

# Operations of Fuel Cell Vehicle-to-Grid Systems: From Rule-based to Supervised Learning

by

Arda Mert Cetin

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada, 2025

©Arda Mert Cetin 2025

## **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

## Abstract

To combat rising greenhouse gas emissions in the transportation sector, hydrogen-powered fuel cell electric vehicles (FCEVs) present a promising alternative. A key advantage of this technology is the ability to use FCEVs as mobile power generation devices in vehicle-to-grid (V2G) stations. This combination of fuel cell electric vehicle-to-grid (FCEV2G) was found to be economically viable, but the profits depend on the station's operation strategy. This thesis investigates an optimal operation strategy for an FCEV2G station, which progresses from baseline operational simulation to an advanced intelligent agent model developed using machine learning.

First, a detailed rule-based operational simulation was developed for a FCEV2G station using historical data from Ontario as an example. This model was improved from the literature by incorporating several key real-world components, including the hydrogen cycle, dynamic FCEV participation patterns, and variations in individual FCEV efficiency due to pre-existing degradation. The analysis concluded that the station's operational performance is limited because it is a rule-based operation strategy that is unable to act optimally within any given hour. This limitation can be explained in three operational failures. First, its non-optimal dispatch logic causes the system to fail to reserve its limited hydrogen for periods of peak value. Second, this mismanagement of hydrogen is then amplified by low round-trip efficiency. Finally, the station's operation is constrained from using high-cost market hydrogen to buffer this deficit.

To overcome these limitations, the second phase focused on developing a machine learning (ML) agent. A behavioral cloning agent was trained to mimic the decisions of an expert i.e., a Mixed-Integer Linear Program (MILP) optimizer, which establishes the theoretical profitability of the system. The trained agent demonstrated definitive success, achieving 93.2% of the expert's optimal profit on training data and a robust 80.4% on an unseen test set. This high performance confirms the feasibility of using ML agent for the FCEV2G operation. This approach also provides a significant advantage in decision speed: the agent makes decisions in milliseconds, replacing the computationally intensive MILP expert. Analysis of the agent's behavior revealed that it successfully learned to navigate volatile market conditions, including extreme price shocks, by mastering the expert's forward-looking strategies.

In conclusion, this research delivers a successful proof-of-concept for an intelligent FCEV2G operational controller. The primary contribution is demonstrating that a fast ML agent can learn the forward-looking operational strategies of a slow optimizer. By mastering these strategies, the agent achieves near-optimal profitability in real-time, proving a viable pathway for deploying intelligent control systems to manage the day-to-day operations of volatile energy assets.

## Acknowledgements

My deepest gratitude is extended to my supervisor, Dr. XiaoYu Wu, whose mentorship was a foundation of my master's journey. His profound wisdom, immense patience, and meticulous guidance not only shaped this research but also inspired my own growth. I am especially thankful for his kindness and steady presence, which empowered me with confidence, and for the remarkable opportunities he created for me, which have been transformative both personally and professionally.

I am also indebted to my thesis committee members, Dr. Roydon Fraser and Dr. Yverick Pascal Rangom, for their critical feedback and valuable contributions, which significantly elevated the quality of this work.

My research was enriched by critical discussions with my peers in the Greener Production Group. These interactions were vital to my learning process.

This project was made possible through the generous financial contributions of the Government of Canada, NSERC, the Transition Accelerator, and MITACS. The advanced computing facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET) and the Digital Research Alliance of Canada were indispensable to this research.

On a personal note, I owe my greatest thanks to my family and friends. Their belief in me has been a constant source of strength throughout my time as an international student. Their encouragement carried me through every challenge, and their unconditional love provided the foundation upon which I could complete my studies.

# Table Of Contents

List of Figures .....	vii
List of Tables .....	viii
Chapter 1 Literature Review .....	- 1 -
1.1 Need for Renewable Energy Solutions .....	- 1 -
1.2 Hydrogen as an Energy Carrier.....	- 2 -
1.2.1 Hydrogen Production.....	- 2 -
1.2.2 Hydrogen Storage and Distribution.....	- 3 -
1.2.3 Hydrogen Usage .....	- 5 -
1.2.4 Proton Exchange Membrane Fuel Cell (PEMFC) Technology .....	- 6 -
1.2.5 PEMFC Degradation .....	- 8 -
1.3 Vehicle-to-Grid (V2G) Integration .....	- 12 -
1.3.1 Overview of V2G Technology.....	- 12 -
1.3.2 Economics of Power Grid .....	- 14 -
1.4 Operation Strategies for FCEV2G Systems.....	- 15 -
1.4.1 Optimization Methods for Energy Management .....	- 15 -
1.4.2 Machine Learning Approaches .....	- 15 -
1.4.3 Reinforcement Learning for FCEV2G Decision Making .....	- 16 -
1.4.4 Behavioral Cloning for FCEV2G Decision Making .....	- 17 -
1.5 Research Gap .....	- 17 -
1.6 Chapter Summary .....	- 18 -
Chapter 2 Rule-based FCEV2G Operational Simulation .....	- 20 -
2.1 Model Description .....	- 20 -
2.2 Electricity Production from V2G.....	- 21 -
2.3 Hydrogen Cycle .....	- 22 -
2.4 Participation of FCEV.....	- 23 -
2.5 Electricity Price Categorization .....	- 26 -

2.6 Base Case Assumptions .....	- 27 -
2.7 Operation Simulation .....	- 27 -
2.7.1 V2G Operation Mode .....	- 27 -
2.7.2 Hydrogen Production Mode .....	- 29 -
2.8 Chapter Summary .....	- 29 -
Chapter 3 Rule-based FCEV2G Operation .....	- 30 -
3.1 Performance Summary.....	- 30 -
3.2 Key Factors Limiting Performance.....	- 32 -
3.3 Chapter Summary .....	- 33 -
Chapter 4 Supervised Learning Model for FCEV2G Operation.....	- 35 -
4.1 Model Description .....	- 35 -
4.2 Participation of FCEV.....	- 36 -
4.3 Supervised Learning Model for FCEV2G .....	- 37 -
4.3.1 Model Action and Observation Spaces.....	- 37 -
4.3.2 Optimization Data as the Expert Dataset.....	- 40 -
4.3.3 Simulation and Evaluation Setup .....	- 43 -
4.4 Training Procedure for the SL Agent .....	- 47 -
4.5 Chapter Summary .....	- 49 -
Chapter 5 Supervised Learning-based FCEV2G Operation .....	- 50 -
5.1 Comparing SL Agent and MILP Expert: Training Dataset.....	- 50 -
5.2 Comparing SL Agent and MILP Expert: Test Dataset.....	- 57 -
5.3 Chapter Summary .....	- 64 -
Chapter 6 Conclusion and Recommendations for Future Work .....	- 65 -
6.1 Conclusion .....	- 65 -
6.2 Recommendations for Future Work.....	- 66 -
References.....	- 67 -

## List of Figures

Figure 1.1 Proton Exchange Membrane Fuel Cell Schematic .....	- 6 -
Figure 1.2 Polarization Curve and Three Distinct Operational Regions .....	- 7 -
Figure 1.3 V2G Overview.....	- 12 -
Figure 1.4 V2G Overview with FCEVs.....	- 13 -
Figure 2.1 The Two Different Operation Modes of the FCEV2G Station: (a) V2G operation mode and (b) Hydrogen production mode .....	- 20 -
Figure 2.2 Polarization and Power curve with Increasing Utilization Time.....	- 22 -
Figure 3.1 The Hourly Average in 24 hours V2G Power Generation, and Hydrogen Consumption; and Electrolyzer Power Input, and Hydrogen Production.....	- 32 -
Figure 4.1 Schematic of Two Different Operation Modes of the FCEV2G Station, a. V2G operation mode and b. H <sub>2</sub> production mode, decided by the Behavioral Cloning Agent.....	- 35 -
Figure 4.2 The Average Truck Count of Actual and Predicted Data.....	- 37 -
Figure 5.1 The Cumulative Profit Difference between SL Agent and MILP Expert on Training data ..	- 50 -
Figure 5.2 Profit Distribution of the Actions Taken by Agent vs Expert.....	- 51 -
Figure 5.3 Frequency Distribution of the Actions Taken by Agent vs Expert .....	- 51 -
Figure 5.4 Comparison of Agent vs. Expert Strategy by Price .....	- 53 -
Figure 5.5 Overall Distribution of Hydrogen Storage Levels.....	- 54 -
Figure 5.6 Agent's Storage Level One Hour before V2G (from Market) Decision.....	- 55 -
Figure 5.7 Distribution of Agent's Idle Hours by Price Category.....	- 56 -
Figure 5.8 The Cumulative Profit Difference between Behavioral Cloning Agent and MILP Expert on Testing data.....	- 58 -
Figure 5.9 Comparison of Agent vs Expert on Profit Contribution.....	- 59 -
Figure 5.10 Comparison of Agent vs Expert on Action Frequency.....	- 59 -
Figure 5.11 Comparison of Agent vs. Expert Strategy by Price in the Test Set .....	- 60 -
Figure 5.12 Overall Distribution of Hydrogen Storage Levels on Test Dataset .....	- 61 -
Figure 5.13 Agent's Storage Level when Deciding to Use Market Hydrogen.....	- 62 -
Figure 5.14 Comparison of Idle Hours by Price Category for Test Data.....	- 63 -

## List of Tables

Table 2.1 Ontario’s Traffic Volume Data for Eastbound and Westbound ONroutes of 401 Highway [54]... - 24 -	
Table 2.2 Daily Congestion Score and Percentage by Time of Day [56] ..... - 25 -	
Table 2.3 Ontario’s Electricity Price Categories and Their Corresponding Potential Hydrogen Costs and V2G Revenues ..... - 26 -	
Table 3.1 Profit Distribution of the FCEV2G Station..... - 30 -	
Table 3.2 Hours of Operation Modes of the FCEV2G Station ..... - 31 -	
Table 4.1 SL Action Space ..... - 38 -	
Table 4.2 SL Observation Space ..... - 39 -	
Table 4.3 MILP Decision Variables ..... - 41 -	
Table 4.4 SL and MILP Constants ..... - 44 -	

## Chapter 1 Literature Review

### 1.1 Need for Renewable Energy Solutions

Electricity is a fundamental necessity, and its continuous supply is critical to support human activities. The growth rate of global electricity demand between 2023 and 2024 was 4.3%, and it is forecasted to remain at 3.9% between 2025 and 2027 [1]. To meet this rising electricity demand, various energy sources are utilized, with fossil fuels being the dominant method. In 2024, fossil fuels, e.g., coal, crude oil, and natural gas accounted for approximately 59.1% of the global electricity supply [1]. When burned, these fuels release carbon dioxide, water, and heat, with the generated heat primarily driving the production of electricity.

The critical challenge associated with fossil fuel usage is the generation of anthropogenic greenhouse gas (GHG) emissions, including carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), and nitrous oxide (N<sub>2</sub>O). These GHG emissions trap heat from the Sun within the Earth's atmosphere, functioning like an insulating blanket and driving global warming and climate change. Compared to pre-industrial levels, the Earth's average surface temperature has already risen by approximately 1.2°C [2]. Such a temperature rise intensifies natural disasters, elevates sea levels, and negatively impacts biodiversity. To combat climate change, the Paris Agreement was adopted in 2015, aiming to limit the average global temperature increase to well below 2°C, with the ultimate goal of capping the rise at 1.5°C above pre-industrial levels [3]. This agreement emphasizes international collaboration, providing financial, technical, and capacity-building support to countries in need of assistance to meet these global targets.

Given these challenges, there is an urgent need to shift toward renewable energy sources, including solar, wind, hydropower, and geothermal, which can sustainably meet global electricity demands. To complement electrification, hydrogen serves as a versatile energy carrier. When produced using electricity from renewable sources, clean hydrogen can store energy to bridge the intermittency of variable renewable energy like solar and wind. Consequently, hydrogen is especially promising for sustainable electricity generation and effective grid integration.

## 1.2 Hydrogen as an Energy Carrier

### 1.2.1 Hydrogen Production

Production of hydrogen is categorized into different colors based on the different types of technology used. The most common method is steam methane reforming, where methane reacts with steam to produce hydrogen along with carbon dioxide as a byproduct. If the carbon dioxide generated in this process is captured and stored, the resulting hydrogen is labeled as blue hydrogen. However, if no carbon capture occurs, the hydrogen produced from methane reforming is known as grey hydrogen, which contributes significantly to greenhouse gas emissions. Another method for producing hydrogen is methane pyrolysis, which breaks down methane into hydrogen and solid carbon. Hydrogen generated this way is known as turquoise hydrogen. Additionally, hydrogen can be produced through gasification of coal; depending on the type of coal used, it is classified as either black hydrogen (from black coal) or brown hydrogen (from lignite or brown coal). These coal-based methods are the least environmentally friendly due to their high carbon emissions and pollution [4].

Yet the most environmentally friendly method for producing hydrogen is through electrolysis using clean electricity. Electrolysis is a process using electricity to split water molecules into hydrogen and oxygen, without generating any carbon dioxide as a byproduct. The specific color classification of hydrogen produced by electrolysis depends on how the electricity is generated. For instance, if the electricity is provided exclusively by solar power, the resulting hydrogen is classified as yellow hydrogen. When nuclear energy supplies the electricity, the hydrogen produced is known as pink hydrogen. Green hydrogen uses the surplus renewable energy sources for electrolysis [4].

Although green hydrogen is the most sustainable method of hydrogen production, it accounted for less than 1% of the total hydrogen produced in the United States in 2024 [5]. In contrast, approximately 95% of hydrogen production came from grey hydrogen. For every kilogram of grey hydrogen produced, around 12 kilograms of carbon dioxide are released [5]. Green hydrogen, by comparison, produces no direct carbon emissions, making it significantly cleaner. The limited adoption of green hydrogen is largely due to the high cost of electrolysis and the intermittent nature of renewable energy sources. While traditional grey hydrogen from fossil fuels costs just 1.6–3.2 CAD per kilogram, green hydrogen produced via renewables is far more expensive, at 4.8–12.7 CAD per kilogram, which creates a major economic barrier [6]. Electrolyzers must become more efficient and less expensive to compete economically with grey hydrogen. Projections indicate this is achievable. Forecasts suggest that green hydrogen production costs will decrease by around 50% by 2030 and continue to fall steadily through 2050 [6]. Additionally, since solar and wind

power aren't available consistently throughout the day, it can be difficult to ensure a steady and reliable electricity supply for continuous hydrogen production.

