	1
II. Literature Review	7
III. Problem Statement	15
3.1 Structural Properties	21
IV. Methodology	37
4.1 Benchmark Policies	39
V. Computational Tests	41
VI. Conclusions	58
Bibliography	60

List of Figures

Figure		Page
1	Diagram outlining the timing of events for the PHEV-SSMP MDP model.....	16
2	Scenarios of demand with respect to inventory for case (b).	30
3	Scenarios of demand with respect to inventory for case (c).	31
4	Charging cost K_t per battery per hour in the Capital Region, NY.	42
5	Mean arrival rate of customers λ_t in a location with 3,000 PHEVs by time of day and day of the week.....	43
6	Interaction plots of significant factors from the first designed experiment.	46
7	Optimal policy by percentage of the charge cost earned for discharging, α	48
8	Expected swap demand and met demand by percentage of the charge cost earned for discharging, α	48
9	State and action over a week time period for three simulated observed demands.	50
10	State and action over one Wednesday for three simulated sample paths.	50
11	Demand throughout a typical Wednesday.....	51
12	Expected demand compared to the number of batteries swapped on a Wednesday for 3 scenarios from the Latin hypercube experiment.	53

List of Tables

Table		Page
1	Factor levels used for the first Latin hypercube designed experiment.	44
2	Factor levels used for the second Latin hypercube designed experiment.	52
3	Results from second Latin hypercube designed experiment.	55

OPTIMAL POLICIES FOR THE MANAGEMENT OF A PLUG-IN HYBRID ELECTRIC VEHICLE SWAP STATION

I. Introduction

Optimizing operations at plug-in hybrid electric vehicle (PHEV) battery swap stations is internally motivated by the movement to make transportation cleaner and more efficient. The U.S. Energy Secretary, Ernest Moniz announced a \$50 million budget in January 2014 for research of vehicle technologies which will also aid the initiative launched in March 2012 to make plug-in electric vehicles more convenient and affordable over the next 10 years [1]. This research initiative is approached by considering the optimal management of PHEV battery swap stations. A PHEV battery swap station allows the PHEV owner to exchange their depleted battery for a fully charged one. By implementing swap stations, not only are PHEV owners offered the convenience to swap their battery, but there is the opportunity to control battery charging and reduce the negative effect of increased demand for electricity on the power grid [2][3] and reduce the difference between high-peak and low-peak energy prices [4].

The concept of battery swap stations for PHEVs was initially developed by the Israeli company Better Place, which financially collapsed in May 2013 [5]. Despite Better Place's collapse, it is still of great interest to examine such swap stations as the manufacturing of PHEVs is on the rise and the motivation to switch from gasoline to battery power has not been diminished. According to the Department of Energy [1], nearly 100,000 plug-in electric vehicles were purchased by Americans in 2013, which is almost twice as many as in 2012.

One of the leading electric car manufacturers, Tesla, first gained worldwide attention when it released the first ever mass produced electric powered sports car in 2010 [6]. The Tesla Model S (sedan) is the current model available for purchase with two battery options and is marked at \$71,070 for the 60 kWh battery option, \$81,070 for the 85 kWh battery option, and \$94,570 for the 85 kWh performance model. The Model X (crossover) has recently been unveiled and is currently available for reservation with delivery expected in Fall 2015 [7]. A third model is said to be released in 2017 at a cost of \$35,000 by the Tesla founder and CEO, Elon Musk [8]. It will be called the Model 3 and will be a direct rival of the current BMW 3 Series electric car. The rolling out of electric vehicles to the market is also occurring for many other vehicle manufacturers. Honda, BMW, Chevrolet, Ford, Nissan, Cadillac, Fiat, Mercedes, Mitsubishi, SMART, Volkswagon, Kia, and Toyota all carry at least one electric vehicle and can cost between \$23,800 for the Mitsubishi i-MiEV to \$137,000 for the 2014 BMW i8 [9].

In addition to being one of the leading electric car manufacturers, Tesla is also the frontrunner when it comes to charging stations. There are currently 129 Tesla supercharge stations in North America, 95 in Europe and 36 in Asia [10]. Electric car owners can plug in their car at a supercharge station and receive 120 kW of charge in just 30 minutes at no cost to the consumer. This provides 170 miles of travel for the Model S 85 kWh battery option. While this is a great option for PHEV owners, it still requires a wait time while the battery is charging and plug-ins may get congested as the number of PHEVs purchased continues to increase. Battery swap stations provide a fast and convenient way to drive away with a fully charged battery. Tesla presented the idea of swap stations in June 2013, but they have not yet come to market [11].

Widely available battery swap stations will help the movement launched in March 2012 by the Department of Energy [1] to make plug-in electric vehicles more conve-

nient and affordable, as well as help control battery charging to avoid loss of power and power quality which can be incurred when batteries are charged during high peak demand for electricity [2]. An ancillary benefit of a swap station is the ability to coordinate discharging back to the power grid through vehicle-to-grid (V2G) technology [12]. When the charging and discharging of batteries is properly coordinated with the power grid, load balancing can occur [13][14][15].

With the significant impact swap stations can have on the growing market for battery powered vehicles, it is valuable to develop a model that optimizes the operations at a swap station. As such, this thesis presents a model which considers uncertainty of battery swap demand and nonstationary charging costs to gain realistic results that are robust to the stochasticity of the system. The PHEV-Swap Station Management Problem (PHEV-SSMP) is considered and a Markov decision process model [16] is developed. Markov decision processes characterize problems with discrete time sequential decision making under uncertainty and can be solved using dynamic programming. They can be modeled using finite or infinite horizons. Infinite-horizon models provide for the determination of a stationary optimal policy, meaning that the optimal action is state dependent and not time dependent. Nonstationary Markov decision processes relax the assumption that problem data does not change with time and are in general unsolvable using infinite-horizon models due to infinite data requirements [17]. A finite-horizon model is considered because the problem data used in the PHEV-SSMP is highly variable with respect to time. The nonstationary variable properties include mean demand for battery swaps, charging price for batteries, and revenue from discharging batteries back to the power grid. In a sequential decision making model, the state of the system is observed at a certain point in time and an action is taken. The action results in an immediate reward to the decision maker and the system transitions to a new state according to a probability

distribution determined by the chosen action.

A Markov decision process contains the following five characteristics: (1) a set of decision epochs or time periods, (2) a state space that is made up of a set of states the system may be in at a given point in time, (3) a set of available actions given the current state of the system, (4) a reward function which is dependent on the set of states and actions, and (5) a transition probability function which is also dependent on the states and actions. The application of Markov decision processes to inventory control models is widely accepted and will provide a framework for the PHEV-SSMP model.

The Markov decision process for the PHEV-SSMP is characterized by the following: (1) decision epochs are a consistent time unit at which a swap station manager needs to determine the number of batteries to charge or discharge, (2) the state of the system is the total number of batteries that are fully charged, where the state of any given battery is either fully charged or depleted, (3) the action space is defined as one dimensional, where the decision maker chooses the total number of batteries to charge or discharge, (4) the reward function is defined using the expected reward criterion which is comprised of revenue from battery swaps, revenue from discharging batteries back to the power grid, and cost from charging batteries, and (5) transition probabilities are determined by customer demand for battery swaps (where demand follows a discrete distribution), the current state, and the chosen action.

A policy consists of decision rules which indicate to the decision maker an action to take in a given state at a given point in time. The objective in solving the Markov decision problem (MDP) is to determine a policy that maximizes the expected total reward criterion. It can be proven that when the demand for swaps follows a discrete nonincreasing distribution, a monotone nonincreasing policy is optimal. The optimal policy, specifically the optimal number of batteries to charge and discharge, for

this finite-horizon model is found using the backward induction algorithm [16]. The optimal policy is compared to two benchmark policies which are easy to implement at the swap station. In the first benchmark policy, which is labeled the stationary benchmark policy, the swap station maintains a single target inventory level of fully charged batteries regardless of time of day and day of week. In the second benchmark policy, which is labeled the dynamic benchmark policy, the swap station maintains a distinct target inventory level for each time period (which captures time of day and day of week information). Each target level is based on the number of batteries at the swap station and the relationship between current and future charging costs. The action for each policy is calculated by taking the difference between the current state of full batteries and the target level. If the swap station has more fully charged batteries than the desired level, they will discharge down to the target and if the swap station has less fully charged batteries than the desired level, they will charge up to the target.

Using realistic data, the optimal solution method and two benchmark policies are computationally tested to gain insight regarding the optimal operations and policies which should take place at a PHEV swap station. Two Latin hypercube designed experiments are performed. The first experiment is conducted to gain overall information for various parameter inputs for the swap station. Specifically, the incentive which should be given by the power company is determined, and other statistically significant factors are analyzed. The second experiment is conducted to gain insight into what the controllable parameters should be set to at a swap station (e.g., number of batteries, swap price) in relationship to the number of PHEVs in a local area and power prices. Further, the results of the second experiment indicate that the dynamic benchmark policy outperforms the stationary benchmark policy, however both exhibit the favorable characteristic of ease of implementation.

Main Contributions. The main contributions of this work are as follows: (i) development of a Markov decision process model to determine the optimal number of batteries to charge and discharge at a PHEV swap station when factoring in stochastic, nonstationary swap demand, nonstationary charging costs, and nonstationary discharging revenues, (ii) proving the existence of a nonincreasing monotone optimal policy when demand is governed by a discrete nonincreasing distribution, (iii) generation of two benchmark policies which are easy to implement by a swap station manager, and (iv) analysis of the results from two designed experiments using realistic data which provide policy insights for a swap station.

The remainder of this thesis is organized as follows. In Chapter II, relevant literature is reviewed in the area of PHEV swap stations, various uses of dynamic programming for energy storage problems, and inventory control Markov Decision Problems. In Chapter III, the PHEV-SSMP is formally defined as an inventory control MDP to include decision epochs, state space, action sets, reward function, and transition probability function. It is theoretically proven that the PHEV-SSMP contains a nonincreasing monotone structure in Chapter III which motivates the optimal and two benchmark policy solution methods presented in Chapter IV. In Chapter V, the proposed model and solution methods are computationally validated by conducting two designed experiments and the results are analyzed to arrive at policy insights. Conclusions and opportunities for future study are provided in Chapter VI.

II. Literature Review

Growing interest in electric powered vehicles has led to extensive research on the topic in both industry and academia. Herein, relevant literature pertaining to the PHEV swap station application and proposed solution approach is discussed. This literature review found no research using an inventory control MDP to model the operations of a PHEV swap station to decide the number of batteries to charge and discharge when factoring in stochastic demand, nonstationary charging costs, and nonstationary revenue from discharging back to the power grid.

The need to examine PHEVs and specifically PHEV swap stations is motivated by a variety of studies. Idaho National Laboratory [18] analyzed the infrastructure requirements for charging of PHEVs in residential settings as well as commercial settings. The report explains that having charging infrastructure available allows the vehicles to require reduced energy storage capability and thus reduces the overall cost of purchasing the vehicles. Transportation system costs can also be reduced by providing rich charging infrastructure rather than using larger batteries to compensate for lesser infrastructure.

Clement-Nys et al. [2] address the issues caused by the increase in demand for large amounts of electrical consumption due to PHEVs. Uncontrolled charging of these batteries in residential areas and charging stations can lead to power losses, reduction in power quality and reliability problems. They use two techniques to model efficient power grid operations, quadratic and dynamic programming. The results from these models indicated that through coordination, which avoids the charging of PHEV batteries during periods of high peak electricity consumption, power quality can be improved and the effects of charging PHEV batteries can be mitigated. Bingliang et al. [3] study the impacts of various charging scenarios on China's power system using data from Shanghai's daily load profile and Monte Carlo simulation. Results from

their study indicate that the level of charging and the increase of charging during high peak hours has a significant effect on the load profile in Shanghai. Investments in battery swapping stations is recommended to control the impact of charging plug-in electric vehicles.

Several approaches have been taken to optimize operations at PHEV swap stations. Worley and Klabjan [19] propose a dynamic programming model to determine the number of batteries for a swap station manager to purchase and the optimal number of batteries to charge at a given point in time. The objective is to minimize the total cost of charging batteries, the opportunity cost of charged batteries that were not used to meet demand, and some penalty defined for unmet demand. The authors do allow for backlogging for unmet past demand. Approximate solutions to the model are obtained by fitting a piecewise linear function to the objective function. The PHEV-SSMP is similar, but it does not look at battery purchasing decisions or backlogging of demand. However, discharging batteries using vehicle to grid (V2G) technology is considered, where this problem does incorporate discharging.

A deterministic integer programming model, considered by Nurre et al. [20], has been used to determine the optimal number of batteries to charge and discharge at a given time. The model presented in this research takes into account a cluster of locations and seeks to optimize operations at multiple swap stations within close proximity to one another such that profit is maximized. In addition to managing the operations at the swap stations to maximize profitability, the authors also examine the impact these policies have on the power grid and seek to minimize the negative impact of wind energy in conjunction with the swap station operations on the power generation curve.

Infrastructure planning of battery swapping has been modeled using robust optimization techniques by Mak et al. [21] for making optimal decisions under limited

and imprecise information. They consider two different objectives; the first focuses on minimizing the expected building and operating costs of the system while the second seeks to maximize a robust estimate of the probability of meeting a return-on-investment target. The decision problem consist of two stages: (1) determining where to locate swap stations with limited information regarding demand, and (2) stocking sufficient number of batteries at each station once uncertain parameters such as demand are observed. Realistic test data is set based on the San Francisco Bay Area freeway network. Results of the two models are similar, suggesting that the two objectives are correlated. Thus, the authors suggest using the retrun-on-investment goal driven model for computational efficiency to produce good solutions for the cost driven model. Finally, they examine how technological advances affect their model and determine that faster recharging technology is critical for increasing profitability.

Tang et al. [22] construct an optimization model seeking to maximize annual profit of electric vehicle battery swap stations that contain photovoltaic power generation. The system they describe has various components that provide charging power including a photovoltaic array which converts solar energy to direct current, and energy storage batteries. These energy storage batteries help regulate and balance the load on the power grid by storing excess generated power and discharging to the system when the system has insufficient generated power.

The adequacy of battery swap stations is assessed by Zhang et al. [23], who examine the ability to have enough fully charged batteries to satisfy battery swapping demands. This is done by analyzing the probability that the amount of fully charged batteries is greater than or equal to the number of electric vehicles that have depleted batteries in any 1 hour interval. They use Monte Carlo simulation over a 10 day period to determine the expected number of electric vehicles that require a battery swap per hour. The results for demand are compared to the current charging plan

to determine the probability that demand does not exceed available supply, which provides valuable insight for charging management and V2G operations.

Plug-in hybrid electric vehicles and PHEV swap stations have been examined in various other contexts as well. Pan et al. [24] present a two-stage stochastic model for locating swap stations with the main objectives being to meet customer demand and reduce variability from renewable technologies on the power grid. The demand for PHEV battery swaps is characterized as a discrete random variable in a transportation network.

Eyer and Corey [4] discuss how increased use of PHEVs can help reduce the significant price difference in electric energy between high peak (on-peak) and low peak (off-peak) prices. At night, energy prices are low because energy use is low. Energy prices are high when energy use is high, which is usually midday on weekdays. If PHEV usage continues to increase, then there will be increased demand for electricity during off-peak periods which will ultimately decrease the price difference and help balance the load on the power grid.

