Optimal Real-Time Battery Scheduling with Reinforcement Learning and Neural Networks

by

Carolina Quiroz Juarez

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

Master of Science

in

Software Engineering and Intelligent Systems

Department of Electrical and Computer Engineering
University of Alberta

Abstract

Climate change concerns have raised awareness about the importance of decarbonizing the power sector. In achieving such a goal, energy storage is a critical operation that is currently done using mostly fossil fuels as a chemical energy storage. The only viable alternatives are battery energy storage systems (BESS) given their portability, scalability, and ease to install when compared with other storage technologies. BESS have been an important subject of research for decades. However, their massification has not been fully realized due to their cost and operational complexity.

The battery scheduling problem has been extensively analyzed and a great variety of algorithms have been proposed as a solution. Nevertheless, considering that BESS operation is highly dependant on the electrolyte chemistry, not all scheduling and control algorithms are useful for every real-time condition and every battery. Moreover, sophisticated high performing BESS control algorithms demand high computational resources that prevent them from being implemented in distributed energy systems. For instance, behind the meter (BTM) applications for residential buildings require real-time BESS control with high time resolution data.

In this work, we propose a real-time BESS control method based on reinforcement learning and neural networks aimed at working with reduced computational resources and independently from battery chemistry, which is then amenable to imbedded systems applications. On the one hand, neural network (NN) algorithms popularity stems from their ability to solve high-dimensional complex problems with minimal computational resources once the model has been trained. On the other hand, the NN training process requires high amounts of good quality labelled data. During this

project, we used 1-min resolution datasets containing photovoltaic (PV) generation, residential demand, and price signals. Notwithstanding, the datasets used did not contain BESS charge and discharge information. We, however, generated charge and discharge data with a reinforcement learning (RL)-based Q-learning algorithm that took into account the system characteristics of a real vanadium redox flow battery experimental setup as well as the technical features of a lithium-ion battery available in the market.

The RL-agent training process uses large amounts of data and takes considerable processing time to obtain an optimal policy for a daily operational period. Therefore, the RL-agent's main function is to generate labels to train different NN models, but not to be deployed on a real-time controller. The RL reward function privileged charge and discharge sequences that minimize final user costs compared to a PV system with no BESS. A positive reward was awarded every time the total electricity cost of a PV system was higher than the cost obtained with a PV-BESS system. All electricity costs were finally compared with the baseline PV system. However, the battery agent was not always able to decrease electricity costs below the baseline as its performance is dependant on battery size and efficiency. In turn, the scheduling labels resulting from the RL-agent operation allowed us to train our NN models with an accuracy such that we were able to abate PV system electricity costs.

Finally, an application of our workflow to the BTM problem is explored, by comparing the electricity costs calculated with both Q-learning and NN algorithms, to a residential flat tariff offered by a local electricity provider in Edmonton. Our simulation suggests that in the current scenario, it is still not economically viable to adopt BESS technology at a large scale in Alberta, Canada.

Preface

This thesis includes work from the article Carolina Quiroz-Juarez and Petr Musilek, "Optimal real-time scheduling of battery operation using reinforcement learning" that will be presented in the *Annual IEEE Canadian Conference of Electrical and Computer Engineering 2021 (CCECE)*. Design of the adaptive battery controller in Chapter 4 is taken from the cited article, and Chapter 5 contains results from a larger amount of experiments than those conducted for the article. The rest of this thesis is the sole original work of the author.

To my parents, Valentin Quiroz Zamora and Rocio Juarez Castro, my sister, Ernestina Quiroz Juarez, my grandfather, Eufemio Juarez Zamora, my late grandmother, Albina Castro Mora, and Bernal Manzanilla Saavedra, who have been my guidance in life.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Petr Musilek who opened the door to me for discovering new opportunities within intelligent systems area, his patience and support along my learning process help me do define new professional pathways. My gratitude extends to Bernal Manzanilla Saavedra, Raja Abou-Ackl and Roberto Vega Romero who helped me in critical moments in pursuing this professional goal and for their guidance from their areas of expertise. I also thank my colleagues within Professor Petr Musilek's research group for helping me in the process of learning about reinforcement learning, Daniel May, Peter Atrazhev, Steven Zhang, Vadim Avkhimenia, and also Kalvin Schofield for all of his help during the Vanadium redox flow battery experimental work. I thank the Fondo Sectorial CONACYT-SENER-Sustentabilidad Energetica of Mexico for providing finacial support to complete this work. Finally, I want to thank the Faculty of Engineering at the University of Alberta.

Table of Contents

1 Introduction **1**

1.1 Problem definition 1

1.2 Thesis Objectives 4

1.3 Thesis Outline 5

2 Background **6**

2.1 BESS for residential applications 6

2.2 BESS control methods 8

2.3 Reinforcement Learning 10

2.3.1 RL learning problem 11

2.3.2 Model-free and Temporal-difference learning (TDL) 12

2.3.3 Q-learning: an off-policy method 13

2.4 Neural Networks 14

2.4.1 Multi layer perceptron architecture 14

2.4.2 Training an MLP for multiclass classification problems 15

2.4.3 Model performance 17

3 Methodology and Experimental Design **19**

3.1 Methodology 19

3.2 VRF BESS setup 21

3.3 VRF BESS experimental characterization 25

4 Model Design **29**

4.1 Input data . . . 29

4.1.1 Demand and PV generation profiles . . . 29

4.1.2 Electricity Prices . . . 31

4.1.3 Battery data . . . 33

4.1.4 Data Preprocessing . . . 34

4.2 RL model: System Optimization . . . 35

4.2.1 Action Space . . . 35

4.2.2 Environment and State Space . . . 36

4.2.3 Reward Function . . . 37

4.3 NN Model: Battery Adaptive Control . . . 38

4.3.1 Offline Training . . . 39

4.3.2 Online Control . . . 40

4.4 Battery controller evaluation: operational costs . . . 40

5 Simulation Results **43**

5.1 Input data to simulations . . . 43

5.2 *Va1* and *Va2* BESS . . . 45

5.2.1 *RL* and *NN* compared to *flat tariff* . . . 46

5.2.2 *RL* and *NN* compared to *PV system* . . . 46

5.3 *Va3* and *Va4* BESS . . . 48

5.3.1 *RL* and *NN* compared to *flat tariff* . . . 48

5.3.2 *RL* and *NN* compared to *PV system* . . . 49

5.4 *Li* . . . 51

5.4.1 *RL* and *NN* compared to *flat tariff* . . . 52

5.4.2 *RL* and *NN* compared to *PV system* . . . 52

6 Conclusions & Future Work **54**

6.1 Conclusions . . . 54

6.2 Future Work 56

Bibliography **58**

List of Tables

3.1 VRF BESS hydraulic system and sensors. *Lpm = Liters per minute 21

3.2 Technical characteristics of VRF battery. 22

3.3 Technical characteristics of Li-ion battery. 22

3.4 Sample of battery charge and discharge cycles and calculation of CE, VE, and EE 26

3.5 60 charge and discharge cycles 27

4.1 Seasonal data for RL model. 31

4.2 Seasonal data for NN model. 32

4.3 Selected NN models. 40

5.1 Details of the data used to train and test the algorithms. 43

5.2 Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va1*. 45

5.3 Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va2*. 45

5.4 PV-BESS cost differences, *RL* and *NN Va1* battery controller. 45

5.5 PV-BESS cost differences, *RL* and *NN Va2* battery controller. 45

5.6 Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va3*. 48

5.7 Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va4*. 48

5.8 PV-BESS cost differences, *RL* and *NN Va3* battery controller. 49

5.9 PV-BESS cost differences, *RL* and *NN Va4* battery controller. 49

5.10 Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Li*. 52

5.11 PV-BESS cost differences, *RL* and *NN Li* battery controller. 52

List of Figures

2.1 Reinforcement learning mechanics. 11

2.2 MLP architecture, minimal number of layers. 15

2.3 Confusion matrix for a multiclass classification problem. 18

3.1 Research approaches, based on [49] 20

3.2 VRF BESS experimental setup 23

3.3 VRF BESS-PV system configuration 24

4.1 Demand and PV generation profiles (RL-summer). 30

4.2 Demand and PV generation profiles (RL-winter). 30

4.3 Demand and PV generation profiles (NN-summer). 31

4.4 Demand and PV generation profiles (NN-winter). 31

4.5 TOU reference rate structure. 32

4.6 Adaptive BESS controller schematics. 41

5.1 Input data used to simulate battery behavior. 44

5.2 Va1 system, one day in **summer**. 48

5.3 Va1 system, one day in **winter**. 48

5.4 Va3 system, one day in **summer**. 51

5.5 Va3 system, one day in **winter**. 51

5.6 Va4 system, one day in **summer**. 51

5.7 Va4 system, one day in **winter**. 51

5.8 Li system, one day in **summer**. 53

5.9 Li system, one day in **winter**. 53

List of Symbols

J	cost function
L	loss
Q	array estimates of action-value function
R_t	reward at time t
S_t	state at time t
V	array estimates of state-value function
$\boldsymbol{\theta}$	neural network parameters
γ	discount rate
$\mathbb{E}$	expectation operator
∇	gradient
π	policy
π_*	optimal policy
g	gradient function
h	neural network hidden layer
m	size of the neural network training set
n	number of neurons in hidden layer
q_*	optimal action-value

q_π	action-value function
v_*	optimal state-value
v_π	state-value function
x	input feature
y	neural network output

Abbreviations

ANN/NN - artificial neural network.

BESS - battery energy storage system.

BTM - behind the meter.

CC-CV - constant current, constant voltage.

DOD - depth of discharge.

DSM - demand side management.

FIT - feed in tariff.

FN - false negative.

FP - false positive.

ML - machine learning.

MLP - multi layer perceptron.

OCV - open circuit voltage.

PV - photovoltaic.

PV-BESS - battery energy storage system connected to a photovoltaic system.

RL - reinforcement learning.

SGD - stochastic gradient descent.

SOC - state of charge.

TDL - temporal-difference learning.

TN - true negative.

TOU - time of use.

TP - true positive.

VRF - vanadium redox flow.

Chapter 1

Introduction

1.1 Problem definition

Awareness about the imminent risks of climate change and global warming leads to creative solutions in the way we generate, transport, and use energy. Global greenhouse gas (GHG) emissions related to energy use account for more than 70% [1]; particularly in Canada, fossil fuel electricity generation plants were responsible for 10.9% of the total CO2e GHG emissions in 2015 [2]. This scenario has strengthened the number of efforts in many jurisdictions to decarbonize electricity power sectors through energy efficiency measures, a higher penetration of renewable energy technologies or the incremental use of nuclear power [1].

Canada's commitment to contribute to the reduction of the risks associated with climate change is centered on the Pan-Canadian framework, December 2016, which establishes 30% GHG reductions below 2005 levels by 2030. The framework consists of a) carbon pricing, b) complementary actions for reducing GHG emissions, c) adaptation measures, and d) support to low carbon technologies [3]. In line with this, the Canada's Energy Futures 2020 outlook (the outlook) establishes two main scenarios, evolving and reference, for the long-term energy sector development. The evolving scenario implies a reduction in the use of fossil fuels and, consequently an increase in the use of low carbon technologies. This scenario relies on the assumption that solar power technology and EV battery costs drop by 75% and 50%, respectively, during 2020 to 2050 [4].

The 2020 pandemic provoked 5.6% energy use decline in 2020, mainly due to the public health measures that instructed people to work from home. In Alberta, the estimation is 5% total electricity decrease in commercial and industrial demand because it shifted to residential consumption. The outlook´s evolving scenario reflects this tendency with an estimated energy use decline in 2050, driven mainly by an assumed higher electrification on the end-user side, which is supplied mainly by renewable power (including hydropower, solar, wind, and renewable fuels) and natural gas. Under this scenario, the current 16% electricity end-use demand increases to 27% by 2050, while wind, solar, and utility-scale battery storage capacities increase up to 40 GW, 20 GW and 3 GW, respectively. Yet, uncertainty in the outlook is related to solar, wind, and battery storage technology capital costs [4].

Four Canadian provinces, including Alberta, rely on coal and natural gas to generate electricity. According to Dolter and Rivers [2], wind and solar energy are able to facilitate the country's GHG reduction. Wind locations in southern Saskatchewan and Alberta would potentially be the least-cost option through the expansion of interprovincial transmission lines with no further requirement of energy storage. Nonetheless, this scenario depends on the transmission infrastructure capital cost; in case such cost is higher than modeled, the use of energy storage increases.

According to the International Energy Agency (IEA), photovoltaic (PV) technology remains as the energy policy backbone to help in the decarbonization of the power sector. Utility scale systems recorded levelized costs of electricity below 1.5 USD cents/kWh, and by the end of 2020, PV supplied more than 3.5% electricity generation globally. In Canada, the addition of new PV capacity concentrates in Ontario with 170.3 MW, Alberta with 31.8 MW, and Manitoba with 20.5 MW. Regarding the 3.3 GW cumulative capacity, 83% is connected to the low voltage distribution infrastructure. This industry added an estimated value to the Canadian economy of 403 MCAD in 2019. As a reference, commercial roof-top system prices varied between 1.80-2.5 CAN/W for 10-250 kW installed capacity systems [3]. Nonetheless, technology costs for residential applications may decrease in the outlook's scenario where end use increases in 11% by 2050.

Both solar and wind energy resources are extensively analyzed to help in reducing the electricity sector carbon footprint; however, its main drawback is the inherent intermittency of renewable resource-dependent electricity generation that limits the reliability of systems with a high penetration of such technologies [5]. Moreover, particularly solar PV for residential applications has an additional disadvantage, the mismatch between generation and demand. Solar power generation peaks during midday, but residential demand mainly is higher during evenings. Given the low demand when most of the PV power is produced, the power profile acquires the so-called 'duck curve' form that negatively impacts the distribution grid [6]. This problem has been tackled through the implementation of net or bidirectional metering programs where a small fee is paid back to the distributed generator. As long as this mechanism is not always economically attractive or nonexistent, the installation of battery storage with PV systems is attracting attention by allowing the shift of electricity from the time it is generated to the time it is required [7].

BESS economic viability is partially given by the optimal operation strategy materialized as savings in the final user's monthly bill, which implies scheduling charge-discharge cycles depending on the time when PV generation is available, retail electricity price, and household residential user's demand [8, 9]. The challenge is then, to find BESS-PV optimal operation strategies that allow the end-residential user to obtain electricity bill savings that contribute to the BESS-PV uptake.

More recently, machine learning methods have been utilized to tackle a wide range of problems within the energy sector given its great potential to solve complex tasks involving multiple highly stochastic variables. Two promising branches of machine learning, deep learning and reinforcement learning (RL), are applied in this work to define the optimal battery scheduling and then to predict the battery operation for residential applications.

1.2 Thesis Objectives

The ability of a RL agent to learn the optimal scheduling of a grid-tied battery and PV system has been shown [10]. However, there is room for improvement using data with higher resolution. Hypothetically, a higher resolution implies a greater amount of data, making the problem more complex and requiring more computation. Using neural networks (NN), once the optimal battery scheduling has been established, computation complexity reduction may be possible. The research question is if a NN real-time BESS controller, potentially implementable in a microprocessor, can reduce the computational complexity of a RL model-based controller.