While green hydrogen is pursued as a manufactured clean energy carrier, recent discoveries have highlighted the potential of white hydrogen (also known as natural or geologic hydrogen), which is hydrogen that exists naturally as a free gas below the Earth's surface. This resource is generated through various geological processes, such as the interaction of water with iron-rich rocks and can accumulate in underground reservoirs. Preliminary estimates suggest that vast quantities of natural hydrogen may exist globally. The primary appeal of white hydrogen lies in its potential for being both clean and remarkably cheap. Since it only requires extraction rather than energy-intensive production like electrolysis, its production costs could be exceptionally low, with some projections as low as \$0.50 to \$1.00 per kilogram [7]. Furthermore, the extraction process is believed to have a significantly smaller environmental footprint compared to the grey hydrogen. With a hydrogen content of 85 per cent and minimal methane contamination, its carbon intensity is at approximately 0.4 kg of CO<sub>2</sub> equivalent per kilogram of hydrogen, which is dramatically lower than that of grey hydrogen, which is 12 kilograms of carbon dioxide per kilograms of grey hydrogen [7]. Although exploration is still in its early stages, white hydrogen is now considered a promising frontier that could potentially provide a clean, low-cost source of hydrogen, complementing the role of green hydrogen in the energy transition.

### **1.2.2 Hydrogen Storage and Distribution**

Hydrogen offers a valuable solution for preventing the curtailment of excess electricity, especially from renewable sources. By converting surplus electricity into hydrogen, energy can be stored and used later as a flexible energy carrier. This concept is being demonstrated in real-world applications; for example, the Advanced Clean Energy Storage hub in Delta, Utah, is a major project designed to produce and store green hydrogen in underground salt caverns, providing seasonal energy storage to balance the grid [8]. This stored hydrogen can then be redeployed to support various sectors, including industrial processes, heavy transportation, and providing power back to the grid during peak demand [9].

While hydrogen's energy storage potential is significant, its physical properties present technical challenges. Although hydrogen has the highest energy content by mass of any common fuel at approximately 120 megajoules per kilogram, which is nearly three times that of gasoline (~44 MJ/kg), it has a very low energy density by volume under ambient conditions. For instance, as a gas at standard temperature and pressure, hydrogen's volumetric energy density is only about 0.01 megajoules per liter. This is thousands of times lower than gasoline (around 32 MJ/L) [9]. This causes advanced storage solutions

like high-pressure compression (up to 700 bar) or liquefaction at cryogenic temperatures ( $-253^{\circ}\text{C}$ ) to make it practical for applications like transportation. Therefore, effective hydrogen storage requires methods that achieve high energy density while remaining cost-efficient.

Hydrogen can be stored using either physically-based or material-based methods. Starting with physically-based storage, the most common method is storing hydrogen as a compressed gas in high-pressure tanks, typically at 350–700 bar. This approach is mature and widely used in the transportation sector due to its practicality and fast refueling times. For instance, the Nikola fuel cell electric trucks, which are being deployed for commercial operations in Alberta, Canada, utilize high-pressure tanks that store hydrogen at 700 bar to maximize their driving range [10]. Similarly, light-duty passenger vehicles like the Toyota Mirai also rely on 700 bar tanks to store enough hydrogen for a competitive range of over 600 kilometers [11].

Another method is storing hydrogen in liquid form at cryogenic temperatures of  $-253^{\circ}\text{C}$ . Storing hydrogen as a liquid increases its volumetric density significantly compared to compressed gas, but this comes at a high cost in terms of both energy and expense. The liquefaction process is highly energy-intensive, consuming between 6 to 12 kilowatt-hours of electricity for every kilogram of hydrogen [12]. This energy penalty is substantial, representing 20% to 40% of the energy content of the hydrogen being stored [12]. Furthermore, even after production, liquid hydrogen requires costly cryogenic tanks and faces continuous energy losses from boil-off, where a portion of the liquid evaporates over time [12].

Cryo-compressed storage, which combines pressure and low temperature (e.g.,  $-220^{\circ}\text{C}$  and 350 bar), aims for higher volumetric density than liquid hydrogen and reduced boil-off losses. However, the engineering complexity of its tanks makes it the most expensive physical storage option. This high cost is because it combines the most significant cost drivers from both conventional methods. High-pressure gas systems are dominated by the high capital cost of the tank, while liquid hydrogen systems are dominated by the high energy cost of the liquefaction process (over  $\$1.20/\text{kg}$ ) [13].

Material-based storage involves absorbing hydrogen into solid materials (like metal hydrides) or adsorbing it onto high-surface-area materials (like metal-organic frameworks). This storage faces significant hurdles that prevent commercial use. The hurdles arise from needing sophisticated thermal management systems. Many materials require high temperatures (often  $>80^{\circ}\text{C}$ ) to release hydrogen at a sufficient rate. Key limitations of these storage types are the slow hydrogen release kinetics, which makes it fail to meet the immediate power demands of a vehicle, fast degradation, and the high cost and weight of the storage materials themselves [14].

Hydrogen also needs to be transported from storage sites to end-use locations. For large-scale demand, pipeline distribution is the most cost-effective method, with a levelized cost potentially below 1 CAD per kilogram, compared to over 2 CAD per kilogram for gaseous tube trailers [15]. However, the high capital cost of pipelines is justified for high-volume corridors. If the distance is short, it typically requires a consistent demand of 1 to 1.2 tonnes of hydrogen per day per kilometer of pipeline to be economical [15]. Additionally, hydrogen's tendency to cause embrittlement in conventional steel demands careful engineering and maintenance. To address this, pipeline designs may include using specialized steel alloys, adding internal polymer liners as barriers, or blending hydrogen into existing natural gas pipelines at low concentrations [15].

For smaller-scale or more flexible distribution needs, gaseous tube trailers and liquefied hydrogen tankers are used, although these methods are typically more expensive and less energy-efficient than pipelines. This inefficiency arises directly from the energy required to prepare the hydrogen for transport. Gaseous hydrogen must undergo intensive multi-stage compression to be stored on tube trailers, while liquid hydrogen requires an even more energy-intensive cryogenic liquefaction process [16], [17]. In summary, while pipelines offer the lowest long-term cost and energy use for steady, high-volume demand, tube trailers and tankers provide greater flexibility at the expense of these higher operational and energy costs.

### **1.2.3 Hydrogen Usage**

Most of the hydrogen produced today is used in industrial processes. These include refining petroleum by removing sulfur from fuels, treating metals, producing ammonia for fertilizers, and manufacturing other chemicals such as methanol. Hydrogen is also used in food processing applications. However, because most hydrogen is currently produced using fossil fuels, it is important to focus on greener end uses to help offset the carbon footprint of its production and move toward a net-zero energy system [18].

As a clean energy carrier, hydrogen has the potential to support emission-free systems for transportation and power generation. It can be converted to power through direct combustion or, more efficiently, through electrochemical conversion in a fuel cell. A Proton Exchange Membrane (PEM) fuel cell, common in transportation, converts hydrogen's chemical energy directly into electricity with an efficiency of 40% to 60% [19]. In contrast, when hydrogen is burned in a hydrogen internal combustion engine (H<sub>2</sub>-ICE), the efficiency is generally lower, typically ranging from 25% to 45% [20]. This performance gap exists because fuel cells are electrochemical devices, not heat engines, and are therefore

not constrained by the same thermodynamic limitations of the Carnot cycle that govern the maximum theoretical efficiency of combustion-based systems [18].

Fuel cells come in different sizes and types depending on their use. They can provide electricity for single buildings and have potential for use in microgrids or remote areas that don't have access to the main power grid. Fuel cell electric vehicles (FCEVs) are being developed due to their higher efficiency and lower carbon emissions compared to conventional internal combustion vehicles. However, the widespread adoption of FCEVs is still limited by high costs and a lack of hydrogen refueling infrastructure [18].

### 1.2.4 Proton Exchange Membrane Fuel Cell (PEMFC) Technology

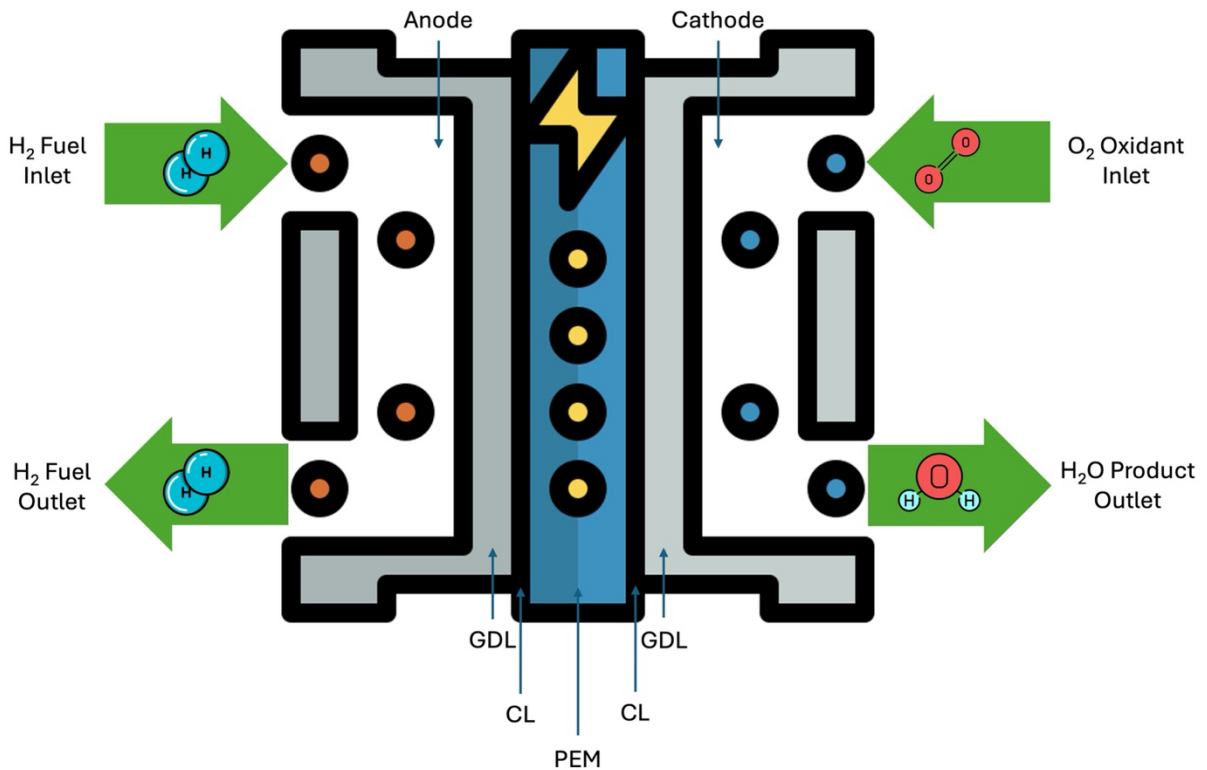


Figure 1.1 Proton Exchange Membrane Fuel Cell Schematic

The most common type of fuel cell used, especially in transportation, is the proton exchange membrane fuel cell (PEMFC), which can be seen in Figure 1.1. PEMFCs use a proton-conducting polymer membrane as the electrolyte, operate at low temperatures (typically 60–80°C), and are capable of rapid response to changes in power demand. While they can also be used for stationary power generation, other

types of fuel cells may be more suitable for those applications due to differences in efficiency, operating conditions, and fuel flexibility [21].

Electricity in a PEMFC is generated through an electrochemical reaction between hydrogen and oxygen. Hydrogen is fed into the anode side, where it meets a platinum-based catalyst layer. This catalyst splits the hydrogen molecules into protons and electrons. The polymer membrane at the center of the fuel cell allows only protons to pass through to the cathode side, while blocking the electrons. As a result, the electrons are forced to travel through an external circuit, generating an electric current before reaching the cathode. At the cathode, the electrons and protons combine with oxygen to form water. The only outputs of a PEMFC are electricity, water, and heat, making the process environmentally clean [22].