The value of PHEV V2G services on the electricity market is estimated by Sioshansi and Denholm [12], who use a unit commitment model. Vehicle to grid technology allows PHEVs to act as energy storage devices thus reducing energy system operators reliance on generators. V2G services include charging during off-peak hours of demand and discharging during high-peak hours of demand, which is commonly referred to as arbitrage. This has the potential to be beneficial not only to the energy system, but to the PHEV owner as well. By allowing their vehicles to be used as an energy storage device, PHEV owners can earn revenue which will reduce the overall lifetime ownership costs. Sioshansi and Denholm use historical data from the Texas electric power system to analyze the benefit of incorporating V2G technology.

Energy storage problems generally involve balancing power from the grid and

stochastic renewable energy sources such as wind or solar power to smooth energy fluctuations. Energy storage problems are being solved using dynamic programming, approximate dynamic programming methods, and other approximation methods. This is of interest since increased demand for electricity due to PHEVs has a direct impact on energy storage and these problems have similar characteristics to the PHEV-SSMP.

A dynamic programming approach is considered by Sioshansi et al. [25] to approximate the capacity value of energy storage devices. Capacity value is the metric used to quantify a resource's effect on system reliability and is used for resource adequacy planning. Using a deterministic profit-maximization dynamic program they model storage operations that contribute to the capacity of the system. Using historical conventional generator, load, and price data to estimate the capacity value on a single storage device, they show that capacity values are sensitive to energy prices with variability up to 40%.

Salas and Powell [26] research the effectiveness of an approximate dynamic programming (ADP) algorithm for stochastic control of multidimensional energy storage problems. Their work primarily focuses on grid-level storage problems with a finite-horizon. Stochastic elements of their model include wind energy supply, demand for electricity and electricity prices. The ADP resulted in near optimal control policies that were within 1.34% of the optimal solution for a variety of stochastic test problems and within 0.08% for various deterministic test problems.

Several approximate policy iteration methods are examined by Scott et al. [27]. They use the least-squares Bellman error minimization and also discuss direct policy search as an alternative method for approximating complex stochastic systems. Their approximate dynamic programming strategies are used for approximating the value function of a class of energy storage problems that require balancing power from the grid and renewable energy sources. Benchmark problems were used to test the perfor-

mance of the algorithms presented in this work. Bellman error minimization methods provided optimal solutions within 60-80% of the optimal solution while direct policy search results averaged within 90% of the optimal solution. The authors conclude that there are advantages to using direct policy search but recognize limitations for time-dependent applications.

An inventory control MDP is used for the PHEV-SSMP, but they also have a wide variety of other applications. Examples of how inventory systems can be modeled as MDPs are examined to gain insight into various applications. Many authors explore whether or not the optimal policy of their system contains structure, which can be valuable due to ease of implementation and the ability to use algorithms with faster computation time. Structured policies could be monotonic or the commonly used $(\sigma, \sum)$ policy. The curse of dimensionality is often mentioned with MDPs and conventional solution methods (e.g., value iteration, policy iteration), thus it is common to see many of these problems being solved using heuristics and newly developed algorithms to approximate optimal solutions.

Inventory control MDPs have been used to model a wide variety of application areas, with the depth of the literature focusing on supply chains. Giannoccaro and Pontrandolfo [28] model a supply chain management problem, which deals with factors such as suppliers, manufacturers, and distributors. Zhang and Cooper [29] model simultaneous seat-inventory control of multiple flights as a customer-choice MDP that specifically looks at how inventory levels effect the distribution of demand. Yin et al. [30] model an inventory control policy for finished products for a large paper manufacturer with stochastic demand. Lewis [31] examines an inventory control model with risk to supply chain disruptions by looking at an example of an international supply chain with the risk of border closures and congestion. ElHafsi [32] examines an inventory allocation model for an assemble-to-order system with multiple demand

classes as a MDP.

Determining if a problem contains structure can provide valuable insight into a problem. Puterman [16] emphasizes the benefits of optimal policies that contain structure such as monotonicity. These structured policies are significant because of their appeal to decision makers, ease of implementation, and faster computation time. One commonly used structured policy for inventory control is the $(\sigma, \sum)$ policy which indicates to order up to a set value $\sum$ once inventory falls below a set value σ . The concept of the $(\sigma, \sum)$ was first presented by Scarf [33], who denotes it (s, S) . ElHafsi [32] determines the structure of their inventory allocation model using a direct application of value iteration [16], rather than determining an optimal solution due to the complex nature of their problem. Lewis [31] also uses value iteration to find an optimal order-up-to level for the international supply chain model.

In the case where structure is not determined, solution methods must be explored for solving large scale MDPs. Giannoccaro and Pontrandolfo [28] use a reinforcement learning (RL) algorithm and average reward criteria to address some of the major issues in supply chain management, specifically focusing on an inventory ordering policy to maximize performance of the supply chain. When tested on the supply chain management problem, the proposed approach proved to be effective and robust enough to deal with changing demand. Das et al. [34] also propose a RL approach in conjunction with a Semi-Markov average reward technique to solve large scale MDPs. Their algorithm uses RL to solve Semi-MDPs using average expected reward criteria. Semi-MDPs are modeled from sequential decision making problems that have probability structures that are not solely characterized by Markov chains. Using RL has an advantage over the traditional methods of solving MDPs as you do not need to compute probability matrices and reward vectors, but instead use discrete event simulation to build a model. Results from the Semi-Markov average reward technique

algorithm developed by Das et al. [34] was tested on a small scale inventory control model and a larger scaled one which resulted in fast and accurate results. Das et al. [34] use discrete event simulation to build their model due to probability matrices and reward functions being difficult to obtain for large scale MDPs.

Simulation techniques are widely used, especially for larger problems. Zhang and Cooper [29] use simulation techniques to solve the stochastic optimization problem where the demand distribution of customer seat choices is dependent on the state of the system. The highly complex model makes the exact solution very difficult to find, thus the authors derive upper and lower bounds for the value function using simulation techniques and heuristics. Chang et al. [35] suggest using simulation in future research with their adaptive sampling algorithm, which approximates the optimal value of a finite-horizon MDP.

The PHEV-SSMP is solved using the backward induction algorithm [16]. This algorithm finds the optimal policy, or specifically the optimal number of batteries to charge and discharge at each decision epoch which maximizes the expected total reward. Structural properties of the system are examined and it is determined that a nonincreasing monotonic structure is present.

III. Problem Statement

The PHEV-SSMP is solved by determining the optimal number of batteries to charge and discharge over time. This problem is modeled as a Markov decision problem (MDP), with stochastic, nonstationary demand for battery swaps, nonstationary charging costs, and nonstationary revenue from discharging. A finite-horizon, single product inventory control model is considered because the problem data is highly variable with respect to time. Nonstationary Markov decision processes relax the assumption that problem data does not change with time and are in general unsolvable using infinite-horizon models due to infinite data requirements [17]. The nonstationary variable properties in the PHEV-SSMP include demand for battery swaps, charging price for batteries, and revenue from discharging batteries back to the power grid. Motivating the decision which comprises the optimal policy is the maximization of profitability at a single swap station.

Within the MDP model the state space is defined as the total number of batteries that are fully charged. The state of the batteries is modeled at a fundamental level where each battery is either fully charged or depleted. A solution where charging and discharging occur simultaneously can be equivalently represented as solely charging or solely discharging when the discharging revenue is less than or equal to the charging price. Thus, the system is modeled such that charging and discharging never occur simultaneously. If the discharging revenue is greater than the charging price, the simplifying assumption is made that the PHEV station solely charges or solely discharges at any point in time. The swap station may discharge up to the minimum of the total number of batteries that are fully charged and the total number of plug-ins available. In this context, what is denoted a plug-in is the physical entity at a swap station that connects a battery to the power grid thereby allowing it to draw from the power grid (i.e., charge) or discharge using V2G. The total number of plug-ins

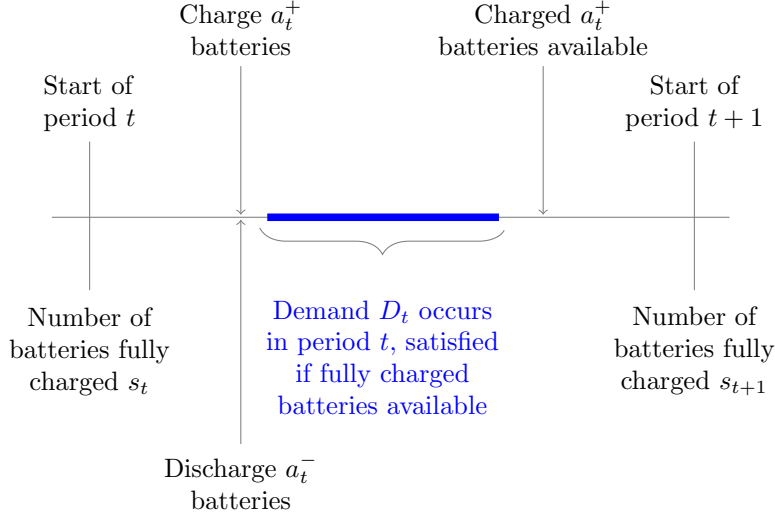


Figure 1. Diagram outlining the timing of events for the PHEV-SSMP MDP model.

or what is denoted charging capacity is assumed constant over time. Similarly, the swap station may charge up to the total number of batteries that are in the depleted state provided that the charging capacity is not exceeded. Thus, the total number of batteries at the swap station is constant over time.

The system is modeled such that batteries charged at time t become full in time $t+1$. Batteries that are discharged take one time period to deplete but are immediately unavailable for exchange. Only fully charged batteries are available for exchange or discharging. Furthermore, batteries that are fully charged are always swapped if available when demand arrives. The cost to charge and revenue from discharging batteries is realized during the time period in which the decision is made. Backlogging of demand is not permitted as it is assumed customers will not wait at the station if batteries are unavailable. The expected reward criterion captures revenue from battery swaps, revenue from discharging batteries back to the power grid through V2G technology, and cost to charge batteries at the swap station. The event timing for the PHEV swap station is outlined in Figure 1.

The MDP for the PHEV-SSMP is mathematically characterized using the following notation:

1. The set of decision epochs¹, $T = \{1, \dots, N-1\}$, $N < \infty$, indicates the discrete time periods in which a decision is made. As previously stated, a finite time horizon is considered due to nonstationary properties.
2. The state of the system at time t , $s_t \in S = \{0, 1, \dots, M\}$ indicates the total number of batteries that are fully charged at decision epoch t , where M is defined as the total number batteries at the swap station, thus $M - s_t$ is the number of depleted batteries at time t .
3. The action at time t , $a_t \in A_{s_t} = \{\max(-s_t, -\Phi), \dots, 0, \dots, \min(M-s_t, \Phi)\}$, $\forall s_t \in S$ indicates the total number of batteries to charge or discharge at time t , where Φ is the charging capacity of the system. A negative action indicates the discharging of batteries and a positive action indicates the charging of batteries. For clarity in the model, the action space is further defined. Let

$$a_t^+ = \begin{cases} a_t & \text{if } a_t \geq 0, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$a_t^- = \begin{cases} |a_t| & \text{if } a_t < 0, \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where a_t^+ is the number of batteries charged and a_t^- is the number of batteries discharged at time t . An assumption of the model is that a_t^+ and a_t^- cannot both be positive during any time interval t .

¹Decision epoch and time period will be used interchangeably throughout this thesis.

4. The immediate reward when action a_t is selected in state s_t at time t which leads to a transition to state s_{t+1} is the profitability of the system, given by

$$r_t(s_t, a_t, s_{t+1}) = \rho(s_t + a_t - s_{t+1}) - K_t(a_t^+) + J_t(a_t^-) \quad (3)$$

for $t = 1, \dots, N - 1$, where $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$, is the number of batteries swapped at time t . Discrete random variable D_t represents the demand for battery swaps at time t , $s_t - a_t^-$ is the number of batteries available for exchange, ρ is the revenue per battery swap, K_t is the charging cost per battery at time t , and J_t is the revenue earned per battery discharged at time t . Specification of K_t and J_t captures the impacts of the nonstationary price for power over time. The terminal reward is calculated as potential swap revenue from fully charged batteries, thus $r_N(s_N) = \rho s_N$.

5. The total number of batteries fully charged at decision epoch $t + 1$ is directly impacted by the batteries charged, discharged, and exchanged during decision epoch t by way of $s_{t+1} = s_t + a_t - \min\{D_t, s_t - a_t^-\}$. The probability of transitioning to state j at time $t + 1$ from state s_t when action a_t is taken, denoted $p_t(j|s_t, a_t)$, is defined by

$$p_t(j|s_t, a_t) = \begin{cases} 0 & \text{if } j > s_t + a_t \text{ or } j < a_t^+ \\ p_{s_t+a_t-j} & \text{if } a_t^+ < j \leq s_t + a_t \\ q_{s_t+a_t-j} & \text{if } j = a_t^+ \end{cases} \quad (4)$$

where $p_j = P(D_t = j)$ and $q_u = \sum_{j=u}^{\infty} p_j = P(D_t \geq u)$. For further clarification, $s_t + a_t - j$ indicates the number of fully charged batteries that are swapped in period t , and $s_t + a_t$ indicates the number of fully charged batteries on hand at the end of the period if none are swapped.

- In the first conditional, state j exceeds the number of fully charged batteries the swap station could possibly have on hand at the end of the period or state j is less than the number of batteries the swap station chooses to charge, which are not available for exchange until after demand is met in that period. In both cases there is a zero transition probability.
- In the second conditional, state j is between the number of batteries the swap station charges and the number of batteries that could possibly be on hand at the end of the period. In this situation, the swap station has enough fully charged batteries to meet demand, hence the probability of transitioning to state j is calculated using the time dependent discrete distribution of demand. It has already been established that j cannot fall below the number of batteries charged in that period, thus the lower bound on j is a_t^+ .
- The last conditional is where $j = a_t^+$, meaning that demand for battery swaps meets or exceeds the supply of fully charged batteries at the beginning of the period. In this situation, the station swaps all batteries on hand but acquires the charged batteries at the end of the period. The transition probability in this case is calculated using the cumulative probability that demand meets or exceeds the number of batteries available for swapping in period t .

To aid the reader, the transition probability function is illustrated using a simple example. Consider the case where there are 15 fully charged batteries (i.e., $s_t = 15$) and the swap station charges 5 (i.e., $a_t = a_t^+ = 5$). If no batteries are swapped the station will have a total of 20 batteries at the end of the period (i.e., $s_{t+1} = j = s_t + a_t = 20$). There is no possible way to have more than $s_t + a_t = 20$ batteries at the end of the period, thus there is a zero transition probability to a state greater

than 20. At the beginning of the period there are $s_t + a_t^- = 15$ batteries available for exchange, thus if all fully charged batteries are swapped, the station still acquires the 5 batteries that were charged at the end of the period. Therefore, the transition probability to a state less than $a_t^+ = 5$ is zero. When $j = a_t^+ = 5$, the 15 batteries that were available at the beginning of the period must have been swapped since the 5 charged batteries are acquired at the end of the period. The transition probability in this case is the probability that demand meets or exceeds $s_t + a_t - j = 15$ batteries, which is captured in the third conditional. Consider the case when the station has 7 batteries at the end of the period (i.e., $j = 7$ which is between a_t^+ and $s_t + a_t$). Since 5 batteries were charged, 2 are remaining from the inventory in the previous period. Since the station started with 15 charged batteries, 13 of them must have been swapped. Thus the transition probability to 7 batteries is the probability that demand for battery swaps was equal to $s_t + a_t - j = 13$.

