NAME: Micah Botkin-Levy
DATE: 4/18/19
Homework: 08
STAT/CS 387

1 Continuing Data Analysis

Significant progress has been made since the last progress update. I have somewhat successfully implemented a Q learning reinforcement algorithm to control the energy storage charging and discharging based on real time prices. Also, I have implemented a random control strategy as well as a centralized perfect forecast optimization problem to compare with the Q learning solution.

To implement the Q learning algorithm I used the Gym python package. In order to do this I needed to create a custom environment which has the ability to take a "step" based on an energy storage system action (charge, discharge, hold on) decided by the Q-learning greedy algorithm. The same environment was used to implement the random control algorithm which randomly charges, discharges or holds on at each time step. In order to implement the optimization problem, I used the cvxpy package for python which is a modeling framework for solving Disciplined Convex Programming problems. I then use the Gurobi optimization solver to get a solution. The total cost for each algorithm is shown in Table 1.

	Central Optimization	Random Action	Q Learning
Total Cost ($)	-29,153	15,934	3,594

Table 1: Total cost for Feb and Mar using three control algorithms

As you can see, the optimization control scheme is able to make money by performing arbitrage (charging when the electricity price is cheap and discharging when it is high) however it has perfect forecast of the electricity prices which is not realistic for a real time control strategy. Q learning is able to perform much better than the random control algorithm but still ends in a net loss of money. This means that operating the energy storage system (ESS) using the current Q Learning does not make sense because if the ESS never charged or discharged it would cost $0.

We can take a closer look at what is happening by looking at the energy and power for each 5 minute time step over the two month period looked at. In Figure 1, the energy of the ESS is plotted for each time step for each of the three control algorithms. The random algorithm has random behavior and tends to stay around a 50% charge. The optimization algorithm has much more drastic changes as it will charge at maximum power while energy is cheap and discharge at maximum power while energy is expensive to maximize profits. The Q learning algorithm does something interesting. Since the reward is only positive when the ESS is discharging, the Q learning arbitrage policy will always want to chose to discharge when it can. Therefore what we end up seeing is it discharges quickly and stays near empty for the whole two months, only charging when prices are low enough.