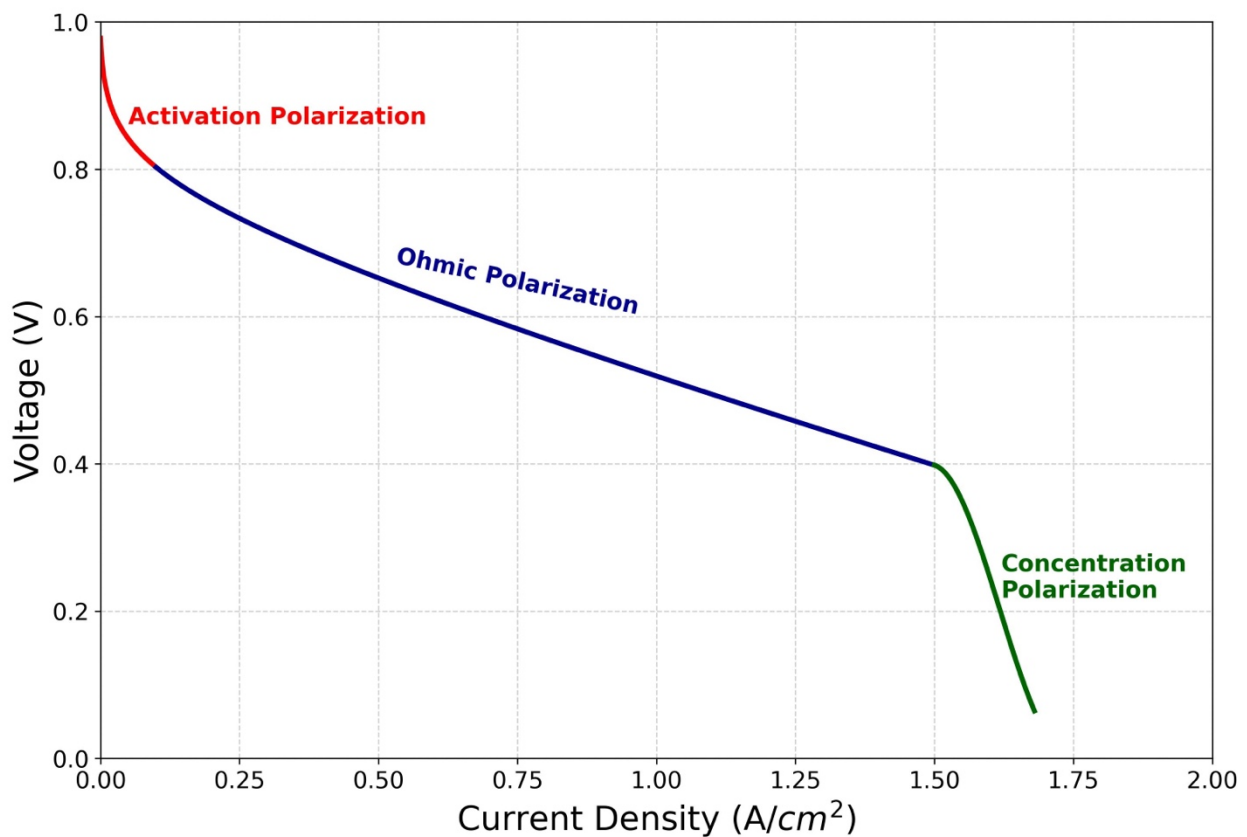


Figure 1.2 Polarization Curve and Three Distinct Operational Regions

To evaluate the performance of a fuel cell, polarization curve experiments are commonly conducted. A polarization curve illustrates the relationship between cell voltage and current density, providing key insights into the electrochemical behavior of the system. This curve can be obtained using a potentiostat, which applies a controlled current while measuring the resulting voltage output. By gradually

increasing the load, different voltage values corresponding to various current densities can be recorded. Polarization curves typically exhibit three distinct operational regions, each dominated by a different type of voltage loss, which can be seen in Figure 1.2. At low current densities, voltage loss is primarily due to activation polarization, which results from the energy barrier associated with the electrochemical reactions on the catalyst surface. At intermediate current densities, ohmic polarization dominates, caused by resistance within the membrane, electrodes, and electrical connections, leading to a roughly linear drop in voltage. At high current densities, the voltage loss is attributed to concentration polarization, which arises from limitations in mass transport—specifically, the inadequate supply of reactants and removal of products from the reaction sites [23].

The main advantages of PEMFCs include high efficiency, low environmental impact, quiet operation, and scalability. They can achieve 40% to 60% efficiency in electricity generation without producing greenhouse gas emissions. Their compact and modular design allows them to be scaled for small and large power applications. However, PEMFCs also face several challenges. The high cost of materials, particularly platinum catalysts, significantly increases system costs. In addition, both the membrane and catalyst layer are sensitive to impurities in the hydrogen and air supply, which can lead to performance degradation over time. The current limited hydrogen infrastructure also restricts widespread adoption [22].

### **1.2.5 PEMFC Degradation**

The long-term durability of a PEMFC is linked to the stability of its core components. The part with the greatest influence on long-term durability is the membrane electrode assembly (MEA), which consists of two gas diffusion layers (GDLs), a proton exchange membrane (PEM), typically Nafion in PEMFCs, and two catalyst layers (CLs). These components can be seen in Figure 1.1, and are vulnerable to degradation over time, especially when operating conditions and water management are not properly controlled [24]. Additionally, factors such as the mechanical design of the fuel cell, the purity of input gases, and the desired voltage output can significantly affect long-term performance [25]. The causes of fuel cell degradation are classified into two main groups: degradation due to system operation and degradation due to contaminants. The first group, operational degradation, is further subdivided into four key modes. The following four subsections are dedicated to analyzing each of these operational modes individually, after which the chapter will proceed to address contaminant-induced degradation.

#### **1.2.5.1 Chemical Degradation**

Under proper operating conditions and sufficient membrane humidification, Nafion offers excellent proton conductivity, making it well-suited for PEMFC applications. However, one of its main limitations

is chemical degradation, primarily caused by the formation of reactive oxygen species within the membrane [26]. During fuel cell operation, oxygen molecules can permeate the membrane from the cathode and interact with the platinum catalyst, generating hydroperoxyl ( $\text{HOO}\bullet$ ) and hydroxyl ( $\text{HO}\bullet$ ) radicals [27]. These radicals attack vulnerable terminal groups in the polymer backbone, particularly  $-\text{COOH}$  and  $-\text{SO}_3\text{H}$  groups, leading to the breakdown of the membrane structure. Two major degradation mechanisms have been identified: main-chain unzipping, which involves the breakdown of the carbon-fluorine backbone, and side-chain scission, typically occurring under low humidity and during extended operation [26]. In addition to radical attack, factors such as hydrogen crossover, contaminant ions, and unstable operating conditions can accelerate membrane degradation and impact PEMFC durability.

### **1.2.5.2 Mechanical Degradation**

Another operational degradation type is mechanical degradation. During the fabrication of the MEA, various mechanical deformations can occur, affecting long-term durability and performance. One common issue is cracking of the CL surface without damaging the membrane itself. This typically results from improper handling during MEA manufacturing or rapid solvent evaporation during CL fabrication. Additional cracking can occur when the MEA is bent or stretched, as the catalyst layer is generally less flexible than the membrane [28], [29]. Thickness variations also arise when the catalyst is unevenly distributed across the CL during manufacturing. Delamination, which is the separation between the CL and membrane, can result from poor lamination conditions, including inappropriate pressure, temperature, catalyst slurry composition, or casting speed [28], [29].

Further defects can be traced to inadequate catalyst ink preparation, where poor mixing or uneven slurry spreading leads to the formation of membrane clusters or agglomerated catalyst particles, negatively impacting performance. Uniform orientation and even distribution of catalyst particles are critical for optimal operation [28], [29]. In addition, like chemical degradation, the membrane is sensitive to changes in humidification and operating conditions, which cause it to swell or shrink. These mechanical stresses lead to material fatigue, creep, wrinkles, pinholes, delamination, and cracks. If these deformations are not addressed, they can compromise the MEA and lead to mechanical failure [28]. Critically, such damage may allow hydrogen and oxygen to come into direct contact within the membrane, generating localized hot spots due to the exothermic reaction, which in turn reduces power output and accelerates degradation.

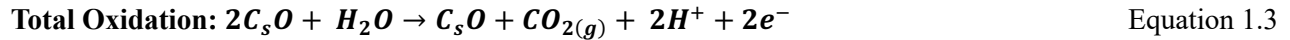
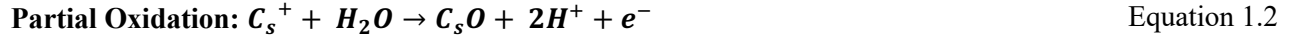
### **1.2.5.3 Catalyst Degradation**

In a PEMFC, the oxygen reduction reaction (ORR) is the rate-determining step, and enhancing its catalytic activity can reduce platinum (Pt) usage and overall system cost. The ORR activity is often

expressed as the mass activity (MA) of Pt, defined as the product of the electrochemically active surface area (ECSA) and the specific activity (SA) normalized to surface area [30]. A decrease in ECSA, which caused by Pt mass loss or catalyst coarsening, lowers the fuel cell's power output [31]. Under high potentials, acidic environments, high temperatures, or dynamic load conditions, Pt can undergo oxidation and dissolution from the CL. Due to hydrogen crossover, dissolved Pt ions may encounter hydrogen molecules and be reduced, leading to Pt deposition within the membrane or ionomer, forming a Pt band. This band reduces hydrogen oxidation capability and increases degradation.

Two primary mechanisms contribute to Pt particle coarsening. The first is Ostwald ripening, where dissolved Pt ions redeposit onto existing Pt particles, increasing their size and accelerating further dissolution. The second is particle coalescence, which occurs when Pt particles migrate from the carbon support and physically merge, altering shape and reducing active surface area [32].

Degradation of carbon support is also a critical issue. Carbon corrosion is primarily driven by water-mediated oxidation reactions, where surface-bound species (denoted by subscript s) react to produce CO<sub>2</sub>, as shown in the following reactions, which are Equation 1.1, Equation 1.2, and Equation 1.3.



These reactions are intensified under high or non-uniform voltages (typically above 0.7 V). Conditions such as fuel starvation, poor water management, and frequent start-up/shutdown cycles exacerbate voltage fluctuations, accelerating carbon corrosion and leading to structural collapse of the CL [26].

#### 1.2.5.4 Thermal Degradation

The Nafion membrane in a PEMFC requires sufficient humidification to maintain effective proton conductivity. Even when ambient temperatures fall below 0°C, the waste heat generated during operation usually prevents water from freezing. However, when the cell stops producing heat under cold conditions, ice formation can occur within the membrane. Since ice has a larger volume than water, this expansion can cause cracking in the membrane and lead to thickness variations in the electrodes [33].

Different scenarios of ice formation under varying thermal conditions can be stated. When heat generation is much lower than heat loss, water freezes in the CL due to limited transfer to the GDL, damaging only the CL over time. If water does reach the GDL, icing begins there first, eventually progressing back into the CL resulting in widespread freezing and operational failure. In contrast, when heat generation equals or exceeds heat loss, ice formation is suppressed. These last two scenarios represent the ideal thermal management strategies for preventing freeze-related degradation in PEMFCs [34].

At the other extreme, high operating temperatures can also harm PEMFC performance. While the optimal temperature range is 60–80°C, excessive heat reduces membrane humidification, leading to a drop in proton conductivity. If the temperature exceeds 280°C, the sulfonic acid side chains of the Nafion membrane may decompose, compromising the membrane's structural integrity [35].

#### **1.2.5.5 Degradation due to Contaminants**

As multiple subsystems are involved in operating a PEMFC, contaminants can originate from different input and output streams. In particular, the purity of the hydrogen fuel and air oxidant plays a critical role in performance. For example, trace gases such as ammonia ( $\text{NH}_3$ ) and nitrogen oxides ( $\text{NO}_x$ ) can react with hydrogen to form ammonium ions ( $\text{NH}_4^+$ ). These ions inhibit proton conductivity by poisoning the catalyst and reacting with the Nafion membrane, resulting in performance degradation [36].

Contaminants can also interfere with the GDL by depositing on carbon fibers, altering the surface characteristics of pores and flow channels. This affects water management and mass transport within the cell. Common examples include contamination by transition metal ions, alkaline earth metals, and sodium ions ( $\text{Na}^+$ ). One notable manifestation is salt precipitation, which blocks GDL pores and channels. This blockage is particularly problematic because it cannot be resolved by simply increasing flow rates, and may cause lasting performance loss [36].

Hydrogen produced via steam methane reforming (SMR) typically contains 40% to 70% hydrogen, 15% to 25%  $\text{CO}_2$ , 1% to 2% carbon monoxide (CO), and trace amounts of sulfur, water vapor, and nitrogen-based contaminants [26]. One of the critical challenges with this composition is that CO has a much stronger affinity for Pt catalyst sites than hydrogen. As a result, CO molecules adsorb onto the Pt surface and block the active sites and preventing the hydrogen oxidation reaction (HOR) from occurring effectively [37].

The concentration of CO significantly impacts cell voltage, with the severity of the effect changing at different current densities. At 25 ppm, the polarization curve slope is increased to that of a CO-free case, indicating that voltage losses are primarily due to CO adsorption blocking reaction sites. However, at 100 ppm and above, the slope increases sharply at higher current densities. This is attributed to the oxidation of

adsorbed CO into CO<sub>2</sub>, which frees up catalyst sites and momentarily increases the reaction rate. At low current densities, the impact of CO poisoning is minimal, but as current density rises, its effects become significantly more pronounced [38].

Compared to the effects of CO, the presence of CO<sub>2</sub> or nitrogen (N<sub>2</sub>) in the hydrogen fuel stream has a much smaller impact on PEMFC performance. CO directly poisons the platinum catalyst by blocking active sites and preventing hydrogen adsorption, whereas CO<sub>2</sub> and N<sub>2</sub> primarily act as diluents, reducing hydrogen's partial pressure without strongly interacting with the catalyst. However, when both CO and CO<sub>2</sub> are present, the situation can worsen due to the reverse water-gas shift reaction, which converts CO<sub>2</sub> and H<sub>2</sub> into additional CO and water. This reaction increases the CO concentration in the system, intensifying catalyst poisoning and further impairing fuel cell efficiency [37].

### 1.3 Vehicle-to-Grid (V2G) Integration

#### 1.3.1 Overview of V2G Technology

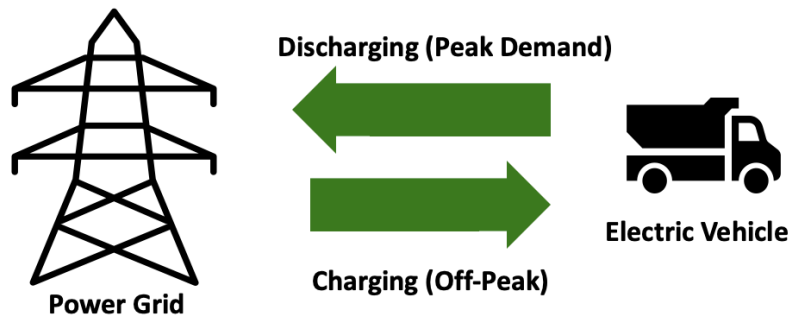


Figure 1.3 V2G Overview

Electric vehicles (EVs) can serve as distributed energy resources by connecting to the power grid in a bidirectional manner, a process known as vehicle-to-grid (V2G). Through V2G, EVs are not only consumers of electricity but can also store energy and return it to the grid when needed. As illustrated in Figure 1.3, this process involves charging the vehicle when demand is low (off-peak) and sending power back to the grid when demand is high (peak demand). This capability enhances grid stability, especially as power systems increasingly rely on renewable energy sources and helps balance fluctuations in supply [39].

