2015

Resource allocation and load-shedding policies based on Markov decision processes for renewable energy generation and storage

Edwards Jimenez
University of Central Florida

RESOURCE ALLOCATION AND LOAD-SHEDDING POLICIES BASED ON MARKOV DECISION PROCESSES FOR RENEWABLE ENERGY GENERATION AND STORAGE

by

EDWARDS JIMENEZ
B.S. University of Central Florida, 2012

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Electrical Engineering and Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2015

Major Professor: George Atia

ABSTRACT

In modern power systems, renewable energy has become an increasingly popular form of energy generation as a result of all the rules and regulations that are being implemented towards achieving clean energy worldwide. However, clean energy can have drawbacks in several forms. Wind energy, for example can introduce intermittency. In this thesis, we discuss a method to deal with this intermittency. In particular, by shedding some specific amount of load we can avoid a total system breakdown of the entire power plant. The load shedding method discussed in this thesis utilizes a Markov Decision Process with backward policy iteration. This is based on a probabilistic method that chooses the best load-shedding path that minimizes the expected total cost to ensure no power failure. We compare our results with two control policies, a load-balancing policy and a less-load shedding policy. It is shown that the proposed MDP policy outperforms the other control policies and achieves the minimum total expected cost.

This thesis is dedicated to my mother and my father, who have supported my academic endeavors and encouraged me to strive for excellence.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

CHAPTER ONE: INTRODUCTION

Renewable energy has become an increasingly popular source of energy, considering the increasing demand for additional electrical energy in the 21st century along with the limited fossil fuel reserves, which has enabled significant development in renewable energy resources. Furthermore, with all the new worldwide regulations and requirements to reduce greenhouse gas emissions, several methods are actively studied for renewable energy. At a recent World Energy Council (WEC), a study was presented that illustrated that without a change in our current energy demand usage, the world energy demand in 2020 will increase by 50-80% from the 1990 levels [1]. Having such a significant increase in energy demand can exert serious pressure on the current energy configuration and cause more environmental health damage by the release of greenhouse gases such as Carbon Oxide, Carbon Dioxide, Sulfur Dioxide, Nitric Oxide, and Nitrogen Dioxide [1, 2].

Energy Pollution

The most effective way to reduce these greenhouse gas emissions is to make radical changes in the technology of energy production, conversion, distribution, and storage [3]. As shown in Figure 1, The US Electric Power produced 38% of all the US CO_2 emissions in 2012 [4]. The majority of the emissions produced are coming from coal plants. The new regulations, Clean Air Act, on clean energy force power plants to have a carbon pollution standard [5]. This is measured as

tons of greenhouse gas emissions per Megawatt-Hour of electricity produced. The concentration of CO_2 in the atmosphere should be kept below 445 and 490 parts per million (ppm). This will ensure a maximum increase in temperature that will be held between 2 and 2.4℃. In 2007 the concentration of greenhouse gases has already reached 450 ppm CO_2. In order to stay in the determined increase of 2℃, the emission of greenhouse gases should peak as soon as possible, the latest being in 2015. Additionally the greenhouse gases should be reduced by 50-80% of the value it was in 2010 to maintain the desired increase in temperature [6].

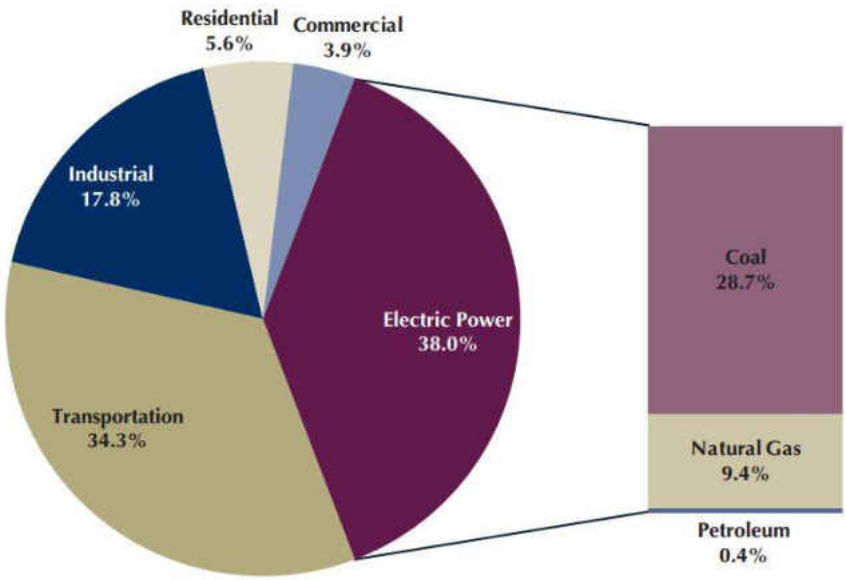

Figure 1: USA 2012 CO_2 Emissions

In order to reduce these greenhouse gas emissions, a significant body of ongoing research is devoted to develop different methods different approaches to address this problem. There have been several solutions proposed including

education and training on how to be more efficient with energy, replacing and upgrading power plants, insulation of houses, and usage of energy efficient lamps [6].

Renewable Energy

Another proposed solution is to make use of renewable energy. Renewable energy can come in various forms, including the sun, wind, biomass, oceans, fuel cells, geothermal, and hydro power. It is estimated that around 19% of the world electricity generation is currently generated from renewable energy. This is projected to reach nearly 23% by 2035 [7]. In Figure 2, one can see that the production of global installed power generation capacity is rapidly increasing at a rate of 3.1% per year for hydro and other renewables generation methods. Solar power is expected to have the largest growth rate at 8.3% per year, followed by wind at 5.3% per year.

In our research, we focus on the wind energy as a renewable source of energy. Since wind energy is the second fastest growing renewable energy, it attracted major attention over the past years. Indeed, the most globally utilized renewable energy technology in power systems at this moment is wind. In 2012, it delivered 2.26% of the world's electricity [8]. One drawback of wind energy, however, is that there is intermittency. In order to cope with this imperfection, there are several proposed solutions. It can be combined with other renewable energy generation to alleviate the effect of intermittency. Other alternatives can be used

such as adding an energy storage system or using load-shedding in situations when there is an excessive energy demand.

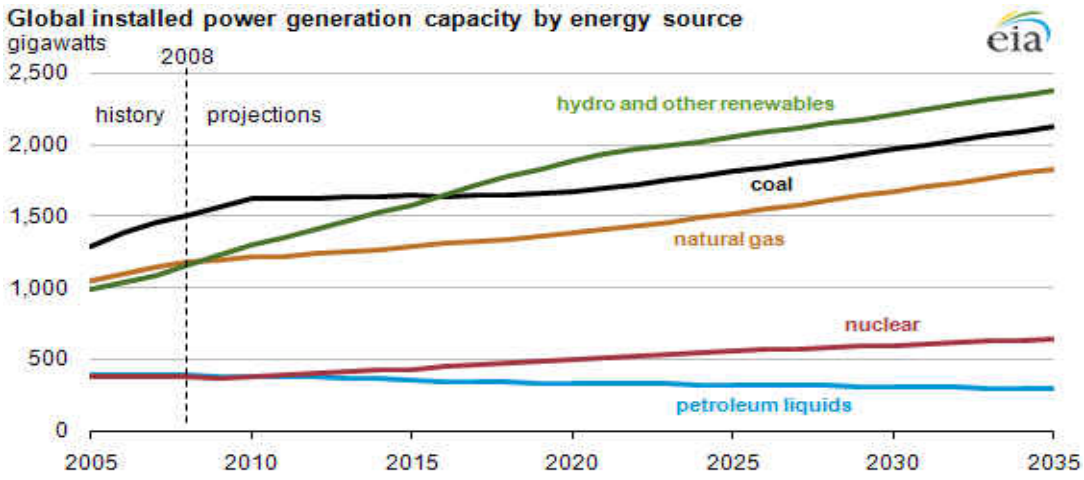

Figure 2: Global installed power generation capacity by energy sources

The occurrence of electric supply failure can have serious financial impact on the systems, customers, and power companies. If no action is taken rapidly, a power company may experience severe damage to expensive equipment. A power company that cannot supply sufficient power to cover the demand can activate back-up generators, energy storage devices, or can shed some load [9], for example, removing household electricity for a certain period of time in order to supply other company electricity during peak energy hours. Utilizing back-up generators is a reliable technique to handle excess energy demand, but requires time for generators to start and get to their optimum running standards. While implementing energy storage devices is a rapid method to cope with

intermittency, one drawback is that the expenses increase rapidly as the storage size increases [10].

Load-shedding has been used to provide reliable and rapid power management in overload or intermittency. The downfall of load-shedding is that a cost is associated with every amount of load that is shed. By integrating an energy storage device with the power system we can reduce the costs associated with load- shedding.

Load Shedding Methods and Applications

There are several reasons that motivate the use of load-shedding. In particular, load-shedding may be used to reduce intermittency, to achieve maximal operation of a thermal unit or battery, and other controllable loads [11-16]. Load shedding can also be used to save money in industries, hospitals, universities, hotels, and other facilities. Objectives such as reducing the operational costs and maximizing the profit were also achieved with load-shedding [17].

Several methods were proposed to efficiently implement load-shedding. In [11], probabilistic modeling was proposed in the context of a study to integrate renewable resources into the electric power generation portfolio of an island, and investigate the feasibility of replacing diesel generation entirely with solar photovoltaics and wind energy, supplemented with energy storage. Another method to implement load-shedding is tabu-search for optimization [12, 13]. In

these papers the authors describe how to reduce the operational cost and maximize the profit by operating generators with higher efficiency. This is done by focusing on the reactive power schedule of the controllable load to determine the optimal operation of the thermal units that satisfy the voltage constraints. Other work focused on particle swarm optimization [14]. There, they modeled the load and generation of two microgrids with wind farms to implement optimal power flow using particle swarm optimization. The objective was to achieve optimal dispatch of controllable loads and generators, as well as to effectively utilize the battery storage of each microgrid. Another method implemented an algorithm using measured frequency and voltage [15]. In this paper, three adaptive combinational load shedding methods are proposed to improve the operation of the conventional under-frequency load shedding scheme in order to enhance power system stability following severe disturbances. In another paper stratified load shedding optimization is studied [16]. There the authors focus on the reactive power schedule of the controllable loads. Or stratifying the controllable loads into different levels based on the degree of load importance and load frequency regulation effect coefficient. Another method is genetic-based under-frequency load-shedding in stand-alone power systems considering fuzzy loads [18]. In this paper the focus is on the use of an under-frequency relay that protects the power system against blackout when the system frequency declines to predetermined settings. The genetic algorithm is employed to minimize the shed load and maximize the lowest swing frequency.

In this thesis, we study a probabilistic framework wherein an operator has to sequentially decide the amount of load shedding and resource allocation. We adopt a Markov Decision Process (MDP) framework, where load-shedding is implemented in order to handle the uncertainty and intermittency of the renewable energy source. In this work, windmills provide the renewable energy source. The MDP framework is leveraged to determine the most cost-effective path to take when considering load shedding in combination with an energy storage model. Our goal is to find an optimal or near-optimal policy for load shedding and resource allocation to minimize the total expected cost. This policy takes into account the probabilistic nature of the renewable energy source. In the scenario that we simulate, we incorporate an energy storage device, a power supply company with a constant load and generation capacity, and a stochastic renewable energy source. The system schematic is illustrated in Figure 3. The stochastic renewable energy corresponds to the time-varying wind energy available at each stage.

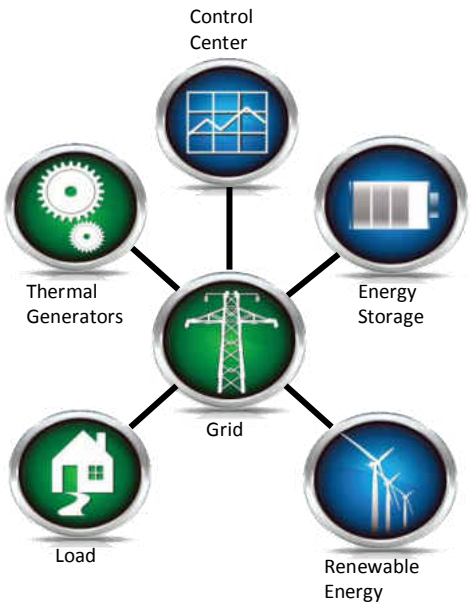

Figure 3: System Schematic

CHAPTER TWO: MARKOV DECISION PROCESSES

In this chapter, we provide some necessary background about Markov Decision Processes (MDPs). The described framework is used to implement load-shedding policies as described in the next chapter. In an MDP we aim to find the most cost-effective decisions to make in an uncertain environment. In what follows, we outline the fundamentals of sequential decision-making and the main components of the MDP framework, namely the state space, the control space, and the costs (or rewards).

Sequential Decision Making

In a sequential decision-making setting we have an agent or a controller interacting with some stochastic environment. At consecutive time steps, the agent has to make sequential decisions. The controller aims to choose the actions (controls) that maximize an expected total reward or minimize an expected total cost, taking into consideration the current and future effect of these actions. The incurred costs depend on the states of the system, the actions taken, as well as randomness from the stochastic environment. The actions taken can influence the evolution of the environment states. This class of problems can be generally efficiently modeled using MDPs.

Various real world situations can be modeled as MDPs. A key characteristic of a perfectly observable MDP is that, *conditioned on the current state of the system and the current action of the controller,* transitions from the present state to future

states are independent of previous states and previous actions [19]. In other words, at any given time k, the state of the system captures all information needed by the controller to compute the best action. In this thesis we focus on discrete time systems.

MDPs have two principal features: (1) a discrete-time dynamic system, and (2) a cost (reward) function that is additive over time [20]. At each time, the controller observes the system state as its input and generates an action as its output. While the outcome of an action is not fully predictable due to the randomness of the environment, there is no uncertainty about the state we are currently in, hence the name “fully observable” MDP.

An MDP is defined by a 4-tuple M = <$\mathcal{X}$, $\mathcal{U}$, T, G> [21], where

- $\mathcal{X}$ is a state space, which has all the possible states the system can be in. In this thesis, we restrict ourselves to a finite state space.
- $\mathcal{U}$ is the control space consisting of a finite set of permissible actions. We also focus on an MDP with a finite control space.
- T is a state transition probability (TP), given for each state and action. The probability distribution T(i,u,j) is the probability of ending in state j ϵ $\mathcal{X}$ given that we start from state i ϵ $\mathcal{X}$, and take action u ϵ $\mathcal{U}$. This is shown in equation (1).

$$T(i,u,j) = P\{X' = j \mid U = u, X = i\} \quad (1)$$

where *X* and *X'* denote the current and next state, respectively. Hence, the TP is used to capture the probabilistic state evolution that depends on the system state and the control actions.

- G is the Reward function. This is the expected immediate reward for taking an action in each state. Hence, G(x,u) is the reward for taking action u, if the system is in state x.

For our research purpose, we focus on discrete MDPs. However, many of the approaches developed herein could be extended to more general settings at the expense of more computational complexity.

MDPs have been extensively used in various applications. For example, MDPs were used in the optimization of grid-connected photovoltaic systems in [22-24]. In [25, 26] MDPs were used in systems with demand response technology. Optimal electricity supply bidding policies using MDPs were developed in [27]. Also, in [28] MDPs were also used to find optimal maintenance policies and in other applications in various fields [29-33]. Next, we provide more details about the mathematical structure and solution methodologies of MDPs.

State Space

As mentioned earlier, the state of the system captures all information needed by the controller to compute the optimal control actions. The state of the system changes in response to the controller actions. The set of all possible states of a

system are known as the state space. Figure 4: MDP Timeline illustrates the structure of an MDP at time k. The state evolution equation is shown in equation (2)

$$X_{k+1} = f_k(X_k, U_k, W_k), \qquad k=0,1, \dots, N\text{-}1, \tag{2}$$

where X_k represents the state of the system at time k, U_k ϵ $\mathcal{U}$ denotes the control, and W_k is an exogenous input known as disturbance or noise that captures the uncertainty in system evolution. N is the total number of stages so that the last control action is taken at stage N-1. It is important to note that (2) is an equivalent representation to the TP introduced earlier in (1). In particular, the state evolution describes the system dynamics through the probabilistic evolution to a future state based on the current state and control action. If the system is in state X_k and we use action U_k, since W_k is random, the state evolution can also be represented using the aforementioned transition matrix T. The uncertainty stems from the disturbance, which is unknown to the controller prior to deciding on a particular control action. Based on the state, action, and the realization of the disturbance, a cost (or reward) is incurred, which is denoted $g_k(X_k, U_k, W_k)$.