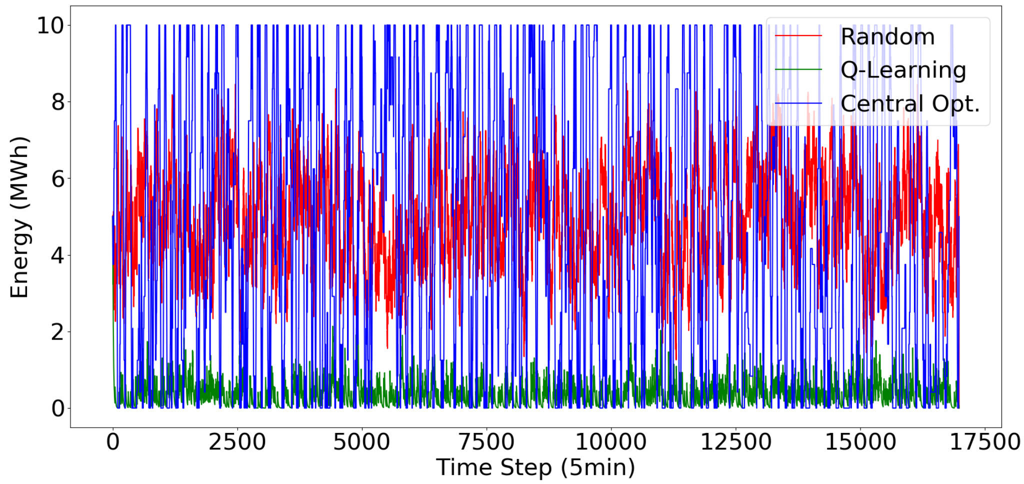

Figure 1: Energy of ESS for Feb and Mar using three control algorithms

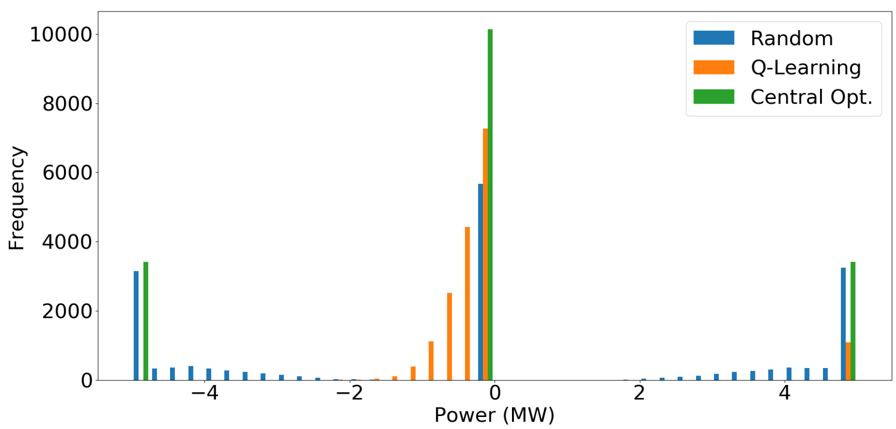

Figure 2: Power Histogram of ESS for Feb and Mar using three control algorithms

We can also take a look at the distribution of charging and discharging prices for each control algorithm. These can be seen in Figure 2 and the log histogram can be seen in Figure 3. As previously stated the Random algorithm has a wider distribution of powers. The optimization always knows the perfect times to charge and discharge and therefore only holds, charges at the maximum power, or discharges at the maximum power. The Q learning control will discharge as much as possible hence the smaller than maximum discharge powers when close to zero and charges at maximum power when it is empty and the price is low enough.

Overall, I believe I meet the forecast in terms of amount of progress but I ended up switching the order of some things. Originally I had tuning the Q learning algorithm before implementing comparison algorithms. I have not finished tuning the Q learning algorithm like

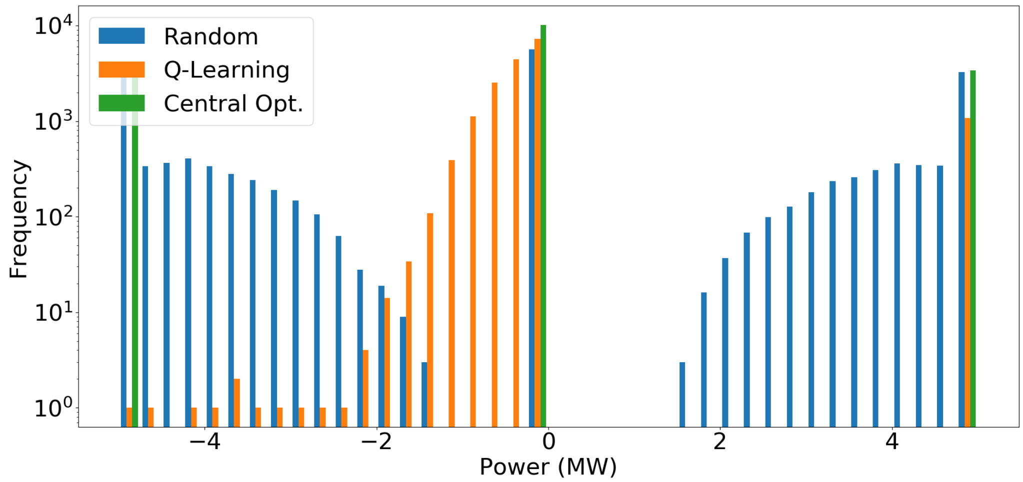

Figure 3: Power Log Histogram of ESS for Feb and Mar using three control algorithms

I originally planned, however, I am ahead of schedule in terms of implementing comparison baseline methods.

Original Forecast:

Week of April 8th: Implement Q-learning Algorithm

Week of April 15th: Debug and tune algorithm performance

Week of April 22nd: Compare Q-learning algorithm to baseline methods, such as full forecast optimization and random actions, and extend methods/application, time permitting

Week of April 29th: Prepare and record presentation, write scientific article and technical report

For an update schedule, I will plan on tuning the Q learning method this week. I will consider it successful if I can get the Q learning method to result in a profit. This is the biggest remaining challenge as it does not appear to be operating as expected and might involve having more data to train the Q lookup table, tuning the Q learning parameters, or changing the Q learning reward. If there is time I will look into more complicated Q learning and reinforcement learnig techniques. Finally, next week I will need to write the reports and prepare the presentation.