Having specified the transition probability function, $p_t(j|s_t, a_t)$, the expected immediate reward function can be expressed in terms of the current state and action only, which is more desirable for subsequent calculations.

$$r_t(s_t, a_t) = \sum_{s_{t+1} \in S} \left[p_t(s_{t+1}|s_t, a_t) (\rho(s_t + a_t - s_{t+1})) - K_t(a_t^+) + J_t(a_t^-) \right]. \quad (5)$$

The decision rules are denoted $d_t(s_t)$, which indicate to the decision maker how to select an action $a_t \in A_{s_t}$ at a given decision epoch $t \in T$ when in state $s_t \in S$. Because the decision rules depend on the current state of the system and not the entire history of states, Markovian decision rules [16] are considered. Furthermore, the decision rules prescribe a single specific action and not a probability distribution on the action set. Therefore the decision rules are deterministic. A policy π is a sequence of decision rules $(d_1(s_1), d_2(s_2), \dots, d_{N-1}(s_{N-1}))$ that specify the decision rule to be used at all decision epochs.

The expected total reward of a policy π , when the initial state of the system is s_1 , denoted $v_N^\pi(s_1)$ is given by

$$v_N^\pi(s_1) = \mathbb{E}_{s_1}^\pi \left[\sum_{t=1}^{N-1} r_t(s_t, a_t) + r_N(s_N) \right]. \quad (6)$$

The objective is to determine the policy π^* with the maximum expected total reward. The optimal value function, $u_t^*(s_t)$, denotes the maximum over all policies of the expected total reward from decision epoch t onward when the state at time t is s_t . Optimality equations, or Bellman equations, that correspond to the optimal value functions are used as a basis for determining the optimal policies. The optimality equations are given by

$$u_t(s_t) = \max_{a_t \in A_{s_t}} \left\{ r_t(s_t, a_t) + \sum_{j \in S} p_t(j|s_t, a_t) u_{t+1}(j) \right\} \quad (7)$$

for $t = 1, \dots, N-1$ and $s_t \in S$. For $t = N$, $u_N(s_N) = r_N(s_N)$. It can be shown that if $u_t(s_t)$ is a solution to Equation (7) then the following hold true:

1. $u_t^*(s_t) = u_t(s_t)$ for all $s_t \in S$, $t = 1, \dots, N$, and
2. $v_N^*(s_1) = u_1(s_1)$ for all $s_1 \in S$.

In other words, the optimality equations are indeed optimal and the solution to the optimality equation at $t = 1$ gives the expected total reward for the entire time horizon. Since S is finite and A_{s_t} is finite for each $s_t \in S$, there exists a deterministic Markovian policy which is optimal [16].

3.1 Structural Properties

Determining if the optimal policy of a MDP contains structure, such as monotonicity, is significant due to the ease of implementation, appeal to decision makers, and

the ability for faster computation time [16]. When an optimal policy has a monotone structure, it can be solved with specialized and more efficient algorithms. Thus, it is advantageous to prove that the system contains a nonincreasing monotonic structure.

A policy π is said to be nonincreasing if for each $t = 1, \dots, N - 1$ and any pair of states $s_i, s_j \in S$ with $s_i < s_j$, it is true that $d_t(s_i) \geq d_t(s_j)$. The existence of an optimal nonincreasing monotone policy can be demonstrated using a series of five properties regarding the reward function and the probability of moving to a higher state [16]. Define

$$g_t(k|s_t, a_t) = \sum_{j \in \{S|j \geq k\}} p_t(j|s_t, a_t), \quad t = 1, \dots, N - 1 \quad (8)$$

as the probability of moving to state $j \geq k$ at decision epoch $t + 1$ when action a_t is chosen in state s_t at decision epoch t . Let $A_{s_t} = A'$ for all $s_t \in S$, where $A' = \{\cup_{s_t \in S} A_{s_t}\}$ is the set of all possible actions independent of the state of the system. Note that a function, $f(x, y)$, is said to be subadditive [16] if for $x \geq \tilde{x} \in X$ and $y \geq \tilde{y} \in Y$,

$$f(x, y) + f(\tilde{x}, \tilde{y}) \leq f(x, \tilde{y}) + f(\tilde{x}, y). \quad (9)$$

First, three lemmas are outlined that are utilized in proving that there exists a nonincreasing monotone policy which is optimal.

Lemma 1 *The function $g_t(k|s_t, a_t) =$*

$$\sum_{j=\max\{a_t^++1, k\}}^{s_t+a_t} p_{s_t+a_t-j} + \left[\sum_{i=s_t+a_t-j}^{\infty} p_i \right]_{j=a_t^+}^{j \geq k}. \quad (10)$$

Proof.

$$g_t(k|s_t, a_t) = \sum_{j \in \{S|j \geq k\}} p_t(j|s_t, a_t) \quad (11)$$

$$= \sum_{\substack{j \geq k \\ a_t^+ < j \leq s_t + a_t}} p_{s_t + a_t - j} + \left[q_{s_t + a_t - j} \right]_{j=a_t^+}^{j \geq k} \quad (12)$$

$$= \sum_{j=\max\{a_t^+ + 1, k\}}^{s_t + a_t} p_{s_t + a_t - j} + \left[\sum_{i=s_t + a_t - j}^{\infty} p_i \right]_{j=a_t^+}^{j \geq k} \quad (13)$$

□

Lemma 2 *The following two summations are equivalent*

$$\sum_{j=k}^{s_t + a_t} p_{s_t + a_t - j} = \sum_{i=0}^{s_t + a_t - k} p_i. \quad (14)$$

Proof.

$$\sum_{j=k}^{s_t + a_t} p_{s_t + a_t - j} = p_{s_t + a_t - k} + p_{s_t + a_t - (k+1)} + \dots + p_{s_t + a_t - (s_t + a_t)} \quad (15)$$

$$= p_{s_t + a_t - k} + p_{s_t + a_t - (k+1)} + \dots + p_0 = \sum_{i=0}^{s_t + a_t - k} p_i \quad (16)$$

□

Lemma 3 *The following two summations are equivalent*

$$\sum_{j=a_t^+ + 1}^{s_t + a_t} p_{s_t + a_t - j} + \sum_{i=s_t + a_t - a_t^+}^{\infty} p_i = \sum_{i=0}^{\infty} p_i. \quad (17)$$

Proof.

$$\sum_{j=a_t^++1}^{s_t+a_t} p_{s_t+a_t-j^+} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i \quad (18)$$

$$= p_{s_t+a_t-(a_t^++1)} + p_{s_t+a_t-(a_t^++2)} + \dots + p_{s_t+a_t-(s_t+a_t)} + \sum_{i=s_t-a_t^-}^{\infty} p_i \quad (19)$$

$$= p_{s_t-a_t^--1} + p_{s_t-a_t^--2} + \dots + p_0 + \sum_{i=s_t-a_t^-}^{\infty} p_i \quad (20)$$

$$= \sum_{i=0}^{s_t-a_t^--1} p_i + \sum_{i=s_t-a_t^-}^{\infty} p_i = \sum_{i=0}^{\infty} p_i \quad (21)$$

□

Utilizing these lemmas, the existence of a nonincreasing monotone policy is proven, which is outlined in Theorem 1.

Theorem 1 *There exists optimal decision rules $d_t^*(s_t)$ for the PHEV-SSMP which are nonincreasing in s_t for $t = 1, \dots, N-1$ when demand D_t is governed by a non-increasing discrete distribution.*

Proof. The claim is shown by demonstrating that the PHEV-SSMP exhibits the following 5 conditions [16].

1. $r_t(s_t, a_t)$ is nondecreasing in s_t for all $a_t \in A'$.

That $r_t(s_t, a_t)$ is nondecreasing in s_t for a fixed a_t means that for a fixed action (i.e., number of batteries charged or discharged), the expected immediate reward will be greater when the number of full batteries is greater. This coincides with intuition as more batteries can be swapped or discharged when there are more full batteries available thereby leading to more reward. Consider $s_t \geq \tilde{s}_t$, using $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$ for any value which D_t can assume. It can be

shown that

$$r_t(s_t, a_t) \geq r_t(\tilde{s}_t, a_t). \quad (22)$$

The expected immediate reward function can be expressed as

$$r_t(s_t, a_t) = \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - a_t^-\}) - K_t(a_t^+) + J_t(a_t^-) \right]. \quad (23)$$

Therefore, it can be show that

$$r_t(s_t, a_t) \geq r_t(\tilde{s}_t, a_t) \Leftrightarrow \quad (24)$$

$$\begin{aligned} & \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - a_t^-\}) - K_t(a_t^+) + J_t(a_t^-) \right] \\ & \geq \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - a_t^-\}) - K_t(a_t^+) + J_t(a_t^-) \right] \Leftrightarrow \end{aligned} \quad (25)$$

$$\sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - a_t^-\}) \right] \geq \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - a_t^-\}) \right]. \quad (26)$$

Therefore, since $P(D_t = j)\rho$ is multiplied by both sides for all values of j , the above can be reduced to

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\}, \quad (27)$$

for all possible values of j . Using a proof by cases, the three possible cases of demand $D_t = j$ with respect to $s_t - a_t^-$ and $\tilde{s}_t - a_t^-$ are considered: (a) $j \leq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$, (b) $j \geq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$, and (c) $j \geq \tilde{s}_t - a_t^-$, $j \geq s_t - a_t^-$. The case where j is greater than $s_t - a_t^-$ and less than $\tilde{s}_t - a_t^-$ does

not need to be considered because it is not possible since $s_t \geq \tilde{s}_t$. In each case, Equation (27) is reduced to a valid statement.

$$(a) \quad j \leq \tilde{s}_t - a_t^-, j \leq s_t - a_t^-$$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow j = j \quad (28)$$

$$(b) \quad j \geq \tilde{s}_t - a_t^-, j \leq s_t - a_t^-$$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow j \geq \tilde{s}_t - a_t^- \quad (29)$$

$$(c) \quad j \geq \tilde{s}_t - a_t^-, j \geq s_t - a_t^-$$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow s_t - a_t^- \geq \tilde{s}_t - a_t^- \Leftrightarrow s_t \geq \tilde{s}_t \quad (30)$$

2. $g_t(k|s_t, a_t)$ is nondecreasing in s_t for all $k \in S$ and $a_t \in A'$.

That $g_t(k|s_t, a_t)$ is nondecreasing in s_t for a fixed a_t and k means that the probability that the number of full batteries in the next state is greater than some threshold k is greater when the number of full batteries in the current state is greater. Consider $s_t \geq \tilde{s}_t$, it can be shown that

$$g_t(k|s_t, a_t) \geq g_t(k|\tilde{s}_t, a_t) \Leftrightarrow \quad (31)$$

$$\sum_{j \in \{S|j \geq k\}} p_t(j|s_t, a_t) \geq \sum_{j \in \{S|j \geq k\}} p_t(j|\tilde{s}_t, a_t) \Leftrightarrow \quad (32)$$

$$\begin{aligned} \sum_{\substack{j \geq k \\ a_t^+ < j \leq s_t + a_t}} p_{s_t + a_t - j} + \left[q_{s_t + a_t - j} \right]_{\substack{j \geq k \\ j = a_t^+}} &\geq \\ \sum_{\substack{j \geq k \\ a_t^+ < j \leq \tilde{s}_t + a_t}} p_{\tilde{s}_t + a_t - j} + \left[q_{\tilde{s}_t + a_t - j} \right]_{\substack{j \geq k \\ j = a_t^+}} &\Leftrightarrow \end{aligned} \quad (33)$$

$$\begin{aligned}
\sum_{j=\max\{a_t^++1,k\}}^{s_t+a_t} p_{s_t+a_t-j} + \left[\sum_{i=s_t+a_t-j}^{\infty} p_i \right]_{j \geq k, j=a_t^+} &\geq \\
\sum_{j=\max\{a_t^++1,k\}}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \left[\sum_{i=\tilde{s}_t+a_t-j}^{\infty} p_i \right]_{j \geq k, j=a_t^+} &. \quad (34)
\end{aligned}$$

Using a proof by cases, all cases of k with respect to a_t are considered. For each case, Equation (34) is reduced to a valid statement. Note that the second term of both the left hand side and right hand side of Equation (34) is only included when both $j \geq k$ and $j = a_t^+$, which represents when demand meets or exceeds supply.

$$(a) \quad a_t^+ \geq k \Rightarrow a_t^+ + 1 > k$$

The second term of each summation appears as both $j \geq k$ and $j = a_t^+$ are satisfied. Using Lemma 3, Equation (35) is reduced to Equation (36).

$$\sum_{j=a_t^++1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i \geq \sum_{j=a_t^++1}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \sum_{i=\tilde{s}_t+a_t-a_t^+}^{\infty} p_i \Leftrightarrow \quad (35)$$

$$\sum_{i=0}^{\infty} p_i = \sum_{i=0}^{\infty} p_i \quad (36)$$

$$(b) \quad a_t^+ < k \Rightarrow a_t^+ + 1 \geq k$$

The second term of each summation does not appear as $j = a^+$ will never be satisfied. Starting from Equation (34), Lemma 2 is utilized to arrive at a known valid statement.

$$\sum_{j=k}^{s_t+a_t} p_{s_t+a_t-j} \geq \sum_{j=k}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} \Leftrightarrow \quad (37)$$

$$\sum_{i=0}^{s_t+a_t-k} p_i \geq \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \quad (38)$$

$$\sum_{i=0}^{\tilde{s}_t+a_t-k} p_i + \sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \geq \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \quad (39)$$

$$\sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \geq 0 \quad (40)$$

3. $r_t(s_t, a_t)$ is a subadditive function on $S \times A'$.