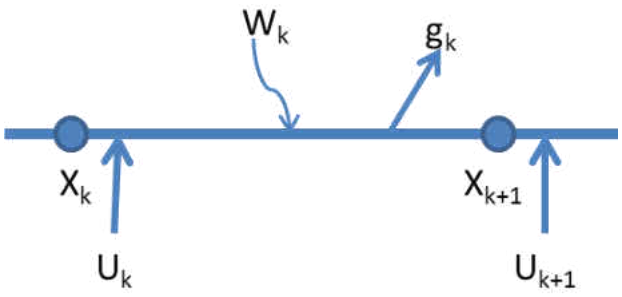

Figure 4: MDP Timeline

Control Space

We also need to define the control space. The control space $\mathcal{U}(X_k)$ is the set of allowable control actions that a controller can take at time k when the system is in state X_k. We denote the control action at time k by U_k.

The cost function $g_k(X_k, U_k, W_k)$ takes into account the current state of the system and the action taken at this state. The total cost is additive, meaning that the cost incurred at time k, accumulates over time. If the MDP has only a finite number *N* of stages then we have a *finite horizon problem*. Also let $g_N(X_N)$ denote the terminal cost that is acquired at the last stage when the problem terminates. Note that, because of the randomness in the system, the total cost is a random variable. Hence, in order to optimize over the choice of controls we use an expected total cost, shown in (3).

$$E\left\{g_N(X_N) + \sum_{k=0}^{N-1} g_k(X_k, U_k, W_k)\right\} \tag{3}$$

The goal of the decision maker is to minimize the total expected cost (or maximize the total reward). The optimization problem is over the choice of *policies*. It is important to note that the actions can have a long term effect since the current control not only affects the current cost, but also the state evolution. As such, the decision-maker has to choose controls that balance the tradeoff between the current stage costs and the future rewards.

Finite Horizon

In finite horizon problems we have a finite number *N* of stages after-which the problem terminates. In this case, the goal is to design *N* control policies $\mu_k, k = 0, 1, \dots, N-1$. A policy at time k is denoted μ_k and is defined as a mapping from the state space to the control space, i.e. μ_k: $\mathcal{X}_k \rightarrow \mathcal{U}_k$. Hence the optimization problem can be written as in (4).

$$J^*(X_0) = \min_{\mu_0,\dots,\mu_{N-1}} E\left\{ g_N(X_N) + \sum_{k=0}^{N-1} g_k(X_k,\ \mu_k(X_k), W_k\ ) \right\} \tag{4}$$

One of the most important features that characterize the optimal solution to a finite horizon problem is that the optimal policy is time-variant. This is rather intuitive since the optimal policy should change as we approach the termination. This is in clear distinction to infinite horizon problems, which will be discussed next.

We note that it is convenient to use transition probabilities for discrete-state systems. The transition probability T(x,u,x') is the probability of going from state x to state x' taking action u. Here the probability of transition is shown in equation (5)

$$P\{X_{k+1} = x' \mid U_k = u, X_k = x\} = P(x'|x, u\ ) \tag{5}$$

In conclusion a discrete-state system can be described in terms of transition probabilities or, equivalently a state evolution equation.

Infinite Horizon

In some cases, we have an infinite number of stages. These problems are known as infinite horizon problems. These could also be used to approximate finite horizon problems involving a finite but extremely large number of stages. In contrast to finite horizon formulations, the optimal solution to infinite horizon problems is an time-invariant policy, i.e., the **optimal rule** for choosing actions does not change from one stage to the next [20].

There are various mathematically viable formulations for infinite horizon problems. One of the most well known formulations is the discounted cost formulation, where we aim to minimize the expected total discounted cost over an infinite number of stages. This formulation is given in equation (6).

$$J_{\mu}(X_0) = \lim_{N\to\infty} E\left\{\sum_{k=0}^{N-1} \alpha^k g\,(X_k, \mu(X_k), W_k)\right\}. \qquad (6)$$

Here, $J_{\mu}(X_0)$ represents the cost associated with the initial state X_0 and a policy μ, and α is a positive number between 0 and 1, called the discount factor. The discount factor is less than one because the future costs have less weight than the cost incurred at the present time [34]. This formulation ensures that the problem is well-defined mathematically, i.e. the total cost does not grow to infinity, provided that the per-stage cost g is bounded above by a constant M and $\alpha < 1$. In this case, the cost $J_{\mu}(X_0)$ is well defined because it is the infinite sum of a sequence of numbers that are bounded in absolute value by the decreasing geometric progress $(\alpha^{\mathrm{k}}\mathrm{M})$ [34].

Dynamic Programming

Thus far, we have described the objective function in both finite and infinite horizon problems. The question remains as of how to solve the corresponding optimization problems given that the optimization is over the choice of policies, which are mappings from the state space to the control space.

One algorithm known as Dynamic Programming transforms this complex problem into a sequence of simpler problems. Dynamic programming is based on the principle of optimality. The principle of optimality states that an optimal policy has the property that whatever the initial state and action are, the remaining actions must constitute an optimal policy with regards to the state resulting from the first action [35]. It is well known that the optimal policy $= \{ _0, _1, \ldots, _{N\ 1}\}$ can be obtained using the recursive dynamic programming algorithm.

For every initial state X_0, the optimal cost $J^*(X_0)$ of the basic problem is equal to $J_0(X_0)$, obtained through the recursive dynamic algorithm equation (7), which proceeds backward in time from period N-1 to 0.

$$J_k(X_k) = \min_{U_k \in \mathcal{U}_k(x_k)} E\{g_k(X_k, U_k, W_k) + J_{k+1}(f_k(X_k, U_k, W_k))\}, \quad (7)$$

$$k = 0,1, \ldots, N-1,$$

$$J_N(X_N) = g_N(X_N) \quad (8)$$

In dynamic programming the combination of the current stage cost with the future cost is shown in equation (8). In the finite horizon setting, we have time-varying policies; hence, for each hour we have a different optimal policy.

In the infinite horizon problem, the optimal solution satisfies the celebrated Bellman equation

(9). This is a fixed-point equation that defines the optimal value function. In an infinite horizon problem, the optimal policy and the optimal value function are stationary (time-invariant). The optimal costs $J^*(1),\ldots,J^*(n)$ are the unique solution of this equation in the discounted cost formulation, where n is the number of states, i.e., the State Space = $\{0,1,\ldots,n\}$.

$$J^*(X) = \min_{U\in\mathcal{U}(X)} E\left[g(X,U) + \alpha \sum_{X'=1}^{n} P(X'|X,U)\, J^*(k)\right], \qquad X = 1,\ldots,n, \tag{9}$$

Also, the costs $J_\mu(1)$, …, $J_\mu(n)$ are the unique solution for the fixed-point equation given in

(10) for an arbitrary stationary policy μ.

$$J_\mu(X) = g(X,\mu(X)) + \alpha \sum_{X'=1}^{n} P(X'|X,\mu(X))\; J_\mu(X'), \qquad X = 1,\ldots,n, \tag{10}$$

CHAPTER THREE: LOAD SHEDDING AND RESOURCE ALLOCATION POLICIES BASED ON MDPs

In this chapter, we formulate the load shedding and resource allocation problem in the presence of a renewable energy source and energy storage using the MDP framework. We show how the different elements of our problem (i.e, the battery, wind energy, and the hourly load demand) are mapped to the aforementioned MDP framework.

State Space

First, we need to define our state space. The state space in our experiment corresponds to the battery level. In our experiment we have a 1MW battery. We split this battery into 100kW levels. Hence, we have a total of 11 battery levels, which correspond to 11 different states. We assume that the battery has a maximal capacity that cannot be exceeded. Hence, the maximum charge the battery can hold is 1MW. The battery can be charged to any level in one time step. In other words, it can go from 0kW to 1MW after one time step if enough power is available to charge the battery. Figure 5, illustrates the battery and all its possible states.

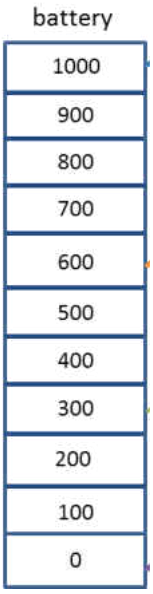

Figure 5: Battery levels shown in kW

One time step is equal to one hour. We will simulate one day, hence we will have 24 time intervals. The control action determines how much of the battery will be used at a given stage. In equation (11), we define the state evolution equation as shown previously in equation (2):

$$X_{k+1} = min[X_k + W_k - E(U), C_{max}] \quad (11)$$

Here, X_k is the current battery state, W_k is the wind energy that we harness, $E(U)$ is the amount of battery power used in 1 hour by taking an action U. The table for $E(U)$ is shown in Table 1, and C_{max} is the maximum battery capacity as mentioned earlier, i.e., 1000kW. In other words, in order to determine the next battery level we need to know the previous battery level that we start with. We take an action that determines how much of the battery we will use. After we take an action the wind realizes itself and we know how much power we can harness

from the wind energy. This gets added to the battery and we obtain our new battery level. Depending on the action we chose we will have to pay a cost.

Table 1: Power Used by Each Action possible

Action (U)	Power Used (kW)
1	500
2	400
3	300
4	200
5	100
6	0

Control Space

As shown in the state evolution equation we have to take an action at the beginning of each time interval. These actions have to satisfy the load demand after each hour. The demand load is fixed at 500kW. With this in mind, we have six actions at each hour. These actions consist of either using the battery, shedding some load or a combination of both. Each action we take has an associated cost. Table 2 describes the six actions and the cost associated with each action. Under each action we can see how much of the battery (B) is used and how much of the load is shed (L). As mentioned before, the battery has 11 states of 100kW increments and each has a cost associated with it. For example, if action 1 is used, we utilize the battery to supply the whole 500kW required for 1 hour to satisfy the demand, hence utilizing 500 kWh. In comparison to action 1, action 5 utilizes only 100kW from the battery in 1 hour and we shed 400kW of load. The actions we take always add up to 500kW.

From each battery state we are currently in there is a probability that we will end up in a new battery state for every permissible action. We calculated these probabilities based on the probability of the wind speed *V*. Since we have 11 battery states we create an 11x11 probability matrix for each action. In order to calculate the transition probability for each state, we first have to determine the wind speed probability. This is shown in equation (12), which has two parameters k and c, which represent the shape and scale parameters, respectively. We calculate the transition probabilities based on a Weibull probability density function in (13), which is typically used to model wind speed probability [36, 37].

$$f(V|k,c) = \frac{k}{c}\left(\frac{V}{c}\right)^{k-1} e^{-(V/c)^k} \tag{12}$$

In order to represent the probability distribution for the wind velocity, we use the Rayleigh distribution, which is a special case of the Weibull probability distribution with k = 2, as shown in equation (13). Here, *V* is the velocity or the wind speed that is changing along the x-axis and c is the mean of the wind speed. Hence, the wind is stochastic and cannot be fully predicted; this corresponds to the disturbance or noise in the MDP model, which affects the state evolution and hence affects our hourly decisions.

$$f(V) = \frac{2V}{c^2} e^{-(V/c)^2} \tag{13}$$

Table 2: Cost Matrix for Each Action

Action 1			Action 2			Action 3	
B=500	L=0		B=400	L=100		B=300	L=200
States	Cost		States	Cost		States	Cost
0	NA		0	NA		0	NA
100	NA		100	NA		100	NA
200	NA		200	NA		200	NA
300	NA		300	NA		300	500
400	NA		400	500		400	400
500	500		500	400		500	300
600	400		600	300		600	200
700	300		700	200		700	100
800	200		800	100		800	100
900	100		900	100		900	100
1000	100		1000	100		1000	100
Action 4			**Action 5**			**Action 6**	
B=200	L=300		B=100	L=400		B=0	L=500
States	Cost		States	Cost		States	Cost
0	NA		0	NA		0	500
100	NA		100	500		100	400
200	500		200	400		200	300
300	400		300	300		300	200
400	300		400	200		400	100
500	200		500	100		500	100
600	100		600	100		600	100
700	100		700	100		700	100
800	100		800	100		800	100
900	100		900	200		900	200
1000	200		1000	200		1000	200

Table 3 and Figure 6 show the hourly average wind speed based on measurements collected in a remote location in Montana over a period of two years [38, 39]. This is the daily average wind speed. By using this 24-hour window we can simulate the power system network on an average day. By knowing the

mean of each hour we can obtain their Rayleigh distribution functions with equation (13).

Table 3: Hourly Wind Speed Mean

Hour	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Avg wind speed	4.45	4.30	4.25	4.35	4.30	4.45	4.20	4.15
Hour	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Avg wind speed	4.45	4.75	5.30	5.80	6.00	6.25	6.65	6.70
Hour	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Avg wind speed	6.30	6.10	5.80	5.45	4.85	4.75	4.70	4.20

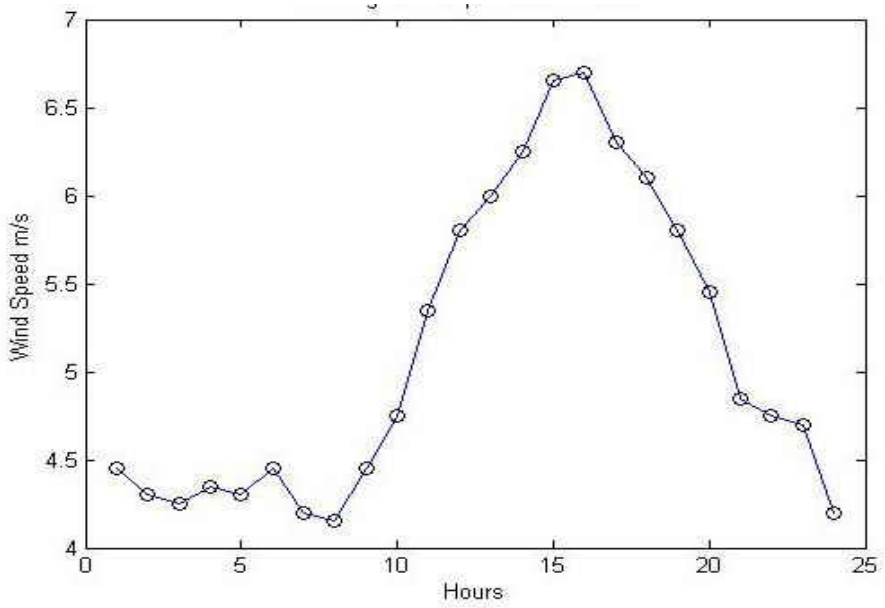

Figure 6: Wind Generation Profile

The power that the wind turbine generates can be calculated using equation (14). The power can be calculated if we know the two parameters *V* and *R, where V* denotes the velocity and *R* the radius of the wind turbine blades. The air density is denoted by ρ, which is a constant 1.2235 kg/m^3. In this research, we chose a wind turbine with a wind blade radius R=45 meters. Using equations (13) and

(14) we can calculate the probability that a certain amount of power is generated at a specific time.

$$P_w = 0.5 \cdot \rho \cdot V^3 \cdot \pi \cdot R^2 \quad (14)$$

Table 4 shows how much power is generated at each wind speed. In Figure 7 one can see that as the wind speed increases the wind power generated increases exponentially.

Table 4: Power generated at different wind speed

Wind speed (m/s)	Power Generated (W)		Wind speed (m/s)	Power Generated (W)
1	3.89E+03		7	1.33E+06
2	3.11E+04		8	1.99E+06
3	1.05E+05		9	2.84E+06
4	2.49E+05		10	3.89E+06
5	4.86E+05		11	5.18E+06
6	8.41E+05		12	6.73E+06

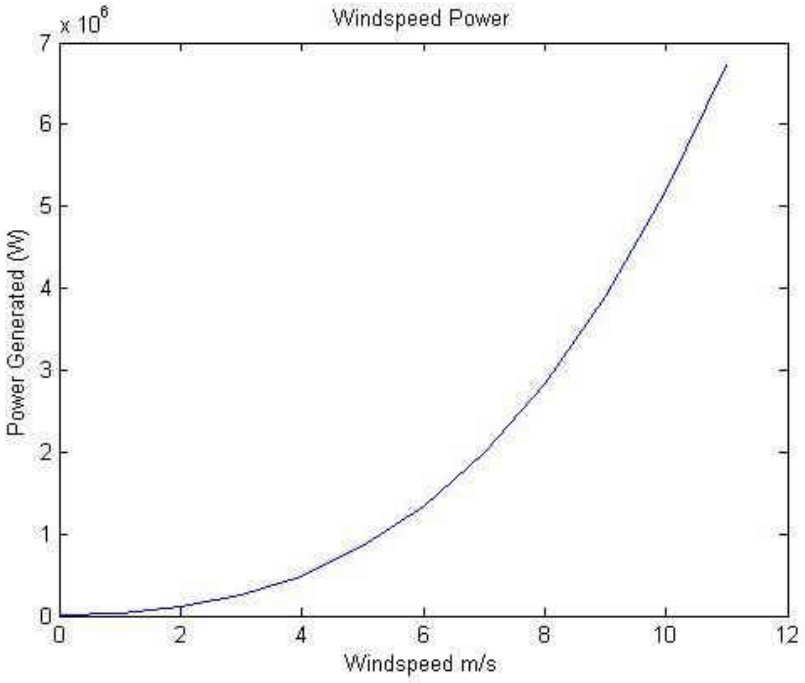

Figure 7: Wind speed Power

Transition Probability

Since the wind behaves in a stochastic manner, depending on the amount of power that we can extract from the wind we can either use power from the battery or store the power harnessed from the wind in the next stage. The wind distribution that we calculated for the first hour is shown in Figure 8.