Updated Forecast:

Week of April 22nd: Tune Q Learning algorithm performance and extend methods, time permitting

Week of April 29th: Prepare and record presentation, write scientific article and technical report

2 Continuing Literature Review

1. R. S. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning*, vol. 3, pp. 9-44, Aug 1988.

   Summary Sutton is consider one of founding fathers of reinforcement learning. This article is one of Sutton's early publications and work on reinforcement learning which he refers to as incremental learning procedures. The paper focuses on a specific type of reinforcement learning that use temporally successive predictions. The main contribution of the paper is to prove the converge and optimality of such methods. Sutton also claims that many supervised learning problems can be approached from a temporal prediction vantage.

   Relevance Sutton's early work provides a strong foundation for the study of reinforcement learning. He argues that temporal difference methods are very well suited for real world problems. The problem that I am considering as a perfect application. The temporal difference methods that Sutton introduces here are the beginning of the vast reinforcement learning research that will eventually create the Q learning techniques.

2. C. J. C. H. Watkins, "Learning from delayed rewards," 1989.

   Summary Watkin's PhD thesis presents some early work on reinforcement learning. He frames the methods in terms of animal behavior and presents Markov decision process as a good model of animal behavior. He states that stochastic dynamic programming could be used to calculate an animals behavior's but an animal would never use it to make decisions or learn optimal policies. Instead, he introduces an incremental Monte-Carlo method for learning optimal values of actions that does not require a model of the environment. This method is latter dubbed Q-Learning.

   Relevance Watkins produces a truly novel piece of work in his PhD thesis. The method he introduces later becomes a prominent statistical learning method which enables learning the value of taking certain actions without a model. This method is perfect for energy storage control using real time prices as formulating a model for these prices is very difficult. His thesis presents a useful method that will be the basic brains of my project.

3. C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279-292, May 1992.

   Summary Walkin and Dayan's work expands on the idea from Watkin's thesis. They present a convergence theorem for Q-learning. The authors state that the Q-learning method will always converge to the optimum action-values as long as all actions are repeatedly sampled in all states and action-values are represented discretely. Additional extensions of the methods are also examined for different cases of Markov environments.

Relevance This piece of work is extremely relevant for the final project. They take ideas from thesis which weren't fully fleshed out and prove that it can work. The convergence guarantee allows us to state that we are taking optimal actions if formulating the problem correctly. This work enables us to be more confident in using the Q-learning method in the project.

4. ADL, "An introduction to Q-Learning: reinforcement learning," *freeCodeCamp*, 03-Sep-2018. [Online]. Available: Link[Accessed: 22-Apr-2019].

   Summary This article gives a clear overview of Q learning and a concrete example. It is written for programmers who are less familiar with the topics and uses an example where a robot has to navigate a maze to get to end. The author goes over the Q lookup table and the intuition of how it gets updated. It uses the robot maze example to run through a few steps of the algorithm so readers understand the process.

   Relevance Since the example is easily understandable and the explanation is simple, I used this article to help me understand Q learning in the beginning. As reading journal papers are often hard to grasp when learning a new subject, a simple example and explanation is very useful to have. In addition the article provides the python code for solving the robot maze problem which is a useful reference.

5. G. A. Rummery and M. Niranjan, *On-line Q-learning using connectionist systems*, vol. 37. University of Cambridge, Department of Engineering Cambridge, England, 1994.

   Summary This paper from Rummery and Niranjan push the boundaries of the early Q-learning literature by comparing modified Q-learning algorithms. The authors argue that the discrete nature of the classical Q-learning is too restrictive and therefore must be expanded to high dimensional continuous state space. The authors compare some recent modifications to Q-learning from Sutton, Peng, Lin as well as introduce a new algorithm, Modified Connectionist Q-Learning. Their work is an early predecessor to deep Q-learning and an important conclusions about robustness and performance are made.

   Relevance This paper provides useful extensions of the Q-learning algorithm for energy storage project. While the energy storage application can be treated as discrete, in the real world it would be much more realistic to use continuous control actions. This work cites some modified Q-learning algorithm to pursue this if time permits.