Model performance will be evaluated by comparison of the electricity costs associated with the PV-system (baseline) and the battery system. To analyze the effect of regulatory incentives on the battery system economics, the battery controller based on time of use (TOU) and feed-in tariff (FIT) incentives is compared to a flat electricity tariff's scheme. Specific objectives are as follows:

1. To apply RL and NN algorithms to battery scheduling and to analyze of the potential improvement in terms of lower computation complexity along with

lower real-time battery operation associated costs.

2. To investigate seasonal BESS operation.

3. To examine BESS economic benefits under the effect of regulatory incentives.

4. To compare two BESS chemistries, lithium-ion and vanadium redox flow, for residential applications to analyze the potential uptake.

1.3 Thesis Outline

This thesis is structured as follows: Chapter 2 contains background information about BESS operation strategies and related research on BESS control methods. It is complemented by a brief introduction of the utilized machine learning methods, Q-learning and NN. In Chapter 3, the vanadium redox flow battery (VRFB) experimental setup and battery characterization are presented. The operational BESS strategy is explained along with the real data utilized in the simulations and the description of the BESS adaptive controller principle in Chapter 4. In this chapter, a brief explanation of TOU/FIT regulations is presented. Simulation results and corresponding analysis are included in Chapter 5, and conclusions as well as potential future work are found in Chapter 6.

Chapter 2

Background

2.1 BESS for residential applications

Electricity demand is variable during different parts of a day; depending on the period of time when electricity is consumed, it is defined as off-peak, mid-peak, or peak. Each tier is defined by the electricity system costs, e.g., at times of peak demand, power system operators must include expensive generation to supply the total demand. To decrease the high variation in demand, programs such as demand side management (DSM) have been implemented together with dynamic pricing tariff schemes such as time-of-use (TOU). The main goal of such programs is to displace demand from peak to off-peak periods through the minimization of the final user electricity bill, which is considered a strong incentive to change the end user's electricity consumption patterns [11]. Nonetheless, DSM programs have not been effectively implemented since final users are required to keep track of their consumption and electricity rates [12].

Load shifting is one of six DSM generic techniques [13–15]; when implemented through BESS systems, it automatically controls charge and discharge cycles by taking advantage of retail tariff differences established by TOU regulations, i.e., the BESS charges during periods when the electricity price is lower and discharges during on-peak periods [6, 16, 17].

Despite the capability of BESS to deal with solar PV intermittency and the mismatch between renewable electricity generation and residential demand, the technology has not yet been broadly adopted in the BTM market due to the technology's prevalent high costs [18]. A remarkable case is Germany, where li-ion BESS declined costs more than 50% between 2013 and 2018 [19]; actually, the net present value of residential BESS-PV systems is there positive [20]. By the end of 2018, 125,000 BTM BESS-PV systems had been installed in Germany [19]. Certainly, regulatory incentives in Germany have played a major role in their mass-market uptake, however, with declining feed-in-tariffs (FIT) and incentives, the question about the required conditions to have economically viable PV-BESS systems persists [21–23].

VRFB has been pointed out as a viable option for large-scale stationary applications because in some aspects they outperform li-ion batteries. For instance, they have a great lifetime, chemical compounds represent a lower risk for human activity, no risk of explosion has been recorded, their unique feature of independent power and energy sizing, larger depth of discharge (DOD) [24], competitive cost for larger capacities [25], and lower cost of manufacturing [26]. However, the technology costs reduction has not been sufficient to penetrate the residential sector.

BESS economic feasibility is highly dependent on the degree PV generation is self-consumed [22], i.e., it implies shifting PV generation for later utilization at a time when no solar energy is available [27]. This, in fact, improves the relative PV system cost by increasing the total amount of PV energy consumed [28]. In accordance to Johann and Madlener [29], and Luthander et al [30], a 0.5-1 kWh BESS system per installed kW PV is able to increase self-consumption by 13-24%. Moreover, Kucevic et al [31] highlights that self-consumption seems to be the main driver for BESS stationary uptake.

Based on the experience in Germany, where retail prices are high and wholesale price access is limited for distributed generators, regulatory incentives are important in the short term to rise investments on residential BESS-PV systems [22]. Especially FIT programs are considered highly effective regulatory instruments for the creation of renewable technology markets by mitigating the costs associated with the technology, among other benefits [32].

In 2009, Ontario provincial government in Canada launched the microFIT program that ended up in December 2017, very likely due to the loss of political support during its implementation [32]. Although this program is already finished, existing contracts received an economical incentive for 20 years of contract duration. Among the five countries in America, Canada and the United States have the highest scores in terms of policy density, which is the number of renewable energy policies related to a specific goal. Canada alone accounted the most acute budget intensity score that indicates that more of the designed energy policies have a budget, and as a result the country will more likely reach its goal in terms of GHG emissions reduction. Ten provinces in Canada created 18 policies between 1998 and 2015 [1]. However, Alberta did not design a FIT program; instead, it established the microgeneration regulation allowing, for instance, up to 5 MW PV systems to receive credits for the electricity sent back to the grid.

Independently of the regulatory incentives in place, the question about the conditions in which BESS-PV systems are an economically viable option for final users persists. This viability is partially supported by the optimal BESS operation strategy that provides savings on the monthly bill to the final user.

2.2 BESS control methods

Four types of algorithms have been used to control battery storage systems coupled with residential PV systems: optimization-based , rule-based, machine learning-based, and model-based [10, 33].

Optimization- and model-based methods usually provide high performance, however, they require high computation resources hardly implementable in microprocessors. Rule-based algorithm implementation is simpler, but they are designed according to the battery operation conditions; in the long term, this type of algorithms may deteriorate if not updated [33]. A review of the existing literature demonstrates the application of each type of methods.

The operation of a residential PV-battery system was simulated with two methods, rule-based and linear programming, to determine if the difference in resulting costs justifies the implementation of the simpler method [28]. Authors found that the rule-based resulted in 5% higher costs with the advantage that it only requires real-time information. In [11], authors employed the sequential quadratic programming method to obtain the maximum profit of a BESS based on real-time prices. The analysis included the comparison of two flow battery chemistries, polysulfide bromine and vanadium. Results demonstrate that the vanadium redox flow battery obtains higher annual revenue, but with a larger payback period based on the optimal operation strategy defined by the optimization algorithm. Bergner et al. [23] explored the linear optimization solved with the simplex algorithm to define the operation strategies of a residential PV-battery system with feed-in power limitation. According to their results, adaptive PV forecasts and intelligent battery operation are economically feasible in the long-term. Residential PV-battery grid-tied systems have demonstrated its potential to reduce the final user's electricity bill by determining the optimal battery scheduling, in this case two different methods were used, stochastic mixed integer nonlinear programming (MINLP) solved with metaheuristic techniques [34] and particle swarm optimization (PSO) [9]. MINLP allowed an annual electricity cost reduction of 58.65%, while with PSO the reduction accounted 4.2%.

Matallanas et al. [35] successfully implemented an active demand-side management and battery load shifting strategy aimed at maximizing the residential self-consumption rate with neural networks (NNs). In this case, the authors tuned NN parameters with a genetic algorithm. At a higher level, to analyze the penetration of BESS in power distribution networks, the optimal scheduling of such devices was modelled with Markov Decision Processes (MDP). The objective of reducing the energy cost and network losses was only reached by the defined optimal policy; the revenue was strictly higher than with the other two defined policies, and the losses reduced a maximum of 1.62% [36]. Another branch of machine learning refers to reinforcement learning (RL) algorithms that rely on an agent with no previous experience to gradually learn an optimal policy. Bijapur et al [37] as well as Kim and Lim [38] simulated energy management systems oriented towards reducing energy costs; in both cases, RL algorithms provided significant cost optimization. Contrastingly, in [10], the economic improvement was marginal when comparing the results of a battery system to a battery-less system, and even lower when compared to a mixed linear programming (MILP) method. In this case, Graham trained the battery-RL agent using the proximal policy optimization in a continuous action space.

Machine learning- and model-based methods are combined by Henri et al [33, 39], and Kazhamiaka et al. [40] for adaptive battery control in household applications. Model-predictive control (MPC) was first employed to define the optimal real-time battery control strategy that maximizes profit depending on the time the electricity is purchased or sold back to the grid i.e., based on a time-of-use pricing scheme. Results obtained with the MPC method were then compared with supervised machine learning algorithms; NNs and support vector machine (SVM) in [33]; NNs, SVM, logistic regression, and random forest in [39]; and NNs in [40]. According to Kazhamiaka et al., NNs approached the MPC-level performance, however, Henri reported in [33] 85% prediction accuracy versus the 72% reached with the MPC algorithm. It is remarkable that the improvement was possible at a fraction of the MPC computational cost when using machine learning-based methods.

2.3 Reinforcement Learning

Reinforcement learning (RL) is considered a third machine learning paradigm, which, different from supervised and unsupervised learning approaches, relies on agent learning through its interaction with the environment based on trial-and-error action selection. Like how human learning functions, the agent is not previously told what action to take to obtain the higher reward from the environment, although its main goal is to get the highest possible accumulated reward in the long run by selecting optimal actions. Given this nature, there are two actions the agent performs; exploration of different actions to 1) observe the reward coming back from the environment and 2) finding possible better actions; exploitation of the actions previously taken to improve what the agent already knows trying to gather more rewards. This is a challenging trade-off between exploration and exploitation in RL.

2.3.1 RL learning problem

The learning agent's problem is captured by Markov decision processes (MDP), where the agent observes the state of the environment and based on the information attained, it takes one of multiple possible actions. During this process, the agent receives a reward, a positive or negative number, depending on the desirability of the transition from state to state, Figure 2.1. When looking at this process from a higher perspective, what the agent is doing when it observes a state and selects an action is that it learns a policy that lets it gather rewards. Reward accumulation modelling depends on the type of task being solved, for instance, a finite horizon model of determined length and state assumes the expected reward is just the sum of the rewards over that horizon [41].

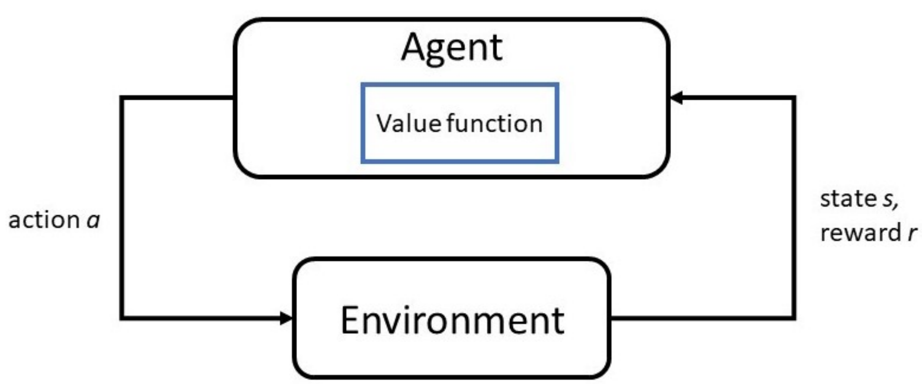


Figure 2.1: Reinforcement learning mechanics.

The reward positively and immediately incentivizes a good action of the agent, but in the long run, the *state-value function* accounts for the rewards the agent is expected to receive when in state s, in other words, the *state-value function* v_π of a state is the expected reward the agent can accumulate under a policy π.

$$v_\pi(s) \doteq \mathbb{E}_\pi\Big[\sum_{k=0}^{\infty} R_{t+k+1}|S_t = s\Big], \tag{2.1}$$

Likewise, the *action-value function* q_π is the expected return of taking action a when starting in state s under a policy π.

$$q_\pi(s,a) \doteq \mathbb{E}_\pi\Big[\sum_{k=0}^{\infty} R_{t+k+1}|S_t = s, A_t = a\Big], \tag{2.2}$$

As the agent's goal is to obtain the maximum possible reward in the long run, it is precise to identify the optimal policy π_* or policies whose expected return is greater when compared to other policies for all states. All optimal policies share the same optimal state-value, v_*, and optimal action-value, q_*, functions; in fact, the state value under an optimal policy equals the maximum expected return for the best action from that state [42].

$$\begin{aligned} v_*(s) &= \max_{a \in \mathcal{A}(s)} q_{\pi_*}(s,a), \\ &= \max_a \mathbb{E}[R_{t+1} + \gamma v_*(S_{t+1})|S_t = s, A_t = a], \end{aligned} \tag{2.3}$$

The optimal action-value, similarly, is given by the expected return when taking the optimal action a' at each following state in the following Q-function.

$$q_*(s,a) = \mathbb{E}[R_{t+1} + \gamma \max_{a'} q_*(S_{t+1}, a')|S_t = s, A_t = a], \tag{2.4}$$

The second part of equation (2.3) and equation (2.4) are the Bellman equations that recursively solve the MDP problem by giving the optimal policy to the agent [42].

2.3.2 Model-free and Temporal-difference learning (TDL)

Optimal policy computing is possible by using either model-based or model-free methods. Model-based methods can be solved with dynamic programming algorithms since the agent is able to estimate the model of the environment for the posterior policy computing. As the term suggests, model-free methods assume the agent has no access neither to the transition model nor the reward model. What the agent can do is sampling and exploring to directly estimate action values, in other words, the agent uses its own experience to solve the prediction problem. Once the agent starts acting, it receives information of the state of the environment and selects an action, it then receives feedback in the form of a reward and updates the estimated value of the next state. This working principle allows value updating along the way without waiting for the update at the end of the process and constitutes the main principle of TDL. As long as TDL is based on the estimation of other values, it is known as a bootstrapping method with the advantage of being implementable in an online fashion [41].

$$V(S_t) \leftarrow V(S_t) + \alpha[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)], \tag{2.5}$$

2.3.3 Q-learning: an off-policy method

Q-learning is considered an off-policy method because it does not depend on a specific policy to approximate the optimal action value function. The only requirement for convergence is that the agent updates continuously all state-action pairs. Equation (2.6) describes the Q-learning update rule that follows the same working principle as TDL. Each time step k the algorithm estimates the Q-value of the action taken in the state where the action is performed, and considering two elements, the reward and the agent's Q-value function, the Q-value of the next state is determined to update Q_k to Q_{k+1}.

$$Q_{k+1}(S_t, A_t) \leftarrow Q_k(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q_k(S_{t+1}, a) - Q_k(S_t, A_t)], \tag{2.6}$$

2.4 Neural Networks

Supervised learning paradigm has been extensively used to solve classification and regression tasks; depending on the problem intricacy, the solving method is selected. Neural network is a popular method because of its flexibility to solve highly diverse problems with different grades of complexity, and its ability to fast process new data once the model has been trained with a large number of labeled examples. For instance, solving a problem with a higher degree of complexity implies the addition of more layers in the network or the number of nodes within a layer.