The subadditivity of $r_t(s_t, a_t)$ implies that the incremental effect on the expected total reward of charging less batteries (or discharging more batteries) is less when the number of full batteries is greater. Consider $a_t \geq \tilde{a}_t$ and $s_t \geq \tilde{s}_t$, using $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$ for any value which D_t can assume. It can be shown that

$$r_t(s_t, a_t) + r_t(\tilde{s}_t, \tilde{a}_t) \leq r_t(s_t, \tilde{a}_t) + r_t(\tilde{s}_t, a_t) \Leftrightarrow \quad (41)$$

$$\begin{aligned} & \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - a_t^-\}) - K_t(a_t^+) + J_t(a_t^-) \right] \\ & + \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - \tilde{a}_t^-\}) - K_t(\tilde{a}_t^+) + J_t(\tilde{a}_t^-) \right] \\ & \leq \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - \tilde{a}_t^-\}) - K_t(\tilde{a}_t^+) + J_t(\tilde{a}_t^-) \right] \\ & + \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - a_t^-\}) - K_t(a_t^+) + J_t(a_t^-) \right] \Leftrightarrow \quad (42) \\ & \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - a_t^-\}) \right] + \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - \tilde{a}_t^-\}) \right] \\ & \leq \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, s_t - \tilde{a}_t^-\}) \right] + \sum_{j=0}^{\infty} \left[P(D_t = j) (\rho \min\{j, \tilde{s}_t - a_t^-\}) \right]. \quad (43) \end{aligned}$$

Therefore, since $P(D_t = j)\rho$ is multiplied by all terms, the above can be reduced

to

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\}, \quad (44)$$

for all values of j . Using a proof by cases, every relevant case of a_t and $\tilde{a}_t$, and each scenario for demand $D_t = j$ with respect to $s_t - a_t^-$, $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, $\tilde{s}_t - a_t^-$ are considered. The case where $\tilde{a}_t \leq 0$ and $a_t \geq 0$ is excluded as this is not possible from the definition of subadditivity that $a_t \geq \tilde{a}_t$. For each case, Equation (44) is reduced down to a valid statement.

$$(a) \quad \tilde{a}_t \geq 0, a_t \geq 0 \Rightarrow \tilde{a}_t^- = a_t^- = 0$$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \quad (45)$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t\} = \min\{j, s_t\} + \min\{j, \tilde{s}_t\} \quad (46)$$

$$(b) \quad \tilde{a}_t \leq 0, a_t \geq 0 \Rightarrow \tilde{a}_t^- \geq 0, a_t^- = 0$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \quad (47)$$

Every possibility for demand j with respect to s_t , $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, and $\tilde{s}_t$ is considered. Figure 2 is provided to aid the reader in visualizing the six possible scenarios. The ranges i-vi in the diagram correspond to the following scenarios i-vi.

$$i. \quad j \leq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq \tilde{s}_t, j \leq s_t - \tilde{a}_t^-, j \leq s_t$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (48)$$

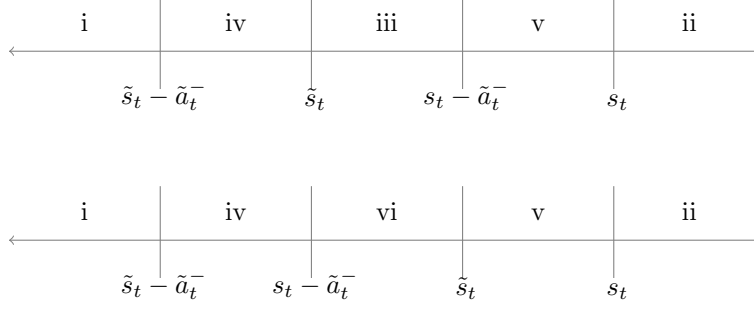


Figure 2. Scenarios of demand with respect to inventory for case (b).

$$j + j \leq j + j \Leftrightarrow 2j = 2j \quad (49)$$

$$\text{ii. } j \geq s_t \Rightarrow j \geq s_t - \tilde{a}_t^-, j \geq \tilde{s}_t, j \geq \tilde{s}_t - \tilde{a}_t^-$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (50)$$

$$s_t + \tilde{s}_t - \tilde{a}_t^- = s_t - \tilde{a}_t^- + \tilde{s}_t \quad (51)$$

$$\text{iii. } j \geq \tilde{s}_t, j \leq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-, j \leq s_t$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (52)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + \tilde{s}_t \Leftrightarrow \tilde{a}_t^- \geq 0 \quad (53)$$

$$\text{iv. } j \leq \tilde{s}_t, j \leq s_t - \tilde{a}_t^-, j \geq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq s_t$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (54)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + j \Leftrightarrow \tilde{s}_t - \tilde{a}_t^- \leq j \quad (55)$$

$$\text{v. } j \geq \tilde{s}_t, j \geq s_t - \tilde{a}_t^-, j \leq s_t \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (56)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + \tilde{s}_t \Leftrightarrow j \leq s_t \quad (57)$$

$$\text{vi. } j \leq \tilde{s}_t, j \geq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-, j \leq s_t$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \quad (58)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + j \Leftrightarrow \tilde{s}_t \leq s_t \quad (59)$$

$$(c) \quad \tilde{a}_t \leq 0, a_t \leq 0 \Rightarrow \tilde{a}_t^- \geq 0, a_t^- \geq 0, \tilde{a}_t^- \geq a_t^-$$

Every possibility for demand j with respect to $s_t - a_t^-$, $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, $\tilde{s}_t - a_t^-$ is considered. Figure 3 is provided to aid the reader in visualizing the six possible scenarios. The ranges i-vi in the diagram correspond to the following scenarios i-vi.

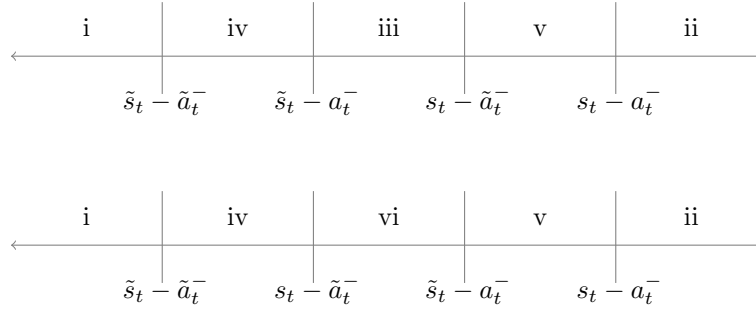


Figure 3. Scenarios of demand with respect to inventory for case (c).

$$\text{i. } j \leq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq \tilde{s}_t - a_t^-, j \leq s_t - \tilde{a}_t^-, j \leq s_t - a_t^-$$

$$\begin{aligned} \min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \\ \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \end{aligned} \quad (60)$$

$$j + j \leq j + j \Leftrightarrow 2j = 2j \quad (61)$$

$$\text{ii. } j \geq s_t - a_t^- \Rightarrow j \geq \tilde{s}_t - a_t^-, j \geq s_t - \tilde{a}_t^-, j \geq \tilde{s}_t - \tilde{a}_t^-$$

$$\begin{aligned} & \min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \\ & \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \end{aligned} \quad (62)$$

$$s_t - a_t^- + \tilde{s}_t - \tilde{a}_t^- = s_t - \tilde{a}_t^- + \tilde{s}_t - a_t^- \quad (63)$$

$$\text{iii. } j \geq \tilde{s}_t - a_t^-, j \leq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-, j \leq s_t - a_t^-$$

$$\begin{aligned} & \min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \\ & \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \end{aligned} \quad (64)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + \tilde{s}_t - a_t^- \Leftrightarrow \tilde{a}_t^- \geq a_t^- \quad (65)$$

$$\text{iv. } j \leq \tilde{s}_t - a_t^-, j \leq s_t - \tilde{a}_t^-, j \geq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq s_t - a_t^-$$

$$\begin{aligned} & \min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \\ & \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \end{aligned} \quad (66)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + j \Leftrightarrow \tilde{s}_t - \tilde{a}_t^- \leq j \quad (67)$$

$$\text{v. } j \geq \tilde{s}_t - a_t^-, j \geq s_t - \tilde{a}_t^-, j \leq s_t - a_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-$$

$$\begin{aligned} & \min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \\ & \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \end{aligned} \quad (68)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + \tilde{s}_t - a_t^- \Leftrightarrow j \leq s_t - a_t^- \quad (69)$$

$$\text{vi. } j \leq \tilde{s}_t - a_t^-, j \geq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-, j \leq s_t - a_t^-$$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$

$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \quad (70)$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + j \Leftrightarrow \tilde{s}_t \leq s_t \quad (71)$$

4. $g_t(k|s_t, a_t)$ is a subadditive function on $S \times A'$ for all $k \in S$.

The subadditivity of $g_t(k|s_t, a_t)$ implies that the incremental effect of charging less batteries (or discharging more batteries) on the probability that the system moves to a state of full batteries above some threshold k is less when the number of full batteries is greater. Consider $a_t \geq \tilde{a}_t$ and $s_t \geq \tilde{s}_t$, it can be shown that

$$g_t(k|s_t, a_t) + g_t(k|\tilde{s}_t, \tilde{a}_t) \leq g_t(k|s_t, \tilde{a}_t) + g_t(k|\tilde{s}_t, a_t) \Leftrightarrow \quad (72)$$

$$\begin{aligned} & \sum_{j=\max\{a_t^++1, k\}}^{s_t+a_t} p_{s_t+a_t-j} + \left[\sum_{i=s_t+a_t-j}^{\infty} p_i \right]_{j=a_t^+}^{j \geq k} \\ & + \sum_{j=\max\{\tilde{a}_t^++1, k\}}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j} + \left[\sum_{i=\tilde{s}_t+\tilde{a}_t-j}^{\infty} p_i \right]_{j=\tilde{a}_t^+}^{j \geq k} \\ & \leq \sum_{j=\max\{\tilde{a}_t^++1, k\}}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \left[\sum_{i=s_t+\tilde{a}_t-j}^{\infty} p_i \right]_{j=\tilde{a}_t^+}^{j \geq k} \\ & + \sum_{j=\max\{a_t^++1, k\}}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \left[\sum_{i=\tilde{s}_t+a_t-j}^{\infty} p_i \right]_{j=a_t^+}^{j \geq k}. \end{aligned} \quad (73)$$

Using a proof by cases, every relevant case of k with respect to a_t and $\tilde{a}_t$ is considered. For each case, Equation (73) is reduced to a valid statement. The function $g_t(k|s_t, a_t)$ is comprised of two terms. The first term calculates the probability when demand never exceeds supply of batteries and the second calculates the cumulative probability that demand equals or exceeds supply. It is indicated in each case of the proof which of the terms are included in the summation based on the relationship between k , a_t , and $\tilde{a}_t$.

$$(a) \quad \tilde{a}_t^+ \geq k \Rightarrow a_t^+ \geq k, \tilde{a}_t^+ + 1 > k, a_t^+ + 1 > k$$

For this case demand for battery swaps may exceed supply, therefore both terms of $g_t(k|s_t, a_t)$ appear.

$$\begin{aligned} & \sum_{j=a_t^++1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i + \sum_{j=\tilde{a}_t^++1}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j} + \sum_{i=\tilde{s}_t+\tilde{a}_t-\tilde{a}_t^+}^{\infty} p_i \\ \leq & \sum_{j=\tilde{a}_t^++1}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \sum_{i=s_t+\tilde{a}_t-\tilde{a}_t^+}^{\infty} p_i + \sum_{j=a_t^++1}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \sum_{i=\tilde{s}_t+a_t-a_t^+}^{\infty} p_i \Leftrightarrow \end{aligned} \quad (74)$$

$$\sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\infty} p_i \leq \sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\infty} p_i \Leftrightarrow \quad (75)$$

$$2 \sum_{i=0}^{\infty} p_i = 2 \sum_{i=0}^{\infty} p_i \quad (76)$$

$$(b) \quad \tilde{a}_t^+ < k, a_t^+ \geq k \Rightarrow \tilde{a}_t^+ + 1 \leq k, a_t^+ + 1 > k$$

For this case, because $a_t^+ \geq k$, the second term of $g(k|s_t, a_t)$ does appear when action a_t is taken as demand can exceed supply. However, because $\tilde{a}_t^+ < k$, demand can never exceed supply when action $\tilde{a}_t$ is taken.

$$\begin{aligned} & \sum_{j=a_t^++1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i + \sum_{j=k}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j} \\ \leq & \sum_{j=k}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \sum_{j=a_t^++1}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \sum_{i=\tilde{s}_t+a_t-a_t^+}^{\infty} p_i \Leftrightarrow \end{aligned} \quad (77)$$

$$\sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\infty} p_i \Leftrightarrow \quad (78)$$

$$\sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i \Leftrightarrow \quad (79)$$

$$\sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i + \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i \Leftrightarrow \quad (80)$$

$$0 \leq \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i \quad (81)$$

$$(c) \ a_t^+ < k \Rightarrow \tilde{a}_t^+ < k, \tilde{a}_t^+ + 1 \leq k, a_t^+ + 1 \leq k$$

For this case, demand for battery swaps never exceeds supply therefore, the second term of $g_t(k|s_t, a_t)$ does not appear when either action a_t or action $\tilde{a}_t$ are taken.

$$\sum_{j=k}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{j=k}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j} \leq \sum_{j=k}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \sum_{j=k}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} \Leftrightarrow \quad (82)$$

$$\sum_{i=0}^{s_t+a_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \quad (83)$$

$$\begin{aligned} \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i + \sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \\ \leq \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i + \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \end{aligned} \quad (84)$$

$$\sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \leq \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i. \quad (85)$$

In Equation (85) the number of terms on each side are exactly the same, however because $a_t \geq \tilde{a}_t$ the start of the summation is greater on the left hand side. Therefore, Equation (85) holds when $p_j = P(D_t = j)$ is governed by a nonincreasing discrete distribution.

5. $r_N(s_N)$ is nondecreasing in s_N .

Consider $s_N \geq \tilde{s}_N$, it can be shown that $r_N(s_N) \geq r_N(\tilde{s}_N)$. This expression is reduced to a known valid statement.

$$r_N(s_N) \geq r_N(\tilde{s}_N) \Leftrightarrow \rho s_N \geq \rho \tilde{s}_N \Leftrightarrow s_N \geq \tilde{s}_N \quad (86)$$

□

Consider two possibilities for the state (i.e., number of full batteries) at a swap station $s_t \geq \tilde{s}_t$. This theorem states that there exists an optimal decision rule where the swap station will never charge less (or discharge more) batteries in state $\tilde{s}_t$ as compared to s_t . Utilizing this result, exact solution methods and two benchmark solution methods are outlined.

IV. Methodology

The objective in solving this Markov decision problem (MDP) is to determine a policy that maximizes the expected total reward criterion expressed in Equation (6). The set of states, S , is finite and the action set, A_{s_t} , is finite for each $s_t \in S$. Therefore there exists a deterministic Markov policy which is optimal. An optimal policy for this finite-horizon model is found using the backward induction algorithm [16]. This dynamic programming algorithm finds the optimal policy, or specifically the optimal number of batteries to charge and discharge at each decision epoch which maximizes the expected total reward. The backward induction algorithm finds sets $A_{s_t, t}^*$ which contain all actions in A_{s_t} which attain the maximum for the optimality equations (7). The algorithm also evaluates the policy and computes the expected total reward from each period to the end of the decision making horizon.

There exists an optimal policy that contains a nonincreasing monotonic structure when demand is governed by a discrete nonincreasing distribution, thus the monotone backward induction algorithm [16] is also used to find an optimal policy, which is outlined in Algorithm 1. The nonincreasing monotone backward induction algorithm modifies the original algorithm by redefining the action set at each iteration of s_t to be limited by the optimal decision rule of $s_t - 1$ for each $t \in T$. For example, if the optimal decision rule at $s_t = 10$ is to charge 20 batteries, then the action space for $s_t = 11$ will now be $A_{11} = \{\max(-11, -\Phi), \dots, 0, \dots, \min(20, \Phi)\}$ instead of $A_{11} = \{\max(-11, -\Phi), \dots, 0, \dots, \min(M - 11, \Phi)\}$. The modifications to the algorithm will result in an optimal policy when demand is governed by a discrete nonincreasing distribution; note however, that there may be alternative optima that are not monotone.