Each transition is based on the amount of wind power that is generated. Depending on that amount of power, the battery level can increase or decrease. With this wind probability we can create the transition probability for the first hour for action 1, as shown in Table 5.

Table 5: Transition Probability

	0	100	200	300	400	500	600	700	800	900	1000
0	0.766	0.034	0.028	0.023	0.019	0.016	0.014	0.012	0.010	0.009	0.069
100	0.724	0.042	0.034	0.028	0.023	0.019	0.016	0.014	0.012	0.010	0.078
200	0.670	0.054	0.042	0.034	0.028	0.023	0.019	0.016	0.014	0.012	0.088
300	0.600	0.070	0.054	0.042	0.034	0.028	0.023	0.019	0.016	0.014	0.100
400	0.503	0.097	0.070	0.054	0.042	0.034	0.028	0.023	0.019	0.016	0.114
500	0.356	0.147	0.097	0.070	0.054	0.042	0.034	0.028	0.023	0.019	0.130
600	0.000	0.356	0.147	0.097	0.070	0.054	0.042	0.034	0.028	0.023	0.149
700	0.000	0.000	0.356	0.147	0.097	0.070	0.054	0.042	0.034	0.028	0.172
800	0.000	0.000	0.000	0.356	0.147	0.097	0.070	0.054	0.042	0.034	0.200
900	0.000	0.000	0.000	0.000	0.356	0.147	0.097	0.070	0.054	0.042	0.234
1000	0.000	0.000	0.000	0.000	0.000	0.356	0.147	0.097	0.070	0.054	0.276

If the wind is blowing below 2.95 m/s we will not be generating enough power to increase the battery energy storage. In Figure 8, one can see that below the 2.95 m/s threshold, we will not receive any additional power from wind power. This implies that in the next hour if we do not have 500kW of power in the battery we

would have to shed some additional load. But in a scenario where the wind blows faster than 2.95 m/s, we can use the additional power generated to charge the battery, since in this scenario we can generate power.

For example, in the zone between 100kW and 500kW we are generating up to 500kW of energy that can be used to store into the battery. As the wind speed increases we can generate more power, which can be stored in the battery for future use.

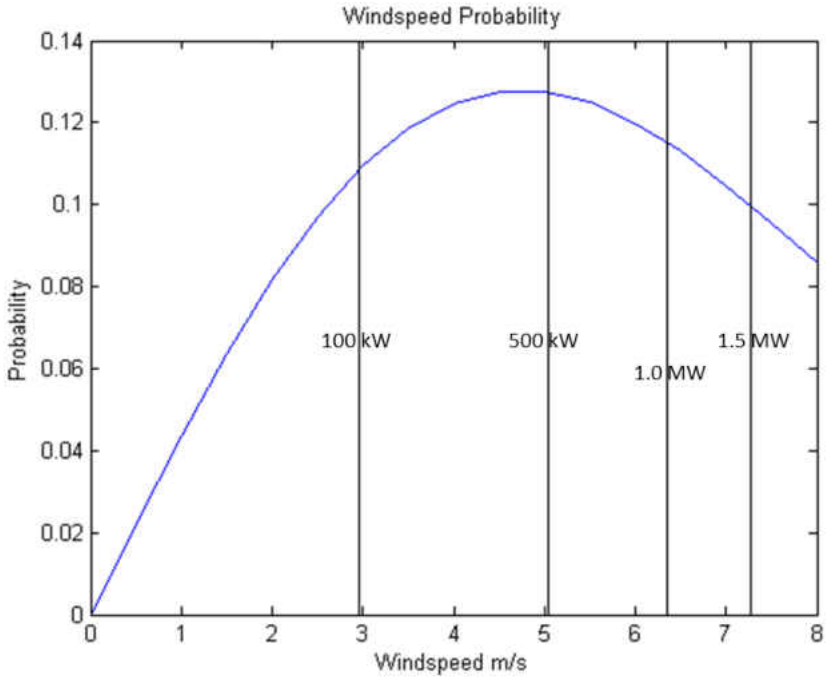

Figure 8: Wind Speed Probability

Finite Horizon

We consider both the finite and infinite horizon formulations. In the finite horizon problem we have a total of 24 stages. When solving the finite horizon problem we implement the dynamic programming algorithm using backward iterations. We start from the 24th hour and work our way backwards towards the 1st hour. At the 24th hour we will have a terminal cost. Based on our previous discussion, in our problem we use equation (15). Here we calculate the cost function of each action. We choose the action that has the lowest cost—note that the total cost consists of the current stage cost and the future cost to go. This is shown in equation (4). We repeat this process for each battery state. Since we are implementing backward dynamic programming, once we are done calculating the cost for each state we proceed backwards to the previous hour.

$$J_k(X_k) = \min_{U_k \in \mathcal{U}(X_k)} [g(X_k, U_k) + E\{J_{k+1}(X_{k+1})\}] \quad (15)$$

Infinite Horizon

In the infinite horizon formulation, we need to find a stationary value function and a stationary control policy. Each stationary policy μ is associated with a linear function $g_\mu + \alpha P_\mu J$ of the vector J, and TJ is the piecewise linear function $min_\mu[g_\mu + \alpha P_\mu J]$. The optimal cost J^* satisfies $J^* = TJ^*$. T_μ and T are contraction mappings that converge to unique fixed points J_μ and J^*, where J^* is the optimal value function.

In order to solve the infinite horizon problem we implement the policy iteration algorithm. The policy iteration algorithm generates a sequence of stationary policies. After each iteration, the cost is improved. This policy iteration algorithm consists of three steps.

Step 1: The Initialization. In this step we start with a guess of an initial policy μ^0.

Step 2: Policy Evaluation. In the second step we compute the corresponding cost function J_{μ^k} from the linear system of equations (16) (initially k = 0). In this equation I, is an identity matrix, P_{μ^k} is the transition probability for each policy, and g_{μ^k} is the cost associated with each decision made at each state. α is the discount factor. The dimension of this system is equal to the number of states. This implies that when there are a large number of states, this approach may be too computational.

$$(I - \alpha P_{\mu^k})\, J_{\mu^k} = g_{\mu^k} \tag{16}$$

Step 3: Policy Improvement. In the final step we obtain a new policy μ^{k+1} that satisfies equation (17).

$$T_{\mu^{k+1}}\, J_{\mu^k} = T J_{\mu^k} \tag{17}$$

If the cost of the new policy μ^{k+1} is different from the previous policy μ^k we go back to step 2 and perform policy evaluation again. We keep iterating over these

2 steps until we attain the optimal policy. When the new policy is equal to the previous policy, the cost will not change. At this point, we would reach the optimal policy.

CHAPTER FOUR: NUMERICAL RESULTS

In this section, we present numerical results using the proposed approach.

Finite Horizon

The MDP evolution schematic that we used is illustrated in Figure 9. This is a simplified schematic showing the first two stages. The battery can increase, decrease or maintain its current power level depending on the action taken by the operator, as well as the power obtained from the wind. For each action we created a transition probability matrix.

Since we are doing backwards iteration, the first step is to choose the policy at the last stage and the cost associated with that policy. The policy chosen is shown in Table 6. The cost for this hour is the terminal cost, which requires no calculation, since each state has its own terminal cost.

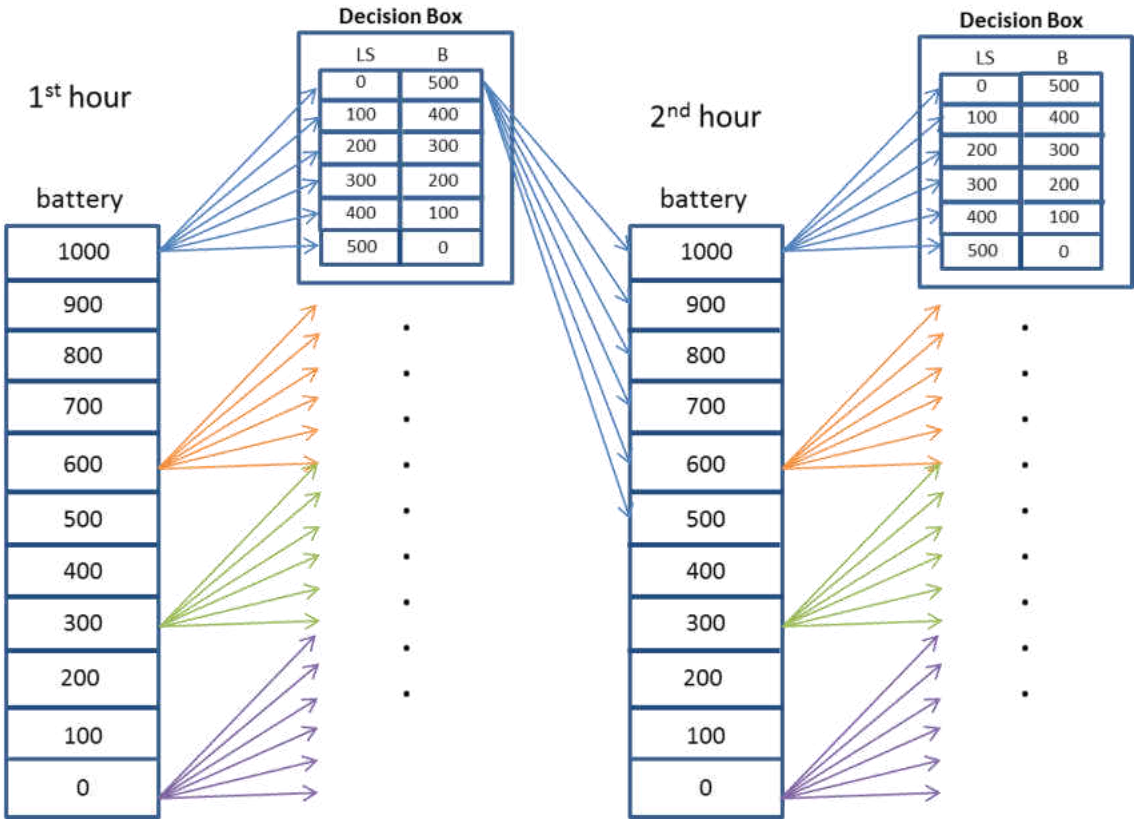

Figure 9: MDP State Diagram

Table 6: Cost and policy for 24th hour

State	0	100	200	300	400	500	600	700	800	900	1000
Cost	500	400	300	200	100	100	100	100	100	100	100

The next step is to calculate the cost for each state from the previous hour. Starting with the first state we compute the expected cost for each action. In Table 7 we can see the expected cost for each action for state 1 (0 battery level). The next step is to apply equation (15) to choose the minimum cost, which in this scenario would be to take action 6. We go through all eleven states and record the action with the lowest cost. Once we go through all eleven states we obtain an optimized policy for that hour. In Table 78 we display the result after

optimizing each state for the 23rd hour. We can also see the expected cost obtained from Table 7 in the first state value function, shown in Table 8.

Table 7: State 1 Expected costs

Action	1	2	3	4	5	6
Cost	NA	NA	NA	NA	NA	798.02

Table 8: Cost and Policy for 23rd hour

State	0	100	200	300	400	500	600	700	800	900	1000
Action	6	5	6	6	6	5	4	3	2	1	1
Cost	798.02	798.02	479.1	332.58	200	200	200	200	200	200	200

These steps are repeated until we reach the first hour using the dynamic programming recursion in equation (17). To illustrate how the actions are chosen for six particular states, Figure 10 shows the optimal control actions for states 500 to 1000 for the different hours. Every hour a new action is chosen. The states 100 to 400 are not illustrated since the main action chosen in these states is action 6. This way we obtain 24 different optimized policies corresponding to the different hours that will minimize the total cost as shown in Table 9. This table shows the optimal actions if we are in any given state at each hour. With this we also generate a cost for each state. By adding up the costs at each hour we obtain the value function, which is a vector of length n, i.e., one entry for each state. The value function (expected total cost) is shown in Figure 11 and in Table 10. Intuitively, the expected total cost is smaller for higher initial battery levels.

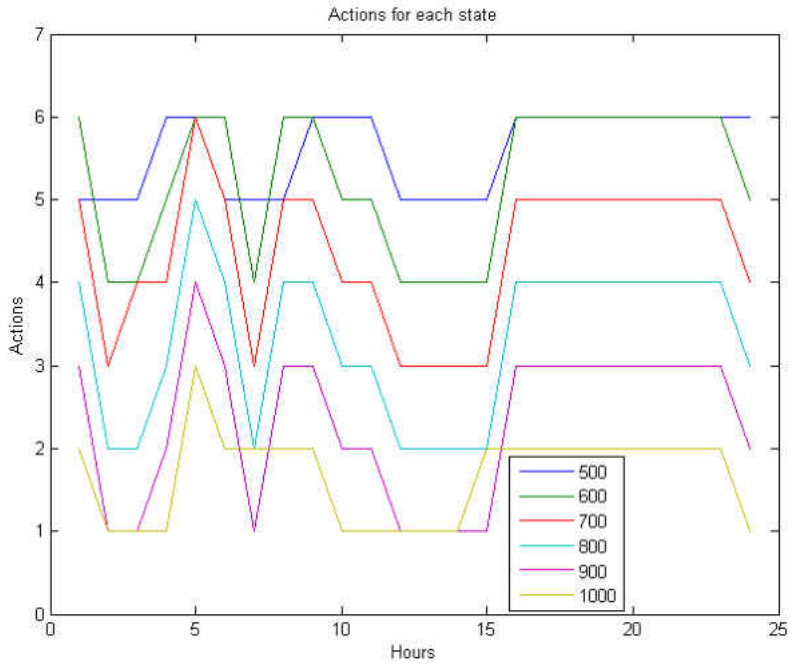

Figure 10: Action Chosen for each State

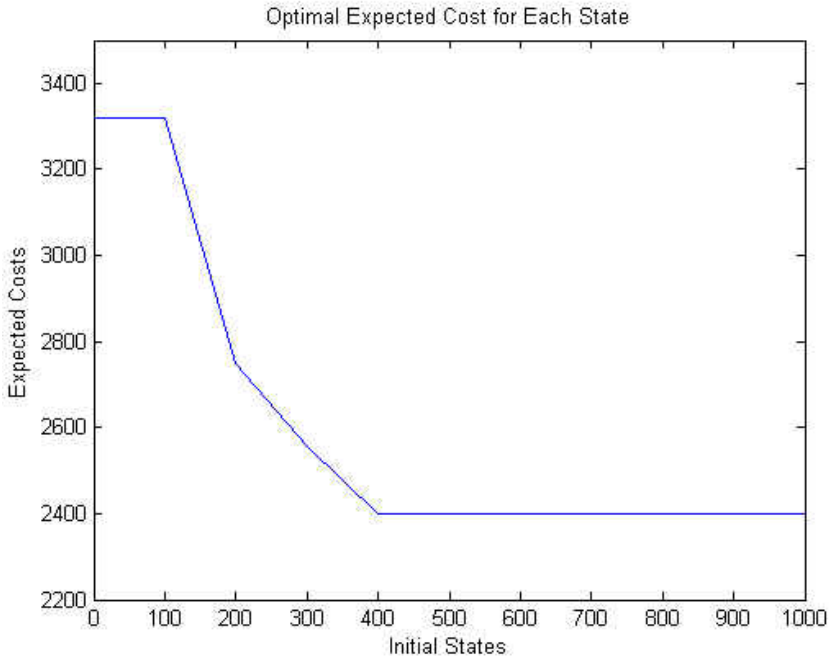

Figure 11: Finite Total Cost per State

Table 9: Policy for each hour

Hour	0	100	200	300	400	500	600	700	800	900	1000
1	6	5	6	6	6	5	6	5	4	3	2
2	6	5	6	6	6	5	4	3	2	1	1
3	6	5	6	6	6	5	4	4	2	1	1
4	6	5	6	6	6	6	5	4	3	2	1
5	6	5	6	6	6	6	6	5	4	3	2
6	6	5	6	6	6	5	6	5	4	3	2
7	6	5	6	6	6	5	4	3	2	1	2
8	6	5	6	6	6	5	6	5	4	3	2
9	6	5	6	6	6	6	6	5	4	3	2
10	6	5	6	6	6	6	5	4	3	2	1
11	6	5	6	6	6	6	5	4	3	2	1
12	6	5	6	6	6	5	4	3	2	1	1
13	6	5	6	6	6	5	4	3	2	1	1
14	6	5	6	6	6	5	4	3	2	1	1
15	6	5	6	6	6	5	4	3	2	1	2
16	6	5	6	6	6	6	6	5	4	3	2
17	6	5	6	6	6	5	6	5	4	3	2
18	6	5	6	6	6	5	4	3	2	1	1
19	6	5	6	6	6	5	4	3	2	1	1
20	6	5	6	6	6	5	4	3	2	1	1
21	6	5	6	6	6	6	5	4	3	2	1
22	6	5	6	6	6	6	5	4	3	2	1
23	6	5	6	6	6	5	4	3	2	1	1

Table 10: Total Expected Cost after 24 hours.