2.4.1 Multi layer perceptron architecture

Cutting-edge neural network research has proposed different network architectures that apply in specific knowledge realms offer superior performance, yet the multilayer perceptron architecture (MLP) is still the backbone of neural network models. MLP is a feedforward neural network whose working principle is based on mapping an input value x to an output category or target value y while learning some parameters θ that result in the best function approximation [43].

Feedforward MLP architecture relies on the neuron as its basic unit to process information, i.e., where a weighted sum of input signals is passed to a transfer function. More than one unit, a node or neuron vertically located one next to another, conforming a layer. The minimal number of layers in a MLP is one hidden and one output layer; the number of neurons in the output layer corresponds to the number of categories we are mapping the inputs to in a classification problem. Neurons in the same layer are not connected to each other, thus they do not communicate; communication is just given between fully interconnected layers. The feedforward term refers to the fact in MLP architecture the information is forward propagated from left to right, with no feedback within the network.

Designing the architecture of an MLP involves the definition of the number of hidden layers, the number of neurons within each hidden layer, and the selection of the appropriate activation function. A diagram of an MLP is shown in 2.2.

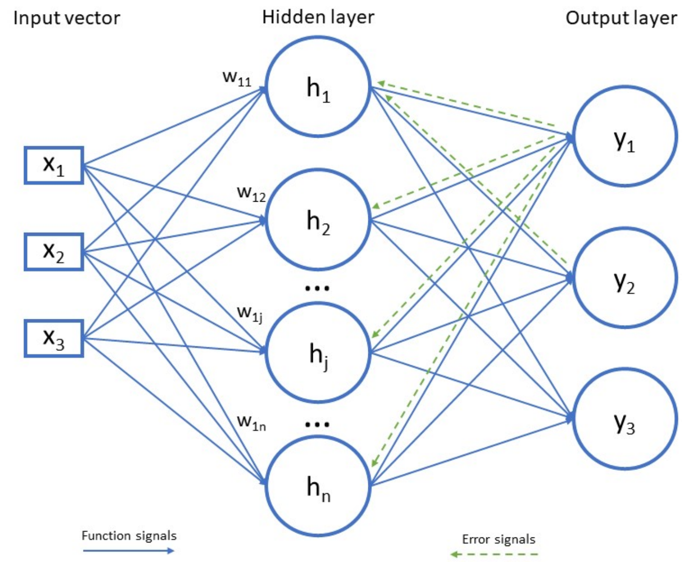


Figure 2.2: MLP architecture, minimal number of layers.

2.4.2 Training an MLP for multiclass classification problems

MLP training or learning refers to the process of finding the value of parameters θ (both bias and weight terms); this process that is carried out with the back-propagation and stochastic gradient descent algorithms. According to Haykin [44], the training process includes two phases denominated forward and backward.

During the forward phase, the input signal is propagated through each hidden layer up to the output layer. First, the input signal is multiplied by the fixed weights associated with each neuron link in the first hidden layer. Second, the sum of all terms in the first step is calculated. Third, the weighted sum is subjected to the activation function and the resulting signal is input to the j^{th} hidden neuron. This process is repeated for all hidden layers throughout the MLP until the output layer.

Supervised learning problems require labeled training datasets to train an MLP. For classification problems, the label of a training example predicted during the forward phase is compared to the label of the corresponding example in the training dataset, if both labels are different, an error signal is generated. The error or loss function expresses the errors as a function of the MLP parameters.

The backward phase, as the name suggests, is the propagation of the loss function through the network in the opposite direction to the forward process, i.e., from the output to the hidden layers (green dashed lines in Figure 2.2). This phase aims at adjusting the values of the parameters associated to each neuron link while looking for the optimization of the classification performance; this is based on the principle that the weight terms modify the network's behavior [45]. Adjusting MLP parameters is a gradual optimization problem driven by the stochastic gradient descent (SGD) algorithm that minimizes a cost function $J(\theta)$ conceptualized as an average over the training set in equation (2.7). L is the error between the model's predicted output $f(x, \theta)$ for an input x, and the target y, and p_{data} is the empirical distribution of the training set [43].

$$J(\theta) = \mathbb{E}_{(x,y)\backsim\hat{p}_{data}} L(f(x;\theta), y), \tag{2.7}$$

Every time the backward phase is carried out, the partial derivative of each parameter is calculated and the gradient updated. The gradient is then used to update parameters θ; in equation (2.8) $\hat{g}$ is the gradient and m is the size of the training set.

$$\hat{g} = \frac{1}{m}\nabla_\theta \sum_i L(f(x^{(i)};\theta), y^{(i)}), \tag{2.8}$$

One forward and one backward pass accomplishes an epoch; particularly for neural network models, updating the MLP parameters is a process that requires minimally thousand epochs.

Multiclass classification problems are a generalization of binary classification problems; as such, they consist of datasets with C distinct classes of data where each label y_p takes a value between 1 and C [46]. Although multiclass classification has been solved with extensions of binary classification methods, algorithms such as neural networks can be naturally applied by using the *softmax* activation function [47] and the *categorical cross entropy* loss function.

2.4.3 Model performance

It is a common practice to divide the dataset into a training, validation, and testing set; the training is the largest set and is utilized to build the model, the validation set is for identifying the best hyperparameters, and the testing set to evaluate the selected model. A model capable of correctly predicting most of the labels in the testing set has a low bias and is considered to generalize well, while the opposite case sheds light on the need of model improvement due to underfitting. Overfitting or high variance happens when the model is too complex and perfectly predicts on the training set, however, when tested with the validation and the testing set, it performs poorly [47]. Underfitting solutions include regularization methods that decrease model complexity at the cost of increasing the models' variance; this is known as the bias-variance trade-off.

Model's performance is assessed in machine learning with the use of metrics, for instance, for classification problems the most popular metrics include the confusion matrix and accuracy.

The confusion matrix summarizes in a table format the actual and the model's predicted labels; the actual labels are shown in one axis while the predicted ones lie in the other axis. The diagonal of the matrix accounts for the correctly predicted labels of each class, while the incorrectly labeled examples are visualized outside the diagonal. For multiclass tasks, the number of rows and columns in the table reflects the number of classes in the classification problem, this is, confusion matrices are squared matrices.

A more refined performance analysis originates in the type of errors defined for binary classification problems. Correct classification is either a true positive (TP) if the actual label is positive and it is correctly classified as positive, or a true negative (TN) in case the label is negative and correctly classified as negative; incorrect classification has also two possibilities, false positive (FP) if the example is negative and classified as positive, or false negative (FN) when a positive sample is classified as negative. According to Figure 2.3, FP for class 1 corresponds to the sum $E_{21} + E_{31}$, while the FN for class 3 equals the sum $E_{31} + E_{32}$ [48].

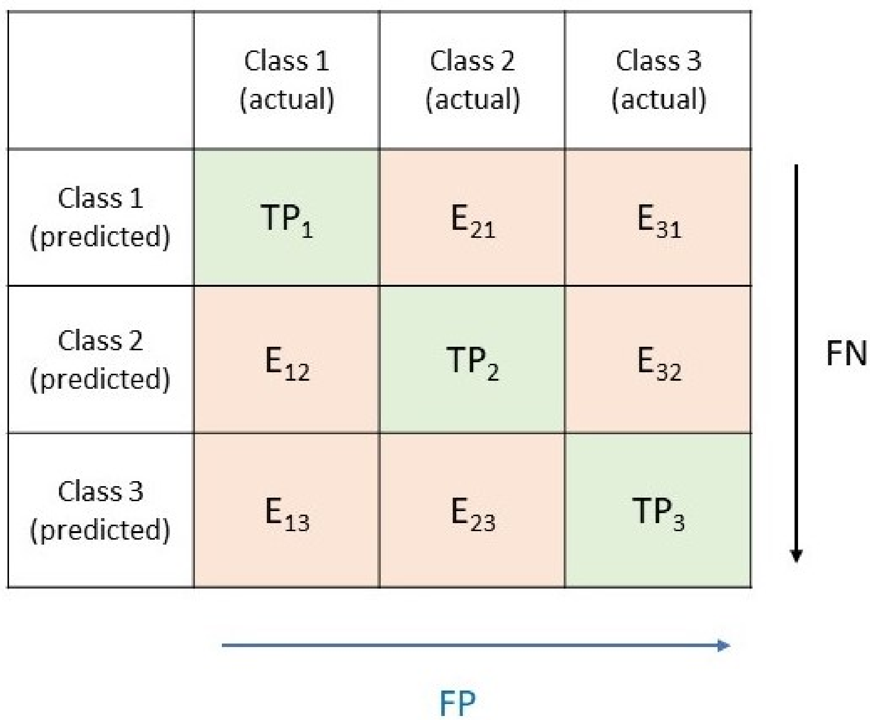


Figure 2.3: Confusion matrix for a multiclass classification problem.

Accuracy in Equation (2.9) accounts for the number of correctly classified examples over the total number of classified examples. This metric ranges between 0 and 1, being 0 in the case when the classifier performed poorly and misclassified all examples, and 1 or 100% corresponds to the perfect case when all examples were perfectly classified.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2.9}$$

Chapter 3

Methodology and Experimental Design

The following chapter defines the methodology followed in developing the research work. It also describes the experimental setup and the tests carried out to characterize the installed VRF BESS.

3.1 Methodology

According to Figure 3.1, this thesis work relies on two research approaches, experimental and simulation. Experimental work was mainly carried out during the first VRF BESS research project's stage, in which a battery system was installed at Nisku industrial park, and then reinstalled at the University of Alberta for its final stage. Research objectives during the experimental work aimed at understanding VRF BESS working principles, starting and operating the battery system, and characterizing the system based on its electrochemical performance. The BESS experimental setup and characterization are correspondingly detailed in sections 3.2 and 3.3.

Simulation work involved the construction of a residential environment using both, real rooftop PV generation and household demand profiles, additional to the VRF BESS information gathered during the experimental phase. The objective during this stage was designing a BESS control strategy based on a load shifting technique in two phases. Firstly, data that allow the observation of the VRF BESS dynamic

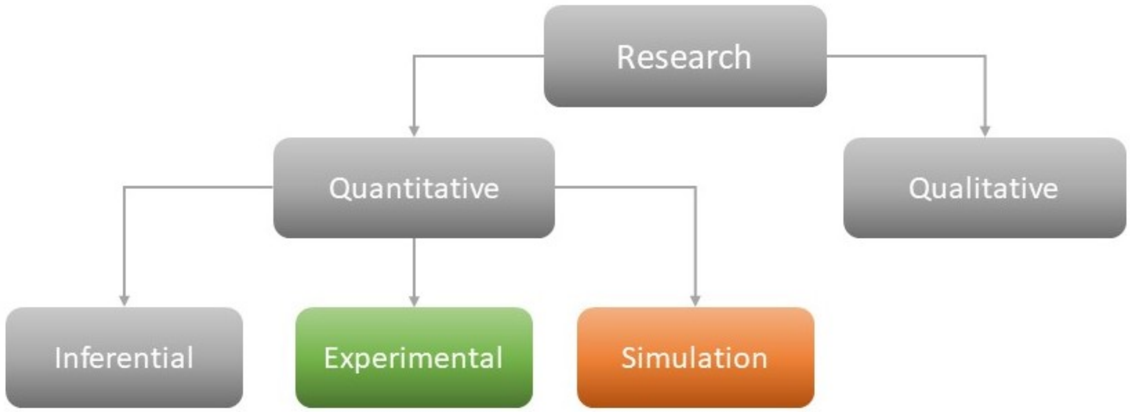


Figure 3.1: Research approaches, based on [49]

behavior for residential applications under real conditions was generated with a RL model-based Q-learning algorithm. Secondly, the generated data were further used in the design of the NN-model based BESS adaptive/intelligent control, as explained in chapter 4. For comparison purposes, the costs associated to the VRF BESS were compared to the costs associated to a lithium-ion BESS in the last section of the following chapter as well.

Engineering research methodology in this work, as agreed by [49] and [50], comprised the following execution process. First of all, the research problem hypothesis and research question are defined in chapter 1 to frame the research work. This step is supported by an extensive review of the existing literature about renewable and sustainable energy technologies in the Canadian context; implemented mechanisms to incentivize the installation of BESS and its impact toward the technology uptake at the residential level; VRF BESS electrochemistry, technology working principles, as well as previous experiences of its characterization and operation; battery technology costs as the main barrier for the broad adoption of the technology and cost decrement strategies through optimal battery operation scheduling; and battery control methods, particularly methods based on machine learning algorithms such as reinforcement learning and neural networks.

Once the problem was framed, the research design involved the elaboration of the outline, highlighting core concepts, and determining the samples to work on. This included the selection of residential load and PV generation profiles used to simulate the operation of a VRF BESS and a li-ion BESS. Data collection involved the use of standard tests to characterize the installed VRF BESS; this information was later used in the referred simulations.

Execution of the project alludes to the two-phase battery control presented in this work. The battery optimized scheduling is obtained in the first phase with a Q-learning algorithm and then utilized in the second phase for the adaptive battery control using neural networks. Finally, the results of both algorithms are statistically analyzed to test the hypothesis in the last part of Chapter **??**.

3.2 VRF BESS setup

The installed VRF BESS for demonstration purposes consists of a 40-cell stack where oxidation-reduction processes occur, two 83 L polyethylene tanks that serve as the positive and negative electrolyte reservoirs, two centrifugal pumps providing the electrolyte flow rates and instrumentation such as inlet and outlet pressure sensors to monitor the inlet and drop pressures in the stack, and flow meters. Technical specifications are shown in Table 3.1

Centrifugal pumps 24 VDC	Max. flow: 28.34 Lpm
Flow meter	Flow range: 0.4 Lpm-18.9 Lpm
Pressure transducer	Max. pressure: 1 bar

Table 3.1: VRF BESS hydraulic system and sensors. *Lpm = Liters per minute

Battery system connection to the low voltage grid is carried out through an inverter able to deliver 3 kW on the AC side, that also works as a battery charger with adaptive

4-stage charging modes (bulk, absorption, float and storage). Details of the VRF battery, the inverter, and a lithium (li-ion) battery for further comparison purposes are shown in Tables 3.2 and 3.3, respectively.

BESS element	**Technical features**
VRF battery stack	Volterion, 40 cells, 2.5 kW Efficiency: 82%; $DOD = 100\%$
Vanadium electrolyte	Oxkem, 40 L
Power inverter	Victron Energy - Quattro 48 Max. Efficiency: 95%

Table 3.2: Technical characteristics of VRF battery.

BESS element	**Technical features**
Li-ion module	Pika Energy – Harbor 6, 6.7 kW Efficiency: 96.5%, DOD = 84%
Total energy capacity	20.3 kWh
Usable energy capacity	17.1 kWh
Power inverter	Pika islanding inverter Max efficiency: 98%

Table 3.3: Technical characteristics of Li-ion battery.

Start up VRF battery control is based on an in-house algorithm programmed in LabView that is connected to a National Instruments data acquisition unit (DAQ 6211) to obtain pressure and flow measurements. Through the LabView program, a voltage command is sent to start up the centrifugal pumps which are usually warmed up for a while before starting charging the VRF BESS.