One of the key reasons V2G technology is viable is that personal vehicles typically remain parked for most of the day. For example, studies in the United States show that personal vehicles spend only 4–

5% of their time on the road, while the remaining time is spent parked at homes or in public lots. During these idle periods, EVs can be plugged into the grid and communicate with grid operators, allowing for real-time decisions that enhance grid stability. This reduces the need for additional grid-scale electricity storage infrastructure and helps manage the variability of renewable energy generation. In return, EV owners benefit by accessing cheaper electricity for charging or earning revenue through participation in grid services. As a result, both the grid operator and the EV owner benefit economically while contributing to lower emissions from electricity production [39].

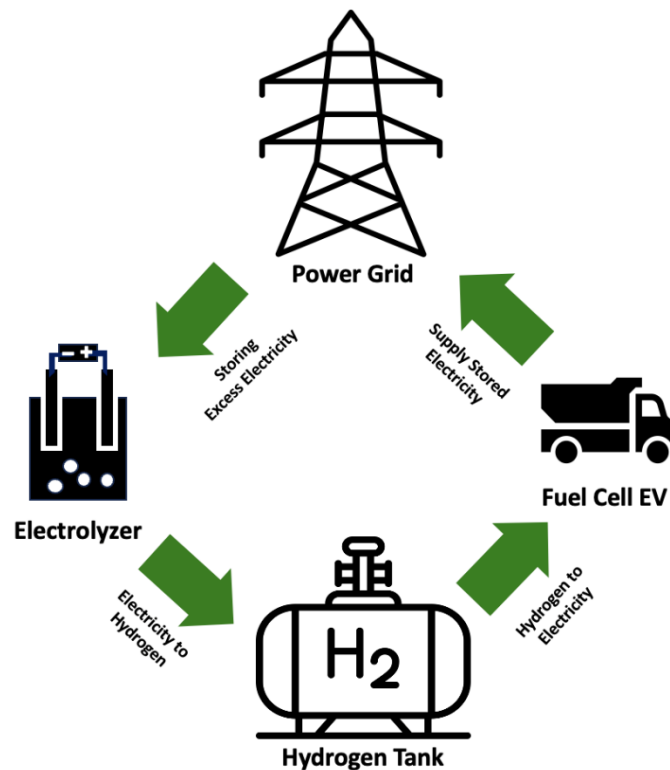


Figure 1.4 V2G Overview with FCEVs

The implementation of V2G services can leverage existing battery charging infrastructure. By upgrading conventional unidirectional chargers to bidirectional systems, charging stations can support battery electric vehicles (BEVs) in providing power back to the grid.

Shifting from a direct charge/discharge cycle, the integration of fuel cell electric vehicles (FCEVs) into V2G operations relies on hydrogen as an intermediate energy carrier. This requires investing in hydrogen storage and electrolyzer systems. This hydrogen-based V2G pathway, illustrated in Figure 1.4, involves two main stages. First, excess grid electricity, often from renewable sources during off-peak times,

powers an electrolyzer to produce hydrogen for storage. Second, during periods of high demand, FCEVs use this on-site stored hydrogen in their onboard fuel cells to generate electricity and supply it back to the grid, thus acting as mobile power plants. Over the long term, this pathway offers a more environmentally sustainable solution. Electrolyzers can convert excess electricity into hydrogen, which can then be stored and later used to generate electricity during peak demand periods.

One of the key challenges in implementing V2G technology is the control of electricity generation and cycling behavior. Both batteries and fuel cells can degrade over time, especially when frequently switching between charging and discharging. These cycles often require different operating conditions, such as varying voltages and temperatures. Rapid shifts between these conditions can accelerate material degradation and shorten the lifespan of the system [40]. Another major challenge is the uncertainty in EV participation. Vehicle availability at V2G stations is often random and unpredictable. Stations must be designed to handle a wide range of participation levels, which makes capacity planning difficult. This is especially important given the high initial investment costs of V2G infrastructure. Without proper planning, underutilization can lead to energy waste and economic losses. To avoid this, careful operational and financial strategies are essential [39].

The fuel cell vehicle-to-grid (FCEV2G) model can be viewed as an alternative to a stationary fuel cell power plant for providing grid services. The fundamental economic trade-off centers on capital investment versus operational certainty. A stationary plant requires a high upfront capital expenditure for its dedicated fuel cell stacks. In contrast, the FCEV2G model avoids this high upfront cost by leveraging the fuel cells already purchased by vehicle owners for transportation. This lower barrier to entry, however, comes at the cost of operational uncertainty. The available power capacity of an FCEV2G station is not fixed but is instead a stochastic function of daily traffic and driver behavior, a challenge not faced by a centralized power plant.

### **1.3.2 Economics of Power Grid**

The balance between electricity supply and demand is a key factor in determining its market price. Electricity is a fundamental necessity that must be produced and consumed in real time. In an ideal system with perfect storage, surplus electricity could be saved and used later without affecting prices. However, when supply exceeds demand without enough storage, electricity prices tend to drop. Since wasting electricity is not desirable, grid operators often lower prices to incentivize consumption. Conversely, when demand exceeds supply, electricity becomes more valuable, and prices rise. These fluctuations are

influenced by daily usage patterns, seasonal variations, and weather conditions, which all affect when and how much electricity is needed [41].

## **1.4 Operation Strategies for FCEV2G Systems**

### **1.4.1 Optimization Methods for Energy Management**

The operation of a FCEV2G system depends on time-dependent and uncertain factors, including electricity prices, grid supply and demand, and the availability of participating FCEVs. To account for the uncertainties, optimization methods are used to determine the best operational strategy. By defining an objective function and constraints, the system can make informed decisions about when to produce hydrogen, and when to dispatch electricity to the grid. Objectives can vary depending on the use case, for example, maximizing profit or minimizing carbon emissions.

Since the objective function and system constraints in FCEV2G operations can be expressed as linear equations, both linear programming (LP) and mixed-integer linear programming (MILP) are suitable methods for system design. The key difference between the two is that LP assumes all decision variables are continuous. In contrast, MILP can incorporate binary and integer variables. LP is used in modeling time-dependent energy systems due to its computational efficiency and ease of interpretation [42]. In the context of FCEV2G, time-dependent variables like hourly electricity prices and vehicle participation can be incorporated into these models to support real-time operational decisions.

However, the lack of reliable past or future data for time-dependent variables creates a challenge of unpredictability. One effective way to address this uncertainty is through stochastic programming, which incorporates the probability distributions of these variables into the optimization process. Stochastic models generate multiple future scenarios, allowing the system to evaluate possible outcomes and identify the best operational strategy under uncertainty. In summary, while linear programming assumes that all inputs are known and fixed, stochastic programming accounts for input uncertainty, making it more suitable for designing FCEV2G systems in dynamic and unpredictable environments [42].

### **1.4.2 Machine Learning Approaches**

Another approach to handling uncertainty in energy systems is the use of machine learning (ML) methods. ML algorithms can learn from historical data and identify underlying patterns between system variables. Once trained, these models can make decisions or predictions to achieve a desired outcome. Unlike traditional optimization methods, which require explicit mathematical formulations, ML models can operate without predefined equations, making them more flexible in complex and dynamic

environments[43]. In the context of FCEV2G, ML can be used to model time-dependent and uncertain variables, enabling adaptive and real-time decision-making in response to changing grid conditions.

ML can be divided into three main categories, which are supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, both the input and output data are labeled, allowing the model to learn the relationship between them. Once trained, the model can predict outputs for new inputs by applying the learned patterns. In contrast, unsupervised learning works with unlabeled input data, enabling the discovery of hidden structures or patterns within the dataset. The model groups the data into clusters based on shared features, which can then be used to classify new inputs according to the identified groupings [43].

An example application of machine learning in energy systems is to create accurate surrogate models that assist larger optimization algorithms. For example, a system was developed to optimize V2G revenue while minimizing battery aging for economic energy management of a plug-in hybrid vehicle [44]. They first trained a deep neural network to accurately predict battery degradation based on factors like charging rate and temperature. This predictive model was then used by a multi-objective optimization algorithm to find the best operational strategy that balanced profit and battery health. In this case, the machine learning model acts as a highly efficient predictor that informs the main decision-making algorithm, rather than acting as the decision-maker itself [44].

### **1.4.3 Reinforcement Learning for FCEV2G Decision Making**

Reinforcement learning (RL) is a form of machine learning in which an agent learns by interacting with its environment, making decisions that affect a cumulative reward over time. The agent's objective is to maximize this reward through real-time decision-making [43]. To tackle the challenge of uncertain real-time electricity prices, a model-free RL strategy for managing electric vehicle charging and discharging was proposed [45]. Their key finding was that the RL agent could learn an effective policy to intelligently manage V2G operations and reduce energy costs without needing to forecast future prices, which then validated with real-world data. This work demonstrates RL's ability to create adaptive control systems capable of responding to dynamic conditions such as fluctuating electricity prices and uncertain driver needs [45]. Focusing on home energy management systems (HEMS), reinforcement learning was used to intelligently schedule the interplay between solar power generation, stationary storage, and vehicle-to-home (V2H) functionalities [46]. Their key contribution was demonstrating that an RL agent could effectively handle the uncertainties of smart microgrid elements, such as variable solar power, to successfully reduce household power costs and smooth the energy load profile [46].

In the context of an FCEV2G system, the reward can be defined as the total profit generated from hourly operational decisions. Revenue may come from selling electricity to the grid using stored hydrogen in the station via V2G, while costs may include the purchase of electricity to produce hydrogen through an electrolyzer. By modeling the operational environment of the FCEV2G system, the RL agent can learn optimal actions through repeated interactions, gradually adjusting its strategy to maximize the cumulative reward. Consequently, RL facilitates the development of adaptive control policies that can make effective decisions in uncertain and dynamic conditions.

#### **1.4.4 Behavioral Cloning for FCEV2G Decision Making**

Supervised learning (SL) offers an alternative machine learning path where a model learns to predict outcomes by training on a labeled historical dataset. Unlike reinforcement learning's trial-and-error approach, SL learns a direct mapping from input features to a known output. An application of this is behavioral cloning. This is a technique where a model is trained to mimic the decisions of an expert system [47]. As a direct alternative to learning-from-scratch methods like RL, a behavioral cloning framework was implemented for EV charging [48]. They used an expert LP model to generate a dataset of optimal demonstrations and then trained a deep neural network with SL to mimic this expert behavior. Their results highlight the primary advantages of this approach, which was the training process was up to 220 times faster than RL. The final policy achieved significantly lower electricity costs and performed close to the theoretical lower bound. This work provides strong evidence that behavioral cloning can outperform RL in both training efficiency and final performance for complex energy scheduling tasks [48].

In the context of FCEV2G energy management, the expert is a sophisticated but computationally intensive optimization model that can calculate the ideal operational strategy for any given set of conditions. These conditions can be electricity price, time of day, stored hydrogen level. A dataset is created by running this expert model across thousands of scenarios to generate optimal state-action pairs. The behavioral cloning model then learns to map these system states to the expert's actions. The primary advantage of this approach is its computational efficiency. When it trained once, the model can replicate the expert's high-quality decisions in near-real-time. This model is exceptionally well-suited for fast-acting control systems.

### **1.5 Research Gap**

The preceding literature review demonstrates that machine learning has emerged as a powerful tool for optimizing the grid integration of electric vehicles. Significant progress has been made using both RL and SL approaches. Studies have successfully employed RL to manage V2G operations for profit maximization, grid services, and residential vehicle-to-home (V2H) applications [43], [45], [46]. More

advanced work has even extended these models to account for component degradation. In parallel, behavioral cloning has been validated as a highly effective SL technique for creating computationally efficient controllers that can outperform standard RL in both training speed and final performance for EV charging scheduling [47], [48].

However, despite this progress, a review of the literature reveals a significant and recurring gap: the overwhelming focus on battery electric vehicles (BEVs). This focus on BEVs overlooks the unique operational challenges and strategic complexities associated with FCEVs in a grid-integrated context. An FCEV2G system introduces multiple interconnected sub-systems, including an electrolyzer for on-site hydrogen production, a hydrogen storage tank, and the fuel cell stack itself. Consequently, the optimal control problem is not a simple charge/discharge decision. The need to simultaneously manage hydrogen inventory, schedule electrolysis during low-price periods, and execute V2G sales during high-price events makes the operational strategy significantly more complicated.

While behavioral cloning has been applied to BEVs, applying behavioral cloning to the operational control of an FCEV2G system remains largely unexplored in the literature. The insights from BEV-focused studies cannot be directly transferred due to the fundamentally different system architecture, decision variables, and the long-horizon, state-dependent strategies required.

Therefore, this thesis aims to bridge this critical research gap. It investigates the application of a behavioral cloning framework to develop a computationally efficient and high-performance operational controller specifically for the energy management of a hydrogen-based FCEV2G system. To the best of our knowledge, this represents a novel application of the behavioral cloning methodology to a complex system whose real-time operational strategy has not been adequately addressed in the existing body of work.

## **1.6 Chapter Summary**

This chapter provided an overview of hydrogen's role in clean energy systems, with emphasis on its production, storage, and application in fuel cell electric vehicles. The principles and degradation mechanisms of PEMFCs were discussed in detail, laying the foundation for understanding their limitations and opportunities in energy systems. The concept of V2G technology was introduced, demonstrating how electric vehicles, particularly FCEVs, can contribute to grid stability and economic efficiency. Operational strategies for managing FCEV2G systems were reviewed, including both mathematical optimization techniques and data-driven methods such as machine learning. This literature review establishes the conceptual foundation for developing intelligent control systems for FCEV2G operations in the chapters that follow. Crucially, this review also identified a significant research gap, which is the lack of studies

applying computationally efficient, expert-guided machine learning methods like behavioral cloning to the architecture of an FCEV2G system.