State	0	100	200	300	400	500	600	700	800	900	1000
Cost	3318.4	3318.4	2750.1	2555.8	2398.8	2398.8	2398.5	2398.5	2398.5	2398.5	2398.5

Simulation – Finite Horizon

Despite the complexity of the DP recursions, it is very important to note that the computational burden is all carried out offline since the controller computes the policies (mappings) not the actions. Once the mappings are determined, the

control actions are simply obtained in real time by mapping the actual state to the optimal action.

Now that we obtained the optimal policy for each hour, we can perform a simulation of multiple trajectories and observe the average cost over a large number of Monte-Carlo runs. In each simulation we simulate one day and observe the cost associated with each action that is taken at each hour. By adding the costs at each hour we obtain the total cost for 24 hours, which in this experiment is the cost of 1 simulation run. In this analysis we ran the experiment 1000 times and took the average of the 24 hours simulated. We ran the simulation for each battery state and observed the average cost, which is demonstrated in Table 11 and in Figure 12. The average cost for each state is compared to the optimal value function obtained in Figure 11. It is shown that the results of analysis and simulation are in agreement.

Table 11: Average Simulation Cost for Each State

	State	0	100	200	300	400	500	600	700	800	900	1000
Sim	Cost	3413.2	3307.2	3091.2	2758.3	2560	2400	2400	2400	2400	2400	2400

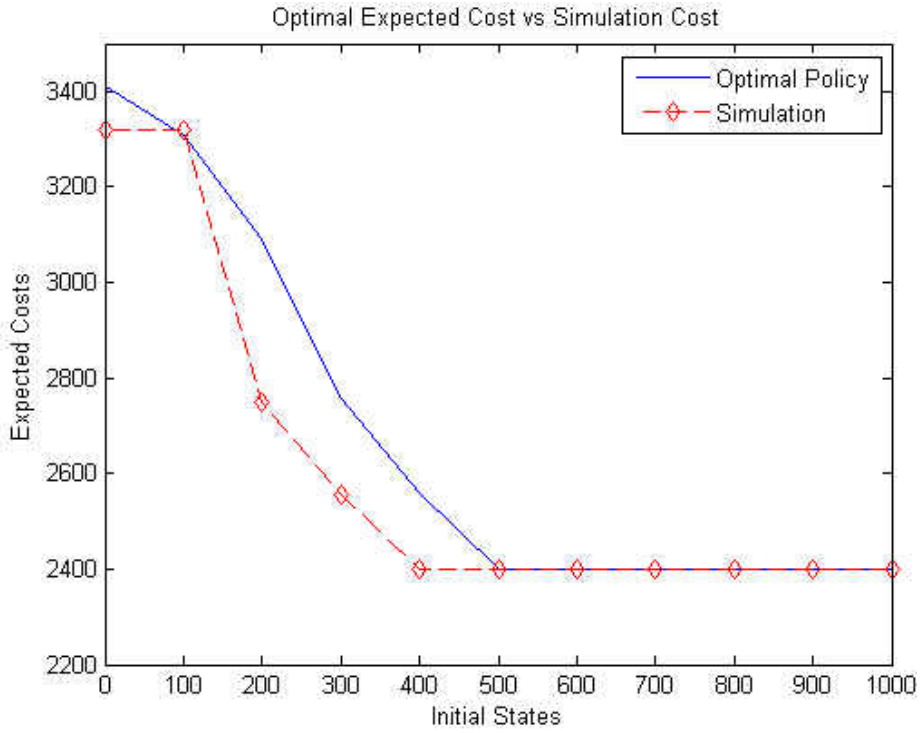

Figure 12: Finite Horizon Simulation vs Optimal Policy Cost

In these simulations we also observed the wind speed with the corresponding wind power generated. Note that the wind speed is not stationary as it depends on the time of the day. Every hour has a different probability distribution. Therefore, the probability to generate a certain amount of energy keeps changing. Figure 13 depicts a realization (a sample path) of the wind speed and generated energy over the course of one day. The blue axis corresponds to the wind speed as it changes throughout the day. The wind speeds at different hours are independent random variables but not identically distributed as they are drawn from different probability distributions. This simulation is consistent with the average wind speed that we obtained in the previous section Figure 6. Between the 10^{th} and the 20^{th} hour there is an increase in wind speed in

agreement with the average behavior of the wind speed that we measured and described in the previous section. The green axis is for the power generated based on these wind speeds.

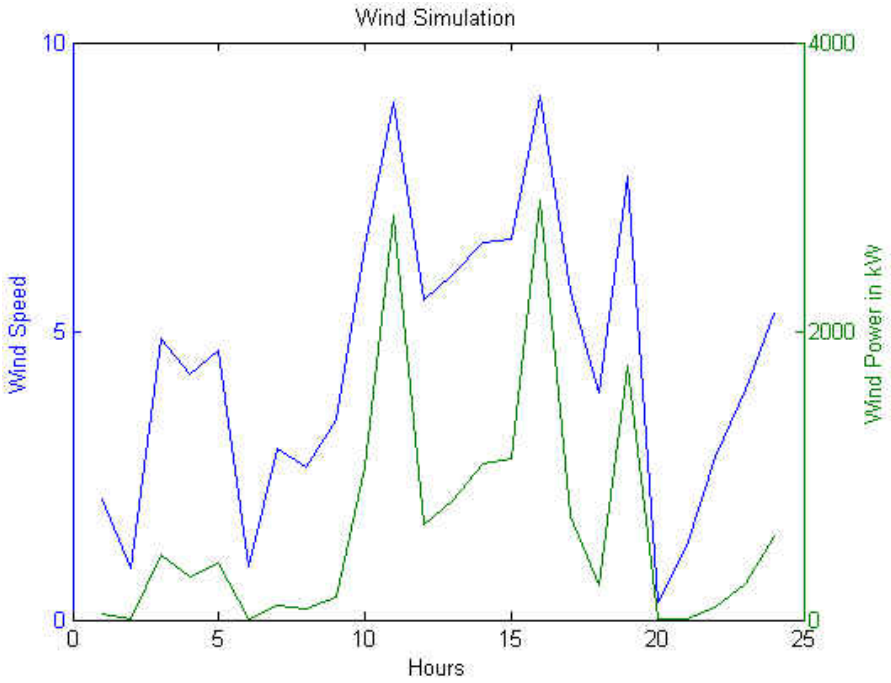

Figure 13: Wind Speed and Power Simulation

As we take an action we recognize that when we store the wind power it is generated back into the battery. During the course of the day one can see the power level fluctuates depending on the usage and storage, as shown in Figure 14. As mentioned before, the maximum storage is 1000kW so there are certain times when the battery stays at its maximum. The majority of those times come between the 10^{th} and 15^{th} hour, where the peak power is generated throughout the day.

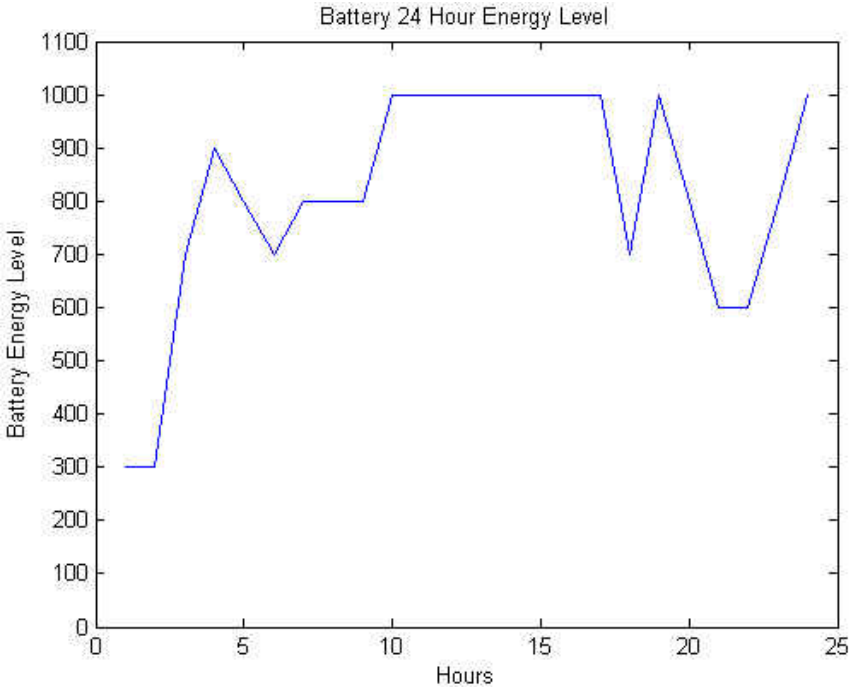

Figure 14: Battery Level during 24 Hours

Policy Iteration – Infinite Horizon

In this section, we carry out more experiments for the infinite horizon setting and adopt a discounted cost formulation. We used a discount factor $\alpha = 0.9$. Solving equation (16) gives the value functions for policy iteration for a fixed policy μ. The result of one iteration is shown in Table 12 for illustration. For the last hour we can start with the policy evaluation, equation (16), and then finish with policy improvement, equation (17).

Table 12: Value Functions for policy Iteration after one iteration

State	0	100	200	300	400	500	600	700	800	900	1000
Expected Cost	4298.8	4083.3	4082.2	1411.3	2360	1000.1	1000.1	1000.1	1000.1	1000.4	1000.1
Policy	2	3	2	6	4	6	5	4	3	2	1

Once we calculate the value function we can start calculating the cost for each state for each action chosen. This is done by utilizing equation (16). Since we are dealing with probabilistic evolutions we get an expected cost for each state. From here we chose the lowest cost per state, which defines the policy that achieves the minimum cost. The costs computed in the second iteration for each state are shown in Table 13. We can observe that the policy has changed and that the costs have reduced. We perform another iteration and obtain new values for the cost as shown in Table 14. Since this is an infinite horizon formulation we calculate the difference between the costs at consecutive iterations until we reach a point where the cost does not change anymore upon which the algorithm has converged to the optimal solution. The optimal action for each state is shown in Figure 15. The error between iterations is calculated until the mean square error falls below a small threshold. The difference between consecutive iterations is shown in Table 15. The mean square error keeps reducing to a point where the difference between iterations is negligible.

Table 13: Second Iteration Results

State	0	100	200	300	400	500	600	700	800	900	1000
Expected Cost	1770.2	1549.6	1352.4	1181.9	1181.9	999.2	999.1	999.1	999.1	999.1	999.1
Policy	6	6	6	6	6	6	6	5	4	3	3

Table 14: Third Iteration Results

State	0	100	200	300	400	500	600	700	800	900	1000
Expected Cost	1744.6	1522.2	1322.1	1146.4	999.3	999.2	999.1	999.1	999.1	999.1	999.1
Policy	6	6	6	6	6	6	6	5	4	3	3

Table 15: Error Mean Square

Iteration	Error
1	6.62E+06
2	1.54E+04
3	3.75E-25

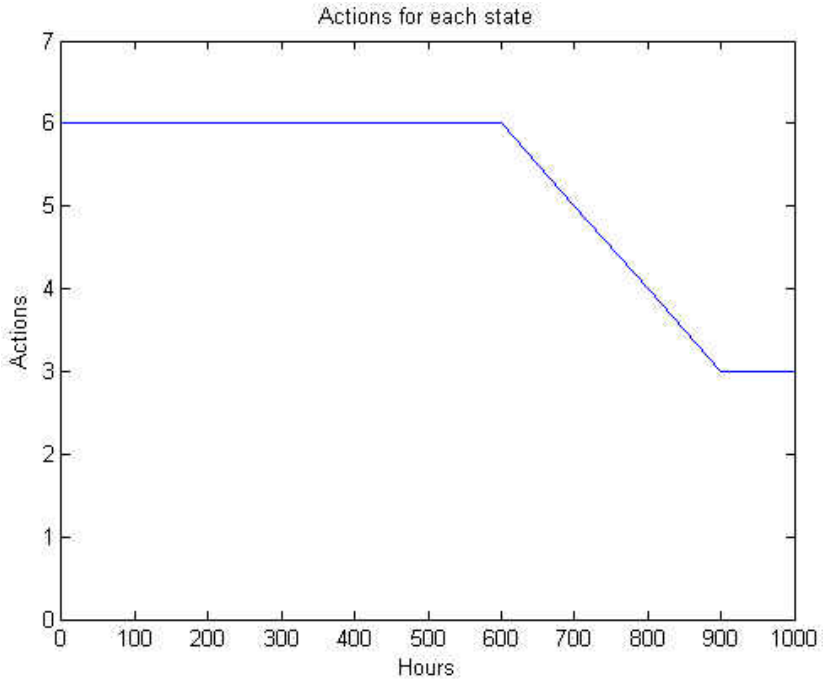

Figure 15: Infinite Horizon each State Action

Simulation – Infinite Horizon

In this section, we simulate trajectories in the infinite horizon setting using the optimal policy obtained through policy iteration. Ideally, we would be simulating the problem for an infinite number of stages. However, since we have a discount factor, equation (10), after a certain number of stages the cumulative cost will increase at a very negligible rate since costs further away in the future are heavily discounted. We used the policy obtained in Table 14.

We ran the simulation for each battery state and recorded the cost incurred averaged over a large number of Monte Carlo simulations. The results are shown in Table 16 and Figure 16. Again, one can observe that as the battery level increases the cost decreases.

Table 16: Infinite Horizon Cost Simulation

State	0	100	200	300	400	500	600	700	800	900	1000
Cost	5000	5000	4000	3000	2000	1000	1000	1000	1000	1000	1000

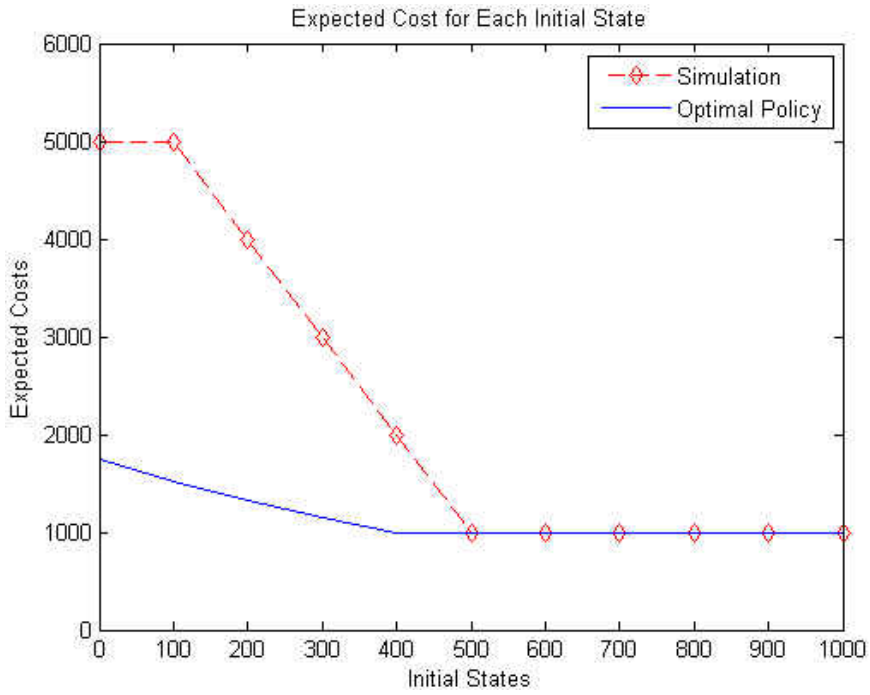

Figure 16: Infinite Horizon Simulation and Optimal Policy Cost for Each State

We can observe how the battery changes every hour as we simulate one day in Figure 17.