Battery characterization procedure explained in the following section is supported by the data acquired during the battery charging and discharging cycles. Battery charge is controlled via a modbus interface with the inverter. As reported in the literature [51, 52], the most utilized constant current – constant voltage (CC-CV) charging strategy was established by firstly setting the charging current in the inverter interface. As long as the battery voltage is not smaller than the charge voltage, the battery charges at constant current; once this condition changes, the battery charges at a constant voltage reducing the charging current to fill the battery charge at a lower pace. To complete the cycle, the battery was discharged by connecting a passive load consisting of a three-phase resistor box that allowed a maximum discharge current of 5.5 A. Further experiments allowed to set higher charge currents and a larger passive load to increase the discharge current up to a maximum of 47 A. In all cases, stack voltage and current data were acquired at a 1-min resolution with the stack controller purchased from the VRF battery provider.

An operational particularity of this technology is to prevent that the charged V^{2+} oxidases to V^{3+} by ensuring the system is air tight; a common practice is purging nitrogen gas into the negative electrolyte container to displace oxygen throughout the system [53–55] as shown in Figure 3.2.

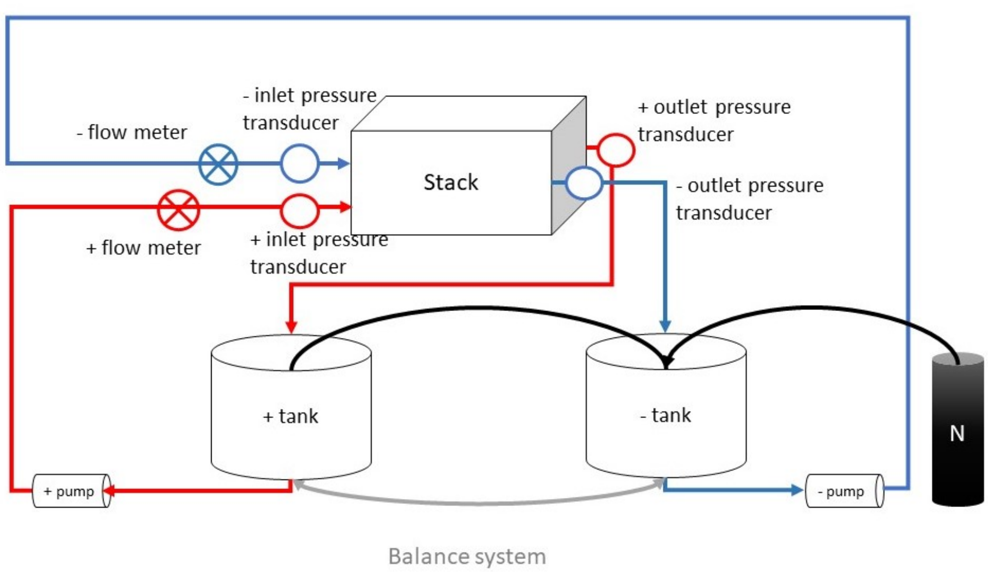


Figure 3.2: VRF BESS experimental setup

Differently to other research works [53, 56–58], in this case the open circuit voltage (OCV) reference cell was not installed.

The simulation work explained in the following sections considers two BESS-PV systems, each with different battery chemistry. One system configuration consists of a VRF BESS, a rooftop PV system, the corresponding inverters, and a residential load. The second system is equally configured, but instead of the VRF BESS, it includes a li-ion battery. In both cases, the BESS-PV systems are grid-tied for the exchange of energy with the distribution electrical grid. Figure 3.3. contains details of the VRF BESS-PV system configuration.

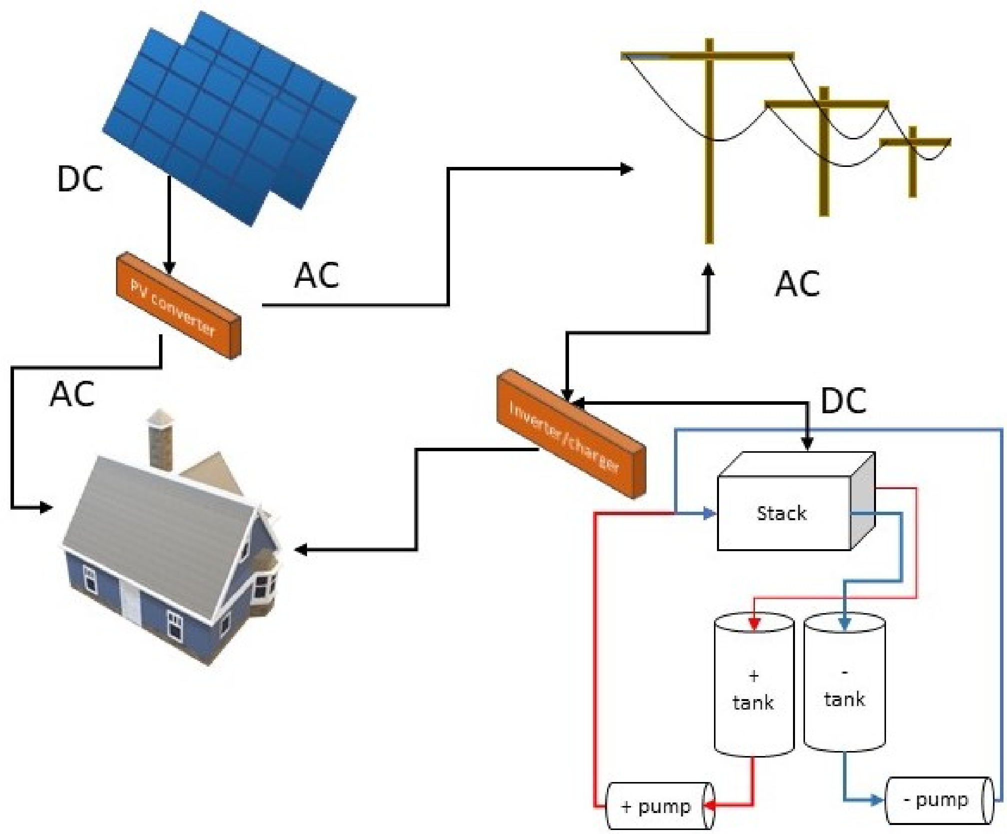


Figure 3.3: VRF BESS-PV system configuration

3.3 VRF BESS experimental characterization

Fully characterizing a VRF BESS, as for any battery chemistry, is a keystone step to understand the potential behavior of the battery under real operation conditions. Particularly, when connected to a PV system, its charge and discharge cycles depend on both external conditions, like PV generation and load, as well as on internal conditions such as its state of charge (SOC). In this order of ideas, the VRF battery system was characterized in terms of its actual capacity, SOC, and efficiency, as it has been previously reported in other research works [53, 56, 58–63].

VRF technology relies on the four possible oxidation states of vanadium (V2+, V3+, V5+, V4+), which means battery charge and discharge processes result when reduction-oxidation (redox) reactions cause electron transfer from one of such vanadium species to another. Two decisive parameters are the electrolyte flow rate and electrical current flowing through each cell in the stack as they define the reaction rate and vanadium concentration inside the cell, i.e., both define the rate at which some vanadium species are produced and others depleted. Indeed, battery operation should guarantee vanadium ion concentrations are not depleted below zero, in other words, the cells in the stack should not overcharge to avoid side reactions such as oxygen or hydrogen evolution that causes loss capacity and negatively affects the battery performance [64–66].

To minimize the likelihood of side reactions, the VRF BESS operation is controlled to charge and discharge within the upper and lower cut-off voltages, with the upper (charge) cut-off voltage being the most critical parameter that may damage the electrode materials if exceeded [66]. Cut-off voltage ranges between 1.26 and 1.41 V, which is slightly lower than previously reported in literature [53, 66] because the charge and discharge currents when characterizing SOC were lower than nominal.

Actual battery capacity was determined by fully charging at 30 A in CC-CV mode up to the upper cut-off voltage and discharging at almost nominal current (47 A) in

Table 3.4: Sample of battery charge and discharge cycles and calculation of CE, VE, and EE

Charge DC energy [Wh]	Max. charge power [kW]	Disch. DC energy [Wh]	Max. disch. power [kW]	CE [%]	VE [%]	EE [%]
905.42	2.16	423.18	2.49	81	52	42
973.92	2.16	489.62	2.54	92	45	41
954.34	2.20	459	2.46	86	49	42
1017.83	1.93	483.44	2.48	84	48	40
1313.65	1.89	348.84	2.35	76	28	21
1112.44	1.91	264.76	2.21	76	27	21

constant current mode to % 0 SOC. In contrast to li-ion batteries, the VRF BESS capacity is determined by the amount of electrolyte in the reservoir tanks. In this set of tests, the battery capacity corresponds to 40 L electrolyte, as shown in Table 3.4, considering the charge voltage reached 62 V.

A common phenomenon when operating VRF BESS is the loss of capacity due to cross-mixing of vanadium species through the membrane in the cell that is appreciated as a shift of the electrolyte from the positive to the negative reservoir. As this condition has been eased before, the electrolyte was manually remixed.

Battery SOC is calibrated based on the OCV-SOC relationship [60, 67]. OCV is the voltage at the battery terminals when the battery is not connected to a load, i.e., there is no external current flowing. For measuring the OCV, the battery was fully charged with the CC-CV strategy up to the charging cut-off voltage and allowed to rest for 45 minutes, then it was discharged in periods of five minutes with 10 minutes rest between measurements down to the discharge cut-off voltage. A second set of experiments reduced the resting time between measurements to 5 minutes. All experiments aimed at OCV-SOC characterization were made within the second-tier measurements in Table 3.5 that resulted in the cut-off voltages above mentioned for

Table 3.5: 60 charge and discharge cycles

Number of cycles	Max. charge current [A]	Max. discharge current [A]	Average CE [%]	Average VE [%]	Average EE [%]
13	16	6.18	69	88	66
23	30	6.18	60	88	53
8	33	6.08	66	82	54
16	39	46.98	75	42	31

a 10% - 90% SOC range.

Certainly, the literature reported up to five hours relaxing time to obtain the OCV measurements for li-ion batteries [67–69] but in this case for a liquid-state battery, a shorter resting time was perceived enough [70]. OCV was finally non-linearly fitted as a function of the SOC with a third-degree polynomial regression [71], equation 3.1.

$$V_{OC}(SOC) = \sum_{i=0}^{n_p} c_i SOC^i, \tag{3.1}$$

VRF BESS efficiency is evaluated in terms of three performance aspects: Coulombic efficiency (CE), energy efficiency (EE), and voltage efficiency (VE). CE is calculated as the quotient between battery discharged capacity (Ah) and charged capacity (Ah) and reflects battery inefficiencies at the electrochemical level, for instance, due to electrolyte species crossover or side reactions. EE assesses the stack energy conversion capability as it corresponds to the ratio of discharged energy (Wh) and charged energy (Wh) and is considered of utmost importance when assessing the battery's performance for real applications since it is an indicator that accounts for energy consumption by auxiliary systems such as pumps, ventilation or cooling system. VE describes the relation between the average voltage at discharge and the average voltage at charging and reflects the ohmic and polarization losses [64, 66].

During a period of almost six months, the VRF BESS was cycled around 60 times at different charge and discharge current rates, where four tiers were defined as shown in

Table 3.5. Of particular importance are the 16 last cycles because they reflect charge and discharge currents close to the nominal battery current.

In 13 out of 60 cycles, the EE of the stack registered values over 65%, which is a reasonable value similar to what is being reported in other research works [53, 62, 63]. Nonetheless, there is a clear need to improve the operation of the system. As it can be seen in Table 3.5, the increase of charge and discharge currents plumbed the average battery EE and VE.

Inherent internal losses in the stack, referred as overpotentials, are associated to the energy required to carry out redox reactions in the cells [65]. Concentration overpotential depends on the current applied to the battery and the electrolyte flow rate, while the ohmic overpotential depends on the temperature [64]; in this case, the electrolyte flow rate was no optimized, so there is an opportunity to improve the system operation. Also important is the reduction of energy losses due to electrolyte leakage.

Chapter 4

Model Design

As described in Chapter 3, subsequent sections describe the input data to the ML model and the algorithms used to simulate both a VRF and a li-ion BESS in a BTM-residential load shifting application.

4.1 Input data

In general terms, ML algorithms are algorithms able to learn from data [43] that require, first of all, large datasets. Whereas there is no agreement about the proper amount of data to train a particular model, it can be said that it depends on the problem's complexity or on the number of instances or features. Given the importance of input data in ML models, the datasets used to model the proposed PV-BESS systems are herein described.

4.1.1 Demand and PV generation profiles

Input data includes 1-min resolution household demand and PV generation profile datasets that correspond to a house building located in Edmonton, Alberta, Canada. Data were collected during a whole year, which means 1440 measurements per day or 525,600 measurements per year. Data included timestamp, total demand, PV generation, electricity imported from the grid, and the electrical energy consumption of the most electricity consuming appliances, although for the model's objective just the timestamp, total demand and PV generation were used.

One year data enabled the analysis of BESS functioning under different weather conditions, in other words, considering different loads and PV generation throughout the year. Given the great granularity of the datasets and to avoid the nonrepresentative problem in the data described in [72], the Q-learning algorithm was trained with 30 days in winter and 30 days in summer; these days exclude weekends according to the daily variation given by the TOU retail electricity tariff later described. Demand and PV generation profiles corresponding to 15 days in summer and 15 days in winter can be appreciated in Figures 4.1 and 4.2. Differences between summer and winter training data are shown in Table 4.1.

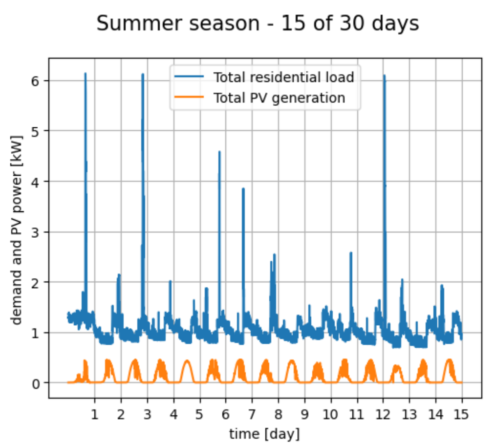


Figure 4.1: Demand and PV generation profiles (RL-summer).

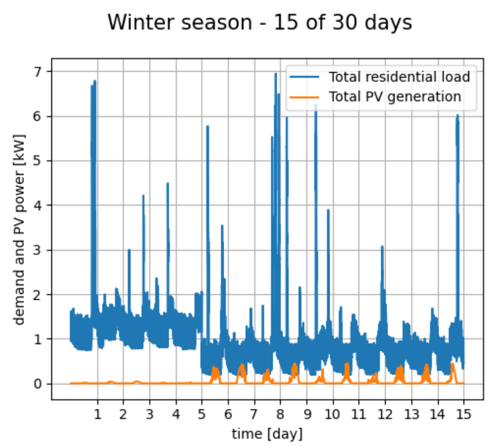


Figure 4.2: Demand and PV generation profiles (RL-winter).