6. R. S. Sutton and A. G. Barto, *Reinforcement learning - an introduction.* MIT Press, 1998

   Summary The book by Sutton and Barto serves as an introduction to reinforcement learning. After an introduction chapter the book is split into three parts. The first part covers tabular solution methods such as multi-armed bandits, Markov Decision Processes, dynamic programming and Q-learning. The second section looks at approximate solution methods such as on-policy prediction and control. The book finishes with a deeper dive into the psychology and case studies of reinforcement learning.

Relevance The book serves as a good refernce of the basics of reinforcement learning. It covers temporal-difference learning and specifically Q-learning which is the algorithm I will use in the project. It also covers many of the motivations, challenges, and tradeoffs with reinforcement learning.

7. G. Hayes, "Getting Started with Reinforcement Learning and Open AI Gym," *Towards Data Science*, 22-Feb-2019. [Online]. Available: Link[Accessed: 22-Apr-2019].

   Summary This article gives an overview of solving a reinforcement learning problem using the gym package. OpenAI Gym is a python package that creates a environment and relevant methods to run reinforcement learning techniques. The article describes in detail how to use the gym package to implement Q learning for the mountain car problem.

   Relevance Since the article is clearly written and provides pseudo and actual code, it is a very useful reference. I used this article to help develop the code to use the gym package to solve my energy storage problem using Q learning.

8. OpenAI, "A toolkit for developing and comparing reinforcement learning algorithms," *Gym*. [Online]. Available: Link[Accessed: 22-Apr-2019].

   Summary This resource covers the OpenAI gym python package and how to use it. The gym package is a toolkit for developing and comparing reinforcement learning algorithms. With gym, you can create an environment where entities can make observations, take an action, check their state, and receive a reward. The documentation goes over how to install the package as well as some useful examples of its use cases.

   Relevance The gym packaged was used to create an environment for the reinforcement learning algorithm specific to the energy storage problem. This documentation helped get me familiar with the package functionality. While I have only scratched the surface of what gym can do it has provided a useful and standard format for creating the reinforcement learning algorithm and implementing it.

9. A. Poddar, "Making a custom environment in gym," *Medium*, 21-Jul-2018. [Online]. Available: Link[Accessed: 22-Apr-2019].

   Summary This article explained how to build a custom environment using the gym python package. OpenAI gym comes with a lot of built-in environments like the ones for solving the frozen lake or cartpole problems. In order to solve other real

world problems with reinforcement learning it is important to create a custom environment. This article breaks down the file and folder structure that OpenAI gym requires and what code should go in the python setup files. It also explains how to install and create an instance of the custom environment.

Relevance In order to implement Q learning on the energy storage system problem I am concerned with I will need to build a custom environment. This article serves as a great reference with how to do that using the gym package in python. The clear references and code examples made it very easy to follow.

10. H. Wang and B. Zhang,"Energy storage arbitrage in real-time markets via reinforcement learning," *CoRR*, abs/1711.03127, 2017.

    Summary The authors of this work perform real-time temporal arbitrage on energy storage using reinforcement learning. The main idea is by repeatedly performing charge and discharge actions under different real-time prices, the proposed reinforcement learning-based policy learns the best strategy that maximizes the cumulative reward. Their strategy uses the Q learning algorithm and applies two reward schemes to the problem for hourly realtime LMP prices for 2 years in ISO New England. They compare the performance of the Q-learning scheme to a baseline algorithm.

    Relevance The paper by Wang and Zhang is extremely relevant to the project as it was my inspiration and forms the basis of what I plan on working on. It provides a useful starting point for applying the Q learning algorithm on a energy storage system. I plan on also using real time market prices as the driving factor of the reward but will be using 5 minute ISO NE data instead of hourly data.

11. A. Juliani, "Simple Reinforcement Learning with Tensorflow Part 0: Q-Learning with Tables and Neural Networks," *Medium*, 25-Aug-2016. [Online]. Available: Link[Accessed: 22-Apr-2019].

    Summary This article gave an introduction to the reinforcement learning algorithm, Q-learning for use in python. First, it goes over the basic Q learning algorithm using a lookup table and how to implement it in python. Next, it expands this idea to real world problems by implementing a neural network with tensorflow in python. It also links the code for a specific example, the frozen lake problem.

    Relevance While this article does not introduce any novel technique, it is a useful reference. It uses the gym package for setting up a reinforcement learning environment and shows an appropriate control loop. The information here was useful when implementing my own problem using Q learning in python.