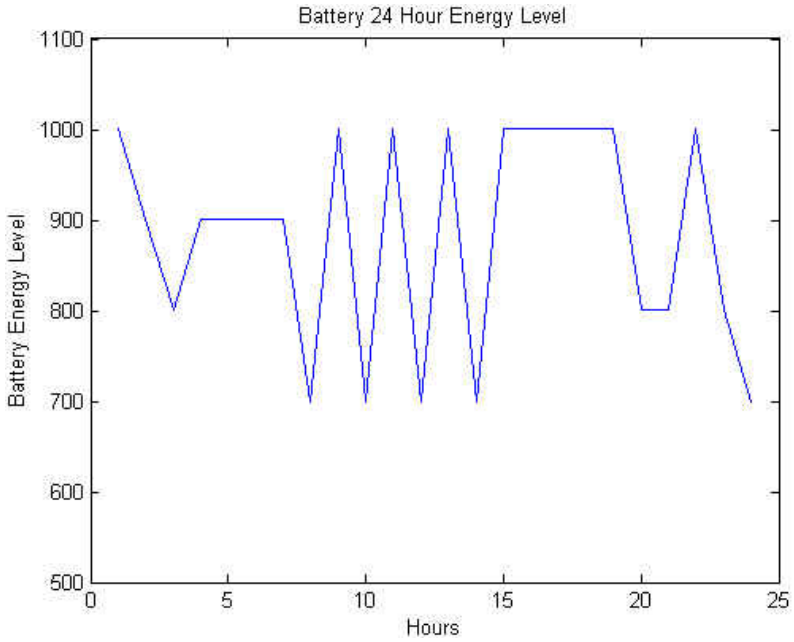

Figure 17: Infinite Horizon Hourly Battery Level

Control Policy – A: Load balancing Policy

In this section, we compare the computed policy to other policies. To this end, we calculated the average cost at each initial battery level for all policies. In order to test multiple scenarios we tested a policy that is inspired from the average daily consumer load curve. As shown in Figure 18 we can see the average daily commercial load curve [40, 41]. From this we can observe that from 8am to 6pm there is a big spike in the load curve. During this time it would be best to not perform load-shedding because it would interrupt businesses that require electrical power in order to continue normal operation.

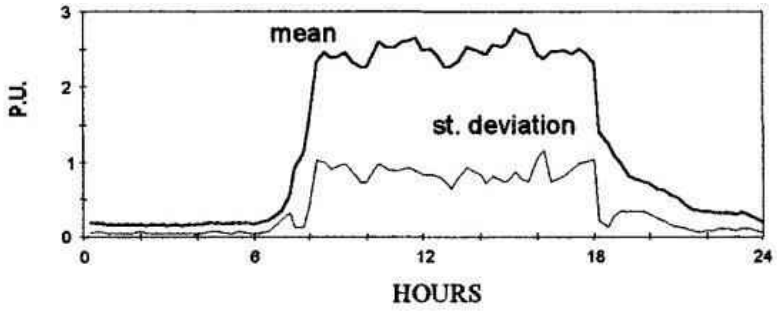

Figure 18: Average Daily Commercial Load

We term this policy, Control policy A. For this policy we take the action to load-shed when the battery is below 500kW. However, when we are between 8am and 6pm, we use the battery as much as possible. This means that if we are at 100kW, we use that remaining 100kW and shed the remaining load. During this time we drain the battery every hour if needed, to load-shed as little as possible. Here we also ran 11 different simulations that represent each battery level. The first simulation we ran started at 0kW battery charge. With this simulation we went through 1000 runs. Subsequently we kept increasing the initial battery level by 100kW for each simulation until we reached 100% battery charge level.

Control policy A consists of two policies of actions. We implement one policy between 8am until 6pm and another policy throughout the rest of the day. The actions taken in each policy for each state is shown in Figure 19. The main difference is that during the mid-day we are more aggressive in utilizing the battery than preserving it for when it is needed the most. One method to observe

these policies in action is to look how the battery level evolves throughout the day.

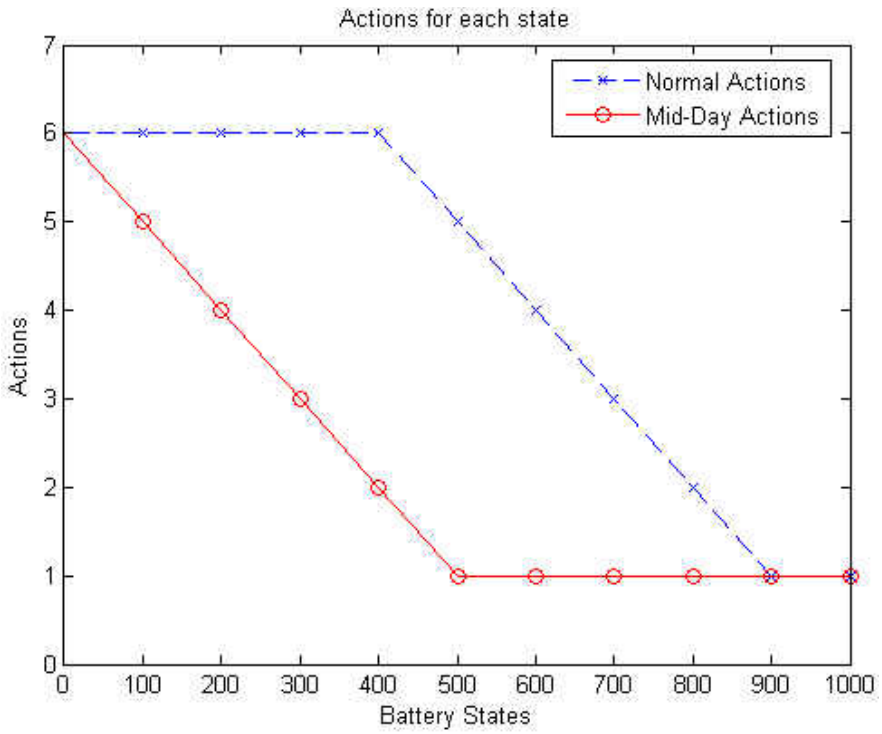

Figure 19: Actions for Normal and Mid-Day Policies

In Table 17 and Figure 20 we can see the result of this simulation in comparison to our MDP optimization policy results. We can see the average cost associated with each battery level. Figure 20 and Table 17 show the expected cost for each battery level of our simulated MDP policy along with that of Control policy A. This figure shows that the MDP policy is significantly better and more cost-effective in comparison to Control policy A. As the battery level increases the gap between the costs of the MDP and Control policy A also increases.

Table 17: Average cost at each state comparison for policy A

State	0	100	200	300	400	500	600	700	800	900	1000
Control Cost	5802	5563	5327	5136	4957	4967	5021	5009	4954	4986	4966
Optimal Policy Cost	3318	3318	2750	2556	2399	2399	2399	2399	2399	2399	2399

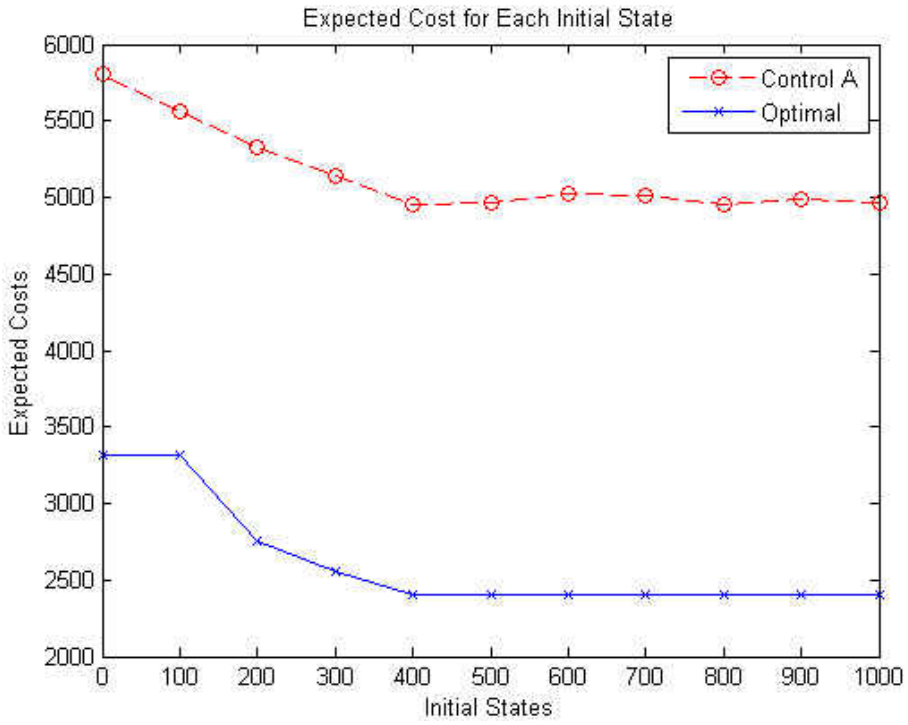

Figure 20: Control Policy A vs Optimal Policy Cost

In Figure 21 we can observe the battery level evolution throughout the day. During the early part of the day we can observe how we are more concerned about maintaining a power level above 500kW. Once we have reached 8am we are more aggressive with the battery utilization until we reach the end of the peak power consumer consumption. After 6pm we become again more conservative

with the battery power storage and try to maintain an average around 500kW or higher.

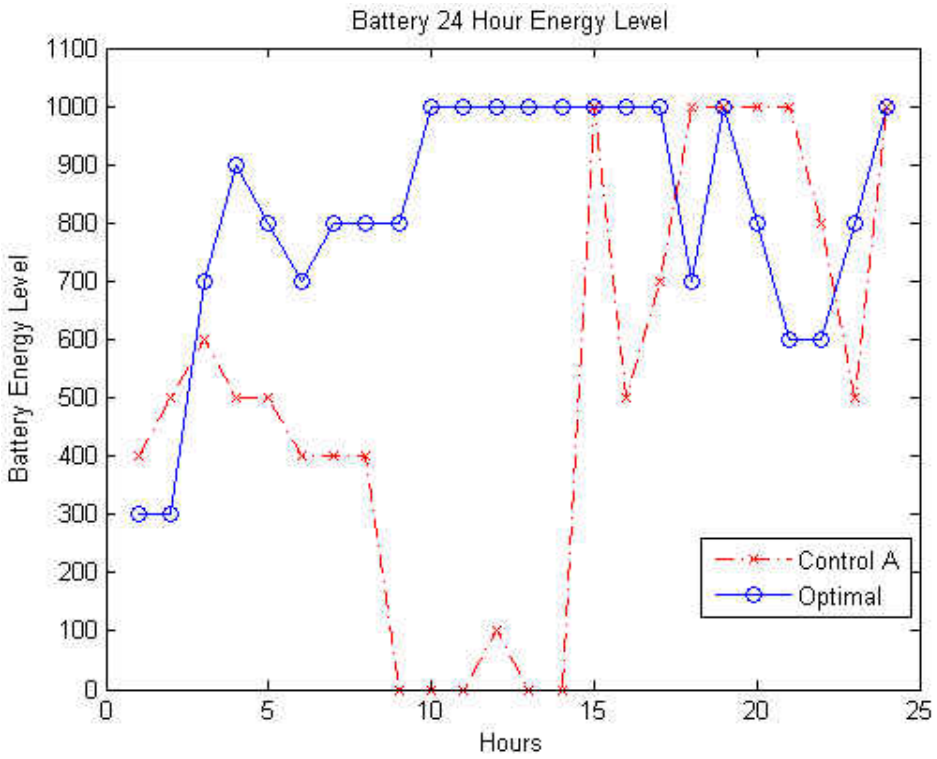

Figure 21: Control Policy Battery Level

Control Policy – B: Less Load-Shedding

The second control policy we compare to aims to maximize the usage of the battery throughout the day and to reduce load-shedding. When we reach a battery level below 300kW we shed the whole load. This policy is shown below in Figure 22. This figure shows that as soon as the battery level reaches 400kW we will only take action two. This is the action that uses 400kW of the battery and sheds 100kW. In this policy we do not worry if the battery is depleted. Our main

focus is *to use load-shedding less frequently*. When we reach a battery level above 500kW we only take action one, which is to use 500kW of battery.

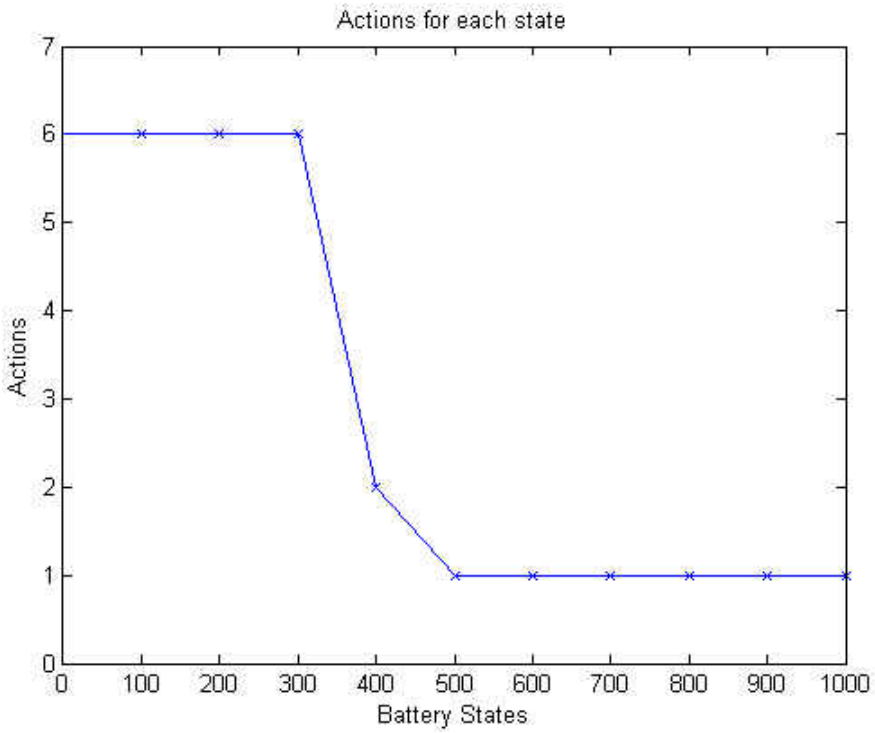

Figure 22: Actions for Control Policy 2

We simulate this policy to observe the average cost in each state in comparison to the optimal MDP policy. After running the simulation for each state around 1000 times we get the average cost that is shown in Figure 23. There is a clear difference in the cost for each state in comparison to our optimal policy. If we observe the costs incurred for the battery states between 400kW and 600kW we can see an increase in costs in comparison with the rest of the states. This is because the actions taken at those states have a higher cost than the other permissible actions. We also compare both policies A and B with the optimal

policy. Control B has a much higher cost than Control A, but both have a much higher cost when compared to the optimal policy.

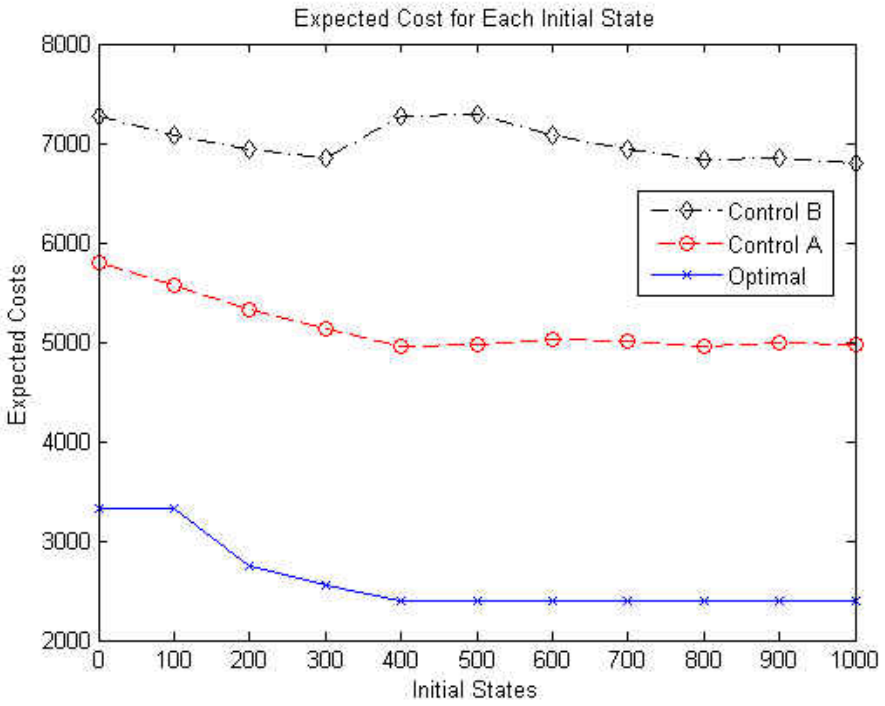

Figure 23: Cost for Control Policy B

Effect of Wind Speed

In this section, we study the effect of wind speed on the performance. To study the effect of wind on the system performance, we analyze the problem with an increase and decrease in wind speed. In order the study the decrease in wind speed we decrease the average wind speed by 1 at each time interval and for the increased wind speed we increase the wind speed average at each hour by 1. We first observe how the wind power changes for all three scenarios. As shown in Figure 24, we can see the amount of power generated under different wind speeds. In comparison to the original setting, the increased wind speed

results in higher wind power on average, especially in the middle of the day. The reduced wind speeds generate much lower average power throughout the day in comparison to the original wind speed.