A second set of 60 days (winter and summer) was used to train and test the NN models, the 15-days profile is shown in Figures 4.3 and 4.4, whereas their corresponding differences can be visualized in Table 4.2.

Finally, an in-depth analysis of the results is driven by the simulation of another set of 67 days in winter and 64 days in summer.

	Winter [kWh]	Summer [kWh]
Max. demand in one day	36.48	34.22
Min. demand in one day	12.25	22.24
Average demand 30-days period	19.29	26.18
Total demand 30-days period	578.59	785.25
Max. PV generation in one day	3.12	3.43
Min. PV generation in one day	0.04	0.51
Average PV generation 30-days period	1.11	2.38
Total PV generation 30-days period	33.16	71.41

Table 4.1: Seasonal data for RL model.

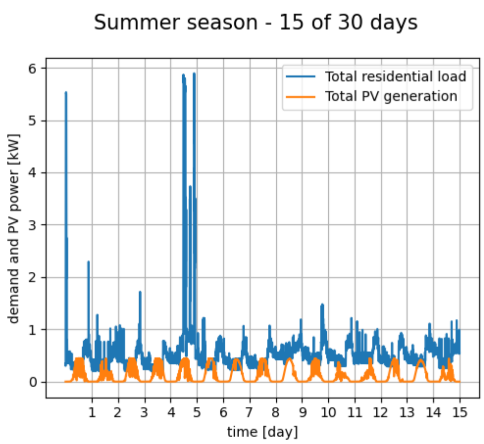


Figure 4.3: Demand and PV generation profiles (NN-summer).

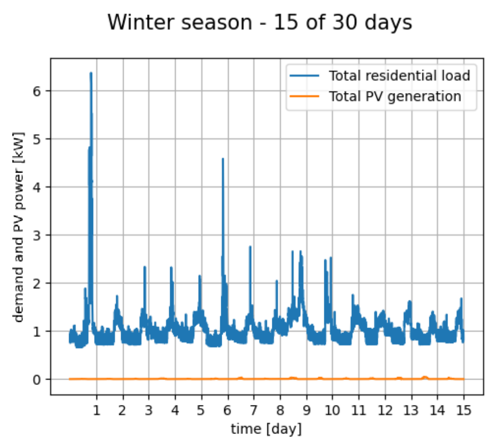


Figure 4.4: Demand and PV generation profiles (NN-winter).

4.1.2 Electricity Prices

As mentioned in Chapter 2, retail electricity prices play a fundamental role in rising investments in new technologies that may ease the negative inherent effects of fossil fuel-based power systems. Indeed, the design and implementation of regulatory incentives is crucial to stimulate household final user investment in PV-BESS. Given the highlighted effectiveness of FIT programs, this work includes the standard price per kWh (63.5 cents/kWh) purchased by the electric utility to the residential user under

	Winter [kWh]	Summer [kWh]
Max. demand in one day	32.04	29.70
Min. demand in one day	22.27	8.73
Average demand 30-days period	24.66	17.24
Total demand 30-days period	739.72	499.99
Max. PV generation in one day	1.0	3.72
Min. PV generation in one day	0.02	0.62
Average PV generation 30-days period	0.33	2.49
Total PV generation 30-days period	9.79	72.30

Table 4.2: Seasonal data for NN model.

the microFIT program launched in 2009 and phased out in November 2016. This program was implemented in Ontario, Canada, for renewable electricity generation up to 10 kW including rooftop PV systems [32, 73].

Moreover, the load shifting strategy to operate the battery is driven by the TOU dynamic pricing tariff scheme also implemented in Ontario. Through this scheme is expected a change in the residential user's consumption pattern along a day in order to mitigate demand power peaks by procuring savings in his monthly electricity bill. TOU reference rate structure can be visualized in Figure 4.5.

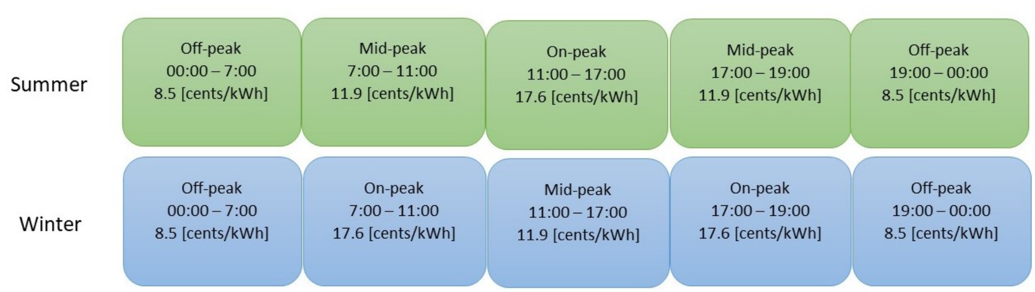


Figure 4.5: TOU reference rate structure.

PV-BESS load shifting operational strategy is detailed in Section 4.2; nonetheless, here it is pointed out that the microFIT tariff corresponds to the remuneration per kWh obtained by the residential user in case of electricity being exported to the grid, while TOU is the rate at which the electricity is purchased from the grid depending on the consumption time.

Finally, a retail electricity tariff in Edmonton is considered for assessing the feasibility and likelihood of a residential user to adopt a PV-BESS system. The best-case scenario would be having a supply system that pays off at a lower electricity cost during its operation than the current cost given by the retail electricity tariff. This tariff is offered by one of the electric utility companies in the city as a two-year fixed electricity rate (6.39 cents/kWh).

4.1.3 Battery data

VRF battery maximum charge $E_{ch,max}$ and discharge $E_{dis,max}$ rates were experimentally determined; the total battery energy charged was calculated and divided by the total time it took the battery to be fully charged, equations 4.1 and 4.2. $E_{dis,max}$ was similarly estimated by determining the total energy discharged and the time it took to reach such state.

$$E_{ch,max} = \frac{\int P_{ch}\ dt}{t}, \tag{4.1}$$

$$E_{dis,max} = \frac{-\int P_{dis}\ dt}{t}, \tag{4.2}$$

Energy conversion losses are included in the charge $E_{ch,t}$ and discharge $E_{dis,t}$ rates as reflected in equations 4.3 and 4.4. η_{conv} is the converter efficiency.

$$E_{ch,t} = E_{ch,max} * \eta_{conv}, \tag{4.3}$$

$$E_{dis,t} = \frac{-E_{dis,max}}{\eta_{conv}}, \tag{4.4}$$

Battery models also include the maximum battery capacity that constraints the charge and discharge actions, i.e., they cannot be charged over their physical maximum capacity E_{max} neither discharged below E_{min}. These values were experimentally obtained for the VRF, however, as li-ion batteries in particular cannot be fully charged or discharged, E_{max} equals the usable energy in Table 3.3 that reflects its depth-of-discharge (DOD). Li-ion SOC ranges 0 to 100% with E_{max}. Moreover, SOC for both battery models is updated at each iteration as defined in equations 4.5 and 4.6.

$$SOC_{ch,t+1} = SOC_{ch,t} + \frac{E_{ch,t}}{E_{max}}, \tag{4.5}$$

$$SOC_{dis,t+1} = SOC_{dis,t} - \frac{E_{dis,t}}{E_{max}}, \tag{4.6}$$

Different $E_{ch,max}$ and $E_{dis,max}$ rates were considered for the simulation of three additional VRF batteries to the already described in Table 3.2, notwithstanding they were calculated based on the same experimental information. Battery ageing and repeated cycling effects over capacity were omitted with the objective of not increasing computational burden.

4.1.4 Data Preprocessing

As stated by Geron [72], the two challenges faced when working with ML are the use of a "bad algorithm" or feeding the model with "bad data", being the second most important. Nonrepresentative training data was dealt in this work by selecting seasonal data in the same proportion for summer as for winter to feed both models. Dataset's quality was investigated to detect outliers and errors in the data; missing data of a couple of hours was copied from a day having similar behavior during the same missing hours. Feature selection process included the selection of the relevant features as explained above.

4.2 RL model: System Optimization

ML algorithms offer an alternative to rethink solutions to the BESS scheduling problem, previously solved with methods such as high-computing optimization-based algorithms that can be hardly run in microprocessors. Paradigms such as smart grid imposes the need to work with real-time BESS control systems that also adapt to fast changing conditions such as demand, renewable generation, and electricity prices.

As overviewed in the first section of Chapter 3, the BESS adaptive control proposed in this work bases on the load shifting technique and is developed in two phases. This section elaborates on the solution to the BESS scheduling problem with the Q-learning algorithm and the following explains the NN-based solution to the adaptive control. The Q-learning solution is an adaptation of the algorithm presented in [10].

4.2.1 Action Space

The action space comprises two possible actions the autonomous battery agent can choose from. At each time step t the agent can charge or discharge up to the maximum charge ($E_{ch,t}$) or discharge rate ($E_{dis,t}$). By default, a third option is neither charging or discharging when the TOU tariff is not enough low from the final user's perspective.

$$A_t \in [E_{ch,t}, E_{dis,t}] \tag{4.7}$$

Action selection process seeks the greatest estimated value action with a ϵ-greedy policy afterwards validated through an action validation function. This function limits charge and discharge actions within the physical battery capacity boundaries; at each iteration, the action validation receives the battery's SOC_t and the action a_t to evaluate if the battery can be further charged or discharged. Any action will result in a battery energy difference ($deltaE$) that feeds the load balancing function described in the following subsection.

4.2.2 Environment and State Space

The main purpose to solve the battery scheduling problem is to supply the normally demanded amount of electricity to the residential consumer at any time with no sacrifice of comfort and at the lowest cost. In other words, the battery should charge and discharge according to the electricity price signals and the balance of energy within the PV-BESS system at each time t. Thereby, the environment comprises all information that allows the battery agent to choose one action from the described set of actions, this is, the exported/imported energy at each moment.

Calculation of the exported/imported energy at each time t considers two energy balancing scenarios, a pure PV system (with no BESS) denominated *PV system* and a PV-BESS system referred to as *PV-BESS system*.

1. *PV system*: the available PV generation exclusively supplies the household demand, any surplus of electricity is exported to the distribution grid. At times when PV generation does no suffice the demand, the lacking amount of electricity is imported from the grid.

2. *PV-BESS system*: demand is firstly supplied by the available PV generation, in case of having any surplus of electricity it is directed to charge the BESS according to the $delta_t$ signal received from the action validation function. Any still remaining electricity is exported to the grid. A second set of options derives from the possibility of the PV system having not enough available energy to supply the load. In this case, if the battery is able to discharge, it supplies the demand in a $delta_t$ amount of energy; any amount of still missing energy is imported from the grid. Finally, at any time step t when the electricity price is low and the balance of energy within the system allows it, the battery is charged with electricity imported from the grid.

Exported and imported electricity information is then delivered to a function that calculates the costs associated with each system. The amount of exported or imported electricity is multiplied by the corresponding tariff (FIT or TOU). This information is vital because it feeds the reward function later explained.

In practical terms, the information received by the agent at each time step t shapes the state space that in this case has been modeled as a quintuple:

$$s_t = [pv_t, demand_t, TOU_t, FIT_t, SOC_t] \tag{4.8}$$

pv_t represents the PV generation, $demand_t$ is the residential demand, TOU_t corresponds to the time of use tariff or import tariff, FIT_t refers to the microFIT tariff or export tariff, and SOC_t describes the battery state of charge; all variables refer to time t.

4.2.3 Reward Function

At the core of the Q-learning algorithm is the reward function that guides the agent throughout the learning process by giving feedback to each action it chooses. The design process of the reward function can lead the agent to the successful completeness of the optimization task or to fail by finding a non-optimal solution; as such, this is a crucial part of the design of the solution that is usually done by trial-and error.

After the agent senses the environment's state at time t, it chooses an action and receives a reward as feedback. The reward is calculated as the difference between the total cost of the *PV-BESS system* and the *PV system*. A reward of $(cost_{PV} - cost_{batt}) * 10$ is awarded to the agent each time it obtains savings for charging and discharging the BESS at times when importing and exporting energy translates into monetary savings for the final user when compared to the PV system. At times when the PV system is more expensive than the PV-BESS system, the agent is punished with a reward calculated as $cost_{PV} - cost_{batt}$. Costs are summed at the end of each episode of 1440 time steps (1 day) and passed to Function 2.6 in order to update the Q-table model.

4.3 NN Model: Battery Adaptive Control

Particularly for residential applications, the BESS market is linked to the installation of rooftop PV systems as its main application is the increase of PV generation self-consumption [19]. Nonetheless, the BESS technology uptake requires to make it accessible to residential users through integrated adaptive controllers that, depending on the real-time conditions, allow the optimal battery charging or discharging. A NN model-based BESS controller is here proposed.

NN algorithms gained great popularity during the last decades due to its ability to approximate solutions of a great variety of high-dimensional nonlinear problems at a fraction of the computational cost other methods require. They have, however, an important disadvantage related to the access to labelled data for the initial training in a supervised setting that is usually not available. Data acquisition for the real implementation of the proposed battery controller is based on two possible cases:

1. **Existing PV system:** assuming the residential user has access to historical PV generation and electricity demand data, otherwise applies the second case.

2. **New PV system:** PV generation and electricity demand data collection process starts with the newly PV installed system, which may imply a delayed initial operation of the BESS.

Either case implies carrying out the collection of data process and the later execution of the Q-learning algorithm to obtain the optimal battery scheduling. As a reference, in this work, the Q-learning execution required 60 working days of 1-min data resolution comprising more than 86,000 data entries. Execution of the NN model-based controller is divided into two phases, the offline training process and the online control.

4.3.1 Offline Training

As explained in Section 4.1, a second dataset of 30 summer and 30 winter days was employed to train and test the NN models. This new dataset was also used to execute the Q-learning algorithm to obtain the corresponding charge and discharge labels at each time step t, this is, the derived dataset includes tuples of six elements: pv_t, $demand_t$, TOU_t, FIT_t, SOC_t and $batt_t$.

$batt_t$ signal is a quantity that describes the optimal battery scheduling in response to the cost difference between the *PV system* and the *PV-BESS system* inscribed in the reward function of the Q-learning algorithm. A positive number corresponds to a battery charge action, a negative number indicates the battery discharge, and, consequently, a zero indicates the battery does neither charge or discharge during each time step.

An 80% - 20% rule was followed to split the dataset into NN model training and testing, respectively, whereas the same proportion of data is taken from winter and summer seasons for training and testing. A total of 40 NN models were trained and tested, with architectures varying between one and seven hidden layers, and 20, 50, 100, 200, 250, and 400 nodes in each hidden layer. The number of nodes in the input layer corresponds to the number of features, six as described above, and the number of nodes in the output layer equals the three possible states the battery can take, charge, discharge or zero.

All NN models were created in Keras deep learning API [74] that is coded with Python, the model arguments were as follows:

1. Adam optimizer to train the models with the gradient descent algorithm

2. Categorical cross-entropy loss function

3. Rectified linear (ReLU) activation functions for the hidden layers and softmax activation function for the output layer

4. Accuracy metric to evaluate the performance of all proposed models

4.3.2 Online Control

Online control includes the selection of the most suitable NN model that is later used in the prediction mode. The performance of each model is assessed with the accuracy metric and the respective confusion matrix, such information assists the preselection model. The final decision is taken once the preselected models are compared in terms of the *PV-BESS system* calculated costs. Selected models are shown in Table 4.3, as it is detailed in the following section, five batteries were simulated and for each one the same model selection process was carried out.