## Chapter 2 Rule-based FCEV2G Operational Simulation

### 2.1 Model Description

The primary motivation behind operating an FCEV2G station is to leverage the hourly variations in wholesale electricity prices, which fluctuate in response to real-time grid supply and demand. Typically, when electricity supply exceeds demand, prices tend to fall, whereas limited supply causes prices to rise. Previous studies have indicated that electricity prices often peak during similar periods as vehicle traffic rush hours [41]. Given the lower efficiency and increased fuel consumption of vehicles during congested traffic periods, we anticipate that participation in the FCEV2G system will be high at these times. This correlation between high traffic volumes and rising electricity prices offers a chance for FCEV2G stations to capitalize on these conditions and boost profitability.

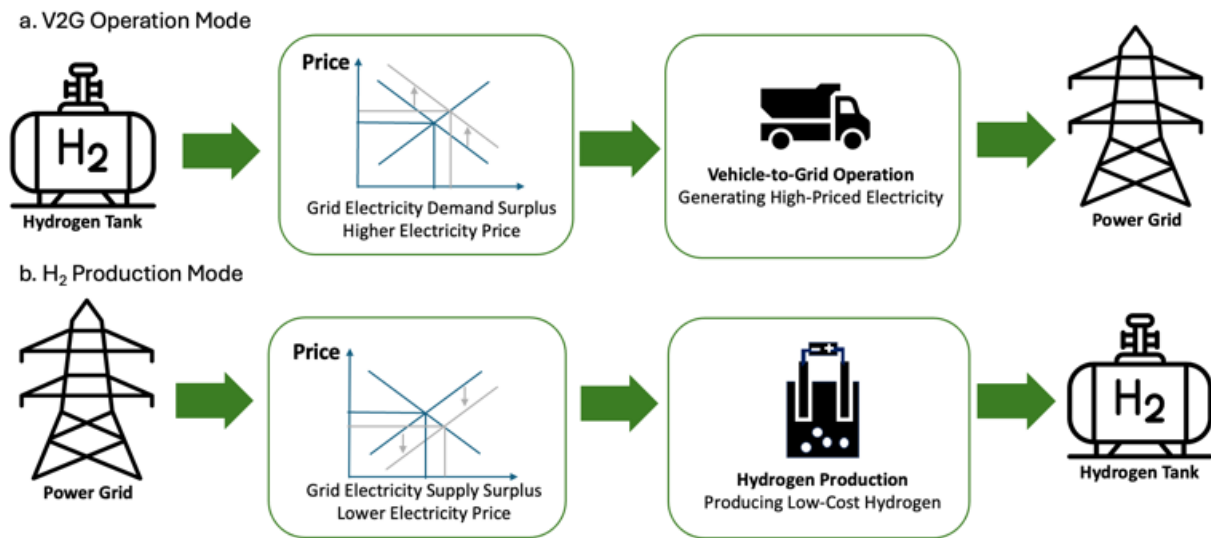


Figure 2.1 The Two Different Operation Modes of the FCEV2G Station: (a) V2G operation mode and (b) Hydrogen production mode

To summarize the problem, the FCEV2G station operates in two distinct modes, illustrated in Figure 2.1. The V2G operation mode generates revenue by supplying electricity to the grid during periods of high electricity prices. Conversely, the second mode, the hydrogen production mode, is activated when there is a surplus of electricity, and prices are low. Hydrogen used in the station can also be sourced externally from the green hydrogen market, where hydrogen is produced exclusively from renewable sources. While market hydrogen provides a reliable and consistent supply, it typically has a higher cost compared to hydrogen produced onsite using electrolyzers during periods of low-priced electricity. In this study, the FCEV2G problem is modeled in three key components: electricity production from V2G, the

hydrogen cycle, and the participation patterns of FCEVs. Each component is simulated using historical data to assess the long-term financial viability and profitability of the FCEV2G system.

## 2.2 Electricity Production from V2G

In this study, we account for the degradation of PEM fuel cells during transportation but assume degradation to be negligible during V2G operation due to stable power output conditions [49]. Therefore, the fuel cell's performance when participating in V2G can be considered as a snapshot representing its cumulative degradation from prior usage. To quantify this snapshot degradation, the mileage ( $X$ ) of each participating FCEV, representing the total distance it has traveled, is used. The mileage of participating vehicles is assumed to follow a normal distribution, allowing each FCEV to be assigned a randomly sampled mileage value. By using the average speed ( $V_{avg}$ ) of long-haul truck driving cycles, which is 76.56 km/h [50], we can relate the mileage directly to the utilization time ( $t$ ). The total mileage,  $X$  (in km) is thus calculated by multiplying the utilization time,  $t$  (in hours), by the average speed,  $V_{avg}$  (in km/h), as given in Equation 2.1.

$$X = t * V_{avg} \quad \text{Equation 2.1}$$

According to the U.S. Department of Energy's 2030 targets for 275 kW heavy-duty fuel cell trucks, the fuel cell's lifetime ( $t_1$ ), initial voltage at beginning-of-life ( $V_0$ ), and voltage at end-of-life ( $V_1$ ) are projected to be 25,000 hours, 0.779 V, and 0.70 V, respectively [51][52]. Using the fuel cell lifetime and the previously stated average vehicle speed, the maximum mileage is calculated as approximately 1,914,000 km, representing the upper boundary of the mileage distribution. By incorporating these parameters into Equation 2.2, an hourly voltage degradation rate ( $r_d$ ), expressed in volts per hour, can be determined. This degradation rate can then be used to estimate the voltage degradation corresponding to specific mileage values.

$$r_d = \frac{V_0 - V_1}{t_0 - t_1} = \frac{0.779 - 0.70}{0 - 25000} = -3.16 * 10^{-6} \frac{V}{h} \quad \text{Equation 2.2}$$

The polarization curve, or current-voltage (I-V) relationship, of the PEMFC was adapted from existing literature [53]. To represent the conditions of an FCEV truck, the active area ( $A$ ) of the fuel cell was scaled to approximately 330 cm<sup>2</sup> [52]. The electrical power output ( $P$ ), measured in watts, is calculated by multiplying the cell voltage ( $V$ ), in volts, by the current ( $I$ ), in amperes. Since current can also be described as the product of current density ( $i$ ), in amperes per square centimeter (A/cm<sup>2</sup>), and the active area ( $A$ ), in square centimeters, the power output is expressed mathematically as follows in Equation 2.3:

$$P = V * I = V * i * A \quad \text{Equation 2.3}$$

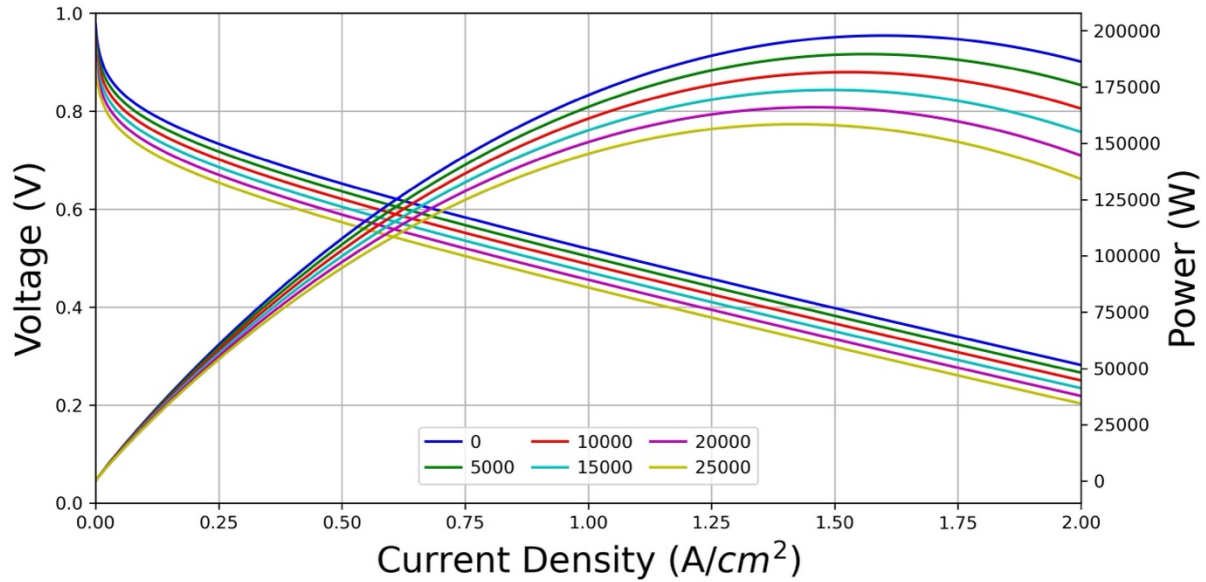


Figure 2.2 Polarization and Power curve with Increasing Utilization Time

For each FCEV participating in the V2G station, the procedure begins by assigning a randomly generated mileage value. Using this mileage, the vehicle’s utilization time is calculated from Equation 2.1. Subsequently, applying the degradation rate provided by Equation 2.2, the corresponding polarization curve of the fuel cell is determined. With this polarization curve, the electrical power curve is derived using Equation 2.3. Examples of the polarization curve and the power curve with different utilization time can be seen in Figure 2.2. Finally, the operational power value for the FCEV is extracted from this power curve based on hydrogen consumption, which is discussed further in the subsequent section. As a result, each participating FCEV is characterized by a unique mileage value, directly corresponding to a distinct operational power output.

### 2.3 Hydrogen Cycle

The hydrogen cycle is modeled to quantify the production, storage, and consumption of hydrogen within the FCEV2G system. When there is surplus electricity and consequently lower electricity prices, hydrogen is produced onsite using electrolyzers and subsequently stored in gaseous hydrogen tanks. The capacity of the electrolyzer and the maximum amount of hydrogen stored are optimization variables that depend heavily on the expected hydrogen consumption in FCEV2G stations.

The hydrogen consumption rate ( $\dot{n}$ ), expressed in moles per second, can be calculated using Faraday's Law (Equation 2.4) assuming a 100% Faraday efficiency, i.e., all reactants fully participate in the electrochemical reaction [38].

$$\dot{n} = \frac{I}{n * F} \quad \text{Equation 2.4}$$

It is important to note that the current produced depends directly on the selected power output from the power curve. Because each participating FCEV is assigned a different mileage, their hydrogen consumption rates will vary accordingly. Operating a PEMFC at its maximum power output typically leads to decreased efficiency. Instead, more efficient operating points exist at lower power outputs. To address this, a model was employed to optimize operation by maximizing the difference between generated power and hydrogen consumption. The operating point on the power curve is found from this optimization. This approach enhances overall efficiency by preventing excessive hydrogen usage for minimal additional power gains during V2G operation.

## 2.4 Participation of FCEV

In this study, we assumed that all 19 ONroute stations located along Highway 401 East and West are operated as FCEV2G stations. To determine the hourly truck traffic volume on this highway, we used the Government of Ontario's Traffic Volume data [54], which includes seasonal average annual daily traffic volumes of conventional trucks collected at various locations across Ontario's highways from 1988 to 2021. The dataset was filtered specifically for Highway 401 in the year 2021, as it represents the most recent data available. Traffic measurement points nearest to the ONroute stations were identified and selected for analysis, including Bainsville, Morrisburg, Mallorytown, Napanee, Trenton, Newcastle, Cambridge, Woodstock, Dutton, and Tilbury.

Table 2.1 Ontario’s Traffic Volume Data for Eastbound and Westbound ONroutes of 401 Highway [54]

Location	FCEV SADT	FCEV WADT	FCT OADT
Bainsville	597.18	372.96	464.30
Morrisburg	289.20	181.20	226.80
Mallorytown	388.02	242.73	302.84
Napanee	360.87	225.63	281.29
Trenton	399.00	249.48	312.68
Newcastle	366.12	286.74	322.80
Cambridge	1514.70	1263.60	1448.70
Woodstock	953.37	679.14	819.45
Dutton	297.48	247.53	281.82
Tilburg	282.45	253.20	268.05

Since the Government of Ontario’s data provides the total number of trucks on both sides of Highway 401 without specifying their directional distribution, we assumed an equal split: half of the trucks on the eastbound side and the other half on the westbound side. The only location that has an ONroute on one side is Bainsville, so the split wasn’t applied there. Given that there were no operational FCEVs on the highway in 2021, we simulated the presence of FCEVs by assuming a market penetration of 6% [55]. Additionally, the dataset presents average daily truck volumes that vary seasonally throughout the year, reflecting realistic traffic patterns. However, since the original data only distinguishes between summer and winter, we calculated averages for the remaining months to estimate traffic volumes. Incorporating this seasonal variability is crucial, as it allows our model to capture fluctuations in truck traffic and better represent real-world operational conditions throughout the year. This data is summarized in Table 2.1, where SADT refers to the Summer Average Daily Traffic, WADT represents the Winter Average Daily Traffic, and OADT denotes the Average Daily Traffic for the remaining seasons.