When there are $|S|$ states, $|A'|$ actions in each state where $A' = \{\cup_{s_t \in S} A_{s_t}\}$, and N time periods, the backward induction algorithm requires $(N - 1)|A'| |S|^2$ multipli-

Algorithm 1 Nonincreasing Monotone Backward Induction [16]

- 1: Set $t = N$ and $u_N(s_N) = r_N(s_N)$ for all $s_N \in S$
- 2: **while** $t > 1$ **do**
- 3: Set $t = t - 1$, set $s_t = 0$ and set $A'_{s_t} = A'_0 = A'$
- 4: **while** $s_t \leq M$ **do**
- 5: Compute $u_t^*(s_t)$ by

$$u_t^*(s_t) = \max_{a_t \in A'_{s_t}} \left\{ r_t(s_t, a_t) + \sum_{j \in S} p_t(j|s_t, a_t) u_{t+1}^*(j) \right\}$$

- 6: Set action that results in $u_t^*(s_t)$

$$A_{s_t, t}^* = \arg \max_{a_t \in A'_{s_t}} \left\{ r_t(s_t, a_t) + \sum_{j \in S} p_t(j|s_t, a_t) u_{t+1}^*(j) \right\}$$

- 7: **if** $s_t < M$ **then**
- 8: Define action space for $s_t + 1$ by

$$A'_{s_t+1} = \{a \in A'_{s_t} : a \leq \min\{a' \in A_{s_t, t}^*\}\}$$

- 9: **end if**
 - 10: Set $s_t = s_t + 1$
 - 11: **end while**
 - 12: **end while**
 - 13: Calculate expected total reward for entire horizon, $v_N^*(s_1) = u_1^*(s_1)$
-

cations to determine the optimal policy, which is a considerable improvement from complete enumeration of all possible solutions, which takes $(|A'|^{|S|})^{(N-1)}(N-1)|S|^2$ multiplications. In the worst case scenario, the monotone backward induction algorithm's computational effort equals that of the backward induction, however when the policy is nonincreasing the action sets decrease in size with increasing s_t and reduce the number of actions that need to be evaluated [16].

4.1 Benchmark Policies

Two benchmark policies are considered such that the swap station charges up to or discharges down to a set target level ζ_t , at each decision epoch t . The first benchmark policy is a stationary benchmark policy which picks a set target level ζ and sets $\zeta_t = \zeta$ for all time periods t . The second is a dynamic benchmark policy and utilizes a distinct ζ_t for each time period t . Utilizing each target level, the policy can be determined by calculating the action for each state and time period with a simple calculation. Thus, this policy can be easily implemented by a swap station manager.

If the state s_t is less than or equal to the target level ζ_t , the swap station does not have as many fully charged batteries as desired, thus they will charge or do nothing. The most that can be charged at any point in time, denoted C , is given by

$$C = \min\{M - s_t, \Phi\}. \quad (87)$$

If s_t is greater than ζ_t the swap station has more fully charged batteries than desired, thus they will discharge. The most that can be discharged at any point in time (i.e., the most negative action), denoted D , is given by

$$D = \max\{-s_t, -\Phi\}. \quad (88)$$

The decision rule $d_t(s_t)$ is given by the following.

$$d_t(s_t) = \begin{cases} \min\{\zeta_t - s_t, C\} & \text{if } s_t \leq \zeta_t \\ \max\{\zeta_t - s_t, D\} & \text{if } s_t > \zeta_t \end{cases} \quad (89)$$

For the first benchmark policy, a stationary target level $\zeta_t = \zeta$ is derived, where ζ is calculated as a percentage of the number of batteries M using some constant

$\mathcal{C}$. Equation (90) calculates ζ using a traditional rounding function. In the second benchmark policy, dynamic target levels ζ_t are derived at each decision epoch as a rounded function of the number of batteries M and charging costs K_t using Equation (91) for constants $\mathcal{C}_\ell, \mathcal{C}_u$ where $\mathcal{C}_u > \mathcal{C}_\ell$.

$$\zeta = \lfloor \mathcal{C}M + 0.5 \rfloor \quad (90)$$

$$\zeta_t = \begin{cases} \lfloor \mathcal{C}_\ell M + 0.5 \rfloor & \text{if } K_t > K_{t+1} \\ \lfloor \mathcal{C}_u M + 0.5 \rfloor & \text{if } K_t \leq K_{t+1} \end{cases} \quad \forall t = 1, \dots, N-1 \quad (91)$$

These policies are validated in Chapter V as usable for real time decision making activities due to their speed of calculation and accuracy.

V. Computational Tests

In this chapter, realistic data is used to computationally test the PHEV-SSMP on a variety of different scenarios. From the optimal policies, insights that would be beneficial to a swap station manager are deduced. Two Latin hypercube designed experiments are also used to gain insights with a focus on the expected total profit and percentage of demand that is met when the optimal policy is implemented. Further, the accuracy and speed of the two benchmark policies is compared to the optimal policy and optimal solution method.

The time horizon examined is a full week in one hour increments, thus the time horizon is $N = (24)(7) + 1 = 169$ and the number of decision epochs is $N - 1 = 168$. The first decision is made on Monday at 0000, the second on Monday at 0100 until the last decision is made on Sunday at 2300. Historical hourly charging cost data from 2013 in the Capital Region, New York is utilized, which is obtained from National Grid [36]. One week from each season is used in this analysis due to the varying climate and drastic variation in prices throughout the year. January 21-27 is used for Winter, April 15-21 for Spring, July 15-21 for Summer, and September 23-29 for Fall. Note that the sum of power prices over every hour of the week is at the maximum for January 21-27 and at a minimum for September 23-29 for 2013. The charging cost per kWh at each time t is multiplied by 60 to calculate the cost to charge one battery K_t , which is consistent with the Tesla Model S 60 kWh battery option [37] and can be completed in an hour with level 2 or 3 charging [18]. The charging cost per battery per hour for the four weeks of interest can be seen in Figure 4. For these computational tests, the discharge revenue J_t , is set equal to a percentage of the charging cost, $J_t = \alpha K_t$ using α between 0.75 and 1.25. The α parameter will give insight into the incentives needed to be placed on the swap station to encourage discharging at favorable points in time.

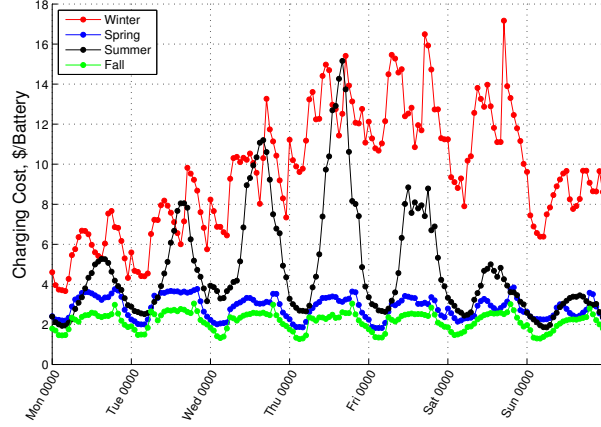


Figure 4. Charging cost K_t per battery per hour in the Capital Region, NY.

A similar methodology is considered to derive the distribution for swap demand at each hour as Nurre et al. [20]. The authors assume that the behaviors for arrivals at a swap station will mimic the currently observed behaviors at a gas station. As such, they calculate the percentage of people who will frequent a gas station for each hour of a day and day of a week based on historical data at Chevron gas stations [38], assuming a customer visits a gas station once per week. This percentage is utilized to calculate the mean arrival rate of customers $\bar{X}_t$, for each decision epoch t . Specifically, the PHEV-SSMP considers an area with γ PHEV users and sets $\bar{X}_t$ equal to the product of γ and the percentage of customers visiting the station at time t from Nurre et al. [20].

Two distributions for modeling swap demand D_t are considered, geometric and Poisson. When swap demand D_t follows a geometric distribution, parameter $\mathcal{P}_t$ is set to $\frac{1}{\bar{X}_t+1}$. When swap demand D_t follows a Poisson distribution, parameter λ_t is set to $\bar{X}_t$. Note that the geometric distribution is a nonincreasing discrete distribution, therefore a monotonic nonincreasing policy is optimal. The mean arrival rate of customers $\bar{X}_t = \lambda_t$ for each hour of each day in a location with $\gamma = 3,000$ PHEVs can be seen in Figure 5. The arrival rate of customers is assumed the same for each

week of the year.

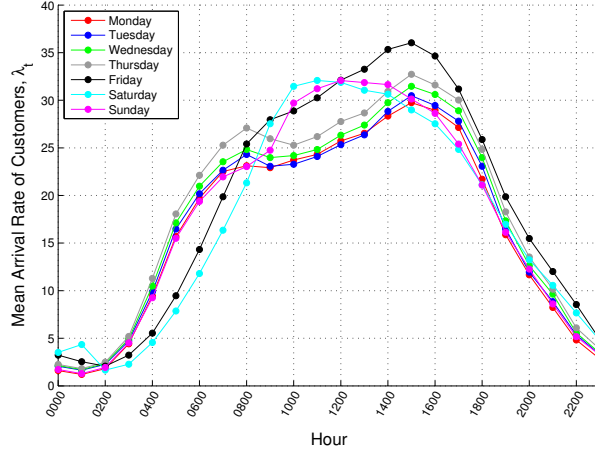


Figure 5. Mean arrival rate of customers λ_t in a location with 3,000 PHEVs by time of day and day of the week.

To computationally test the PHEV-SMMP, two designed experiments are conducted. The first designed experiment is used to gain general insights when a wide range of inputs are considered. The second designed experiment is conducted with more targeted values based on the results of the first experiment. With this second experiment, values for the controllable parameters at a swap station are able to be determined. With both, the expected total reward, percentage of met demand, and policies are utilized to infer valuable policy insights.

For the first designed experiment the expected total reward is used as the response variable. This is found using the monotone nonincreasing backward induction algorithm [16] when demand follows the geometric distribution. When demand follows a time dependent Poisson process, two policies with corresponding expected total rewards are found: the optimal policy is found using the backward induction algorithm, and a heuristic policy is found using the monotone backward induction algorithm. Note that the monotone policy is not always optimal, however empirically it has been verified to be optimal in almost all cases.

A 50-scenario Latin hypercube designed experiment is preformed, which is a widely used design for deterministic computer simulation models [39]. This space filling design spreads the design points nearly uniformly to better characterize the response surface in the region of experimentation. Because four separate weeks for charging cost data is considered, K_t is a categorical factor with four levels representing the four weeks extracted from the year. The 50-scenario design is conducted for each of the four seasons and each of the two demand distributions, resulting in a total of 400 scenarios. Factors that are used in the design include the total number of batteries M , the charging capacity Φ , the total number of PHEVs in the local area γ , the revenue per battery swap ρ , and the percentage of revenue earned from discharging with respect to the charging cost α . Using JMP11Pro software, a 50-scenario design is generated with various levels of each factor ranging between two values. The high and low levels used for this experiment can be seen in Table 1. The charging costs for the four weeks of interest, $K_t^W, K_t^{Sp}, K_t^{Su}$, and K_t^F , are representative of Winter, Spring, Summer, and Fall, respectively. The low value for the swap revenue ρ , is set less than the minimum charging cost over the four weeks and the high value for ρ is set greater than the maximum charging cost.

Table 1. Factor levels used for the first Latin hypercube designed experiment.

Factor		Low	High
Total Number of Batteries	M	50	200
Charging Capacity	Φ	$\lfloor 0.25M \rfloor$	M
Swap Revenue (\$)	ρ	1	20
Percent Discharge Revenue ($\%K_t$)	α	0.75	1.25
PHEVs in the Local Area	γ	1,000	6,000

When considering the time dependent Poisson process for demand, the monotone policy was optimal in all but 22 scenarios. Of these 22 scenarios, the largest percentage gap in expected total reward when compared to optimal was 0.77%. Therefore,

while the monotone policy is not always optimal when demand does not follow a non-increasing distribution, it is empirically observed to provide a good approximation. Further, very similar optimal policies are seen when using the Poisson and geometric distributions to model demand. Only 41 scenarios resulted in a different expected total reward with the largest gap being 2.7%. Discharging is often favored when demand follows a Poisson process, however discharging does occur when demand is governed by a geometric distribution. Due to the similarities seen, the results presented herein apply to both distributions unless otherwise stated.

Results from the designed experiment indicate that all factors have a significant effect on the expected total reward at a 95% confidence level, except for the charging costs K_t . This indicates that even though there is a drastic variation in seasonal charging prices, it does not affect the swap station's profit. As expected, the swap revenue ρ , has the greatest impact on the expected total reward. Thus, the most effective way to increase the expected total reward would be to increase the swap cost, however this is based on the assumption that demand for swaps is independent of the swap cost which is unrealistic. Future work should consider the sensitivity of customers to the price for swapping as utilizing a charging station can occur instead of swapping. Next, the significant interaction terms with M are examined: $M\Phi$, $M\rho$, and $M\alpha$. The interaction plots produced by JMP11Pro Software for the second order terms can be seen in Figure 6.

When M is at the low level, the charging capacity Φ does not have a significant effect on the expected total reward and the revenue earned from discharging, α , has only a small effect. While increasing the swap revenue ρ significantly increases the expected total reward when M is low, it has a greater effect when M is high. Furthermore, when M is high, Φ and α at the high level result in a significantly higher expected total reward than Φ and α at the low level. From this examination

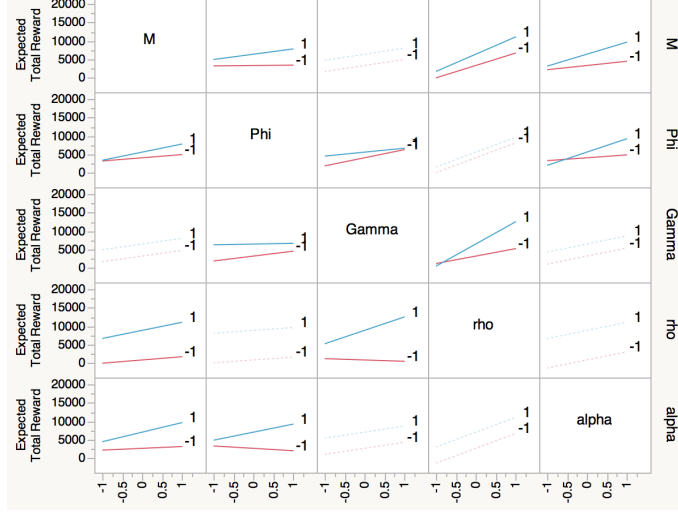


Figure 6. Interaction plots of significant factors from the first designed experiment.

the following policy insights are deduced. Having a correct number of batteries M is an integral part of optimally managing the swap station. When M is too low for the demand, even higher charging capacity and greater percentage earned from discharging cannot make up for the lack of revenue earned from not being able to exchange due to too few batteries. Further, if it is desirable for the swap station to serve a dual purpose by both satisfying swap demand and aiding the power grid via discharging, having a sufficient number of batteries M is essential. The second designed experiment looks at what M should be with respect to the number of PHEVs in the local area γ to serve this dual purpose.