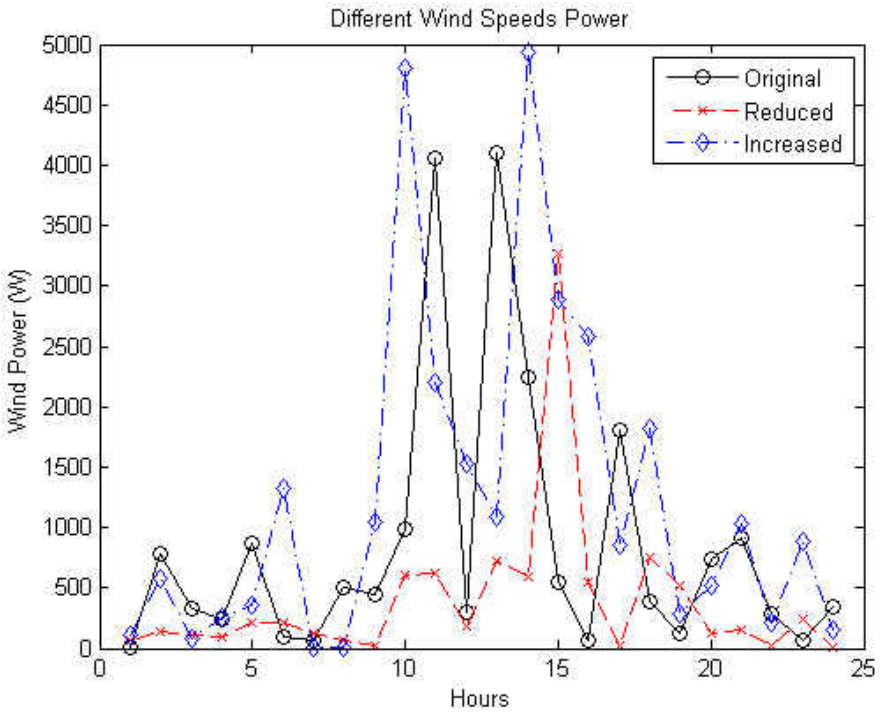

Figure 24: Different Wind Speeds

Next, we analyze the effect of wind speed on the control policies by looking at the actions for different battery levels. We observe the battery at 100kW, 300kW, 600kW, and 1MW charge level. In Figure 25, the hourly action taken for each battery level is shown. One main observation is that as the battery level increases the actions taken are more aggressive towards utilizing the battery.

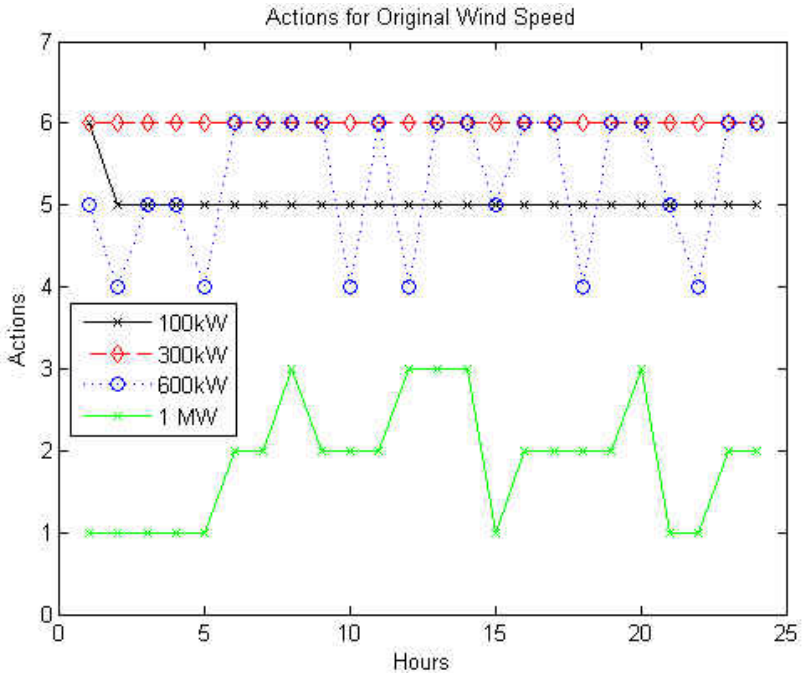

Figure 25: Hourly Action for Original Wind Speed

Now we consider the actions at reduced wind speed, which is shown in Figure 26. In comparison to the original wind speed the actions taken in a reduced wind speed region are less aggressive towards utilizing the battery. Since the wind does not deliver as much power we utilize the battery less often, especially during the hours of low wind probability, which run from 4 to 10 in the morning. Observing the 600kW battery line, it is seen that we shed more load throughout the day, especially in the morning hours in comparison to the original wind speed.

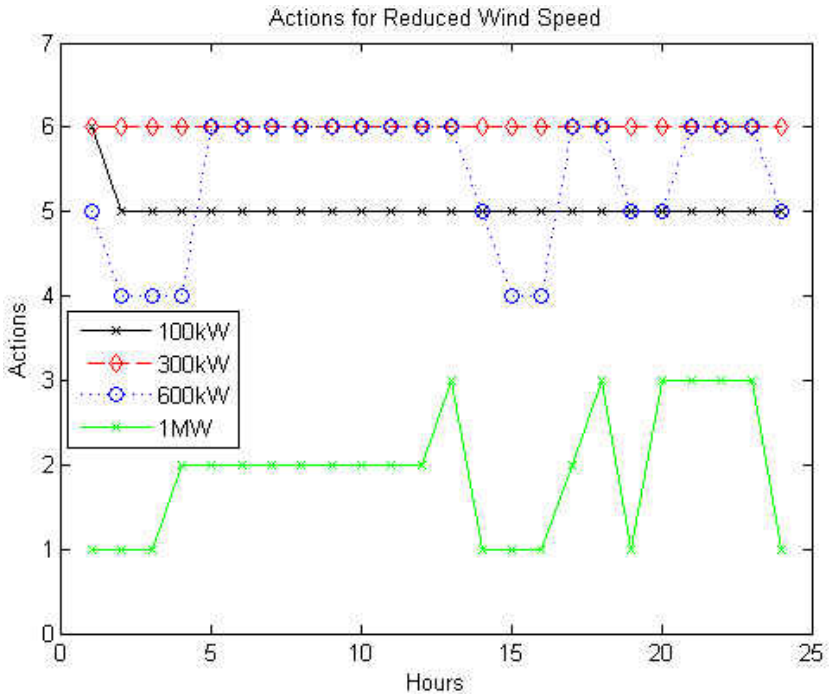

Figure 26: Hourly Action for Reduced Wind Speed

We conduct similar experiments to compare to the case where we have an increase in wind speed. As shown in Figure 27, since we have a higher probability to receive wind power at each hour we can be more aggressive at certain battery levels. At levels 100kW and 300kW, the actions do not undergo any change due since it is not recommended to utilize the remaining battery at low battery levels given that shedding some load is more cost-effective. The main difference shows at higher battery levels. When the wind probability is higher throughout the day we can utilize more battery at almost every hour when comparing to the original wind speed.

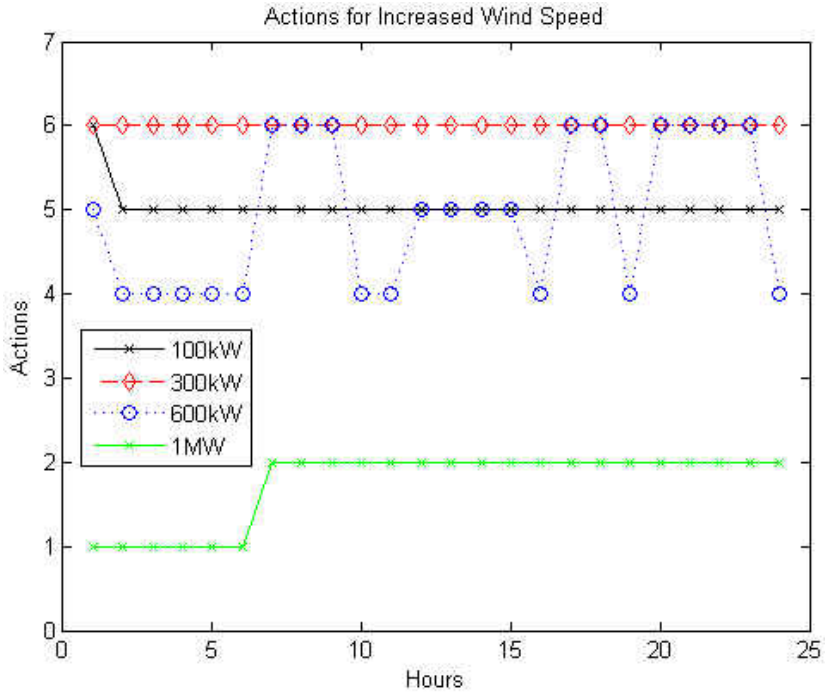

Figure 27: Hourly Action for Increased Wind Speed

If we take a closer look at the actions for the 600kW and 1MW battery level, shown in Figure 28, we can observe how the action changes at different wind speed regions. At 600kW in the reduced wind setting we utilize action 6 more often in comparison to the other two wind regions. This action corresponds to shedding the entire load. As the wind probability increases we are more disposed to use actions that use more of the battery.

At the 1MW battery level we observe that with the increase in wind speed we tend to utilize one action throughout the day. As we decrease the wind probability we start switching actions more frequently to minimize the cost.

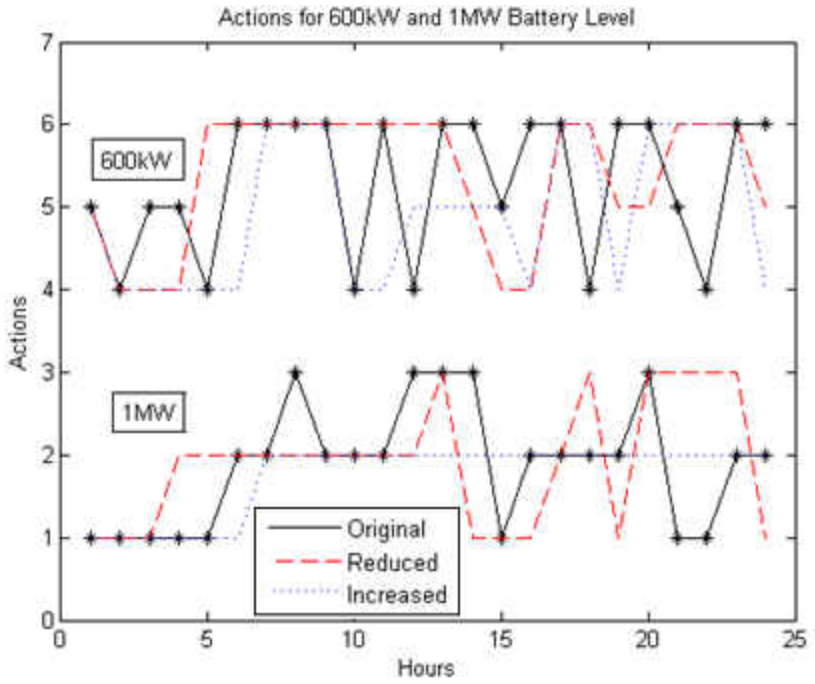

Figure 28: Actions for 600kW and 1MW

In Table 18, we can see the number of times a particular action is chosen at two different battery levels at different times of the day. These frequencies are obtained from Figure 28. At the 600kW battery level we focus on actions 4-6 and at 1MW we observe actions 1-3. We can see that at a reduced wind speed more load-shedding is used and as the wind speed increases we utilize more of the battery. At 1MW we can clearly see that load shedding is very minimal. We only load shed 100kW at an increased wind speed and at a reduced wind speed we chose to load shed more than at original wind speed.

Table 18: Frequency of Action Taken at Different Battery Levels

Actions Chosen at 600kW	**4**	**5**	**6**		**Actions Chosen at 1MW**	**1**	**2**	**3**
Original	6	5	13		Original	8	11	5
Reduced	5	5	14		Reduced	8	10	6
Increased	10	5	9		Increased	6	18	0

Another way to examine how the wind affects performance is to observe the battery charge level at each hour. As we observed the actions at each wind region, in this section we observe the four initial battery levels. At the first battery level, i.e., 100kW, we obtain the result shown in Figure 29. We start the battery level at 100kW in each wind region and simulate the battery level change throughout the day for one sample trajectory. We can observe that in the reduced wind region we are less likely to be aggressive in using the battery compared to the original wind speed. Since there is less probability of receiving wind we take actions that use the least amount of battery as possible. On the other hand when we compare the increased wind region to the original we can take actions that utilize more of the battery than in the original region. With a higher probability of wind for the coming hours we can take actions that deplete half of the battery while being certain that later on enough wind power would be generated to replenish the used power.

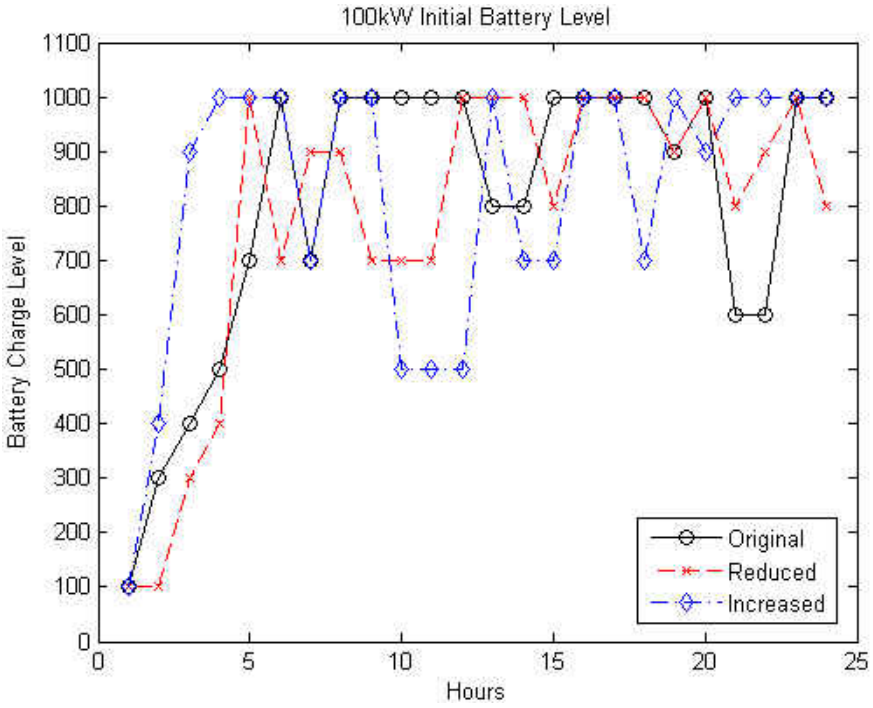

Figure 29: Hourly Battery Level at 100 kW for Different Wind Speed

We also consider an initial battery level of 300kW. In Figure 30 we show the results of simulation for the different wind regions. It is evident that at the early hours the reduced wind region has a slower starting wind power generation. This is due to more load-shedding being used in the early stages. When the wind is at its peak, i.e., around mid-day, the actions that are taken are not as aggressive as for the original wind speed. With increased wind speed, we can be more aggressive again in using the battery. We tend to utilize more battery charge each hour knowing that getting power back from the wind is more likely.

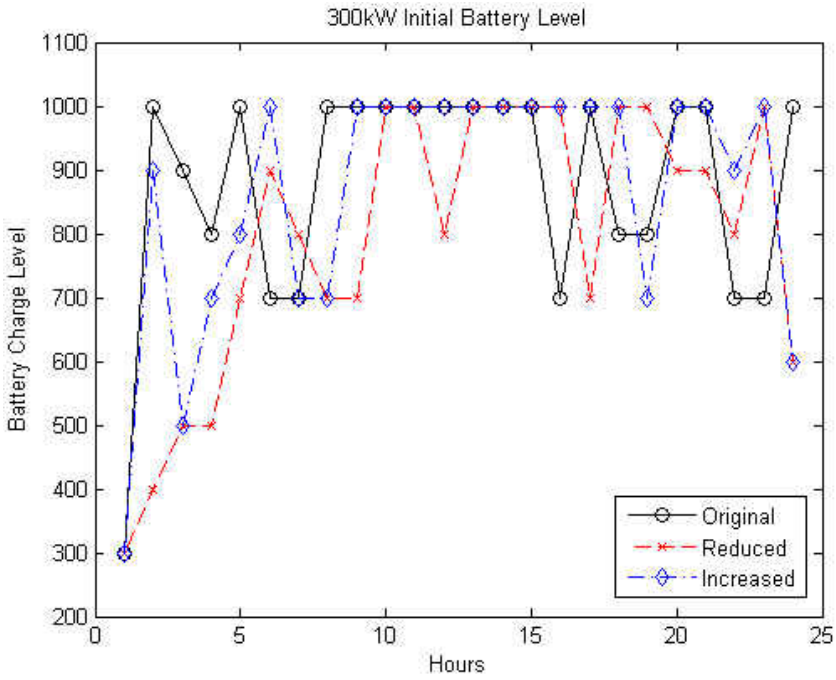

Figure 30: Hourly Battery Level at 300 kW for Different Wind Speed

Going to higher initial battery levels we observe that the optimal policy is increasingly more aggressive in battery usage in comparison to the previously observed battery level. In Figure 31, we notice that the battery is more aggressively utilized for all wind speeds. However, if we compare the reduced and the original we can see that when we take action to utilize the battery we use 400kw of the battery rather than in the reduced wind regions, where we use mostly 300kW. On the other hand comparing the increased wind region with the original we utilize 500kW more frequently with the battery than the original wind region. This shows again the same intuitive trend that as we get into a region with a higher wind probability we can be more aggressive utilizing the battery.