Battery system	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
Va1	50	300	300	50	
Va2	50	300	300	50	
Va3	50	300	100	300	50
Va4	50	200	200	200	50
Li	50	300	300	50	

Table 4.3: Selected NN models.

In a real application, a selected NN model should be installed in the battery microprocessor; after an initial set of vectors with information from the PV generation system, smart household meter, battery SOC, time of the day, and electricity import and export tariffs are given, the controller can predict the BESS action. Figure 4.6 shows the adaptive battery controller schematic.

4.4 Battery controller evaluation: operational costs

A second set of 30 summer and 30 winter days were simulated with the selection of NN models, five models in total. *Va* refers to VRF batteries, correspondingly, *Li* refers

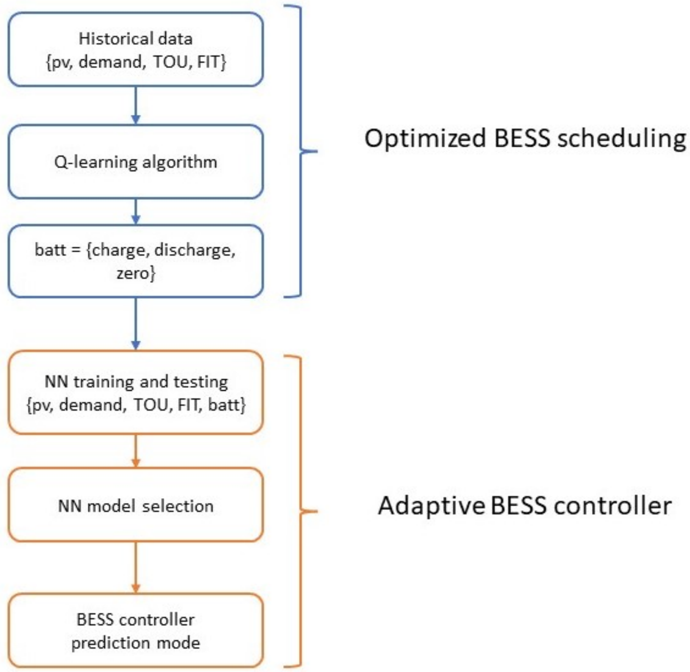


Figure 4.6: Adaptive BESS controller schematics.

to a li-ion battery. The efficiency of battery *Va1* is given in Table 3.2, it represents the installed battery when operated at a low charge and discharge constant current. *Va2* also represents the installed battery but in this case operated at higher charge and discharge currents with a maximum current of 47 A at discharging. Efficiency in this case dropped because battery operation was not optimized. Two larger batteries were simulated to further analyze the proposed battery controller. *Va3* and *Va4* are two high-efficiency batteries sized as 49 kWh and 28 kWh. *Li* battery is modelled with information from Table 3.3.

The NN-model adaptive controller *NN* and the Q-learning default control system *RL* are compared with the two systems in terms of daily electricity costs reflected in the residential final user's monthly bill.

1. *PV*: baseline system, costs are calculated when optimizing the *PV system* with the Q-learning algorithm.

2. *Flat tariff*: calculated costs with no *PV system* neither *PV-BESS system*.

PV, *RL* and *NN* systems include the effect of both, the microFIT regulatory incentive applied when exporting electricity surplus from *PV system*/*PV-BESS system* to the grid and the TOU policy applied when importing missing electricity to the *PV system*/*PV-BESS system* based on the differences in electric retail tariffs in periods of 24 hours.

An in-depth analysis of the results obtained when comparing the described systems can be found in Chapter 5.

Chapter 5

Simulation Results

5.1 Input data to simulations

Data were divided into summer and winter seasons mimicking the TOU policy. Table 5.1 contains details of the data used to train and test both the Q-learning and the NN algorithms, as well as data for final simulations. Figure 5.1 only shows PV generation and demand used for the simulation of the battery behavior. It is highlighted that PV generation in winter is about 40% of the generation in summer, while demand is 126% higher in winter than in summer.

No. Days	Dataset	Season	Month
30	Training RL	Summer	August, September
30	Training & Testing NN	Summer	July, August
64	Simulation NN	Summer	May, June, September, October
30	Training RL	Winter	February, March
30	Training & Testing NN	Winter	November, December
67	Simulation NN	Winter	March, April, December, January

Table 5.1: Details of the data used to train and test the algorithms.

Simulation data used in the paper "Optimal real-time scheduling of battery operation using reinforcement learning" referred in the Preface section is in this work extended in 131 summer and winter days. Results shown in this section correspond

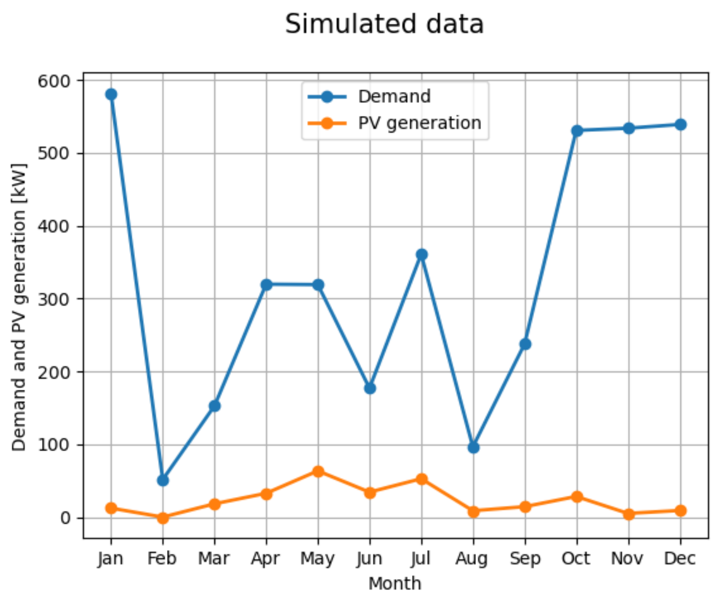


Figure 5.1: Input data used to simulate battery behavior.

to a total of 93 days in summer and 97 days in winter, additionally to the 60 days used to train the Q-learning algorithm.

Regarding the retail electricity tariff in Edmonton used to compare the battery controller electricity costs, since March 2015 up to July 2021 includes variations between 7.29 cents/kWh that decreased during 53 months down to a minimum of 5.49 cents/kWh. This tariff later sustained a cost per kWh above 6 cents during the last 21 months up to now. Indeed, this trend does not reflect large variations from the residential consumer point of view, but in the long term nothing prevents from larger variations that may relate, for example, to green taxes aimed at decarbonizing the power sector.

The following analysis of the results includes a comparison of electricity costs calculated as follows:

1. *RL* and *NN* vs the *flat tariff*
2. *RL* and *NN* vs the baseline *PV system*

5.2 *Va1* and *Va2* BESS

Va1 and *Va2* correspond to the experimental VRF battery setup with a capacity of 1.03 kWh and efficiencies of 82% and 40%.

Season	Month	*Flat tariff*	*PV*	*RL*	*NN*
Winter	Jan-Apr, Nov,Dec	136.909	236.226	236.656	234.984
Summer	May-Aug, Sep, Oct	108.264	159.218	159.667	158.828

Table 5.2: Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va1*.

Season	Month	*Flat tariff*	*PV*	*RL*	*NN*
Winter	Jan-Apr, Nov,Dec	136.909	236.225	237.648	237.095
Summer	May-Aug, Sep, Oct	108.264	159.218	160.772	159.468

Table 5.3: Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va2*.

	RL			*NN*		
Season	*Costs*	*Savings*	*Revenue*	*Costs*	*Savings*	*Revenue*
Winter	-0.431	0.000	0.000	0.000	1.242	0.000
Summer	-0.449	0.000	0.000	0.000	0.390	0.000

Table 5.4: PV-BESS cost differences, *RL* and *NN* *Va1* battery controller.

	RL			*NN*		
Season	*Costs*	*Savings*	*Revenue*	*Costs*	*Savings*	*Revenue*
Winter	-1.422	0.000	0.000	-0.869	0.000	0.000
Summer	-1.554	0.000	0.000	-0.250	0.000	0.000

Table 5.5: PV-BESS cost differences, *RL* and *NN* *Va2* battery controller.

5.2.1 RL and NN compared to flat tariff

As shown in Tables 5.2 and 5.3, installing a battery system with technical characteristics similar to those of Va1 or Va2 is not an option that translates into savings for a residential user in Edmonton when taking the *flat tariff* scenario as a benchmark, electricity costs are higher with any of these systems. Additional to the calculated electricity cost, a residential user would have to consider investments in equipment.

As mentioned above, *Va1* and *Va2* are different just in terms of efficiency. The most efficient battery (*Va1*) reflects slightly lower costs with respect to the *flat tariff* than battery *Va2*, in the order of 1 CAD during both seasons. This result suggests the importance of battery operation optimization in terms of charge and discharge currents and flow rate for VRF batteries. Consistently, *NN* battery controller resulted in lower electricity costs than the *RL* controller with respect to the *flat tariff*.

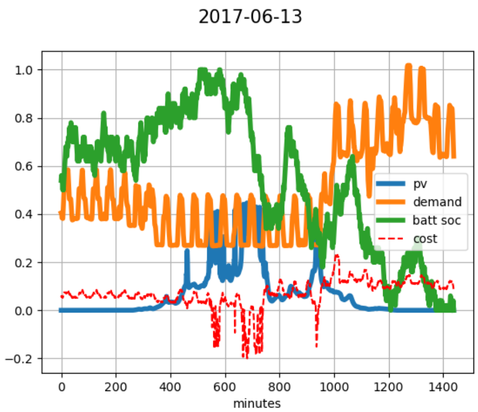
5.2.2 RL and NN compared to PV system

Having a Va1 battery controlled with the suggested *NN* controller results in slightly lower electricity costs than costs derived from the *RL* controller when compared to the baseline *PV system*. Contrastingly, *NN* and *RL* controllers are not able to schedule *Va2* battery to decrease electricity costs as the PV system renders slightly lower costs during both seasons. Calculated electricity costs with *NN* and *RL* are further analyzed in Tables 5.4 and 5.5. TOU and microFIT regulations are used to solve the BESS scheduling problem with the Q-learning algorithm in two systems, *PV* and *PV-BESS*, able to export and import energy from the power grid. Therefore, total electricity costs are broken down in potential costs, savings and revenues for the residential user. *NN* calculated costs are also broken down into potential costs, savings and revenues, since *Va1* and *Va2* models were trained with labels generated with the Q-learning algorithm, indirectly containing the effect of both regulations.

The operation of *Va1* system during winter and summer seasons with the *RL* controller is not economically feasible for a residential user because it would end paying more for the operation of the battery. In any of both seasons, *RL* controller results in the user paying more for operating the battery than just letting operate the *PV system*. *NN* controller offers potential savings in the electricity bill to the final user by optimally scheduling the BESS. *Va2* reflects the same tendency than *Va1* when comparing costs calculated with the *RL* and the *NN* controllers to the baseline. Yet, *Va2* economic feasibility's is lower than *Va1* from the final user's perspective because both controllers result in electricity costs when operating the BESS with no savings. The *NN* controller, in this case, is able just of decreasing the BESS electricity costs with no possibility of offering potential savings in the electricity bill. As mentioned before, differences in electricity costs are due to the *Va2* system lower efficiency. BESS performance in summer and winter is remarkably different. Electricity costs derived from the *RL* action for *Va1* results in lower costs in winter than in summer, likewise, *NN* controller potential savings in winter are three times higher than in summer. Electricity costs related to the *RL* controller for *Va2* are higher in summer than in winter, and *NN* action translates into higher winter than summer costs.

Furthermore, the battery sizing effect can be visualized in Figures 5.2 and 5.3. *Va1* battery system is required to fully charge and discharge in a period of 24 hours to satisfy electricity needs within the residential system in summer and in winter. As shown in Figure 5.2, during a day in summer season, *Va1* with PV system are able to deal with the residential demand. However, during a day in winter (Figure 5.3) the demand is too high in the evening hours with no local generation during that time, so the battery is required to fully discharge to supply demand during a very short period of time up to its minimum capacity.

Units in these graphs are not shown since all displayed variables are different, however, they can be understood as [kWh] for PV and *demand*, [%] for *batt soc*, and [CAD/kWh] for the *cost*.



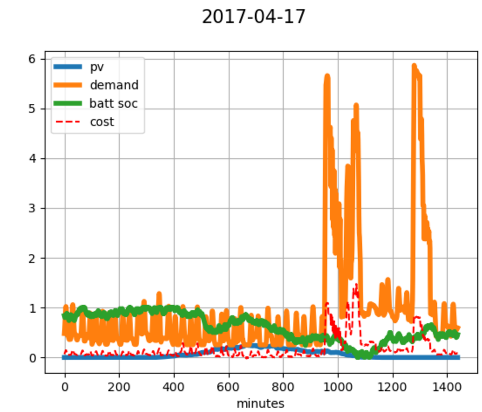


Figure 5.2: Va1 system, one day in **summer**.

Figure 5.3: Va1 system, one day in **winter**.

5.3 Va3 and Va4 BESS

Va3 and *Va4* systems were defined with capacities of 49 and 28 kWh, and 82% efficiency for comparison purposes with *Va1*, *Va2* and *Li* systems.

Season	Month	*Flat tariff*	*PV*	*RL*	*NN*
Winter	Jan-Apr, Nov,Dec	136.909	236.226	225.946	133.418
Summer	May-Aug, Sep, Oct	108.264	159.218	111.032	54.147

Table 5.6: Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va3*.

Season	Month	*Flat tariff*	*PV*	*RL*	*NN*
Winter	Jan-Apr, Nov,Dec	136.909	236.226	238.977	221.804
Summer	May-Aug, Sep, Oct	108.264	159.218	148.837	127.231

Table 5.7: Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Va4*.

5.3.1 RL and NN compared to flat tariff

In contrast with *Va1* and *Va2*, in this case we have two larger systems that represent potential monetary benefits for residential users in form of costs reduction and

Season	*RL*			*NN*		
	Costs	*Savings*	*Revenue*	*Costs*	*Savings*	*Revenue*
Winter	0.000	10.087	0.192	0.000	82.633	4.911
Summer	0.000	32.000	13.304	0.000	56.484	22.646

Table 5.8: PV-BESS cost differences, *RL* and *NN* *Va3* battery controller.