Upon analysis of the remaining interactions, the interaction between Φ and α was the only one found to be insightful. When α is high, a higher charging capacity Φ results in a greater expected total reward. However, when α is low the charging capacity does not have a significant effect on the expected total reward. This is predominantly driven by the lack of discharging when α is low thereby causing less need for charging capacity Φ . Upon further inspection of the policies for the different levels of Φ and α there were some interesting trends in relationship to ρ . When $\alpha < 1$,

discharging will only be desirable when swapping is not desirable (i.e., ρ is below some threshold). However, when ρ is above this same threshold, discharging never occurs when $\alpha < 1$ even when Φ is high. For this experiment, the thresholds for ρ were \$3.71, \$2.16, \$2.16, and \$1.78 for Winter, Spring, Summer, and Fall, respectively. These all fall below the mean charging costs which are \$9.58, \$2.86, \$4.80, and \$2.17. Further, an oscillation between charging and discharging occurs when $\alpha > 1$ regardless of the charging capacity Φ or the swap revenue ρ , and little demand for swaps is met. These trends should be particularly informative to the power company. Even if the swap station has sufficient charging infrastructure they are not incentivized to discharge if they are earning a discounted rate, as long as ρ is set appropriately. Further, a negative behavior occurs possibly furthering the fluctuations seen in the load on the power grid when the incentive to discharge is too high, regardless of the charging infrastructure at the swap station.

The analysis is proceeded by further examining the optimal policies for different scenarios. Figure 7 illustrates the optimal policies for a scenario with $M = 50$, $\Phi = M$, $\rho = 15$, $\gamma = 3,000$, and $K_t = K_t^{Su}$ differentiated by three values for α . For a typical Wednesday, Figures 7a and 7b show the optimal policies in 4 hour increments and Figure 7c shows two consecutive hours. It can be visually seen that the swap station never discharges when $\alpha = 0.75$ as the policy never drops below zero in the grayed area of the Figure. When $\alpha = 1$, discharging does occur when it appears that the number of full batteries at the swap station is above some threshold (between 25 and 35 full batteries). When $\alpha = 1.25$, the optimal policy alternates charging and discharging every hour when the swap station has between about 10 and 45 fully charged batteries.

Taking a closer look at this phenomenon, the impact of α on the amount of swap demand that is met at a swap station is examined. Figure 8 depicts the ceiling of

the expected demand $\lceil \lambda_t \rceil$ when demand follows a Poisson process as compared to the number of batteries the swap station is able to swap when the optimal policy is implemented and the initial state is M . When $\alpha = 1.25$, the oscillating behavior between charging and discharging that was seen in Figure 7c prevents the satisfaction of most demand. Further, even when discharging never occurs ($\alpha = 0.75$) much demand is left unsatisfied. The next designed experiment is performed to identify the relationship between the total number of batteries M and the demand in a local area to ensure some level of demand is met.

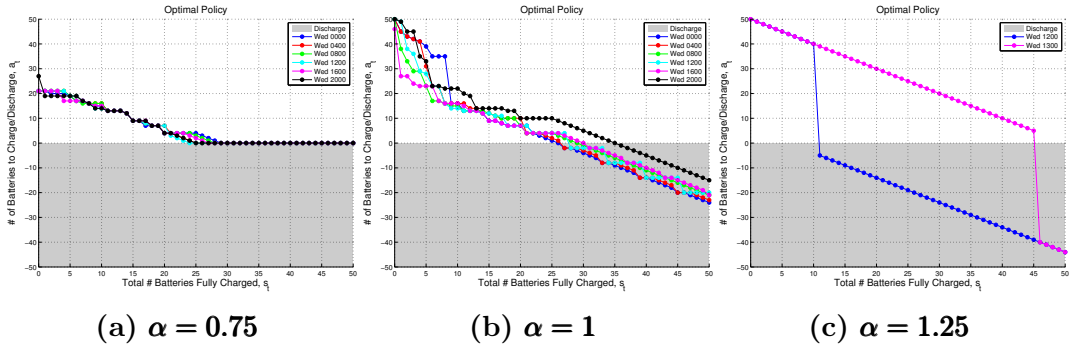


Figure 7. Optimal policy by percentage of the charge cost earned for discharging, α .

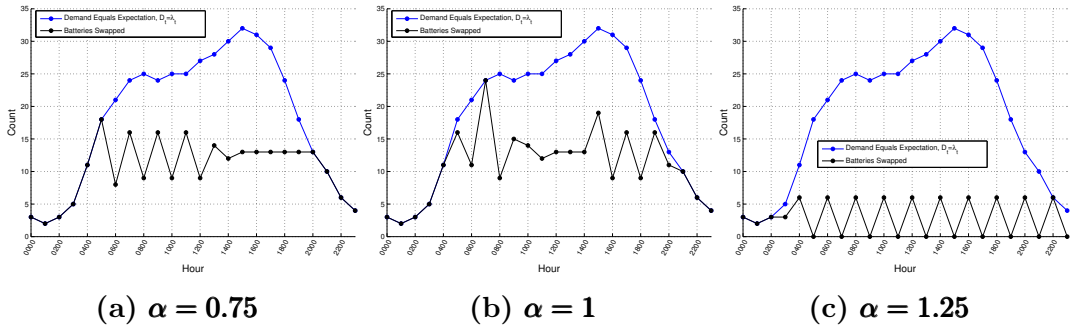


Figure 8. Expected swap demand and met demand by percentage of the charge cost earned for discharging, α .

From this analysis into α it is decided that to maintain the dual purpose of the swap station of meeting swap demand and still exhibiting some favorable V2G discharging behavior, $\alpha = 1$ is best. With $\alpha = 1$ the money the swap station earns

from discharging is exactly the cost for charging a battery. Thus, further analysis will focus on the scenarios when $\alpha = 1$ to arrive at policy insights.

Next, the state of the system is illustrated when operating using the optimal policy, or the number of fully charged batteries the swap station has on hand throughout a typical week and day for a swap station with $M = 50$, $\Phi = M$, $\rho = 5$, $\gamma = 3,000$, $\alpha = 1$, and $K_t = K_t^F$. To do this three sample paths for observed demand at the swap station are generated. In the first sample path, the demand observed at the swap station is exactly the mean arrival $\lceil \lambda_t \rceil$ when demand follows a Poisson process. Monte Carlo simulation is used to generate two more sample paths for observed demand at each decision epoch. A Meresenne Twister pseudorandom number generator is used to generate random numbers R_t , between 0 and 1 for each decision epoch and then the battery swap demand is calculated using the cumulative probability distribution of demand. The probability that demand is less than or equal to $\hat{X}_t$, $P(D_t \leq \hat{X}_t)$ is set equal to R_t , where $D_t \sim \text{Poisson}(\lambda_t)$. The state at the next decision epoch $t + 1$ is calculated using the optimal decision rule $d_t^*(s_t)$ for the current state s_t , and the observed demand, denoted $\hat{X}_t$, by way of

$$s_{t+1} = s_t + d_t^*(s_t) - \min \{ \hat{X}_t, s_t - |\min\{d_t^*(s_t), 0\}| \}. \quad (92)$$

Assuming the swap station starts with all full batteries, an entire week is examined and then a specific day in more detail. The state of the system at each decision epoch and the corresponding optimal action can be seen in Figure 9 for an entire week. From this figure, note that the assumption that the swap station starts with all full batteries at the start of a time horizon is not a simplifying assumption as the number of full batteries naturally increases at the start of each day. Also note that the state and action taken is relatively consistent for each of the three observed sample paths of demand. This is a nice result as it appears the action taken in relation to the state

balances. Similar results for Wednesday can be seen in Figure 10.

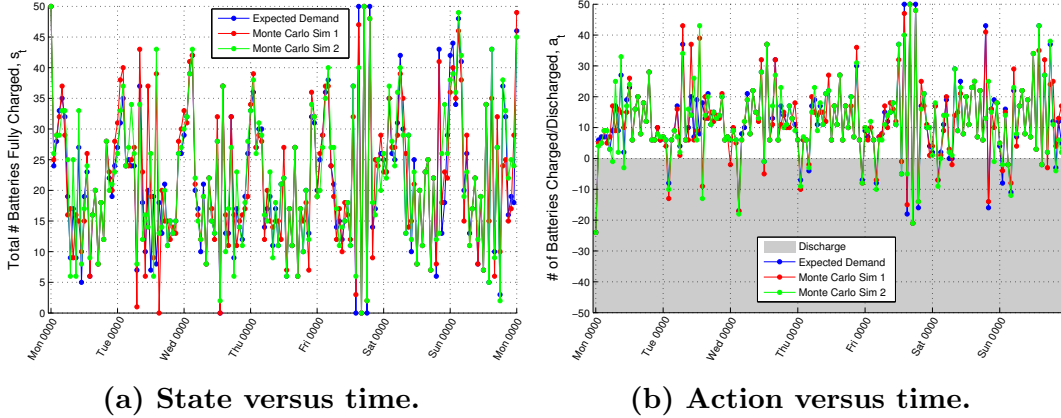


Figure 9. State and action over a week time period for three simulated observed demands.

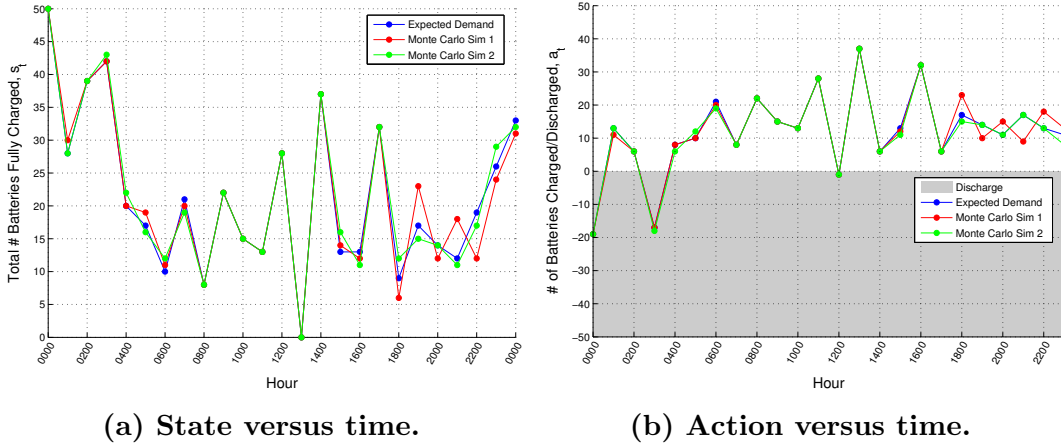


Figure 10. State and action over one Wednesday for three simulated sample paths.

For Wednesday, the number of swaps occurring in relation to each sample path of demand is examined further. Figure 11 shows this relationship for the first sample path of observed demand which equals the expected demand and the two Monte Carlo simulations. The number of batteries swapped are consistent even as the sample path of demand is different. Further, it can again be seen that for this particular scenario there is not a sufficient number of batteries at the swap station to consistently meet

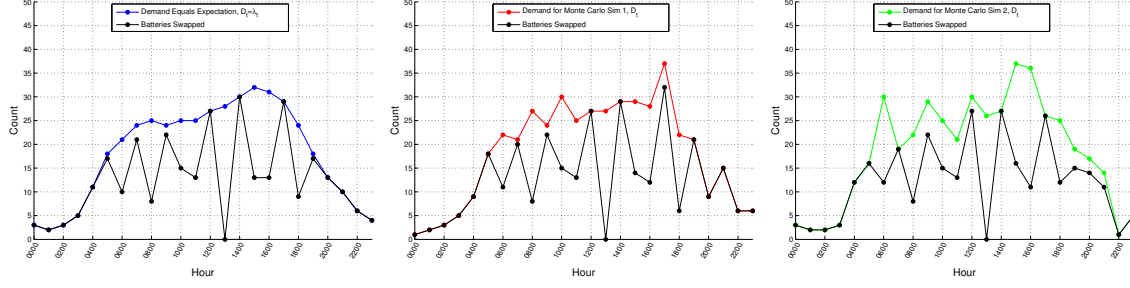


Figure 11. Demand throughout a typical Wednesday.

demand. This is compounded with the fact that unmet demand is not penalized in this model.

Based on the insights drawn from the first experiment, the next experiment is performed to gather insight into what the swap station should use for its controllable parameters when $\alpha = 1$ and one season is considered. Specifically, the aim is to determine the number of batteries M , charging capacity Φ , and swap cost ρ in relation to the non-controllable parameters. With this, the focus transitions from the expected total reward to the amount of demand that is met. A second Latin hypercube designed experiment with 40 scenarios is performed. The response variable for this experiment is the percentage of demand being met over the entire week when the optimal policy is implemented, the initial state of the system is M fully charged batteries, and the demand is set to $\lceil \lambda_t \rceil$. For this experiment, only $\alpha = 1$ is considered and for simplification only look at charging costs for the week of April 15-21 (Spring). The seasonal charging cost K_t is not statistically significant with respect to the expected total reward and the percentage of demand that is met. Thus, the design consists of four factors. Using JMP11Pro software, a 40 scenario design is generated with various levels of each factor ranging between two values. The high and low levels used for this experiment are shown in Table 2. This experiment is again run when demand follows a geometric distribution and Poisson process where both a monotone policy and optimal policy are found when demand follows a Poisson process. For

these 40 scenarios the monotone policy is always optimal, and that the expected total reward found for all scenarios are identical regardless of the demand distribution used. The policies differ indicating there are multiple optimal solutions. The results are presented when demand follows a Poisson process and the policy is monotone nonincreasing.

Table 2. Factor levels used for the second Latin hypercube designed experiment.

Factor		Low	High
Total Number of Batteries	M	25	100
Charging Capacity	Φ	$\lfloor 0.25M \rfloor$	M
Swap Revenue (\$)	ρ	2	20
PHEVs in the Local Area	γ	500	3,000

The results from the second Latin hypercube designed experiment can be seen in Table 3. Statistically significant factors at the 95% confidence level with respect to the percentage of met demand include the number of batteries, M , and the number of PHEVs in the local area, γ . This supports the intuition that the number of batteries must be sufficient based on the number of PHEVs in the local area to meet demand. Note that the charging capacity, Φ , and the swap revenue, ρ , are not significant factors with respect to the percentage of demand that is met, as long as ρ is above some threshold.

In scenario 34 when $\rho = 2$, the swap station meets only 0.08% of demand. In this scenario, the optimal policy indicates to charge only when there are zero fully charged batteries. Thus, if ρ is set too low the swap station does not have enough incentive to have fully charged batteries available for swapping, but rather discharges to earn a profit. In all other scenarios at least 59.07% of demand is met, even when $\rho = \$2.92$ in scenario 7 which meets 98.45% of demand. This indicates that there is some threshold that ρ must be set to with respect to the charging costs K_t , for the

swap station to have monetary incentive to meet demand over the opportunity cost to discharge batteries. Once this threshold is met, it appears that increasing ρ does not increase the percentage of demand that is met.