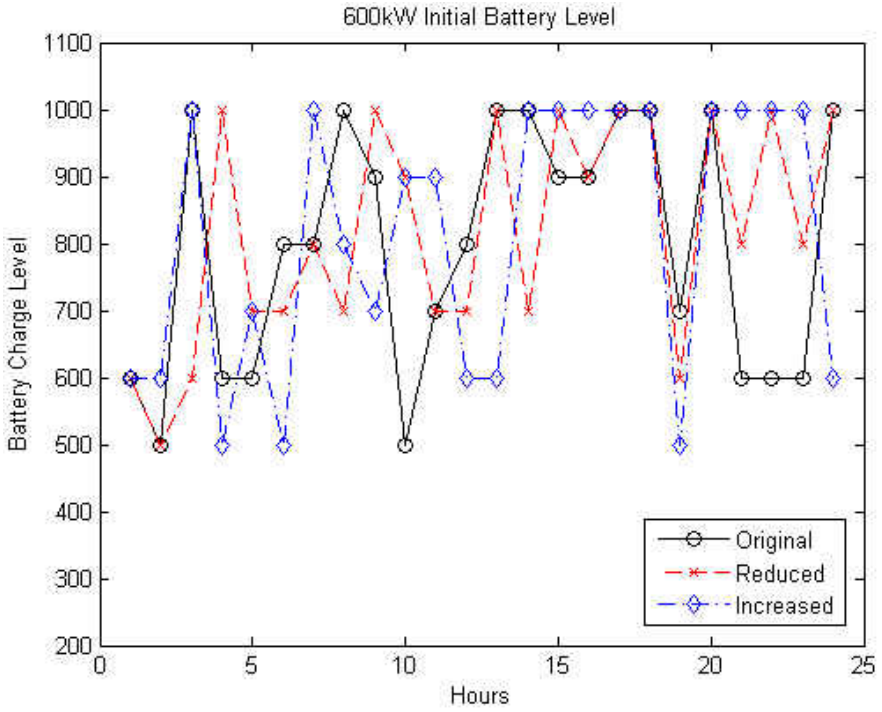

Figure 31: Hourly Battery Level at 600 kW for Different Wind Speed

Finally, we repeat the experiment when the initial battery level is the maximum capacity, i.e., 1MW and the results of one trajectory are shown in Figure 32. Since we began with higher initial battery storage we begin using the battery regardless of the wind setting. Again, at the original wind speed we are more aggressive that at reduced wind speeds in using the battery. On the other hand, at the higher initial battery level even for the reduced wind region, the battery is more utilized that at lower battery levels. In other words, we take the same amount of risk utilizing the battery as the original wind region in this level. We tend to take more risks at increased wind speeds.

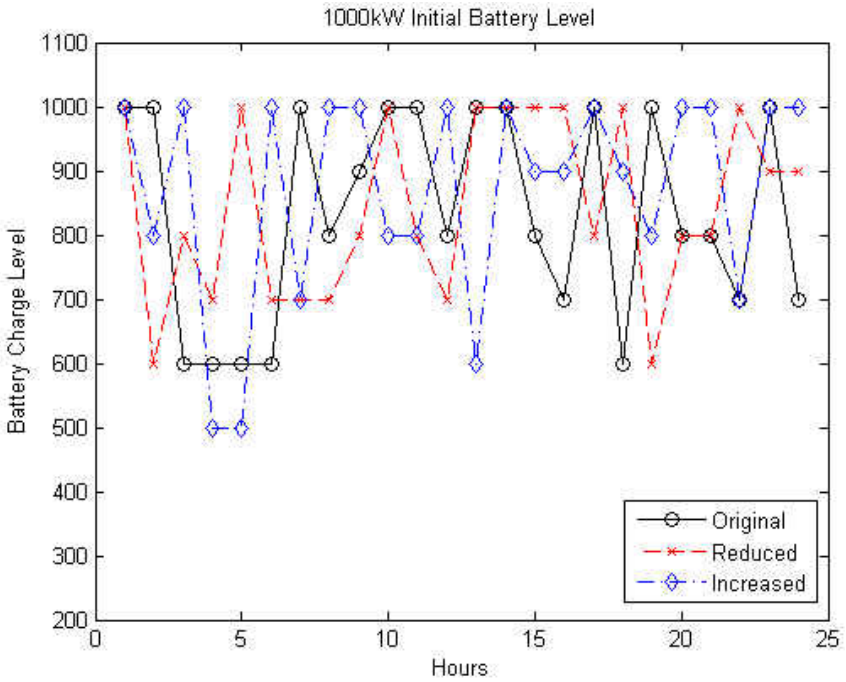

Figure 32: Hourly Battery Level at 1000 kW for Different Wind Speed

In Figure 33 we study the effects of different wind speeds on the total cost incurred. We show the total average cost for different initial battery level for the different wind settings. At 100kW, we notice that in the reduced wind region we have a higher total average cost than the original wind region. This is mostly because it costs more to shed some load than to use the battery, and since we do not have enough wind power to replenish the battery we have to resort to load-shedding more often. On the other hand, we can use the battery more frequently at increased wind speeds, hence the total cost is lower than the original wind region. The same pattern is observed at 300kW. However, in this case the differences between the costs corresponding to different wind regions are substantially reduced. As we move to higher battery levels these differences

are further reduced until the total average costs are equal. This is because we have more freedom in using the battery when we start from higher initial battery levels, thereby producing a reduced total cost at the end of the day. This starts around an initial battery level of 500kW.

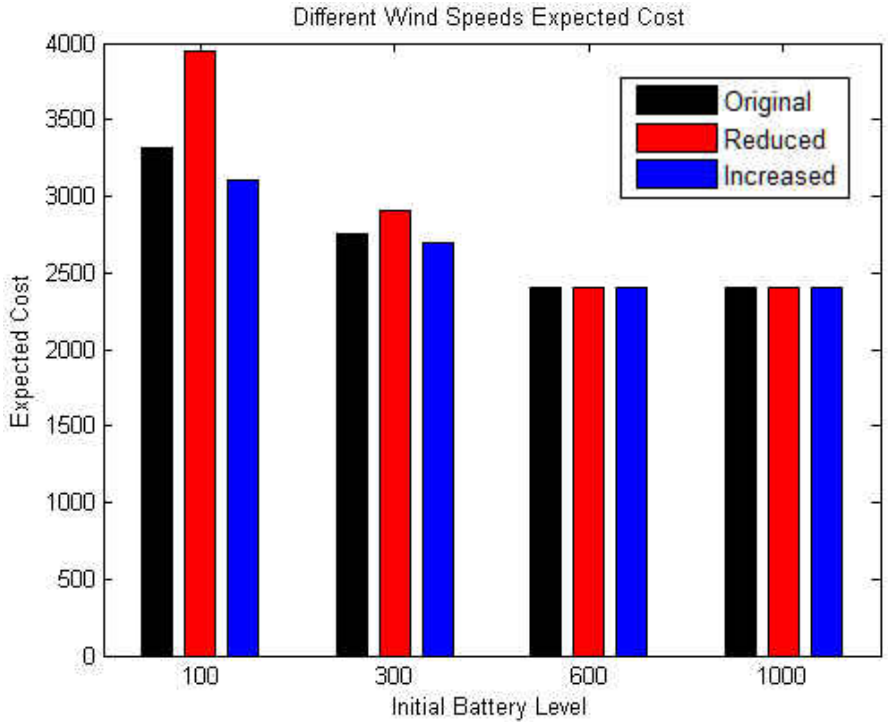

Figure 33: Total Cost for Different Wind Speed and Initial Battery Level

Effects of Demand

In this section, we study the effects of the daily demand on the control policy and the associated costs. We compare three different demands. The original demand was 500kW, so now we also consider 400kW and 300kW. Note that changing the demand requires us to also change the control space. In particular, originally we had 6 actions for a demand of 500kW. For the 400kw and 300kW demands, we will have a control space consisting of a total of 5 and 4 actions, respectively.

In Figure 34 we show the costs for the different demands. When comparing the cost in the 100kW battery level we can see that the 500kW demands has a slightly higher cost than the other two demands. But, as we increase the battery charge level the costs do not seem to depend much on the demand. This is due to the fact that the costs for the actions do not change. Even with less demand we can see that we have an equal cost at higher battery levels. Since at each hour we have to pay a minimum of 100, over a 24-hour period the minimum total cost will be 2400. This is the reason we get those values at battery levels 600 and 1000kW.

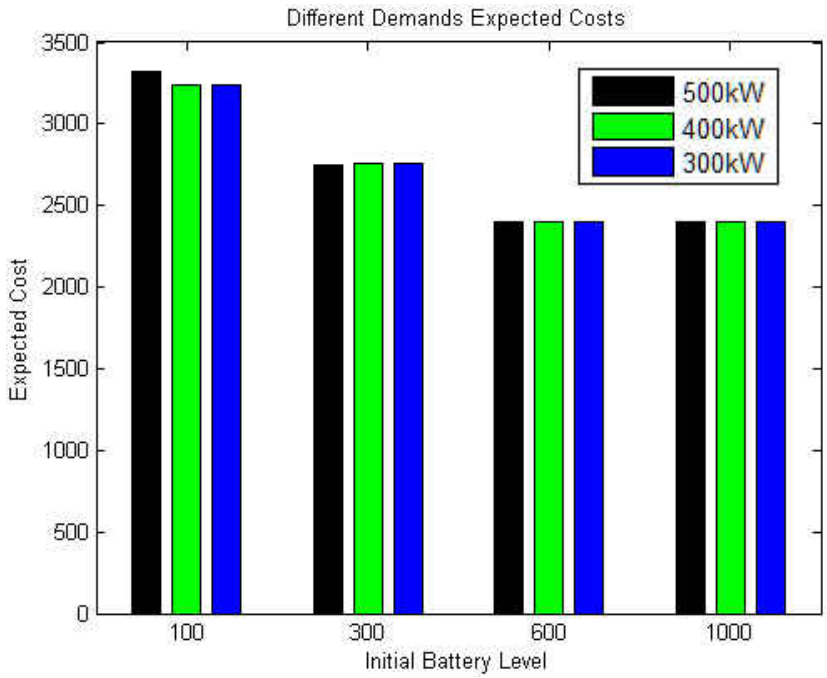

Figure 34: Different Demand Costs

Effect of Load Shedding Costs

In this section, we study the change in performance as we increase the cost of load-shedding. We compare the total cost of the optimal policy to the other control policies. We test the four previous battery states again, namely, 100 kW, 300kW, 600 kW, and 1MW. At each battery state we simulate our policy a large number of times to get an estimate of the average cost. We change the load-shedding cost two times to study its effect on the optimal policy. We have three load-shedding costs. The first cost, termed original, is the load-shedding cost that was initially used. The second cost is calculated from the percentage of the load we are shedding and multiply the cost by 1.2. Finally, the third cost is computed from the percentage of the load we are shedding, but this time we multiply the cost by 1.4. In Table 19, for example, for action 2 we load-shed 100kw of the 500kW total demand. This means, we load-shed 1/5 of the load. By multiplying the original cost time 1.2 we get the new cost. When it comes to multiplying the cost 1.4 times we have to make sure that we put a heavy penalty on the first four stages of the battery. Since these actions are not possible we must increase the value in the first four battery states.

Table 19: New Costs Example

Action 2			
B=400 & L=100			
States	Cost	1.2x LS Cost	1.4x LS Cost
0	1000	1200	3400
100	1000	1200	3400
200	1000	1200	3400
300	1000	1200	3400
400	500	600	700
500	400	480	560
600	300	360	420
700	200	240	280
800	100	120	140
900	100	120	140
1000	100	120	140

If we compare our original policy with the two control policies costs we can clearly see the difference in total costs in each battery state. In Figure 35 we can see this difference. The original cost shown in black is the one with the lowest total cost in each battery level. In green we can see the cost of control policy one, here we use the daily load demand curve as policy rule. During business hours we implement load-shedding as minimal as possible. In the blue control policy two, we implement load-shedding whenever we are below 300kW power level. Otherwise we use as much of the battery as possible.

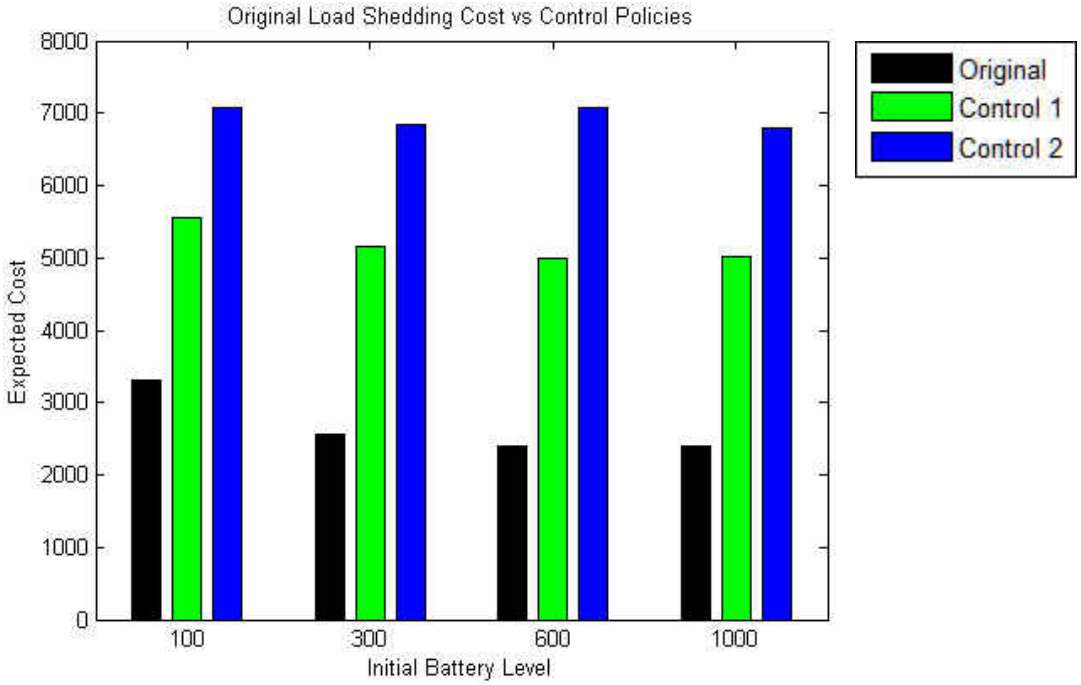

Figure 35: Original Load Shedding Cost Comparison

In our second comparison where we start increasing the cost of load-shedding, we first obtain a new cost for each action in each state as shown previously. In Figure 36 we can see the result of this simulation. By increasing the load-shedding cost we can see an increase in the overall total costs. When comparing the two control policies with the original policy it shows that the original policy has a total cost that is a lot less than the control policies. By increasing the cost of load-shedding our method of solving is still providing optimal results. Not only is it less costly in comparison to the other two control policies, but it is a significant difference in cost.