Table 2.2 Daily Congestion Score and Percentage by Time of Day [56]

Hour	Day						
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
0	10 (2.7 %)	19 (1.6%)	7 (1.3 %)	3 (0.4 %)	2 (0.4 %)	5 (0.8 %)	2 (0.6 %)
1	11 (3.0 %)	22 (1.8 %)	10 (1.8 %)	5 (0.7 %)	3 (0.6 %)	4 (0.7 %)	2 (0.6 %)
2	9 (2.4 %)	20 (1.7 %)	7 (1.3 %)	3 (0.4 %)	3 (0.6 %)	4 (0.7 %)	3 (1.0 %)
3	7 (1.9 %)	21 (1.8 %)	7 (1.3 %)	4 (0.6 %)	2 (0.4 %)	3 (0.5 %)	2 (0.6 %)
4	6 (1.6 %)	25 (2.1 %)	7 (1.3 %)	4 (0.6 %)	2 (0.4 %)	2 (0.3 %)	2 (0.6 %)
5	5 (1.4 %)	32 (2.7 %)	9 (1.7 %)	6 (0.8 %)	4 (0.8 %)	3 (0.5 %)	1 (0.3 %)
6	4 (1.1 %)	55 (4.6 %)	18 (3.3 %)	17 (2.4 %)	11 (2.1 %)	10 (1.7 %)	2 (0.6 %)
7	6 (1.6 %)	75 (6.3 %)	34 (6.3 %)	31 (4.3 %)	24 (4.7 %)	20 (3.4 %)	3 (1.0%)
8	7 (1.9 %)	87 (7.3 %)	44 (8.1 %)	38 (5.3 %)	29 (5.7 %)	28 (4.7 %)	4 (1.3 %)
9	11 (3.0 %)	76 (6.3 %)	34 (6.3 %)	29 (4.0 %)	25 (4.9 %)	27 (4.5 %)	9 (2.9 %)
10	18 (4.9 %)	65 (5.4 %)	26 (4.8 %)	26 (3.6 %)	21 (4.1 %)	29 (4.9 %)	18 (5.8 %)
11	24 (6.5 %)	58 (4.8 %)	24 (4.4 %)	28 (3.9 %)	24 (4.7 %)	32 (5.4 %)	27 (8.7 %)
12	29 (7.9 %)	53 (4.4 %)	26 (4.8 %)	27 (3.8 %)	26 (5.1 %)	39 (6.5 %)	31 (10.0 %)
13	29 (7.9 %)	52 (4.3 %)	26 (4.8 %)	32 (4.5 %)	29 (5.7 %)	40 (6.7 %)	29 (9.4 %)
14	28 (7.6 %)	64 (5.3 %)	31 (5.7 %)	40 (5.6 %)	39 (7.6 %)	48 (8.0 %)	30 (9.7 %)
15	28 (7.6 %)	82 (6.9 %)	48 (8.9 %)	58 (8.1 %)	53 (10.3 %)	64 (10.7 %)	29 (9.4 %)
16	26 (7.1 %)	97 (8.1 %)	55 (10.2 %)	67 (9.3 %)	58 (11.3 %)	70 (11.7 %)	26 (8.4 %)
17	24 (6.5 %)	89 (7.4 %)	48 (8.9 %)	70 (9.8 %)	53 (10.3 %)	61 (10.2 %)	21 (6.8 %)
18	20 (5.4 %)	68 (5.7 %)	29 (5.4 %)	60 (8.4 %)	37 (7.2 %)	41 (6.9 %)	18 (5.8 %)
19	17 (4.6 %)	46 (3.8 %)	13 (2.4 %)	40 (5.6 %)	22 (4.3 %)	26 (4.4 %)	13 (4.2 %)
20	15 (4.1 %)	30 (2.5 %)	7 (1.3 %)	30 (4.2 %)	15 (2.9 %)	15 (2.5 %)	8 (2.6 %)
21	10 (2.7 %)	24 (2.0 %)	8 (1.5 %)	31 (4.3 %)	12 (2.3 %)	10 (1.7 %)	8 (2.6 %)
22	10 (2.7 %)	21 (1.8 %)	12 (2.2 %)	39 (5.4 %)	11 (2.1 %)	9 (1.5 %)	11 (3.6 %)
23	14 (3.8 %)	16 (1.3 %)	11 (2.0 %)	29 (4.0 %)	8 (1.6 %)	7 (1.2 %)	10 (3.2 %)
Total	368	1197	541	717	513	597	309

The next step involved converting the average annual daily truck volumes into hourly estimates for each week hour by utilizing congestion scores from the Ministry of Transportation [56]. These scores reflect factors such as traffic jams, delays, travel speed, and jam length. By summing the daily congestion scores and calculating hourly percentages, we distributed the daily FCEV estimates across each hour. Hourly percentages can be seen in Table 2.2. Multiplying these percentages by the daily FCEV totals resulted in hourly truck counts. This method captures both seasonal trends and hourly traffic variations throughout the week, offering a more realistic representation of FCEV presence on the highway.

With the hourly FCEV truck counts estimated for each location near an ONroute station, the final step is to determine the number of FCEV trucks that will stop and participate in FCEV2G operations. Congestion scores are used as a proxy for traffic efficiency, where a higher score indicates worse traffic conditions. It is more rational for FCEV drivers to stop and idle when congestion is high and traffic is inefficient, which makes those vehicles available for V2G participation. This percentage of idling vehicles is assumed to reflect the number of FCEVs that can engage in V2G activities at the station. An example of the calculation of the FCEVs available on a summer Sunday at 10 a.m. in Morrisburg can be seen below.

$$\text{Participating FCEVs} = 289.20 * 0.049 * 0.18 = 2.55 (\sim 3)$$

## 2.5 Electricity Price Categorization

The hourly electricity price influences the operation modes, i.e., V2G operation mode and hydrogen production mode. Due to significant fluctuations in price, a categorization step is required to guide operational decisions. In this study, K-means clustering was used to classify electricity price data into categories. Before clustering, outliers were removed using the interquartile range (IQR) method, which defines lower and upper bounds based on the first and third quartiles. Values outside these bounds were identified as outliers. After clustering, these outliers were reintroduced into the lowest or highest price categories, depending on their values. To account for seasonal variation in electricity prices, clustering was performed separately for each season.

Table 2.3 Ontario’s Electricity Price Categories and Their Corresponding Potential Hydrogen Costs and V2G Revenues

Price Category	Hydrogen Production Cost (CAD/kg)	V2G Revenue (CAD/kg)
0	0.33	0.08
1	0.98	0.24
2	1.30	0.41
3	1.67	0.53
4	3.71	1.28

Table 2.3 presents the mean hydrogen production cost and V2G revenue values for Ontario in 2023. The hydrogen production cost is calculated for hours with electricity supply surplus, while V2G revenue is assessed during periods of excess demand. The hydrogen production cost is obtained by multiplying the electricity price by the energy required to produce 1 kg of hydrogen (43 kWh/kg) [57]. V2G revenue is

calculated by dividing the power generated by each truck, multiplied by the electricity price, by the corresponding hydrogen consumption in kilograms. These computations form the basis for the values reported in Table 2.3.

In this study, we implemented a simplified decision-making framework to maximize profit by utilizing electricity price categories. Based on the categorized data, operational decisions can be made for each hour. For instance, a straightforward strategy involves operating the electrolyzer when the price category equals 0 and engaging FCEVs in V2G activities when the price category is 2 or higher. As hydrogen production costs are evaluated only during hours of electricity surplus and V2G revenues during hours of excess demand, profitability requires that V2G revenue exceed hydrogen production costs. This logic serves as an illustrative example of FCEV2G station operation, further examined in the results section.

## **2.6 Base Case Assumptions**

The simulation assumes that each station is equipped with a maximum electrolyzer capacity of 1 MW and a hydrogen storage capacity of 10 MWh. The electrolyzer operates at 77% efficiency, which is the ultimate goal of electrical efficiency of a PEM electrolyzer, requiring 43 kWh of electricity to produce one kilogram of hydrogen [57]. The capital expenditure (CAPEX) for a 1 MW electrolyzer was taken from [58], estimated at 1381 CAD/kW. With a lifetime of 20 years, and a discount rate of 8% [41], the yearly CAPEX for a 1 MW electrolyzer is 140,910 CAD/year. The simulation begins with the hydrogen tank fully charged, under the assumption that the station operates on a yearly cycle and ends each year with a full storage level. The cost of market hydrogen, including production, processing, and delivery, is approximately 7 CAD/kg [59]. The 2023 electricity demand and supply data for Ontario were obtained from IESO [60]. Two types of demand data are available: Ontario demand, which reflects in-province consumption, and market demand, which includes both domestic consumption and electricity trade (exports and imports). Market demand was used in this study as it offers a more comprehensive representation of overall electricity usage. In contrast, only one supply dataset is available, representing total electricity generation in Ontario. Analysis of the 2023 data shows that of the 8760 hours in the year, 7110 hours (81%) experienced excess demand, while 1650 hours (19%) had a supply surplus.

## **2.7 Operation Simulation**

### **2.7.1 V2G Operation Mode**

The simulation operates in two distinct modes, each requiring three conditions to be met. The V2G operation mode activates during price categories 2, 3, or 4, when there is excess demand and adequate stored hydrogen. Conversely, the hydrogen production mode initiates during price category 0, under

conditions of supply surplus, and when storage capacity is available. If any of the three required conditions for either mode is not satisfied, the system remains inactive for that hour.

The simulation begins by evaluating whether the conditions for either operation mode are satisfied for a given hour. This assessment uses electricity data, including supply surplus status and the electricity price ( $p_t$ ). If the conditions for the V2G operation mode are met, random power values ( $P_n$ ) and corresponding current values ( $I_n$ ), based on FCEV lifetime degradation, are assigned to the participating vehicles. The potential hourly revenue (Rev) and hydrogen consumption rate ( $\dot{n}$ ) are calculated from the assigned values by using the equations (Equation 2.5, Equation 2.6, Equation 2.7, and Equation 2.8) below.

$$P_{total} (MW) = \sum_{n=1}^n P_n \quad \text{Equation 2.5}$$

$$I_{total} (A) = \sum_{n=1}^n I_n \quad \text{Equation 2.6}$$

$$\text{Rev} \left( \frac{CAD}{h} \right) = P_{total} (MW) * p_t \left( \frac{CAD}{MWh} \right) \quad \text{Equation 2.7}$$

$$\dot{n}_{consumed} \left( \frac{kg}{h} \right) = \frac{I_{total} * 2.016 \frac{g}{mol} * 3600 \frac{s}{h}}{n F * 1000 \frac{g}{kg}} \quad \text{Equation 2.8}$$

The next step is to determine how to supply the required hydrogen. First, the model checks whether using market hydrogen is profitable. If the potential revenue from V2G exceeds the cost of market hydrogen, then market hydrogen is used. The corresponding equations, which are Equation 2.9, and Equation 2.10, shown below.

$$\text{Cost} \left( \frac{CAD}{h} \right) = \dot{n}_{consumed, MH_2} \left( \frac{kg}{h} \right) * 7 \left( \frac{CAD}{kg} \right) \quad \text{Equation 2.9}$$

$$\text{Profit} \left( \frac{CAD}{h} \right) = \text{Rev} \left( \frac{CAD}{h} \right) - \text{Cost} \left( \frac{CAD}{h} \right) \quad \text{Equation 2.10}$$

When the current stored hydrogen is below 10% of maximum capacity and market hydrogen is not profitable, V2G operation is skipped for that hour, resulting in zero profit.

If utilizing market hydrogen is not profitable and sufficient stored hydrogen is available, the system will use stored hydrogen to support V2G operation. The profit calculation will simplify to Equation 2.11.

$$\text{Profit} \left( \frac{CAD}{h} \right) = \text{Rev} \left( \frac{CAD}{h} \right) \quad \text{Equation 2.11}$$

## 2.7.2 Hydrogen Production Mode

The first condition for activating the hydrogen production mode is that the current storage must be under 90% of the maximum capacity. If it is currently higher than the threshold, hydrogen production is halted, and no profit is generated for that hour. When all conditions for hydrogen production are met, the electrolyzer operates using the available electricity supply surplus. However, its power consumption is limited by the maximum electrolyzer capacity, so any surplus beyond this cap cannot be utilized. The equations (Equation 2.12, and Equation 2.13) for this process are presented below.

$$\dot{n}_{produced} \left( \frac{kg}{h} \right) = \frac{\min(\text{Supply Surplus (MW)}, 1 \text{ MW} * 19 \text{ Stations})}{43 \frac{kWh}{kg} * \frac{1}{1000} \frac{MW}{kW}} \quad \text{Equation 2.12}$$

$$\text{Profit} \left( \frac{CAD}{h} \right) = -\min(\text{Supply Surplus (MW)}, 1 \text{ MW} * 19 \text{ Stations}) * p_t \left( \frac{CAD}{MWh} \right) \quad \text{Equation 2.13}$$

## 2.8 Chapter Summary

This chapter presented the methodology used to model and simulate an FCEV2G station. The system model incorporated PEMFC behavior, with electricity output and hydrogen consumption dependent on truck power curves and varying lifetimes. Hydrogen supply was modeled through both market purchases and on-site production, constrained by the efficiency of the electrolyzer and storage limits. To estimate FCEV participation, traffic and congestion data from Ontario highways were used to derive hourly truck availability near ONroute stations, incorporating assumptions about market penetration and driver behavior. Electricity price volatility was addressed by applying K-means clustering to classify prices into categories, with outliers handled using the interquartile range method, and seasonal patterns preserved through separate seasonal clustering. Base case assumptions, such as electrolyzer capacity and hydrogen costs, were defined to ground the simulation. The simulation itself was divided into two operation modes: V2G, activated during high-price, demand-excess hours, and hydrogen production, activated during low-price, supply-surplus hours.

## Chapter 3 Rule-based FCEV2G Operation

### 3.1 Performance Summary

The performance of the rule-based FCEV2G operational model was evaluated to establish a baseline for its effectiveness and limitations. The analysis focused on key quantitative outcomes, including the overall profit distribution, the frequency of each operational mode, and the average hourly energy dynamics.

Table 3.1 Profit Distribution of the FCEV2G Station

Price Category	Hydrogen Production Cost (CAD)	V2G Revenue with Stored H <sub>2</sub> (CAD)	V2G Revenue with Market H <sub>2</sub> (CAD)
0	-35,992	0	0
1	0	0	0
2	0	32,986	0
3	0	13,444	0
4	0	1,474	3,546
Total Profit (CAD)	15,458		

Table 3.1 presents a breakdown of the system's profit by the price categories, which consists of revenue from V2G operations, the cost of on-site hydrogen production. V2G revenue is generated using hydrogen from two potential sources: low-cost hydrogen produced and stored on-site, or hydrogen purchased from the external market. If the CAPEX for the components of the FCEV2G station is not considered, the FCEV2G station demonstrates profitability for the 2023 period in Ontario. However, factoring in the CAPEX of even an electrolyzer is sufficient to make the station unprofitable. One thing to note is that the system only used market hydrogen instead of the produced hydrogen on a single occasion throughout the year. This decision was profitable, as the cost of using market hydrogen at that moment was lower than the average V2G revenue achievable within the highest electricity price category.