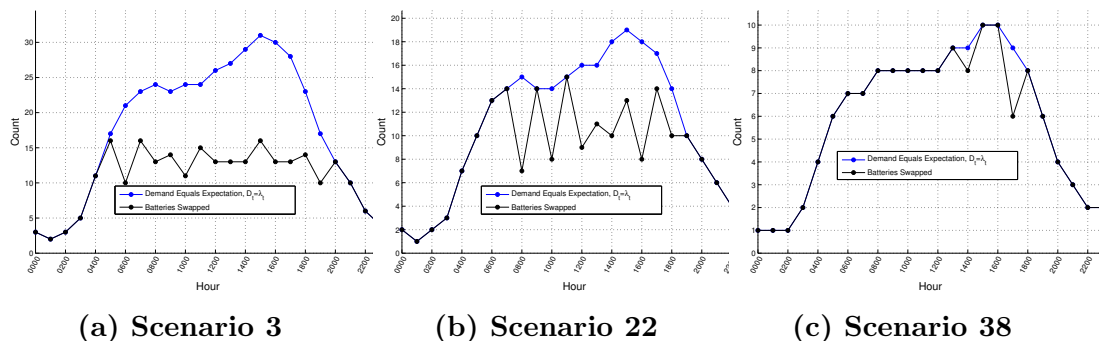


Figure 12. Expected demand compared to the number of batteries swapped on a Wednesday for 3 scenarios from the Latin hypercube experiment.

The demand and battery swaps on a typical Wednesday for three scenarios are illustrated in Figure 12. Demand is met 59.07% of the time in scenario 3. This indicates that 60 batteries is not enough to meet demand in a location with 2,872 PHEVs. Scenario 22, where 77.32% of demand is met, indicates that 81 batteries isn't quite enough to meet demand when there are 1718 PHEVs in the local area. Demand is met 99.07% of the time in scenario 38, indicating that 73 batteries is sufficient to meet 99.07% of demand in a location with 885 PHEVs. Examining all scenarios, if $M \geq 6\%\gamma$ then consistently above 95% of demand is met.

Next, the benchmark policies are examined to assess their accuracy and speed. For all scenarios in the second Latin hypercube experiment, the stationary benchmark policy (SBM) is tested with $\mathcal{C} = 0.5$ and the dynamic benchmark policy (DBM) with $\mathcal{C}_\ell = 0.25$ and $\mathcal{C}_u = 0.75$. These values were selected with the aim that any point in time the swap station should have approximately half of the batteries full and available for swapping. For all tests, the computation time, optimal expected total reward, and expected percentage of met demand is compared to an optimal policy

found via the monotone backward induction algorithm (BI). An optimality gap is calculated using the optimal expected total reward $v_N^*(s_1)$ and found expected total reward $v_N^\pi(s_1)$ for policy π using Equation (93), where an optimality gap of 0.00% indicates an optimal solution has been found.

$$\text{Optimality Gap} = \frac{v_N^*(s_1) - v_N^\pi(s_1)}{v_N^*(s_1)} \quad (93)$$

The expected percentage of demand met is compared by calculating a demand gap equal to the subtraction of the value found in the benchmark policy from the value found in the optimal policy. With this value a positive number indicates that the optimal policy is meeting more demand, whereas a negative number indicates the benchmark policy is meeting more demand. The optimality gaps, demand gaps, and elapsed computation time needed to arrive at policies can be found in Table 3.

All three solution methods require the use of probability transition matrices and reward vectors. The average computation time for creating the probability matrices and reward vectors was 249.40 and 4.67 seconds, respectively. Computations were done using MATLAB R2014a software on a 2.4 GHz Intel Core i5 processor laptop with 4GB 1600 MHz DDR3 of memory.

Disregarding scenario 34 with the unrealistically low swap revenue $\rho = 2$, the stationary benchmark policy is on average 13.08% from optimal with a range of 1.11% to 28.85%. The demand gaps indicated that on average this benchmark policy increases the percentage of met demand by 5.16%. At best, the stationary benchmark policy increases met demand by 29.57% and at worst it decreases met demand by 26.11%. For the dynamic benchmark policy, the policy is on average 6.45% from optimal, with the best case being 0.56% and worst 14.42%. This benchmark policy meets on average only 0.28% less demand than the optimal policy, where in the best case met demand is increased by 20.34% and at worst met demand is decreased by

28.58%.

Table 3. Results from second Latin hypercube designed experiment.

Scenario	M	Φ	γ	ρ	Met	Time (s)			Optimality Gap			Demand Gap (%)		
					Demand	BI	SBM	DBM	BI	SBM	DBM	BI	SBM	DBM
1	58	20	1846	4.77	74.20 %	2.24	0.15	0.16	0.00%	10.28%	4.96%	0.00	-13.47	-4.66
2	48	22	1974	18.15	71.64 %	2.03	0.42	0.14	0.00%	10.34%	5.17%	0.00	-1.75	3.88
3	60	18	2872	13.08	59.07 %	2.30	0.19	0.17	0.00%	9.08%	4.57%	0.00	-6.72	0.98
4	27	21	1077	15.38	93.33 %	1.41	0.07	0.08	0.00%	10.12%	4.67%	0.00	17.23	23.29
5	62	45	756	14.46	99.64 %	3.20	0.17	0.19	0.00%	21.82%	10.56%	0.00	-0.36	0.24
6	42	32	692	7.54	99.36 %	2.21	0.11	0.12	0.00%	16.99%	8.09%	0.00	-0.64	3.86
7	63	32	821	2.92	98.45 %	2.47	0.17	0.17	0.00%	7.24%	3.50%	0.00	-1.55	-0.89
8	44	28	2936	17.69	59.69 %	2.77	0.12	0.13	0.00%	7.01%	3.51%	0.00	9.23	12.47
9	75	31	628	16.77	99.86 %	2.19	0.20	0.23	0.00%	25.73%	12.52%	0.00	-0.14	-0.14
10	92	63	2615	10.77	69.31 %	7.47	0.36	0.27	0.00%	13.05%	6.52%	0.00	-24.31	-15.38
11	65	32	2808	5.69	62.15 %	3.48	0.19	0.20	0.00%	8.80%	4.40%	0.00	-9.77	-4.66
12	56	47	2744	8.92	65.46 %	4.11	0.17	0.16	0.00%	8.69%	4.34%	0.00	0.92	5.40
13	38	23	1590	3.85	79.83 %	1.99	0.10	0.11	0.00%	7.80%	3.69%	0.00	7.70	13.72
14	94	36	2679	14.92	63.99 %	3.93	0.27	0.26	0.00%	13.36%	6.54%	0.00	-29.57	-20.34
15	85	63	1269	20.00	92.96 %	5.38	0.24	0.25	0.00%	19.98%	9.99%	0.00	-7.04	-5.56
16	100	63	1782	6.15	79.94 %	5.94	0.29	0.36	0.00%	15.45%	7.73%	0.00	-20.06	-15.84
17	50	24	1141	9.85	90.66 %	2.21	0.15	0.18	0.00%	14.86%	7.13%	0.00	-8.69	0.00
18	54	48	1462	19.54	85.77 %	4.41	0.16	0.19	0.00%	14.03%	6.76%	0.00	-9.25	0.32
19	79	75	2038	12.62	77.43 %	6.81	0.21	0.24	0.00%	13.93%	6.78%	0.00	-19.93	-11.10
20	96	79	564	14.00	100.00 %	6.16	0.26	0.29	0.00%	28.85%	14.42%	0.00	0.00	0.00
21	40	10	1910	11.69	66.32 %	1.30	0.11	0.12	0.00%	9.04%	4.98%	0.00	0.85	10.00
22	81	29	1718	19.08	77.32 %	3.01	0.22	0.23	0.00%	16.23%	8.12%	0.00	-22.63	-16.31
23	98	52	1397	13.54	89.24 %	5.05	0.38	0.29	0.00%	20.06%	9.83%	0.00	-10.76	-9.20
24	29	11	1013	16.31	92.50 %	1.56	0.13	0.08	0.00%	11.23%	5.61%	0.00	9.41	16.09
25	88	50	500	8.46	100.00 %	3.54	0.31	0.25	0.00%	26.24%	13.12%	0.00	0.00	0.00
26	33	14	2551	4.31	61.95 %	1.65	0.09	0.10	0.00%	5.28%	2.65%	0.00	16.32	20.03
27	87	58	2359	18.62	70.85 %	5.98	0.25	0.25	0.00%	13.82%	6.75%	0.00	-25.01	-16.03
28	83	71	2423	3.38	77.56 %	7.55	0.23	0.25	0.00%	8.15%	3.97%	0.00	-15.29	-6.51
29	25	22	2231	15.85	66.80 %	1.52	0.07	0.07	0.00%	5.24%	2.70%	0.00	26.11	28.58
30	77	20	949	8.00	95.15 %	1.91	0.20	0.23	0.00%	19.49%	9.80%	0.00	-4.85	-4.66
31	31	8	1205	5.23	79.08 %	1.44	0.08	0.09	0.00%	9.08%	4.63%	0.00	1.80	12.02
32	90	29	2167	7.08	70.04 %	3.25	0.25	0.25	0.00%	13.63%	6.66%	0.00	-28.76	-19.88
33	46	42	1654	10.31	81.69 %	3.05	0.13	0.15	0.00%	11.06%	5.29%	0.00	1.10	8.72
34	37	29	2487	2.00	0.08 %	0.32	0.10	0.11	0.00%	109.77%	113.87%	0.00	-51.05	-48.17
35	71	66	3000	17.23	62.39 %	5.66	0.20	0.20	0.00%	9.97%	4.84%	0.00	-10.77	-5.09
36	35	20	2295	11.23	65.90 %	1.91	0.10	0.10	0.00%	6.80%	3.21%	0.00	13.86	17.26
37	67	37	2103	12.15	74.21 %	3.65	0.19	0.20	0.00%	12.14%	5.89%	0.00	-15.02	-6.57
38	73	73	885	9.38	99.07 %	5.06	0.20	0.20	0.00%	20.30%	10.15%	0.00	-0.93	-0.72
39	69	49	1526	6.62	84.73 %	4.55	0.20	0.21	0.00%	13.88%	6.94%	0.00	-15.15	-8.19
40	52	50	1333	2.46	93.40 %	5.04	0.19	0.15	0.00%	1.11%	0.56%	0.00	-3.51	5.62

These results indicate that the dynamic benchmark policy outperforms the stationary benchmark policy due to decreased optimality gaps and comparable amount of met demand. Further, these results indicate that the dynamic benchmark policy could be a viable option for implementation at a swap station. This benchmark policy allows for an easy calculation of the number of batteries to charge and discharge over

time based off a target level for each hour of a week. Therefore, all that is needed is a table with 168 numbers, one for each hour of a week. In contrast, implementation of the optimal policy would require a very large look up table by state and time. The results and analysis of these computational tests are summarized in the following policy insights for a PHEV swap station manager and the power grid.

1. It is integral to have the number of batteries at a swap station M in line with the PHEVs in the local area γ for meeting demand, maximizing expected total reward, and allowing for discharging back to the power grid using V2G. From the results, it is observed that $M \geq 6\%\gamma$ was a sufficient value for M .
2. To ensure that the swap station is meeting demand and not solely focused on discharging to earn revenue, the swap revenue ρ must be set appropriately. There is a threshold level which ρ must be greater than to ensure demand is met. After this threshold, increasing ρ did not seem to incentivize meeting demand over discharging. For the experiments conducted, this threshold was less than the average charging cost K_t for a week.
3. When the incentive to discharge is too high, the negative behavior of oscillating between charging and discharging in consecutive time periods occurs at the swap station thereby leading to further variability in the power grid. Further, when the incentive is too low and ρ is set appropriately, discharging never occurs. When the revenue earned from discharging is exactly the cost for charging, a good balance of some discharging but limited oscillating behavior occurs.
4. The dynamic benchmark policy which calculates a target level for each time period in a time horizon was superior to a stationary benchmark policy. The action for the dynamic benchmark policy is to charge up to or discharge down to this time dependent target level based on the number of full batteries on hand.

In addition to the advantage over a stationary benchmark policy, this could be a viable policy to implement at a swap station due to its accuracy in the regards to expected total reward and met demand, and ease of implementation.

5. With all scenarios considering different number of batteries M , charging capacity Φ , swap revenue ρ , charging cost by week K_t , incentive to discharge α , and PHEVs in a local area γ , the swap station was always able to remain profitable with the model. Certain combinations of these factors led to greater profitability, but this result indicates that in all circumstances considered a swap station is a viable, profitable option for PHEVs.

VI. Conclusions

Motivated by the movement to make transportation cleaner and more efficient, the PHEV-SSMP is introduced. This problem considers the management of operations at a plug-in hybrid electric vehicle (PHEV) swap station facing stochastic, nonstationary demand for battery swaps, nonstationary prices for charging depleted batteries, and nonstationary prices for discharging fully charged batteries utilizing V2G technology. With this, the optimal number of batteries that the swap station should charge and discharge over time is determined using sequential decision making over a fixed time horizon, which results in the maximum expected total profit.

A Markov decision process model is used when demand follows a discrete probability distribution. A finite-horizon model is considered because the problem data is highly variable with respect to time. In the model, the state of the system, or the number of fully charged batteries on hand is observed at a certain point in time and the swap station manager chooses the number of batteries to charge or discharge. The action results in an immediate reward and the system transitions to a new state.

It has been proven that there exists an optimal nonincreasing monotone policy when demand follows a discrete nonincreasing distribution. Therefore, both the backward induction and monotone backward induction algorithms can be utilized to find the optimal policy. Two easy to implement benchmark policies were created and empirically compared to an optimal policy. In the stationary benchmark policy, the swap station maintains a single target inventory level of fully charged batteries regardless of time of day and day of week. In the dynamic benchmark policy, the swap station maintains a distinct target inventory level for each time period which takes into account current and future charging costs.

Two Latin hypercube designed experiments were performed to computationally test the optimal solution method and two benchmark policies. The first experiment

is conducted to gain overall information for various parameter inputs for the swap station. Specifically, the incentive which should be given by the power company is determined and other statistically significant factors are analyzed. The second experiment is conducted to gain insight into what the controllable parameters should be set to at a swap station (e.g., number of batteries, swap price) in relationship to the number of PHEVs in a local area and power prices.

From this analysis, it is determined that the dynamic benchmark policy is best, the number of batteries M is an integral parameter, α needs to be appropriately set by the power company to encourage discharging and not oscillating behavior, and other policy insights. Following the culmination of this work and Widrick et al. [40], future work should consider how the swap price ρ impacts the demand for swaps in comparison to using at home charging or a charging station. Further, the uncertainties regarding the power prices, power load, and other renewables should be incorporated into the state space of the MDP to fully capture the load balancing potential of a PHEV swap station.

Bibliography

1. U.S. Department of Energy, “Secretary Moniz announces nearly \$50 million to advance high-tech, fuel efficient American autos,” January 2014. Last accessed on November 20, 2014 at <http://energy.gov/articles/secretary-moniz-announces-nearly-50-million-advance-high-tech-fuel-efficient-american-autos>.
2. K. Clement-Nyngs, E. Haesen, and J. Driesen, “The impact of charging plug-in hybrid electric vehicles on a residential distribution grid,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 371–380, 2010.
3. Z. Bingliang, S. Yutian, L. Bingqiang, and L. Jianxiang, “A modeling method for the power demand of electric vehicles based on Monte Carlo simulation,” in *2012 Asia-Pacific Power and Energy Engineering Conference*, (Shanghai, China), pp. 1–5, March 2012.
4. J. Eyer and G. Corey, “Energy storage for the electricity grid: Benefits and market potential assessment guide,” tech. rep., Sandia National Laboratories, February 2010.
5. D. Pearson and S. T. Stub, *Better place’s failure is blow to Renault*. The Wall Street Journal, May 2013. Last accessed on November 20, 2014 at <http://online.wsj.com/articles/SB10001424127887323855804578507263247107312>.
6. P. Abreu, “The world’s only electric sports car: 2010 tesla roadster,” April 2010. Last accessed on November 20, 2014 at http://www.motorauthority.com/news/1044161_the-worlds-only-electric-sports-car-2010-tesla-roadster.