Season	*RL*			*NN*		
	Costs	*Savings*	*Revenue*	*Costs*	*Savings*	*Revenue*
Winter	-2.752	0.000	0.000	0.000	14.421	0.000
Summer	0.000	9.904	0.477	0.000	25.042	4.078

Table 5.9: PV-BESS cost differences, *RL* and *NN* *Va4* battery controller.

even obtaining savings and revenues. Although electricity costs calculation for *Va3* with *RL* system accounts for a reduction of about 10 CAD in winter and 48 CAD in summer when compared to *Va1* also with *RL*, it still does not improve the *flat tariff* scenario. This maintains the premise that it would be cheaper for a residential user not having a PV-BESS system in Edmonton. However, when comparing electricity costs derived from the action of the proposed *NN* controller, the scenario may change and the operation of a PV-BESS system could pay off the final user by offering lower electricity costs in summer and winter, as shown in Table 5.6. *Va4* system decreases the difference between *Va1-RL* and *flat tariff* electricity costs but does not achieve the goal of economical feasibility for the final user, Table 5.7. Moreover, *NN* controller improves over cost decrease of *RL* controller, but costs are still higher when compared to the *flat tariff* scenario.

5.3.2 RL and NN compared to PV system

The comparison between the *Va3-RL* system and the baseline scenario accounts lower estimated electricity costs due to the action of the RL agent in winter and summer. As expected, this result is even improved by the proposed *NN* battery controller. Total electricity costs resulting from the action of the *RL* and *NN* controller are further broken down in costs, savings and revenues in Tables 5.8 and 5.9. The effect of the TOU and microFIT regulations are in this case clearly visualized. *Va3-RL* electricity cost dropped to zero while the amount of savings and revenue is significant when compared to those obtained with *Va1* and *Va2*. *NN* controller, as mentioned before, improved the amount of savings and revenue for *Va4*, savings were almost two times higher in summer and around eight times higher during winter. *Va4-RL* electricity costs during winter registered savings during some months, however, the higher costs of the remaining months during the season did not result in lower total season electricity costs. In summer, differently to winter, resulted in potential savings and a low revenue for the residential user. This can be explained by the RL agent having a limited margin for optimal battery scheduling with lower PV generation and higher demand to supply during winter. *NN* controller in both seasons resulted in some savings and even revenue in summer for the residential user. In general terms, although *RL* and *NN* for *Va3* and *Va4* represent potential savings and revenues for the householder, PV-BESS economical feasibility should consider even higher equipment investments than the required for *Va1* and *Va2*.

Sizing effects of *Va3* and *Va4* can be visualized in Figures 5.4, 5.5, 5.6, and 5.7. For each system one day in summer and one day in winter is presented. As expected, contrary to the requirement of fully charging and discharging *Va1* and *Va2* batteries, *Va3* and *Va4* do not fully cycle throughout the day. Lower summer demand, compared to winter demand, can be supplied by the local generation and both battery systems without requiring them to fully charge or discharge. In fact, the SOC of both batteries is kept around 50% along the day. In winter, although the demand is higher and the PV generation is lower, battery systems are still not required to fully charge and discharge. This battery behavior suggests that battery capacities are underutilized for the given residential demand and the residential user would have more benefits by properly sizing the battery capacity.

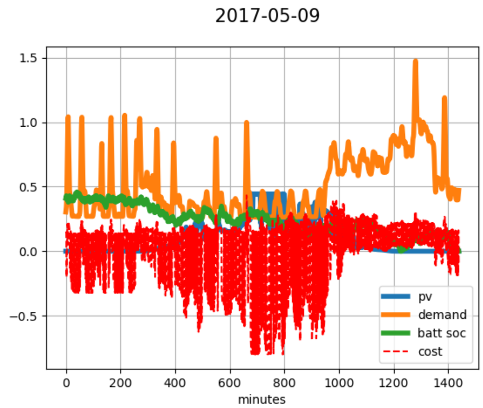


Figure 5.4: Va3 system, one day in **summer**.

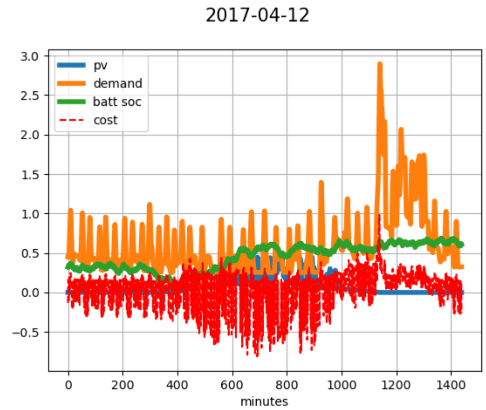


Figure 5.5: Va3 system, one day in **winter**.

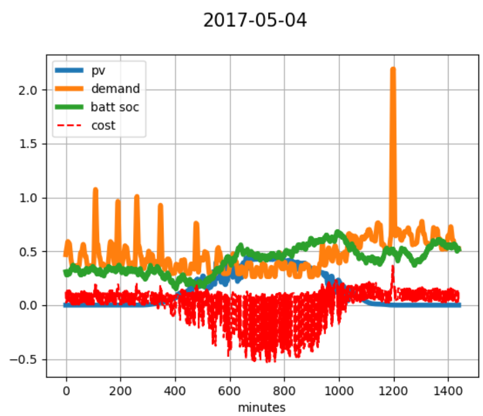


Figure 5.6: Va4 system, one day in **summer**.

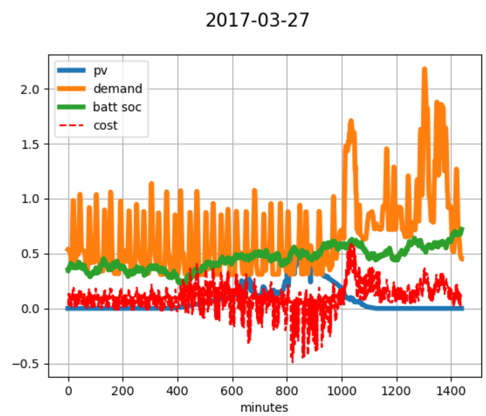


Figure 5.7: Va4 system, one day in **winter**.

Units can be understood as [kWh] for *PV* and *demand*, [%] for *batt soc*, and [CAD/kWh] for the *cost*.

5.4 Li

Li-ion battery was defined as 17.1 kWh capacity and efficiency of 96.5%, according to vendor information.

Season	Month	*Flat tariff*	*PV*	*RL*	*NN*
Winter	Jan-Apr, Nov,Dec	136.909	236.226	235.503	234.454
Summer	May-Aug, Sep, Oct	108.264	159.218	156.050	156.153

Table 5.10: Total *flat tariff*, *PV* and *PV-BESS* systems costs for *Li*.

	RL			*NN*		
Season	*Costs*	*Savings*	*Revenue*	*Costs*	*Savings*	*Revenue*
Winter	0.000	0.722	0.000	0.000	2.002	0.000
Summer	0.000	3.106	0.063	0.000	2.967	0.098

Table 5.11: PV-BESS cost differences, *RL* and *NN* *Li* battery controller.

5.4.1 RL and NN compared to flat tariff

Li system is closer in magnitude to *Va4* capacity but is characterized by a higher efficiency and lower DOD. Lower li-ion DOD with respect to VRF BESS impacts on the amount of energy the battery can charge and discharge. which could probably increase electricity costs derived from the action of *RL* and *NN* controllers. Despite of this, the DOD effect can be compensated with the li-ion higher efficiency.

Electricity costs that result from the *Li-RL* controller are slightly lower than costs from *Va1-RL*, which is expected because *Va1* capacity is lower. Though, when *Li-RL* associated costs are compared to the *flat tariff* scenario, no *Li* system neither *Va1* represent a viable option from the householder perspective. It is more expensive for the user to keep a *Li* system working in comparison to just purchasing electricity from the local retailer throughout the year. Information is found in Table 5.10.

5.4.2 RL and NN compared to PV system

By comparing *Li-RL* electricity costs to the baseline system, it was found that *Li* system renders subtly lower costs that are comparable to *Va1-RL* costs. However, differently to all other battery systems, *Li-NN* case in summer does not outperforms *RL* by a negligible amount of 103 cents. A more detailed analysis of the total costs in Table 5.11 shows that savings obtained with *RL* in summer are higher than savings obtained with *NN* controller. This cost difference suggests that *NN* model selection method should be improved.

Li battery size behaves as *Va4* battery to supply demand throughout a day in summer and winter. In other words, battery does not fully charge or discharge in a period of 24 hours. For instance, for a day in summer, *Li* battery does not fully charge at any moment when PV generation is maximum, but it discharges up to 40% SOC at evening to supply demand when there is no more local generation available. Similarly, in a day in winter, *Li* battery maintains a high SOC during the day to discharge up to 60% at evening. By considering both elements, electricity costs derived from the *RL* and *NN* functioning, as well as battery SOC being not fully depleted, the following is suggested. *NN* model selection process can include a more detailed analysis of electricity costs for the selection of the best model. Battery sizing, particularly for li-ion batteries, may imply higher costs than for VRF technology. Scaling the VRF battery capacity solely relies on the acquisition of extra quantities of electrolyte with no need of acquiring a new battery stack, while scaling li-ion capacity implies the acquisition of new battery modules.

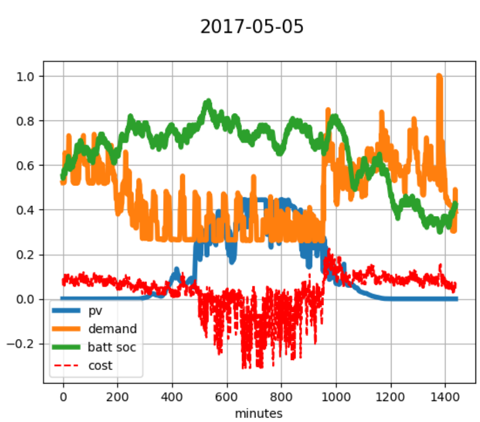


Figure 5.8: Li system, one day in **summer**.

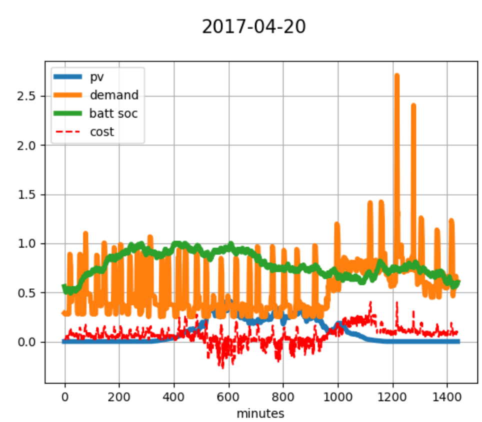


Figure 5.9: Li system, one day in **winter**.

Chapter 6

Conclusions & Future Work

6.1 Conclusions

The battery scheduling problem is investigated in this thesis by proposing an adaptive battery controller that relies on reinforcement learning (RL) and neural networks (NN). NN algorithms working in a supervised setting require large amounts of high-quality data containing labels for each sequence of input features during the training process. In our case, as battery charge and discharge labels were not available to train the NN, the optimal scheduling problem was firstly solved with a RL-based Q-learning algorithm based on Graham's [10] work. The RL-agent is supposed to find the optimal scheduling policy by comparing in the reward function the electricity costs derived from two supply residential systems, a purely photovoltaic (PV) system, and a PV-battery energy storage system (BESS). At any time step when the electricity costs from the PV-BESS system are lower than those from the PV system, the RL-agent is rewarded. RL algorithm's performance is first investigated by comparing PV-BESS and PV electricity costs. Labels that result from the battery scheduling process are then used to train some NN models. NN model validation is analyzed through the model's accuracy and inspection of the confusion matrix, as well as from the comparison of PV-BESS and PV electricity costs. To verify the functioning of RL and NN models, experiments over five BESS were performed. BESS models vary in terms of size, efficiency, and chemistry. The individual cases are named Va1, Va2, Va3, Va4 and Li. Va refers to vanadium redox flow (VRF) battery and Li to lithium ion (li-ion) battery. VRF technical features were obtained from the characterization of a real battery setup while li-ion features were taken from manufacturer information.

Results obtained with the NN battery controller from a residential user's perspective, favor the use of the PV-BESS. Total electricity costs from Va1, Va3, Va4, and Li decreased in comparison to PV system (baseline) electricity costs in the winter and summer seasons, except for Li whose results improved just in winter. For all batteries, a variable reduction of electricity costs and an increment in savings were observed. Likewise, depending on the size of the batteries, different revenues were observed. That is the case of the larger battery system (Va3) that earned the highest revenue throughout the year. Va4 and Li systems obtained modest revenues in summer.

Battery efficiency is another factor with an important impact on the results. A clear example is Va2 system that equals Va1 capacity but has a lower efficiency. Va2 could not improve on the baseline electricity costs, although the NN battery controller was able to decrease the loss registered by the RL algorithm in comparison to the baseline. The exceptional case is the Li system during summer. Li-NN electricity cost decreased in comparison to the baseline system, however, it did not improve on the reduction reached by the RL when compared to the baseline. In this case, a more detailed analysis of the NN model is recommended.

Finally, the contributions made in this thesis are summarized. First, the RL algorithm reached, depending on the BESS size and efficiency, reductions in electricity costs when compared to the baseline system. The proposed NN controller decreased electricity costs with respect to the baseline system in a larger proportion than the cost reduction obtained by the RL initial labeling for charge and discharge. Another benefit of using NN controllers is that they require reduced computational resources once the models have been trained. In this case, NN controllers required a couple of seconds to compute an episode of 24 hours, while the RL algorithm computes the same information in about 3.5 minutes. This processing time difference is important when dealing with a real-time BESS controller. Second, the house demand in winter season is 79% higher than summer season, while winter PV generation represents just the 38% of the summer PV generation. Larger battery systems, Va3, Va4, and Li, reflected the described seasonal differences by obtaining higher savings and revenues during the summer. A higher PV generation and lower demand conditions in summer give a broader margin for larger batteries satisfying local demand and trading the remaining energy with the local electricity supplier. In the case of smaller batteries, as their capacities are below the total demand of energy, the impact on savings due to seasonal changes is negligible. Third, TOU and microFIT regulatory incentives are included in the design of the battery controllers. We analyzed its impact on the Edmonton area market. For this case study, the electricity cost associated with the action of RL and NN battery controllers was compared to the electricity costs calculated with the flat tariff offered by a local electricity supplier. Except for the Va3-NN system, all other BESS are not economically feasible for a residential user in Edmonton based on the regulatory incentives, as retail electricity rates are below the baseline with incentives. Fourth, the comparison of simulation results of VRF versus a li-ion battery suggests that vanadium-based battery technology is a potential competitor of li-ion batteries for the residential market in terms of operational performance, despite their lower efficiency. However, an important disadvantage for VRF BESS that falls out of the scope of this work, is the requirement of space for installing the equipment as well as the cost of installation that should include qualified technicians to put the battery into operation.

6.2 Future Work

Utilization of reinforcement learning and neural networks to configure an adaptive battery controller has shown that real-time operation of PV-BESS systems is feasible for residential applications. Nonetheless, based on the opinion of the author, a continuation of this project should aim at answering the following questions. Firstly, what is the impact of the battery size on the economic feasibility for the installation of PV-BESS systems. Secondly, what is the minimum amount of data to train the proposed controller. In the third place, an in-depth analysis of regulatory incentives and their impact on economic viability for the end user is necessary.