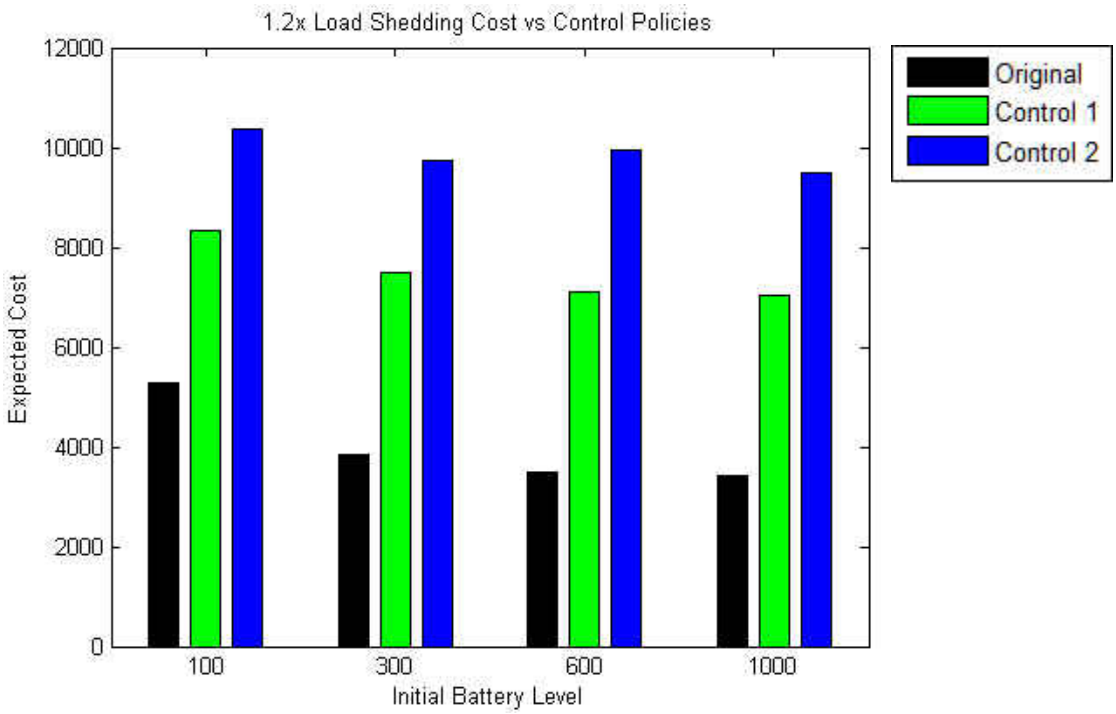

Figure 36: 1.2x Load-Shedding Cost Comparison

Finally for our final simulation, we simulate the cost of load-shedding with an increase in load-shedding cost of 1.4 times. The result of simulation is shown in Figure 37. In these results we can clearly see that our optimal policy that we obtain is performing a lot better than our control policies. Our total costs keep increasing due to the increase in cost of load-shedding. But if we compare the total cost of our optimal policy to the other two control policy the control costs do not even come close to the original policy for all costs that we simulated.

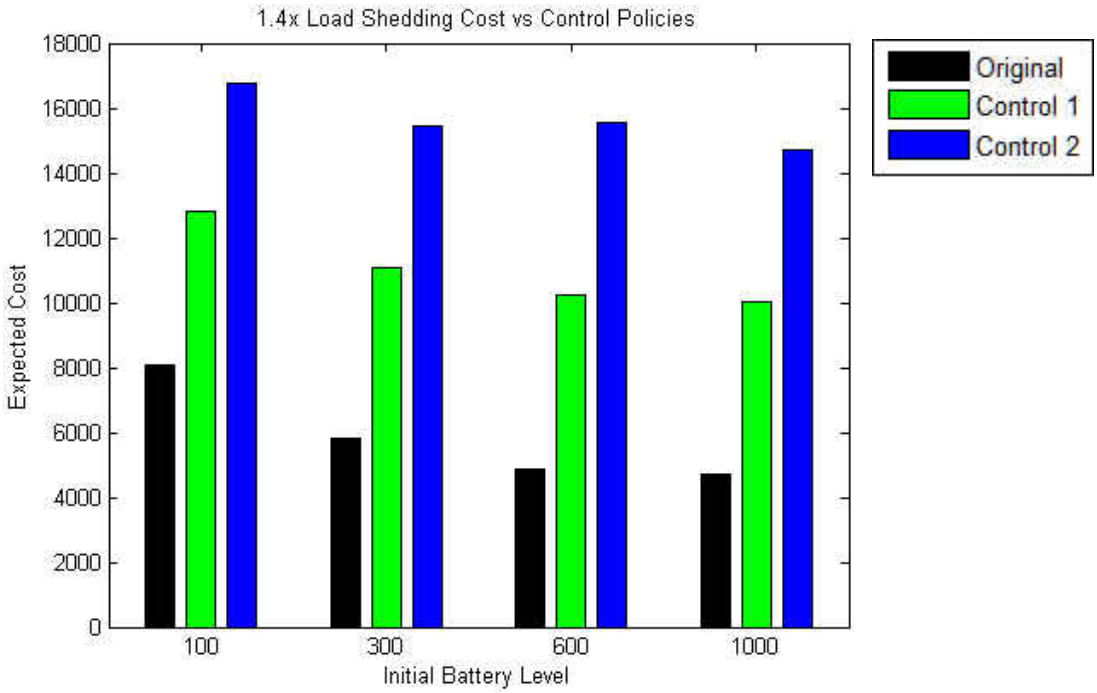

Figure 37: 1.4x Load-Shedding Cost Comparison

CHAPTER FIVE: CONCLUSIONS & FUTURE WORKS

Due to the recent rapid expansion in renewable energy sources, various challenges arise when we try to integrate renewable energies with the current generation methods from power plants that use, for example, diesel engines or nuclear energy. In this thesis, the renewable energy source is the wind energy. This can cause intermittency on the power generation that could lead to various problems, such as increased costs for power generation companies or severe equipment failures. In this thesis, we proposed a method to minimize the cost stemming from this problem.

We proposed a methodology based on an MDP formulation whereby we obtained optimal load-shedding policies that can be potentially reduce the total cost incurred by the power company. In order to study the stochastic behavior and intermittency of the wind energy, we assumed a constant demand. We also used an energy storage model to help reduce the load shedding by supplying power when needed and storing power if we were generating a surplus.

To this end, we first obtained the average wind speed from data collected over a two-year period. The average wind speed was then used to generate a probabilistic model for the power generation at various wind speeds. Specifically, the wind velocity probability was modeled using a Rayleigh distribution, which is a special case of the Weibull probability distribution.

Next, we considered the finite horizon formulation wherein the optimal control policies are obtained for 24 hours. The problem was mapped to an MDP framework with suitably defined state, space, control space, state evolution and cost functions. Using Dynamic programming, we started from the 24^{th} hour of the day, obtained our cost function and worked our way backwards towards the first hour of the day generating the optimal time-varying policy. By solving the MDP problem with backwards dynamic programming, we calculated at each hour the optimal policy defined as a mapping from the state space to the control space. The optimal value function obtained from dynamic programming was compared with the total average cost from simulations where we average over a number of Monte Carlo runs and these were shown to match for the different states.

We also considered an infinite horizon formulation with discounted cost. For this setting, we implemented policy iteration to obtain the optimal policy as a solution to the discounted Bellman equation. The costs from simulations and policy iteration were shown to closely match. A major difference from the finite horizon counterpart is that the policy in this setting is time-invariant (stationary).

We compared the finite horizon optimal policy to other control policies, namely a load balancing policy and another policy that uses minimal load shedding. The

proposed MDP policy was shown to significantly outperform the other policies, i.e., produces a significantly lower average cost for each state.

We studied the effect of various system parameters on the performance of the optimal policy. In particular, we investigated the effect of different wind speeds on the average cost and the control policy. It was shown that higher wind speeds result in lower costs. At lower wind speeds, the average costs increase since we have a reduced capability of replenishing the power used from the battery. We also concluded that the optimal policy tends to be more aggressive in utilizing the battery at higher average wind speeds.

We also studied the effect of demand, and observed that the optimal cost is not too sensitive to the demand level especially at higher battery levels where the cost cannot be reduced any further.

Finally, we investigated the effect of the load-shedding cost on the system performance. First, it was observed that the total cost increases for the optimal policy, as well as for the other control policies. Nevertheless, the optimal policy was shown to be significantly better than the aforementioned control policies in terms of total costs. Control policy B, which uses less load shedding, has a much higher cost than control policy A, which uses load balancing.

In the future, we would like to consider problems with a larger state space by having finer battery levels, as well as a larger control space with more permissible actions. However, in such settings we have to deal with the so-called curse of dimensionality. One way to approach this problem is to resort to techniques of approximate Dynamic programming [42]. Moreover, in this thesis we considered states that are perfectly observable. The same techniques can be extended to situations with a richer state space, wherein parts of the states may be only partially observed. In this case, we can invoke tools from the theory of Partially Observable Markov Decision Processes (POMDPs). In such settings, the state can be replaced with a sufficient statistic known as the belief, which is the posterior distribution of the state given all past information. We can then leverage solution methodologies such as point-based value iteration to find the optimal allocation and load-shedding policies. Finally, we could consider scenarios where certain parameters of the wind distribution are unknown. This requires combining DP with learning using techniques of reinforcement learning which are known to strike favorable tradeoffs between exploration and exploitation.

REFERENCES

[1] A. M. Omer, "Energy, environment and sustainable development," *Renewable and Sustainable Energy Reviews,* vol. 12, pp. 2265-2300, 2008.

[2] A. Demirbas, "Global Renewable Energy Projections," *Energy Sources, Part B: Economics, Planning, and Policy,* vol. 4, pp. 212-224, 2009/10/30 2009.

[3] M. I. Hoffert, K. Caldeira, A. K. Jain, E. F. Haites, L. D. Harvey, S. D. Potter, M. E. Schlesinger, S. H. Schneider, R. G. Watts, and T. M. Wigley, "Energy implications of future stabilization of atmospheric CO2 content," *Nature,* vol. 395, pp. 881-884, 1998.

[4] K. Aarons, "CARBON POLLUTION STANDARDS FOR EXISTING POWER PLANTS: ISSUES AND OPTIONS," *Center for Climate and Energy Solutions,* 2014.

[5] W. Livoti, "Survival of the Efficient: Updating Power Plants to Meet the Challenges of the Future," *Power and Energy Magazine, IEEE,* vol. 12, pp. 112-116, 2014.

[6] B. V. Mathiesen, H. Lund, and K. Karlsson, "100% Renewable energy systems, climate mitigation and economic growth," *Applied Energy,* vol. 88, pp. 488-501, 2011.

[7] U. D. o. Energy. (2011). *Renewable energy shows strongest growth in global electric generating capacity*. Available: http://www.eia.gov/todayinenergy/detail.cfm?id=3270#

[8] N. Research. (2013). *International Wind Power Development World Market Update 2012*. Available: http://www.btm.dk/reports/world+market+update+2012

[9] W. F. Pickard and D. Abbott, "Addressing the intermittency challenge: Massive energy storage in a sustainable future," *Proceedings of the IEEE,* vol. 100, p. 317, 2012.

[10] F. Díaz-González, A. Sumper, O. Gomis-Bellmunt, and R. Villafáfila-Robles, "A review of energy storage technologies for wind power applications," *Renewable and Sustainable Energy Reviews,* vol. 16, pp. 2154-2171, 2012.

[11] X. Huan, T. Ufuk, S. H. Low, C. R. Clarke, and K. M. Chandy, "Load-shedding probabilities with hybrid renewable power generation and energy storage," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 2010, pp. 233-239.

[12] T. Goya, T. Senjyu, A. Yona, N. Urasaki, T. Funabashi, and K. Chul-Hwan, "Optimal operation of controllable load and battery considering transmission constraint in smart grid," in *IPEC, 2010 Conference Proceedings*, 2010, pp. 734-739.

[13] T. Goya, T. Senjyu, A. Yona, N. Urasaki, T. Funabashi, and K. Chul-Hwan, "Thermal units commitment considering voltage constraint based on controllable loads reactive control in smart grid," in *Power Electronics and Drive Systems (PEDS), 2011 IEEE Ninth International Conference on*, 2011, pp. 793-798.

[14] E. Sortomme and M. A. El-Sharkawi, "Optimal Power Flow for a System of Microgrids with Controllable Loads and Battery Storage," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1-5.

[15] A. Saffarian and M. Sanaye-Pasand, "Enhancement of Power System Stability Using Adaptive Combinational Load Shedding Methods," *Power Systems, IEEE Transactions on,* vol. 26, pp. 1010-1020, 2011.

[16] S. Xing, "Microgrid emergency control based on the stratified controllable load shedding optimization," in *Sustainable Power Generation and Supply (SUPERGEN 2012), International Conference on*, 2012, pp. 1-5.

[17] T. Goya, K. Uchida, T. Senjyu, A. Yona, T. Funabashi, and K. Chul-Hwan, "Optimal operation of controllable load and battery considering forecasted error," in *Electrical Machines and Systems (ICEMS), 2010 International Conference on*, 2010, pp. 1759-1762.

[18] H. Ying-Yi and C. Po-Hsuang, "Genetic-Based Underfrequency Load Shedding in a Stand-Alone Power System Considering Fuzzy Loads," *Power Delivery, IEEE Transactions on,* vol. 27, pp. 87-95, 2012.

[19] C. C. White III and D. J. White, "Markov decision processes," *European Journal of Operational Research,* vol. 39, pp. 1-16, 1989.

[20] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, ed. vol. 2: Athena Scientific, 2005.

[21] W. L. Winston and J. B. Goldberg, "Operations research: applications and algorithms," 1994.

[22] L. Yingzi, N. Jincang, L. Ru, and Y. Yuntao, "Research of multi-power structure optimization for grid-connected photovoltaic system based on Markov decision-making model," in *Electrical Machines and Systems, 2008. ICEMS 2008. International Conference on*, 2008, pp. 2607-2610.

[23] L. Ying-zi, R. Luan, and N. Jin-cang, "Benefit analysis for grid-connected photovoltaic system based on markov decision processes with expected total reward criterion," in *Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on*, 2009, pp. 2188-2191.

[24] L. Yingzi, R. Luan, and N. Jincang, "Power structure optimization for grid-connected photovoltaic system based on Markov decision processes," in *Industrial Electronics and Applications (ICIEA), 2011 6th IEEE Conference on*, 2011, pp. 2688-2692.

[25] A. Kashyap and D. Callaway, "Estimating the probability of load curtailment in power systems with responsive distributed storage," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2010 IEEE 11th International Conference on*, 2010, pp. 18-23.

[26] K. Turitsyn, S. Backhaus, M. Ananyev, and M. Chertkov, "Smart finite state devices: A modeling framework for demand response technologies," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, 2011, pp. 7-14.

[27] H. Song, L. Chen-Ching, J. Lawarree, and R. W. Dahlgren, "Optimal electricity supply bidding by Markov decision process," *Power Systems, IEEE Transactions on,* vol. 15, pp. 618-624, 2000.

[28] G. K. Chan and S. Asgarpoor, "Optimum maintenance policy with Markov processes," *Electric Power Systems Research,* vol. 76, pp. 452-456, 2006.

[29] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts, "Modeling medical treatment using Markov decision processes," in *Operations Research and Health Care*, ed: Springer, 2004, pp. 593-612.

[30] M. Mundhenk, J. Goldsmith, C. Lusena, and E. Allender, "Complexity of finite-horizon Markov decision process problems," *Journal of the ACM (JACM),* vol. 47, pp. 681-720, 2000.

[31] E. Altman, "Applications of Markov decision processes in communication networks," in *Handbook of Markov decision processes*, ed: Springer, 2002, pp. 489-536.

[32] P. Xuan, V. Lesser, and S. Zilberstein, "Communication decisions in multi-agent cooperation: Model and experiments," in *Proceedings of the fifth international conference on Autonomous agents*, 2001, pp. 616-623.

[33] R. Becker, "Solving transition independent decentralized Markov decision processes," *Computer Science Department Faculty Publication Series,* p. 208, 2004.

[34] D. P. Bertsekas, *Dynamic Programming and Optimal Control* vol. 1: Athena Scientific, 2005.

[35] R. Bellman, "The theory of dynamic programming," DTIC Document1954.

[36] D. Weisser, "A wind energy analysis of Grenada: an estimation using the 'Weibull' density function," *Renewable Energy,* vol. 28, pp. 1803-1812, 2003.

[37] M. Savenkov, "On the Truncated Weibull Distribution and its Usefulness in Evaluating the Theoretical Capacity Factor of Potential Wind (or Wave) Energy Sites," in *University Journal of Engineering and Technology* vol. 1, ed, 2009, p. 5.

[38] W. D. Kellogg, M. H. Nehrir, G. Venkataramanan, and V. Gerez, "Generation unit sizing and cost analysis for stand-alone wind, photovoltaic, and hybrid wind/PV systems," *Energy Conversion, IEEE Transactions on,* vol. 13, pp. 70-75, 1998.

[39] NREL. (2014). *Western Wind Resources Dataset*. Available: http://wind.nrel.gov/Web_nrel/

[40] J. A. Jardini, C. M. V. Tahan, M. R. Gouvea, A. Se Un, and F. M. Figueiredo, "Daily load profiles for residential, commercial and industrial low voltage consumers," *Power Delivery, IEEE Transactions on,* vol. 15, pp. 375-380, 2000.

[41] J. A. Jardini, H. P. Schmidt, C. M. V. Tahan, C. C. B. De Oliveira, and A. Se Un, "Distribution transformer loss of life evaluation: a novel approach based on daily load profiles," *Power Delivery, IEEE Transactions on,* vol. 15, pp. 361-366, 2000.

[42] D. P. Bertsekas and J. N. Tsitsiklis, "Neuro-dynamic programming: an overview," in *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, 1995, pp. 560-564.