12. J. Eyer, G. Corey, and S. N. Laboratories,*Energy Storage for the Electricity Grid: Benefits and Market Potential Assessment Guide : a Study for the DOE Energy Storage*

Systems Program. SAND (Series) (Albuquerque, N.M.), Sandia National Laboratories,2010.

Summary This study which is performed by Sandia National Laboratories for the Department of Energy assesses the potential benefits and economic market potential for energy storage. There is quite a lot of information covered in this work including different energy storage technologies and the models used to represent them. The work also looks at all the different storage applications and benefits that can be gained from them. Also included are the challenges that energy storage faces and areas for further research.

Relevance While there is certainly more information than needed in this work it is a very useful reference for this project. The first important contribution is serving as a reference for the energy storage technology and discharging and charging models to be used. Secondly, it lists the possible energy storage applications including real-time arbritrage which will shape the reward function for Q-learning.

13. R. H. Byrne and C. A. Silva-Monroy, "Estimating the maximum potential revenue for grid connected electricity storage: Arbitrage and regulation," Sandia National Laboratories, 2012.

    Summary This study which is performed by Sandia National Laboratories for the Department of Energy looks in depth into two potenetial sources of income for energy storage: energy arbitrage and frequency regulation. Energy arbitrage referes to using control to buy energy when prices are low and sell when prices are high. Frequency regulation is related to an ancillary services market which attempts to maintain the system frequency. The work includes explanations of energy storage dynamics, optimization formulation, and case study results.

    Relevance This study is quite important for the project. First, it validates the linear dynamic model of energy storage which includes the charging and discharging power at each time step. Secondly, it provides a formulation for a linear optimization problem to determine the maximum the potential revenue. This could come in useful to provide a comparison to the Q-learning approach.

14. D. Mears, H. Gotschall, H. Kamath, E. P. R. Institute, U. S. D. of Energy, T. I. (Firm), and E. P. Corporation, *EPRI-DOE Handbook of Energy Storage for Transmission and Distribution Applications*. EPRI, 2003.

    Summary The goal of this report was to provide the benefits and costs of currently available energy storage technologies. In the beginning of the handbook the authors give an introduction and a perspective of the benefits of electric storage on the national level. Then, the authors go over the potnetial transmission and distribution applications for energy storage as well as their benefits, costs, and barriers. Chapters 6-15 each cover a different energy storage technology and details their description,

current development status, potential applications as well as their cost benefit analysis.

Relevance While this is a massive report with lots of unrelated information, it does serve as a trusted technology resource for energy storage. The report motivates the different revenue or cost avoided benefits of installing energy storage. One of these is arbitrage where an energy storage system can charge or discharged based on real time or time of use prices. The report justifies the financial reasons why solving the energy storage problem in my final report is important.

15. V. Kurama, "Reinforcement Learning with Python," *Towards Data Science*, 25-Nov-2018. [Online]. Available: Link[Accessed: 22-Apr-2019].

    Summary This article begins by introducing reinforcement learning and all of its potential applications. The article explains important reinforcement learning terminology such as states, actions, and rewards. Then it introduces the Q learning method and applies it to a taxi pickup problem. Finally, it provides python code and options for other reinforcement learning techniques.

    Relevance The article gives a great explaination of reinforcment learning and provides pseudo and actual code, so it is a very useful reference. I used this article to help my understanding of reinforcement learning as well as a guide to develop the code to use the gym package to solve my energy storage problem using Q learning.

16. T. Hubert and S. Grijalva, "Modeling for residential electricity optimization in dynamic pricing environments," *IEEE Transactions on Smart Grid*, vol. 3, pp. 2224-2231, Dec 2012.

    Summary The goal of this paper was to solve an optimal scheduling problem for the purchase, storage, and generation of electricity of a residential consumer. The model used was a residential electricity consumer with non-interruptable appliances, an energy storage system, and a controllable HVAC system. The authors form a mixed-integer linear programming problem with the objective to minimize the total electricity cost to the consumer while maintaining all system limits and dynamics. Additionally the authors claim their optimization solution is robust as it considers the impact of stochastic inputs. The paper compares the optimization solution to a reference algorithm where users obey some ruleset based on the state of the system.

    Relevance The paper serves as a good reference for a couple of aspects for my final project. First, the problem they are solving is very similar to mine and their formulation of the model including the energy storage system dynamic is helpful. Also, they present alternative methods to the Q learning technique to solve the scheduling problem and compare them. These will be useful baselines to implement if time allows.