Table 3.2 Hours of Operation Modes of the FCEV2G Station

Price Category	Hydrogen Production Hours (h)		V2G Operation Hours (h)	
	Insufficient Storage	Sufficient Storage	Insufficient Hydrogen	Sufficient Hydrogen
0	8	264	-	-
1	-	-	-	-
2	-	-	2,784	150
3	-	-	1,523	46
4	-	-	246	8
Total	8	264	4,553	204

Table 3.2 provides a detailed breakdown of the operational hours for the FCEV2G station, categorized by the operation modes. For the hydrogen production mode, the hours are classified into two conditions: Insufficient Storage, representing periods when the electrolyzer could not operate because the hydrogen tank was at maximum capacity, and Sufficient Storage, indicating that capacity was available. Similarly, for the V2G operation mode, the hours are divided into Insufficient Hydrogen for periods when V2G was not possible due to a lack of available hydrogen, and Sufficient Hydrogen, indicating that the necessary fuel was on hand to participate in V2G.

Out of a total of 8,760 hours in 2023, 5,029 hours were identified as potential operating hours, comprising 272 hours for hydrogen production and 4,757 hours for V2G sales. Despite this potential, the station's actual operational time was limited to 468 hours. An analysis of the unutilized hours reveals a critical imbalance in the hydrogen cycle. The hydrogen production system demonstrated high utilization, operating for 264 of its 272 potential hours (97%). Conversely, the V2G system was severely constrained, operating for only 204 of its 4,757 potential hours (4.3%). This underperformance was almost entirely due to hydrogen shortages, which accounted for 4,553 missed V2G hours. In total, 4,561 potential operating hours were lost to system idling, driven overwhelmingly by this production-demand mismatch. These results demonstrate that the system's hydrogen demand for V2G significantly outpaces its production capacity, making storage sizing and replenishment strategies the central challenge for financial viability.

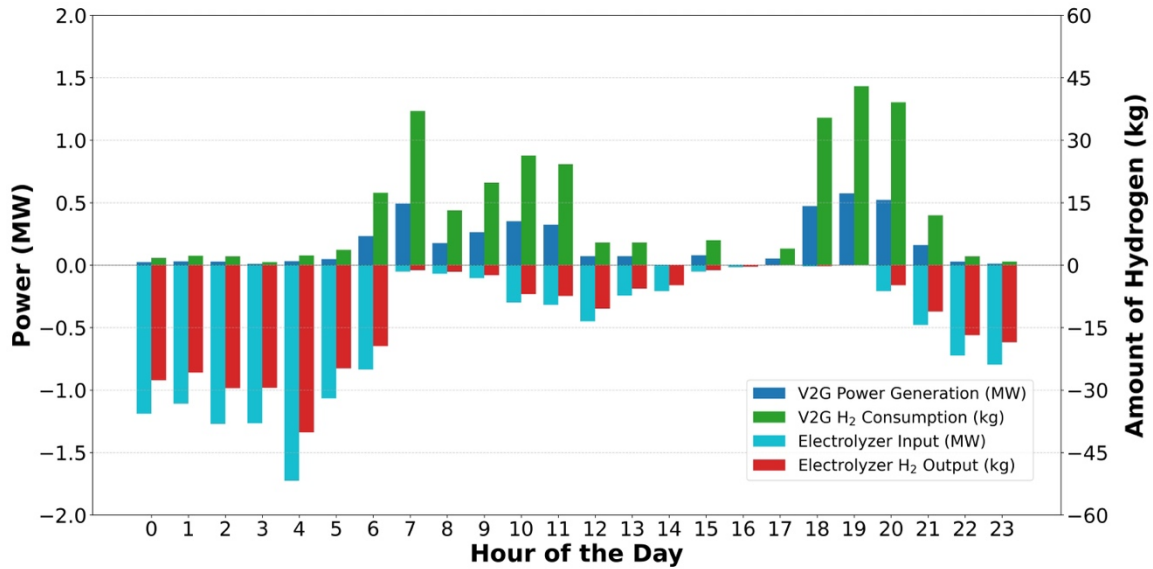


Figure 3.1 The Hourly Average in 24 hours V2G Power Generation, and Hydrogen Consumption; and Electrolyzer Power Input, and Hydrogen Production

Figure 3.1 presents the 24-hour operational profile of the station, averaged over its active hours. The plotted values represent the mean hourly V2G power output, electrolyzer power consumption, and hydrogen usage. The data used for these averages is derived from the 264 hours of hydrogen production and 204 hours of V2G operation that occurred when the system was not limited by physical constraints (i.e., full storage or insufficient hydrogen).

### 3.2 Key Factors Limiting Performance

There are three factors contributing to these results, the first being the difference in energy conversion efficiency. The primary factor constraining the system's profitability is the low round-trip efficiency of the hydrogen energy cycle. A significant energy discrepancy exists between hydrogen production and its subsequent use for power generation. Specifically, the on-site electrolyzer consumes 43 kWh of electricity to produce 1 kg of hydrogen [57], while the FCEV system, in V2G mode, generates only 13.34 kWh of electricity from that same kilogram. This comparison mixes the ultimate target efficiency for PEM electrolyzers with recent measured efficiencies for PEM fuel cells, so Figure 3.1 should be viewed as an indicative estimate rather than a like-for-like measurement. The calculation of the efficiency of V2G mode can be found in Section 2.3. Figure 3.1 represents a round-trip efficiency of approximately 31%, indicating that nearly 70% of the initial energy input is lost in the conversion process. This inherent inefficiency means that the revenue generated from selling power will almost always be less than the cost

of the electricity used to create the hydrogen, unless the V2G sale occurs at a price that is more than three times the price of the electricity used for production. As the hydrogen produced must support V2G operations over multiple hours, this efficiency gap creates a fundamental economic hurdle that is difficult to overcome.

The second limiting factor, directly linked to the first, is the simulation's lack of an optimized dispatch strategy. As established in Table 3.1 and Table 3.2, the high cost of hydrogen production necessitates a highly selective V2G strategy that exclusively targets the peak-price periods to overcome the low round-trip efficiency. However, the current operational model does not incorporate such a priority system. It engages in V2G operations whenever hydrogen is available and prices are favorable, without distinguishing between moderately profitable and highly profitable opportunities. This non-optimized approach leads to the depletion of scarce and valuable hydrogen on lower-margin sales, thereby omitting the chance to maximize revenue during true price spikes. This highlights the critical need for an intelligent control policy to improve profitability.

Finally, the system's inability to overcome its on-site production deficit is due to the prohibitive cost of external market hydrogen. While market hydrogen could theoretically supplement supply and enable greater V2G participation, its price makes it economically unviable for this purpose. For a market-based V2G operation to be profitable, the revenue generated (in CAD per kg of H<sub>2</sub>) must exceed the market's asking price of 7 CAD/kg. However, the simulation shows that even under the most favorable electricity prices in the base case, the maximum average revenue generated from one kilogram of hydrogen is only 1.28 CAD. This five-fold gap between cost and potential revenue confirms that, at current market rates, purchasing hydrogen for electricity sales is a consistently loss-making proposition. Consequently, a significant reduction in green hydrogen market prices is a prerequisite for it to become a viable component of the FCEV2G operational model.

### **3.3 Chapter Summary**

This chapter evaluated the performance of a rule-based FCEV2G operational model to establish a performance baseline. The analysis of the system's profitability, operational mode frequency, and hourly energy flows revealed a significant gap between potential and realized profit. The subsequent discussion identified three fundamental factors that constrain the system's economic viability. The primary limitation is the low round-trip efficiency (~31%) of the hydrogen production and power generation cycle, which creates a substantial economic hurdle. This is compounded by the lack of an optimized dispatch strategy, causing the system to deplete its valuable stored hydrogen on moderately profitable opportunities instead

of conserving it for true price spikes. Finally, the prohibitively high cost of external market hydrogen makes it an economically unviable option for supplementing V2G operations. Together, these factors demonstrate that a simple, rule-based approach is insufficient and highlight the critical need for an advanced, intelligent control policy to navigate these complex trade-offs and unlock the station's full profitability.

# Chapter 4 Supervised Learning Model for FCEV2G Operation

## 4.1 Model Description

The rule-based system detailed in the previous chapter, while providing a functional baseline, is fundamentally limited by its static logic. To develop a more adaptive and high-performance operational controller, this chapter introduces a Supervised Learning (SL) approach. We employ Behavioral Cloning (BC), which is a technique where a deep neural network is trained to mimic the sophisticated operational strategy of an expert MILP optimizer. The core advantage of this framework is its ability to distill the optimizer's computationally intensive, forward-looking plan into a rapid and reactive policy. This trained agent can make near-optimal decisions in milliseconds, learning to navigate complex market conditions and respect physical constraints with a level of adaptability unattainable by fixed-rule systems.

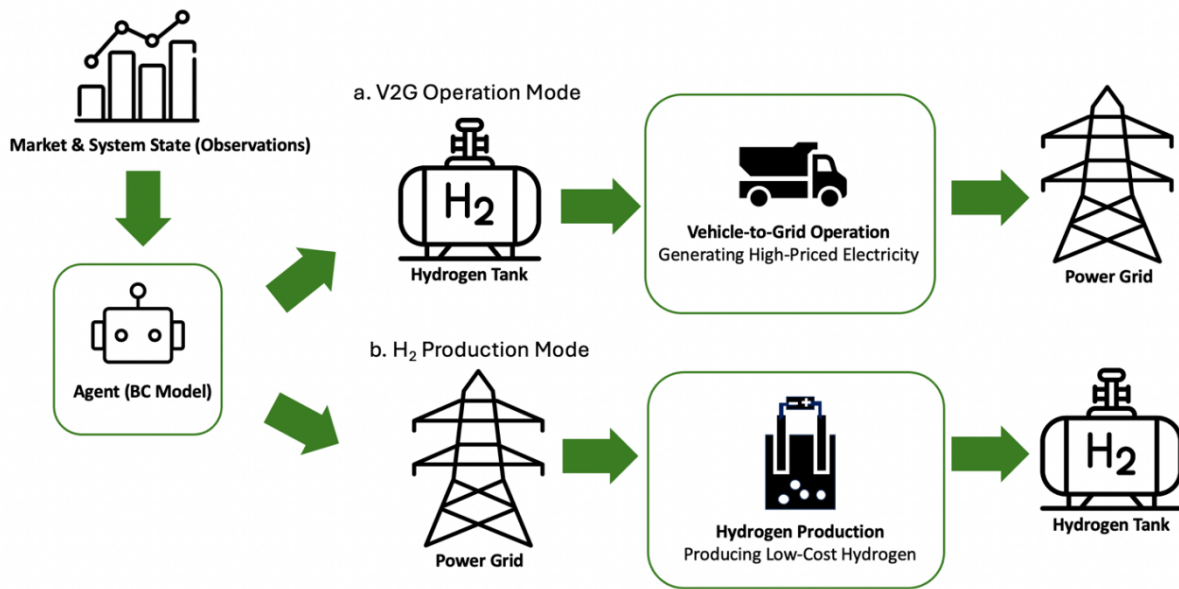


Figure 4.1 Schematic of Two Different Operation Modes of the FCEV2G Station, a. V2G operation mode and b. H<sub>2</sub> production mode, decided by the Behavioral Cloning Agent

In the behavioral cloning model, this study proposes a framework where an autonomous agent makes dynamic operational decisions for the FCEV2G station, as illustrated in Figure 4.1. The agent's decision-making process begins by assessing market and system state observations, which include data such as real-time electricity prices, current hydrogen storage levels, and vehicle availability. Based on these inputs, the agent selects one of two primary operational modes. First, it can initiate the V2G operation

mode, where it supplies hydrogen to FCEVs to gain high-priced electricity to the grid for revenue generation. Alternatively, it can initiate the hydrogen production mode, which uses periods of low electricity prices to produce and store low-cost hydrogen for future use. In addition to the two primary modes, there is also an idling mode for the system, when the agent decides to do nothing. This approach replaces the fixed rules simulation presented in the previous chapter with a dynamic policy. The aim is to maximize long-term profitability by intelligently navigating the trade-offs between immediate revenue and strategic energy storage.

## 4.2 Participation of FCEV

Following the methodology outlined in [41], a traffic model is incorporated that models FCEV participation using Alberta traffic data. The 2022 hourly traffic counts were sourced from a monitoring location near Calgary, along Alberta Provincial Highway No. 2. The number of participating FCEVs was estimated by multiplying the market penetration rate of FCEVs among all trucks, the willingness of drivers to participate in V2G, and the proportion of trucks within total vehicle traffic. To reflect the stochastic nature of traffic flow, a Poisson distribution was applied to these estimates, introducing variability into the model.

The electricity data used in this study spans from 2015 to 2025, while the traffic data is only available for 2022. To estimate FCEV participation for the remaining years, traffic predictions were made using patterns observed in 2022. Specifically, an SL approach modeled the total number of trucks based on temporal features: hour, month, and season. The data was divided into training (80%) and testing (20%) sets, and an XGBoost regression model with a Poisson objective was trained to predict truck counts. The model utilized 200 estimators, a learning rate of 0.1, and a maximum tree depth of 3. Model performance was evaluated using mean absolute error (MAE) on the test set, where the best parameters were selected by finding the minimum MAE. Once trained, it was applied to extrapolate truck volumes for the remaining years based on their corresponding hour, month, and season. Figure 4.2 shows the actual and forecasted hourly average of participating FCEVs in Alberta for the corresponding years.