7. Tesla motors, “Model X, Utility Meet Performance,” August 2014. Last accessed on November 20, 2014 at <http://www.teslamotors.com/modelx>.
8. S. Fowler, “Tesla model 3 to challenge BMW 3 series - World exclusive,” July 2014. Last accessed on November 20, 2014 at <http://www.autoexpress.co.uk/tesla/87867/tesla-model-3-to-challenge-bmw-3-series-world-exclusive>.
9. Plug-In Cars, “Detailed list of electric cars and plug-in hybrids,” August 2014. Last accessed on November 20, 2014 at <http://www.plugincars.com/cars>.
10. Tesla motors, “Road trips made easy,” 2014. Last accessed on November 20, 2014 at <http://www.teslamotors.com/supercharger>.
11. Tesla motors, “Battery swap,” 2014. Last accessed on November 20, 2014 at <http://www.teslamotors.com/batteryswap>.
12. R. Sioshansi and P. Denholm, “The value of plug-in hybrid electric vehicles as grid resources,” *The Energy Journal*, vol. 31, no. 3, pp. 1–23, 2010.
13. M. Peng, L. Liu, and C. Jiang, “A review on the economic dispatch and risk management of the large-scale plug-in electric vehicles (PHEVs)-penetrated power systems,” *Renewable and Sustainable Energy Reviews*, vol. 16, pp. 1508–1515, April 2012.
14. J. Wang, C. Liu, D. Ton, Y. Zhou, J. Kim, and A. Vyas, “Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power,” *Energy Policy*, vol. 39, pp. 4016–4021, July 2011.
15. L. Göransson, S. Karlsson, and F. Johnsson, “Integration of plug-in hybrid electric vehicles in a regional wind-thermal power system,” *Energy Policy*, vol. 38, pp. 5482–5492, October 2010.

16. M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, 2005.
17. A. Gbate and R. L. Smith, “A linear programming approach to nonstationary infinite-horizon markov decision processes,” *Operations Research*, vol. 61, no. 2, pp. 413–425, 2013.
18. K. Morrow, D. Karner, and J. Francfort, “Plug-in hybrid electric vehicle charging infrastructure review,” tech. rep., U.S. Department of Energy, Idaho National Laboratory, November 2008.
19. O. Worley and D. Klabjan, “Optimization of battery charging and purchasing at electric vehicle battery swap stations,” in *2011 IEEE Vehicle Power and Propulsion Conference*, (Chicago, IL), pp. 1–4, September 2011.
20. S. G. Nurre, R. Bent, F. Pan, and T. C. Sharkey, “Managing operations of plug-in hybrid electric vehicle (PHEV) exchange stations for use with a smart grid,” *Energy Policy*, vol. 67, pp. 364–377, 2014.
21. H.-Y. Mak, Y. Rong, and Z.-J. Shen, “Infrastructure planning for electric vehicles with battery swapping,” *Management Science*, vol. 59, pp. 1557–1575, July 2013.
22. X. Tang, N. Liu, J. Zhang, and S. Deng, “Capacity optimization configuration of electric vehicle battery exchange stations containing photovoltaic power generation,” in *2012 7th International Power Electronics and Motion Control Conference*, (Harbin, China), pp. 2061–2065, June 2012.
23. C.-H. Zhang, J.-S. Meng, Y.-Z. Cao, X. Cao, Q. Huang, and Q.-C. Zhong, “The adequacy model and analysis of swapping battery requirement for electric vehicles,” in *2012 IEEE Power and Energy Society General Meeting*, (San Diego, CA), pp. 1–5, July 2012.

24. F. Pan, R. Bent, A. Berscheid, and D. Izraelevitz, “Locating PHEV exchange stations in V2G,” in *2010 First IEEE International Conference on Smart Grid Communications*, (Gaithersburg, MD), pp. 173–178, October 2010.
25. R. Sioshansi, S. H. Madaeni, and P. Denholm, “A dynamic programming approach to estimate the capacity value of energy storage,” *IEEE Transactions on Power Systems*, vol. 29, pp. 395–403, January 2014.
26. D. F. Salas and W. B. Powell, “Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of multidimensional energy storage problems,” tech. rep., Department of Operations Research and Financial Engineering, Princeton, NJ, 2013.
27. W. R. Scott, W. B. Powell, and S. Moazehi, “Least squares policy iteration with instrumental variables vs. direct policy search: Comparison against optimal benchmarks using energy storage,” tech. rep., Department of Operations Research and Financial Engineering, Princeton University, January 2014. Submitted to *INFORMS Journal on Computing*.
28. I. Giannoccaro and P. Pontrandolfo, “Inventory management in supply chains: a reinforcement learning approach,” *International Journal of Production Economics*, vol. 78, pp. 153–161, July 2002.
29. D. Zhang and W. L. Cooper, “Revenue management for parallel flights with customer-choice behavior,” *Operations Research*, vol. 53, pp. 415–431, May-June 2005.
30. K. K. Yin, H. Liu, and N. E. Johnson, “Markovian inventory policy with application to the paper industry,” *Computers & Chemical Engineering*, vol. 26, pp. 1399–1413, October 2002.

31. B. M. Lewis, *Inventory control with risk of major supply chain disruptions*. PhD thesis, Georgia Institute of Technology, 2005.
32. M. ElHafsi, “Optimal integrated production and inventory control of an assemble-to-order system with multiple non-unitary demand classes,” *European Journal of Operational Research*, vol. 194, pp. 127–142, April 2009.
33. H. Scarf, “The optimality of (s,S) policies in the dynamic inventory problem,” *Stanford University Press*, 1960.
34. T. K. Das, A. Gosavi, S. Mahadevan, and N. Marchallick, “Solving semi-markov decision problems using average reward reinforcement learning,” *Management Science*, vol. 45, no. 4, pp. 560–574, 1999.
35. H. S. Chang, M. C. Fu, J. Hu, and S. I. Marcus, “An adaptive sampling algorithm for solving markov decision processes,” *Operations Research*, vol. 53, no. 1, pp. 126–139, 2005.
36. National Grid, “Hourly electric supply charges,” 2013. Last accessed on October 8, 2014 at http://www.nationalgridus.com/niagaramohawk/business/rates/5_hour_charge.asp.
37. Tesla motors, “Specs,” 2014. Last accessed on November 23, 2014 at <http://www.teslamotors.com/models/specs>.
38. Nexant, Inc., A. Liquide, A. N. Laboratory, C. T. Venture, G. T. Institute, N. R. E. Laboratory, P. N. Laboratory, and T. LLC, “H2A hydrogen delivery infrastructure analysis models and conventional pathway options analysis results,” 2008. Last accessed on November 26, 2014 at http://www1.eere.energy.gov/hydrogenandfuelcells/pdfs/nexant_h2a.pdf.

39. D. C. Montgomery, *Design and analysis of experiments*. New Jersey: John Wiley & Sons, 8th ed., 2008.
40. R. S. Widrick, S. G. Nurre, and M. J. Robbins, “Optimal policies for the management of a plug-in hybrid electric vehicle swap station,” tech. rep., Air Force Institute of Technology, 2014.

OPTIMAL POLICIES FOR THE MANAGEMENT OF A PLUG-IN HYBRID ELECTRIC VEHICLE SWAP STATION

REBECCA S. WIDRICK*, SARAH G. NURRE, MATTHEW J. ROBBINS
SPONSOR: LOS ALAMOS NATIONAL LABORATORY



BACKGROUND

- PHEV battery swap stations can have a significant impact on the growing market for battery powered vehicles.
- A model that optimizes the operations at a swap station is valuable as implementing swap stations can offer:
 - PHEV owners the convenience to exchange their depleted battery for a fully charged one.
 - Opportunity to control battery charging and reduce the negative effect of increased demand for electricity using vehicle-to-grid technology (V2G) (i.e. discharging batteries back to the power grid to balance load).

METHODOLOGY

- Operations of a PHEV station are modeled as a finite horizon, single product inventory control Markov Decision Problem (MDP) [1] facing stochastic, nonstationary demand for battery swaps, nonstationary charging costs, and nonstationary revenue earned from discharging.
- The model is mathematically characterized as follows:
 1. **Decision Epochs:** time unit at which the number of batteries to charge/discharge is decided.
 2. **State:** total number of fully charged batteries.
 3. **Action Space:** total number of batteries to charge or discharge.
 4. **Reward Function:** expected reward criteria is used and calculated by the revenue from exchanges and discharging minus the cost from charging.
 5. **Transition probabilities:** determined by demand (discrete random variable), state and action chosen.
- The optimal policy is found using backward induction and monotone backward induction [1]. Two benchmark policies are evaluated and compared to the optimal policy.

REFERENCES

- [1] Martin L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, 2005

CONTRIBUTIONS

- The main contributions of this work are:
1. Development of a MDP model to determine the optimal number of batteries to charge and discharge when factoring in stochastic, nonstationary swap demand, nonstationary charging costs, and nonstationary discharging revenues.
 2. Proving the existence of a nonincreasing monotone optimal policy when demand is governed by a discrete nonincreasing distribution.
 3. Generation of two benchmark policies which are easy to implement by a swap station manager.
 4. Analysis of the results from two designed experiments using realistic data.

MODEL

- 
- **Decision Epochs:** $T = \{0, 1, \dots, N\}$, $N < \infty$
 - **State:** $s_t \in S = \{0, 1, \dots, M\}$, $M = \text{total \# batteries}$
 - **Action:** (ϕ) is charging capacity; $a_t \in A_t$
 $A_t = \{\max(-s_t, -\Phi), \dots, 0, \dots, \min(M - s_t, \Phi)\}$
 - **Reward:**
 $r_t(s_t, a_t) = \rho(s_t + a_t - s_{t+1}) - K_t(a_t^+) + J_t(a_t^-)$
 $\rho = \text{revenue from swap}$, $K_t = \text{charging cost}$, $J_t = \text{discharge revenue}$
 - **Transition Probability function:**

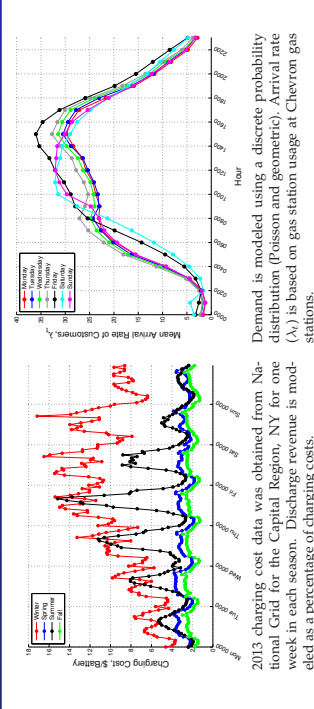
$$p_t(j|s_t, a_t) = \begin{cases} 0 & \text{if } j > s_t + a_t \text{ or } j < a_t^+ \\ p_{t+1|t-j} & \text{if } a_t^+ < j \leq s_t + a_t \\ q_{t+1|t-j} & \text{if } j = a_t^+ \end{cases}$$

$$p_t = \mathbb{P}(D_t = j) \text{ and } a_t = \sum_{j=0}^{\infty} p_j = \mathbb{P}(D_t \geq u)$$
 Random variable, D_t is demand

FUTURE STUDY

Future work should consider how the swap price impacts the demand for swaps in comparison to using at home charging or a charging station. Further, the uncertainties regarding power load, and other renewables

DATA



RESULTS AND CONCLUSIONS

- This model was tested with decision epochs in one hour increments over one week (i.e. $N = 168$) with varying levels of # batteries, M , charging capacity Φ , swap revenue ρ , charging/discharging cost/revenue, K_t and J_t , and demand, D_t .
- Proving the existence of a nonincreasing monotone policy allows for easier to implement optimal policies and faster computation time.
- Easily implementable benchmark policies were evaluated and compared to the optimal policy showing promising results.
- Valuable policy insights were gained from two designed experiments. They include:
 - Having the correct number of batteries is essential to meet demand and have discharging occur at desirable points in time.
 - Discharge revenue equal to charging costs is desirable, which provides valuable insight to the power companies who would be setting discharging incentives.
 - Considering current and future charging costs in the benchmark policy resulted in better solutions than the benchmark policy with a stationary target level.
 - In all circumstances considered a swap station is a viable, profitable option for PHEVs.

CONTACT INFO

*Rebecca S. Widrick, 2d Lt, USAF
Department of Operational Sciences, AFIT
rebecca.widrick@us.af.mil
(845) 978-0578

REPORT DOCUMENTATION PAGE					<i>Form Approved</i> <i>OMB No. 0704-0188</i>	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY) 26-03-2015		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From — To) Oct 2013 — Mar 2015		
4. TITLE AND SUBTITLE Optimal Policies for the Management of a Plug-In Hybrid Electric Vehicle Swap Station				5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER 		
6. AUTHOR(S) Widrick, Rebecca S., Second Lieutenant, USAF				5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER 		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENS) 2950 Hobson Way WPAFB OH 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENS-MS-15-M-135		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Russell Bent Los Alamos National Laboratory Risk and Decision Analysis Division Los Alamos, NM, USA Email: rbent@lanl.gov				10. SPONSOR/MONITOR'S ACRONYM(S) LANL		
12. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Statement A. Approved for Public Release; distribution unlimited.				11. SPONSOR/MONITOR'S REPORT NUMBER(S) 		
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.						
14. ABSTRACT Optimizing operations at plug-in hybrid electric vehicle (PHEV) battery swap stations is internally motivated by the movement to make transportation cleaner and more efficient. A PHEV swap station allows PHEV owners to quickly exchange their depleted PHEV battery for a fully charged battery. The PHEV-Swap Station Management Problem (PHEV-SSMP) is introduced, which models battery charging and discharging operations at a PHEV swap station facing nonstationary, stochastic demand for battery swaps, nonstationary prices for charging depleted batteries, and nonstationary prices for discharging fully charged batteries. Discharging through vehicle-to-grid is beneficial for aiding power load balancing. The objective of the PHEV-SSMP is to determine the optimal policy for charging and discharging batteries that maximizes expected total profit over a fixed time horizon. The PHEV-SSMP is formulated as a finite-horizon, discrete-time Markov decision problem and an optimal policy is found using dynamic programming. Structural properties are derived, to include sufficiency conditions that ensure the existence of a monotone optimal policy. A computational experiment is developed using realistic demand and electricity pricing data. The optimal policy is compared to two benchmark policies which are easily implementable by PHEV swap station managers. Two designed experiments are conducted to obtain policy insights regarding the management of PHEV swap stations. These insights include the minimum battery level in relationship to PHEVs in a local area, the incentive necessary to discharge, and the viability of PHEV swap stations under many conditions.						
15. SUBJECT TERMS Green logistics; Markov Decision Processes; Monotone Policy; Plug-In Hybrid Electric Vehicles						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT		18. NUMBER OF PAGES	
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU		77	
U	U	U	19a. NAME OF RESPONSIBLE PERSON Dr. Sarah G. Nurre (ENS)			
						19b. TELEPHONE NUMBER (include area code) (937) 255-3636 x4319 sarah.nurre@afit.edu