17. J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Online modified greedy algorithm for storage control under uncertainty," *IEEE Transactions on Power Systems*, vol. 31, pp. 1729-1743, May 2016.

Summary The goal of this paper was to find a method to operate energy storage under uncertainty. The uncertainty they introduced was stochastic price and demand for a system. The authors propose a simple online deterministic optimization with an offline component to determine the optimal control of the energy storage system. The paper also includes a derivation of performance guarantees.

Relevance This work is relevant as it serves an example of a non-statistical learning method for solving a similar problem to the one I would like to solve. I will not have time to actually program and run their algorithm but it is useful to know the advantages and downfalls of other approaches.

18. G. Henri and N. Lu, "A multi-agent shared machine learning approach for real-time battery operation mode prediction and control," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 15, Aug 2018.

Summary The paper by Henri and Lu aims to perform real time battery operation across multiple energy storage devices using a machine learning approach. The general idea is that all devices share data to a learning aggregator which trains algorithms for the devices to use. The learning aggregator uses a neural network with one layer and 20 neurons. The paper compares this approach with an economic model predictive controller method and finds the new approach was faster, more accurate, and returned higher savings.

Relevance The paper is relevant to the final project because it shows an alternative statistical learning method to a similar problem. A slightly different application and data set are used but some pieces of information are still helpful. They define incremental energy storage modes besides just idle, charge, and discharge which I could use in my model.

19. T. Matiisen, "Demystifying Deep Reinforcement Learning," *Computational Neuroscience Lab*, 15-Dec-2015. [Online]. Available: Link[Accessed: 22-Apr-2019].

Summary The main goal of this article is to provide background and motivation for deep reinforcement learning. The author starts by explaining what task reinforcement learning strategies are trying to serve and their approach. They introduce the ideas of an agent and environment and represent the set of their actions, states, and rewards as a Markov decision process. Then, they formulate the objective of the algorithm in terms of the discounted future reward. This is then formulated into the Q learning algorithm. The article then goes on to cover the basics of the motivation and implementation of deep Q learning.

Relevance While the reasoning behind this article was to motivate and introduce deep Q learning, I found it as a great reference for learning about reinforcement learning and Q learning. The clear explanation made it easy to grasp the new ideas of agents and how they interact with there environment. The introduction to deep Q learning also opened up the possibility of expanding on the Q learning algorithm current developed if time permits.

20. A. S. Zamzam, B. Yang, and N. D. Sidiropoulos,"Energy Storage Management via Deep Q-Networks," *arXiv e-prints*, p. arXiv:1903.11107, Mar 2019.

Summary The work published just last month looks at how deep Q-networks can serve to operate energy storage unit in a microgrid with renewable generation. The paper enhances Q learning by using neural network to approximate the action-value function for continuous states. This approach avoids the discretization needed for classic Q learning. The authors compare results to an optimal solver approach with full knowledge of future prices and energy demand and generation as well as compare to a model predictive controller approach with knowledge of the short future.

Relevance The work provides a extension of the Q-learning method. While it is likely be too complicated an adaptation for the final project, it is still a useful reference. I liked the two baseline algorithms that they used as comparison and will try and use these as a comparison to the Q-learning algorithm I choose if there is time.

21. T. Hubert and S. Grijalva, "Modeling for residential electricity optimization in dynamic pricing environments," *IEEE Transactions on Smart Grid*, vol. 3, pp. 2224-2231, Dec 2012.

Summary The goal of this paper was to solve an optimal scheduling problem for the purchase, storage, and generation of electricity of a residential consumer. The model used was a residential electricity consumer with non-interruptable appliances, an energy storage system, and a controllable HVAC system. The authors form a mixed-integer linear programming problem with the objective to minimize the total electricity cost to the consumer while maintaining all system limits and dynamics. Additionally the authors claim their optimization solution is robust as it considers the impact of stochastic inputs. The paper compares the optimization solution to a reference algorithm where users obey some ruleset based on the state of the system.

Relevance The paper serves as a good reference for a couple of aspects for my final project. First, the problem they are solving is very similar to mine and their formulation of the model including the energy storage system dynamic is helpful. Also, they present alternative methods to the Q learning technique to solve the scheduling problem and compare them. These will be useful baselines to implement if time allows.