Bibliography

[1] E. C. Pischke, B. Solomon, A. Wellstead, A. Acevedo, A. Eastmond, F. De Oliveira, S. Coelho, and O. Lucon, "From kyoto to paris: Measuring renewable energy policy regimes in argentina, brazil, canada, mexico and the united states," *Energy Research & Social Science*, vol. 50, pp. 82–91, 2019.

[2] B. Dolter and N. Rivers, "The cost of decarbonizing the canadian electricity system," *Energy Policy*, vol. 113, pp. 135–148, 2018.

[3] I. E. Agency. (2020). "Photovoltaic power systems programme annual report." M. J. Brunisholz, Ed., [Online]. Available: https://iea-pvps.org/annual-reports/.

[4] C. E. Regulator. (2020). "Canada's energy future 2020," [Online]. Available: https:https://www.cer-rec.gc.ca/en/data-analysis/canada-energy-future/.

[5] M. S. Ziegler, J. M. Mueller, G. D. Pereira, J. Song, M. Ferrara, Y.-M. Chiang, and J. E. Trancik, "Storage requirements and costs of shaping renewable energy toward grid decarbonization," *Joule*, vol. 3, no. 9, pp. 2134–2153, 2019.

[6] H. Gong, O. Akeyo, V. Rallabandi, and D. M. Ionel, "Real time operation of smart homes with pv and battery systems under variable electricity rate schedules and transactive power flow," in *2018 7th International Conference on Renewable Energy Research and Applications (ICRERA)*, IEEE, 2018, pp. 1392–1395.

[7] E. Tervo, K. Agbim, F. DeAngelis, J. Hernandez, H. K. Kim, and A. Odukomaiya, "An economic analysis of residential photovoltaic systems with lithium ion battery storage in the united states," *Renewable and Sustainable Energy Reviews*, vol. 94, pp. 1057–1066, 2018.

[8] H. C. Hesse, M. Schimpe, D. Kucevic, and A. Jossen, "Lithium-ion battery storage for the grid—a review of stationary battery storage system design tailored for applications in modern power grids," *Energies*, vol. 10, no. 12, p. 2107, 2017.

[9] M. A. Hossain, H. R. Pota, and C. M. Moreno, "Real-time battery energy management for residential solar power system," *IFAC-PapersOnLine*, vol. 52, no. 4, pp. 407–412, 2019.

[10] P. R. Graham, *Control of residential battery charge scheduling using machine learning*, 2018.

[11] W. Hu, Z. Chen, and B. Bak-Jensen, "Optimal operation strategy of battery energy storage system to real-time electricity price in denmark," in *IEEE PES general meeting*, IEEE, 2010, pp. 1–7.

[12] R. Barzin, J. J. Chen, B. R. Young, and M. M. Farid, "Peak load shifting with energy storage and price-based control system," *Energy*, vol. 92, pp. 505–514, 2015.

[13] T. Logenthiran, D. Srinivasan, and K. Vanessa, "Demand side management of smart grid: Load shifting and incentives," *Journal of Renewable and Sustainable Energy*, vol. 6, no. 3, p. 033 136, 2014.

[14] C Tiganasu, C Nemes, F Munteanu, D Munteanu, and C Scinteie, "The load management concept applied on the residential sector," in *2019 International Conference on Electromechanical and Energy Systems (SIELMEN)*, IEEE, 2019, pp. 1–5.

[15] A. Pechmann, F. Shrouf, M. Chonin, and N. Steenhusen, "Load-shifting potential at smes manufacturing sites: A methodology and case study," *Renewable and Sustainable Energy Reviews*, vol. 78, pp. 431–438, 2017.

[16] E. Telaretti, L Dusonchet, F Massaro, L Mineo, G Pecoraro, and F Milazzo, "A simple operation strategy of battery storage systems under dynamic electricity pricing: An italian case study for a medium-scale public facility," in *3rd Renewable Power Generation Conference (RPG 2014)*, IET, 2014, pp. 1–7.

[17] M. Zheng, C. J. Meinrenken, and K. S. Lackner, "Electricity storage in buildings for residential sector demand response: Control algorithms and economic viability evaluation," *US: National Institute of Standards and Technology*, 2014.

[18] S. Agnew, C. Smith, and P. Dargusch, "Causal loop modelling of residential solar and battery adoption dynamics: A case study of queensland, australia," *Journal of cleaner production*, vol. 172, pp. 2363–2373, 2018.

[19] J. Figgener, P. Stenzel, K.-P. Kairies, J. Linßen, D. Haberschusz, O. Wessels, G. Angenendt, M. Robinius, D. Stolten, and D. U. Sauer, "The development of stationary battery storage systems in germany–a market review," *Journal of energy storage*, vol. 29, p. 101 153, 2020.

[20] A. Malhotra, B. Battke, M. Beuse, A. Stephan, and T. Schmidt, "Use cases for stationary battery technologies: A review of the literature and existing projects," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 705–721, 2016.

[21] V. Muenzel, I. Mareels, J. de Hoog, A. Vishwanath, S. Kalyanaraman, and A. Gort, "Pv generation and demand mismatch: Evaluating the potential of residential storage," in *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, IEEE, 2015, pp. 1–5.

[22] J. Hoppmann, J. Volland, T. S. Schmidt, and V. H. Hoffmann, “The economic viability of battery storage for residential solar photovoltaic systems–a review and a simulation model,” *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 1101–1118, 2014.

[23] J. Bergner, J. Weniger, T. Tjaden, and V. Quaschning, “Feed-in power limitation of grid-connected pv battery systems with autonomous forecast-based operation strategies,” in *29th European Photovoltaic Solar Energy Conference and Exhibition*, 2014, pp. 2363–2370.

[24] A. Clemente and R. Costa-Castelló, “Redox flow batteries: A literature review oriented to automatic control,” *Energies*, vol. 13, no. 17, p. 4514, 2020.

[25] M. Uhrig, S. Koenig, M. R. Suriyah, and T. Leibfried, “Lithium-based vs. vanadium redox flow batteries–a comparison for home storage systems,” *Energy Procedia*, vol. 99, pp. 35–43, 2016.

[26] S. Ha and K. G. Gallagher, “Estimating the system price of redox flow batteries for grid storage,” *Journal of Power Sources*, vol. 296, pp. 122–132, 2015.

[27] G. Merei, J. Moshövel, D. Magnor, and D. U. Sauer, “Optimization of self-consumption and techno-economic analysis of pv-battery systems in commercial applications,” *Applied Energy*, vol. 168, pp. 171–178, 2016.

[28] E. S. Ruiz, J. H. Cubillo, J. de la Vega Rodríguez, and S. Martín, “Sizing and operation of a residential pv-battery system: Rule-based and real time vs optimization and perfect foresight,” in *2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON)*, IEEE, 2020, pp. 476–481.

[29] A. Johann and R. Madlener, “Profitability of energy storage for raising self-consumption of solar power: Analysis of different household types in germany,” *Energy Procedia*, vol. 61, pp. 2206–2210, 2014.

[30] R. Luthander, J. Widén, D. Nilsson, and J. Palm, “Photovoltaic self-consumption in buildings: A review,” *Applied energy*, vol. 142, pp. 80–94, 2015.

[31] D. Kucevic, B. Tepe, S. Englberger, A. Parlikar, M. Mühlbauer, O. Bohlen, A. Jossen, and H. Hesse, “Standard battery energy storage system profiles: Analysis of various applications for stationary energy storage systems using a holistic simulation framework,” *Journal of Energy Storage*, vol. 28, p. 101 077, 2020.

[32] L. C. Stokes, “The politics of renewable energy policies: The case of feed-in tariffs in ontario, canada,” *Energy Policy*, vol. 56, pp. 490–500, 2013.

[33] G. Henri, N. Lu, and C. Carreio, “A machine learning approach for real-time battery optimal operation mode prediction and control,” in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, IEEE, 2018, pp. 1–9.

[34] R. Hemmati and H. Saboori, “Stochastic optimal battery storage sizing and scheduling in home energy management systems equipped with solar photovoltaic panels,” *Energy and Buildings*, vol. 152, pp. 290–300, 2017.

[35] E Matallanas, M. Castillo-Cagigal, A Gutiérrez, F Monasterio-Huelin, E. Caamaño-Martín, D Masa, and J Jiménez-Leube, "Neural network controller for active demand-side management with pv energy in the residential sector," *Applied Energy*, vol. 91, no. 1, pp. 90–97, 2012.

[36] S. Grillo, A. Pievatolo, and E. Tironi, "Optimal storage scheduling using markov decision processes," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 755–764, 2015.

[37] P. Bijapur, P. Chakradhar, M Rishab, S Srinivas, and B. Nagabhushana, "Reinforcement learning for energy storage optimization in the smart grid," in *2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*, IEEE, 2020, pp. 1–6.

[38] S. Kim and H. Lim, "Reinforcement learning based energy management algorithm for smart energy buildings," *Energies*, vol. 11, no. 8, p. 2010, 2018.

[39] G. Henri and N. Lu, "A supervised machine learning approach to control energy storage devices," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 5910–5919, 2019.

[40] F. Kazhamiaka, S. Keshav, and C. Rosenberg, "Adaptive battery control with neural networks," in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, 2019, pp. 536–543.

[41] M. A. Wiering and M. Van Otterlo, "Reinforcement learning," *Adaptation, learning, and optimization*, vol. 12, no. 3, 2012.

[42] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[43] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[44] S. Haykin, *Neural networks and learning machines, 3/E*. Pearson Education India, 2010.

[45] M. Kubat, *An introduction to machine learning*. Springer, 2017.

[46] J. Watt, R. Borhani, and A. K. Katsaggelos, "Machine learning refined," 2019.

[47] A. Burkov, *The hundred-page machine learning book*. Andriy Burkov Canada, 2019, vol. 1.

[48] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, 2020.

[49] K Nallaperumal and A Krishnan, *Engineering research methodology: A computer science and engineering and information and communication technologies perspective*, 2013.

[50] D. V. Thiel, *Research methods for engineers*. Cambridge University Press, 2014.

[51] M. P. Akter, Y. Li, J. Bao, M. Skyllas-Kazacos, M. Rahman, J. E. Fletcher, and X. Zhang, "Charging control of vanadium redox battery based energy storage systems with variable input power," in *2018 IEEE 27th International Symposium on Industrial Electronics (ISIE)*, IEEE, 2018, pp. 489–494.

[52] E. Banguero, A. Correcher, Á. Pérez-Navarro, F. Morant, and A. Aristizabal, "A review on battery charging and discharging control strategies: Application to renewable energy systems," *Energies*, vol. 11, no. 4, p. 1021, 2018.

[53] M. Guarnieri, A. Trovò, G. Marini, A. Sutto, and P. Alotto, "High current polarization tests on a 9 kw vanadium redox flow battery," *Journal of Power Sources*, vol. 431, pp. 239–249, 2019.

[54] D. Aaron, Q Liu, Z Tang, G. Grim, A. Papandrew, A Turhan, T. Zawodzinski, and M. Mench, "Dramatic performance gains in vanadium redox flow batteries through modified cell architecture," *Journal of Power sources*, vol. 206, pp. 450–453, 2012.

[55] R. Gundlapalli, S. Kumar, and S. Jayanti, "Stack design considerations for vanadium redox flow battery," *INAE Letters*, vol. 3, no. 3, pp. 149–157, 2018.

[56] A. Trovò and M. Guarnieri, "Battery management system with testing protocols for kw-class vanadium redox flow batteries," in *2020 2nd IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES)*, IEEE, vol. 1, 2020, pp. 33–38.

[57] L. Fialho, T. Fartaria, L. Narvarte, and M. Collares Pereira, "Implementation and validation of a self-consumption maximization energy management strategy in a vanadium redox flow bipv demonstrator," *Energies*, vol. 9, no. 7, p. 496, 2016.

[58] H. Bindner, C. Ekman, O. Gehrke, and F. Isleifsson, "Characterization of vanadium flow battery," 2010.

[59] D. Bryans, V. Amstutz, H. H. Girault, and L. E. Berlouis, "Characterisation of a 200 kw/400 kwh vanadium redox flow battery," *Batteries*, vol. 4, no. 4, p. 54, 2018.

[60] F. Baccino, M. Marinelli, P. Nørgård, and F. Silvestro, "Experimental testing procedures and dynamic model validation for vanadium redox flow battery storage system," *Journal of Power Sources*, vol. 254, pp. 277–286, 2014.

[61] T. A. Nguyen, X. Qiu, J. D. Guggenberger II, M. L. Crow, and A. C. Elmore, "Performance characterization for photovoltaic-vanadium redox battery microgrid systems," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 4, pp. 1379–1388, 2014.

[62] E. García-Quismondo, I. Almonacid, M. Á. Cabañero Martínez, V. Miroslavov, E. Serrano, J. Palma, and J. P. Alonso Salmerón, "Operational experience of 5 kw/5 kwh all-vanadium flow batteries in photovoltaic grid applications," *Batteries*, vol. 5, no. 3, p. 52, 2019.

[63] P. Zhao, H. Zhang, H. Zhou, J. Chen, S. Gao, and B. Yi, "Characteristics and performance of 10 kw class all-vanadium redox-flow battery stack," *Journal of Power Sources*, vol. 162, no. 2, pp. 1416–1420, 2006.

[64] C. Blanc, "Modeling of a vanadium redox flow battery electricity storage system," Ph.D. dissertation, Verlag nicht ermittelbar, 2009.

[65] C. Blanc and A. Rufer, “Understanding the vanadium redox flow batteries,” *Paths to Sustainable Energy*, vol. 18, no. 2, pp. 334–336, 2010.

[66] H. Zhang, X. Li, and J. Zhang, *Redox Flow Batteries: Fundamentals and Applications*. CRC Press, 2017.

[67] R. Xiong, *Battery management algorithm for electric vehicles*. Springer, Singapore, 2020.

[68] R. Xiong and W. Shen, *Advanced battery management technologies for electric vehicles*. John Wiley & Sons, 2019.

[69] R. Zhang, B. Xia, B. Li, L. Cao, Y. Lai, W. Zheng, H. Wang, W. Wang, and M. Wang, “A study on the open circuit voltage and state of charge characterization of high capacity lithium-ion battery under different temperature,” *Energies*, vol. 11, no. 9, p. 2408, 2018.

[70] Z. Wei, A. Bhattarai, C. Zou, S. Meng, T. M. Lim, and M. Skyllas-Kazacos, “Real-time monitoring of capacity loss for vanadium redox flow battery,” *Journal of Power Sources*, vol. 390, pp. 261–269, 2018.

[71] Z. Wei, T. M. Lim, M. Skyllas-Kazacos, N. Wai, and K. J. Tseng, “Online state of charge and model parameter co-estimation based on a novel multi-timescale estimator for vanadium redox flow battery,” *Applied Energy*, vol. 172, pp. 169–179, 2016.

[72] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, 2019.

[73] *Feed-in tariff program*. [Online]. Available: https://www.ieso.ca/en/Sector-Participants/Feed-in-Tariff-Program/FIT-Archive.

[74] F. Chollet *et al.*, *Keras*, https://keras.io, 2015.