22. C. Guan, X. Lin, Y. Wang, , S. Nazarian, and M. Pedram, "Reinforcement learning-based control of residential energy storage systems for electric bill minimization," in *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 637-642, Jan 2015.

    Summary This work studies energy storage control using TD(λ)-learning algorithm under renewable generation uncertainty. The model of the system includes an energy storage system and a photovoltaic system connected to the grid. The authors introduce the TD(λ)-learning as a reinforcement learning algorithm to perform cost savings. They perform a case study and compare the results to a baseline strategy which charges the energy storage during off-peak hours and discharges during peak hours.

    Relevance This work provides another slightly variation to the Q-learning style reinforcement learning algorithms. The algorithm and results will be a useful comparison the classic Q-learning algorithm and can be implemented if time permits.

23. E. Kuznetsova, Y.-F. Li, C. Ruiz, E. Zio, G. Ault, and K. Bell, "Reinforcement learning for microgrid energy management," *Energy*, vol. 59, pp. 133 146, 2013

    Summary The authors in this work attempt to solve the optimal control of a microgrid using reinforcement learning. They model a consumer with load, renewable gen, and storage under market prices. They base the control strategy on Q-learning but extend the method to 2 time steps look-ahead. Additionally, they include a sensitivity analysis of learning parameters. A further contribution their approach has is that the model can account for uncertainty such as forecast error in renewable generation.

    Relevance The work of Kuznetsova et al. expands the ability of Q learning. This paper uses a more complicated model and calculations with multiple time steps that likely will not be incorporated in my final project. However, the paper provides further evidence of Q learning application and usefulness in the control of energy storage systems.

24. L. Xiao, X. Xiao, C. Dai, M. Peng, L. Wang, and H. V. Poor, "Reinforcement learning-based energy trading for microgrids," *CoRR*, vol. abs/1801.06285, 2018.

    Summary The paper covers deep Q-learning for connected microgrids. The authors propose a stochastic energy trading game to determine the optimal energy trading schedule. Each microgrid can choose amount of energy to sell or purchase from other microgrids or from powerplants. They authors also provide the Nash equilibrium for the game which alludes to the performance bound of the scheme used. Finally, the paper preseents results from simulations and shows how the scheme improves performance over a benchmark strategy, "hotbooting Q".

    Relevance The paper shows an extension of Q learning where netural network

is used. It also expands on the decision of one entity to a network of entities. The exact algorithm will not be used but the work presents a good reference and potential data acquisition strategies.

25. S. Kim and H. Lim, "Reinforcement learning based energy management algorithm for smart energy buildings," *Energies*, vol. 11, p. 2010, 08 2018.

    Summary The paper presents an application of Q learning for smart energy building actions to minimize cost. The work uses a model which includes utility prices, vehicle to grid technology, photovoltaic generation and an energy storage system. A slight adaptation of the classical Q-learning scheme is used to reduce convergence time by initializing the Q-table. The authors then show simulation result which even outperforms hourly optimization.

    Relevance The paper is a very relevant application of Q learning. It is a clearly written paper and a great reference for defining actions, states, and rewards. While the setup and model will be slightly different it was the only other reference that is using 5 minute datasets. The adaptation to improve convergence and their data sources are particularly useful for the project.

26. T. Orfanogianni and G. Gross, "A general formulation for lmp evaluation," *IEEE Transactions on Power Systems*, vol. 22, pp. 1163-1173, Aug 2007.

    Summary The goal of the paper is to give a general formulation for locational marginal prices. Locational marginal prices (LMPs) are defined as the least cost to supply an additional unit of load at specific location. The paper describes that LMPs are used in congestion management and hedging and their decomposition is needed for market settlement purposes and hedging. The formulation used breaks the LMPs into energy, loss, and congestion terms. The work then gives some examples of the decomposition on a six-bus network.

    Relevance While not directly related to Q-learning or energy storage systems, this paper gives good insight into LMPs. The paper provides some intuition into what factors make up a LMP and why they are an important number to use as a price. This work gives some justification for using LMPs as real-time energy market prices and to define our rewards for the energy storage system.

27. ISO New England, "FAQs: Locational Marginal Pricing," *iso-ne.com*, 21-Jul-2018. [Online]. Available: Link[Accessed: 22-Apr-2019].

    Summary This informational article from ISO New England gives an overview of locational marginal prices. First, they explain what location marginal pricing is and what a locational marginal price represents. Then they break down the components of a locational marginal price and explain what they are used for. Finally, they explain the difference types of locational marginal prices and how they are calculated.

Relevance This website is published directly from ISO New England which is the grid operator for the New England region and directly operates and oversees the wholesale electricity markets. Since I am using the locational marginal prices as my dynamic pricing for the energy storage problem this information is very useful for understanding what those values represent. This websites serves as a justification for using locational marginal prices as my price set point and as an input to my reward and objective functions for the energy storage problem.

28. ISO New England, "Pricing Reports," *iso-ne.com*, 21-Jul-2018. [Online]. Available: Link[Accessed: 22-Apr-2019].

    Summary This information which is published by ISO New England is the locational marginal prices for the New England real time electricity market. Avaialable on the site are the most recent 36 hours of data to be downloaded as csv. There is also a historical data lookup which allows for data to be downloaded as zip files for past locational marginal prices. The data downloaded includes the locational marginal prices as well as the three components that make up the locational marginal price for each five minute timestamp broken up by locational id. An auxilliary "pnode" table maps these locational ids to physical locations on the grid.

    Relevance This website provided the dataset that I will use for the prices of electricity in the energy storage problem. I used the historical data portal to download five minute data for February through March 2019.

29. B. Kim, Y. Zhang, M. van der Schaar, and J. Lee, "Dynamic pricing and energy consumption scheduling with reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 7, pp. 2187-2198, Sep. 2016.

    Summary The research in this work tried to answer the challenges of dynamic pricing and energy consumption scheduling of electricity for utility providers. Previously dynamic pricing schemes have been difficult do to the lack of customer information and uncertainties in forecasts. In this work, the authors use a reinforcement learning technique to try and solve this problem. Their results show that their proposed algorithm reduces system costs without future forecasts of system information. The authors develop a slightly modified Q learning method and apply it to a microgrid model and claim that their method performs better than an optimization approach.

    Relevance This paper solves a similar but not exactly the same problem as what I am looking at. Also they use a slightly modified Q learning algorithm. All the same, it is another example of how reinforcement learning algorithms are effective at temporal forecasting problems in the application of energy systems and dynamic pricing.

30. K. Rahbar, J. Xu, and R. Zhang, "Real-time energy storage management for renewable integration in microgrid: An off-line optimization approach," *IEEE Transactions on*

Smart Grid, vol. 6, pp. 124-134, Jan 2015.

Summary The goal of this paper is to develop an offline algorithm that can minimize the system cost for a microgrid under uncertainty. The model that is used is a single microgrid system with an energy storage system, renewable generation, and background load. Their approach is to assume a perfect forecast and solve an offline optimization problem and then use a sliding window online algorithm for feedback of the actual stochastic processes. The paper then compares this methods performance with other offline and online algorithms.

Relevance The problem that this paper is trying to solve is essential the same as the problem of my final project. The uncertainty of the energy system, including but not limiting to dynamic energy prices, make solving a scheduling problem difficult. While they use an offline optimization approach, I will be taking a reinforcement learning approach. Both make certain assumptions about how our past and predicted future should affect our current decisions. The work serves as a useful reference of another approach to a similar problem.

31. Y. Xu and L. Tong, "Optimal operation and economic value of energy storage at consumer locations," *IEEE Transactions on Automatic Control*, vol. 62, pp. 792-807, Feb 2017.

    Summary The goal of this work is to develop operation policies for maximizing a energy storage system owners payoff. The authors use a dynamic programming approach and incorporate stochastic electricity prices and fluctuations in energy demand. They then state that under special cases, inelastic energy demand, the dynamic programming algorithm can be simplified to a basic ruleset that allows the energy storage to use arbitrage to maximize the value of storage. This ruleset become more easily solvable by greedy algorithms if deterministic electricity prices are also assumed.

    Relevance This paper is useful for the final project as it provides an alternative approach to the reinforcement learning methods for a similar problem. While it starts off with dynamic programming which is very different than Q learning, with some assumptions the control algorithm becomes a simple ruleset which is not too different than Q learning after the Q matrix has been learned.

References