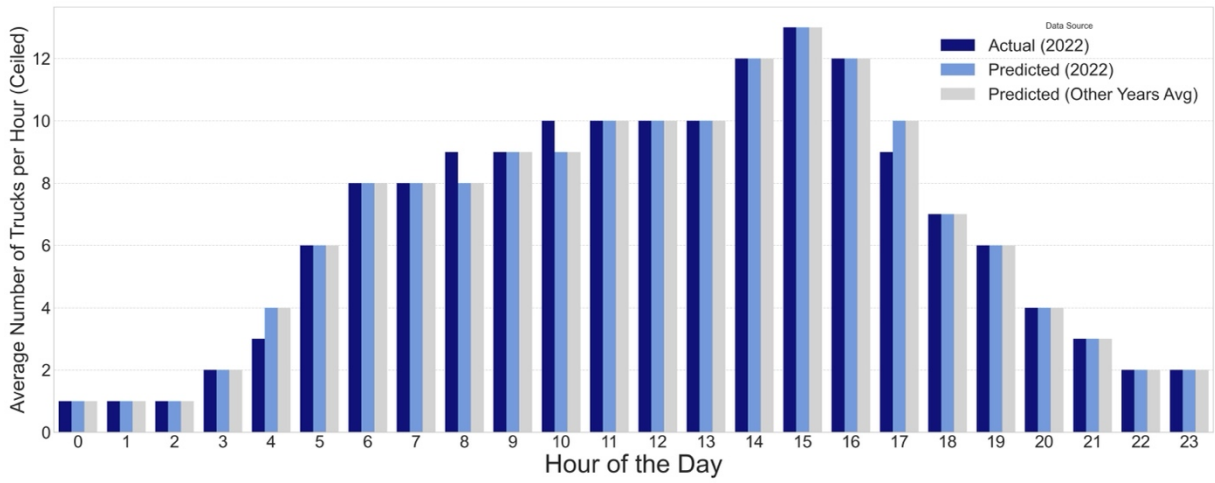


Figure 4.2 The Average Truck Count of Actual and Predicted Data

## 4.3 Supervised Learning Model for FCEV2G

### 4.3.1 Model Action and Observation Spaces

The action space defines the set of possible outputs from our policy at each time step, which can be seen in Table 4.1. For this FCEV2G problem, we designed a two-dimensional continuous action space. The primary decision ( $D(t)$ ) variable, which ranges from -1.0 to 1.0, controls the station's main operational mode. A positive value initiates V2G operation mode. This continuous value is then translated into a discrete number of trucks by scaling it to the maximum fleet size and rounding up to the nearest integer. A negative value activates the hydrogen production mode. Its magnitude represents the fraction of the electrolyzer's maximum capacity to use. Finally, if the decision value is zero, none of the operational modes are selected. This decision causes the station to be in an idling mode. The secondary alpha ( $\alpha(t)$ ) variable, ranging from 0.0 to 1.0, is used exclusively during V2G operation mode to manage the hydrogen source. It determines the proportion of hydrogen demand to be met by purchasing from the market, with the rest being drawn from the station's internal storage.

Table 4.1 SL Action Space

Variable Name	Lower Bound	Upper Bound	Unit	Type	Description
$D(t)$	-1	1	-	Float	Operational Mode Selection
$\alpha(t)$	0	1	MW	Float	Proportion of Market Hydrogen Usage

To make effective decisions, our SL model receives a set of observations that describe the system at each hour as shown in Table 4.2. These features are selected to provide the model with a comprehensive understanding of the current situation and outlook. We can categorize these features into direct and indirect observations. Direct observations are variables used to calculate the station's profit, such as the current electricity price,  $p(t)$ . Indirect features provide valuable context that enables the model to make more informed choices, e.g., time-based indicators such as the season  $S(t)$ , and forecasted prices. The agent learns to anticipate future conditions by including these forecasting features, which enables it to make more informed decisions about the present. The methodology used to generate these price forecasts is detailed in the following section. Observations are included to help the agent generalize and identify patterns across different operating conditions.

Categorical features are converted into a binary format using one-hot encoding. This creates a separate feature for each category (e.g., 'winter', 'spring') that is either active (1) or inactive (0). This prevents the model from assuming a false numerical hierarchy between the categories. For example, without this technique, the model might incorrectly learn that hour 23 is mathematically greater than hour 1. By giving each hour its own distinct on/off feature, we ensure the model treats each hour, e.g., '23:00' and '01:00' as an independent category. Additionally, the electricity price categories were based on the Alberta electricity price data, considering different seasons. The same clustering methodology for price categories described in Section 2.5 is used.

Table 4.2 SL Observation Space

Variable Name	One-hot Encoded	Lower Bound	Upper Bound	Unit	Data Type	Description
$p(t)$	No	0	1050	$\frac{CAD}{MWh}$	Float	Electricity Price at $t=t$
$p(t + 1)$	No	0	1050	$\frac{CAD}{MWh}$	Float	Electricity Price at $t=t+1$
$p(t + 2)$	No	0	1050	$\frac{CAD}{MWh}$	Float	Electricity Price at $t=t+2$
$n(t)$	No	0	8	-	Float	Total FCEVs in the Station
$H(t)$	No	0	10	MWh	Float	Amount of Stored Grid Hydrogen
$p_{min}^{24h}$	No	0	1050	$\frac{CAD}{MWh}$	Float	Min Electricity Price in the span of 24h
$p_{max}^{24h}$	No	0	1050	$\frac{CAD}{MWh}$	Float	Max Electricity Price in the span of 24h
$p_{max}^{72h}$	No	0	1050	$\frac{CAD}{MWh}$	Float	Max Electricity Price in the span of 72h
$S(t)$	Yes	0	1	-	Binary	Season of the year (encoded as 4 binary features)
$h(t)$	Yes	0	1	-	Binary	Hour of day (encoded as 24 binary features)
$d(t)$	Yes	0	1	-	Binary	Day of week (encoded as 7 binary features)
$c(t)$	Yes	0	1	-	Binary	Price category (encoded as 5 binary features)

Forecasted data is included in the observation space of the SL model to enable forward-looking decision-making. The Alberta Electric System Operator (AESO) provides Application Programming Interfaces (APIs) that grant real-time access to electricity data, including forecasted system marginal prices, pool prices, and supply surplus levels for the next two hours [61]. These forecasted price values are integrated into the observation space as  $p(t + 1)$  and  $p(t + 2)$ .

To provide the agent with a longer-term outlook, we engineered three additional hybrid forecast features: the minimum price over the next 24 hours ( $p_{min}^{24h}$ ), the maximum price over the next 24 hours ( $p_{max}^{24h}$ ), and the maximum price over the next 72 hours ( $p_{max}^{72h}$ ). For historic data points, these features are populated with the actual values. For current data where the future is unknown, these three features are populated by predictions from a set of three LightGBM forecasting models. These models were trained on historical time-series features, such as rolling price averages and seasonal indicators. This data population approach ensures the agent is trained on a perfect signal while getting evaluated with a forecasted signal.

The inclusion of these forecast features is critical for overcoming the inherent myopia of an agent. The short-term and long-term features allow the agent to learn complex strategies, such as conserving stored hydrogen during moderately high prices in anticipation of a higher price event several days in the future. By providing a forward-looking context, these features empower the SL agent to replicate the strategies of the perfect-foresight optimizer. This leads to substantially improved performance in the long run.

### 4.3.2 Optimization Data as the Expert Dataset

To train our supervised learning agent, a dataset representing an optimal operational strategy is required. We generated this expert dataset using a rolling horizon optimization solved by the Gurobi optimizer. This approach was chosen to create an expert whose strategy is both profit-maximizing and realistically learnable by a real-time agent. A deployed agent would not have perfect foresight over many years; its decisions must be based on a limited forecast window. Therefore, we constrained our expert to a similar worldview, forcing it to produce a dynamic, moment-to-moment policy that our agent could effectively imitate.

The system state was simulated hour-by-hour, and at the start of each day, the MILP model was solved for a forward-looking planning horizon of seven days. From this 7-day optimal plan, only the decisions for the first action horizon of 24 hours were recorded and implemented. The simulation then advanced 24 hours to the next day, where the process was repeated with updated information. This rolling approach is crucial because it forces the expert to generate a strategy that is optimal given a realistic lookahead, which makes its behavior much more suitable for our supervised learning agent to imitate.

We integrated a lookahead mechanism to prevent myopic behavior near the end of each 7-day planning window. Before each optimization, the system scanned the next 30 days immediately following the current planning horizon. If a high price event (i.e., a Category 4 price) was detected in this distant future, the terminal value of hydrogen in the optimizer's objective function was significantly increased. This incentivized the expert to conserve hydrogen during its current 7-day plan in anticipation of the future high-profit opportunity, leading to more strategically patient behavior.

#### 4.3.2.1 Decision Variables

The decision variables for the mentioned MILP approach are detailed below at Table 4.3.

Table 4.3 MILP Decision Variables

Variable Name	Lower Bound	Upper Bound	Unit	Data Type	Description
$P_{G2E}(t)$	0	1	MW	Float	Electrolyzer Input Power
$use_{truck}(n, t)$	0	1	-	Binary	Truck Participation Indicator
$P_{E2HT}(t)$	0	inf	MW	Float	Electrolyzer to Storage
$P_{HT2FC}(t)$	0	inf	MW	Float	Storage to FCEV
$P_{MH2FC}(t)$	0	inf	MW	Float	Market Hydrogen to FCEV
$H(t)$	0	10	MWh	Float	Amount of Stored Grid Hydrogen
$P_{FC2G}(t)$	0	inf	MW	Float	FCEV to Grid

#### 4.3.2.2 Constraints

To set up the optimization framework, we begin by defining constraints that reflect the operational characteristics of the FCEV2G system. One key constraint is on the total power output from participating FCEVs ( $P_{FC2G}(t)$ ), which must equal the sum of the power generated by each selected truck, which can be seen in Equation 4.1. Because each truck's power output ( $P_{FC}(t)$ ) is based on a randomly assigned lifetime, the optimization model determines the best subset of FCEVs to use in each hour to maximize profit within  $n$  FCEVs. This selection is handled through binary decision variables that control which trucks are selected at each time ( $t$ ) step. Additionally, sum of the selected FCEV number can't be higher than total FCEVs in the station (Equation 4.2).

$$P_{FC2G}(t) = \sum_{n=0}^n use_{truck}(n, t) * P_{FC}(t) \quad \text{Equation 4.1}$$

$$n(t) \geq \sum_{n=0}^n use_{truck}(n, t) \quad \text{Equation 4.2}$$

The hydrogen consumption of each FCEV ( $\dot{n}_{H_2}$ ) depends on the IV curve of the fuel cells in the participating truck at a randomly assigned lifetime. This consumption must be fully met using either stored grid hydrogen ( $P_{HT2FC}(t)$ ) or market hydrogen ( $P_{MH2FC}(t)$ ). This equality constrain is shown in Equation 4.3. The model applies the same selection mechanism to determine how this demand is satisfied at each hour.

$$P_{HT2FC}(t) + P_{MH2FC}(t) = \sum_{n=0}^n use_{truck}(n, t) * \dot{n}_{H_2}(t) \quad \text{Equation 4.3}$$

Another key constraint, Equation 4.4, is that as the efficiency of the electrolyzer is fixed, the amount of hydrogen produced and stored in the hydrogen tank ( $P_{E2HT}(t)$ ) will be equal to the product of the efficiency ( $\eta_{el}$ ) and power input of electrolyzer ( $P_{G2E}(t)$ ).

$$P_{E2HT}(t) = P_{G2E}(t) * \eta_{el} \quad \text{Equation 4.4}$$

The storage needs to be updated at each time step. At  $t = 0$ , the hydrogen storage ( $H(t)$ ) is initialized at its full capacity ( $H_{max}$ ). For each subsequent hour, the storage level is updated by adding the hydrogen produced from the electrolyzer ( $P_{E2HT}(t)$ ) and subtracting the hydrogen consumed during V2G operation ( $P_{HT2FC}(t)$ ). Storage updates can be seen in Equation 4.5, and Equation 4.6. This ensures that the model captures the dynamic behavior of hydrogen availability over time. An additional constraint (Equation 4.7) ensures that the hydrogen stored in the tank does not exceed its maximum capacity at any hour.

$$H(t) = H_{max} + P_{E2HT}(t) * 1h - P_{HT2FC}(t) * 1h \quad \text{where } t = 0 \quad \text{Equation 4.5}$$

$$H(t) = H(t - 1) + P_{E2HT}(t) * 1h - P_{HT2FC}(t) * 1h \quad \text{where } t > 0 \quad \text{Equation 4.6}$$

$$H_{max} \geq H(t) \quad \text{Equation 4.7}$$

### 4.3.2.3 Objective Function

The primary goal of the optimization model is to maximize the total profit ( $\Pi(t)$ ) over the entire planning horizon. This objective function is composed of three main financial components: revenue from the V2G operation mode ( $R_{V2G}(t)$ ), the cost of the hydrogen production mode ( $C_{el}(t)$ ), and the cost of

purchasing market hydrogen ( $C_{MH2}(t)$ ). To accurately incorporate the fixed market hydrogen cost ( $c_{MH2}$ ), which is typically priced per kilogram, it is converted to a per-MWh equivalent using the lower heating value of hydrogen. The main financial components are calculated in Equation 4.8, Equation 4.9, and Equation 4.10, respectively, following the order stated earlier.

Furthermore, to incentivize long-term strategic behavior, a terminal storage value ( $V_H(t + 168)$ ) is included in the objective function. This component adds a reward ( $p_H(t + 168)$ ) for each unit of hydrogen remaining in the tank at the end of the 7-day planning horizon ( $H(t + 168)$ ). The value of this reward is dynamically adjusted by a mechanism to encourage the expert to conserve hydrogen when significant future profit opportunities are identified. Specifically, the system performs a long-term check of a 30-day lookahead window that begins after the 7-day planning horizon ends. If a very high price event is detected within this future window, the terminal value is set to a high figure, which is the average price of the highest category. If no such opportunity is found, the terminal value is set to a more conservative figure, which is a price between the averages of third and fourth highest categories. The final objective function can be seen in Equation 4.11.

$$R_{V2G}(t) = P_{FC2G}(t) * p(t) \quad \text{Equation 4.8}$$

$$C_{el}(t) = P_{G2E}(t) * p(t) \quad \text{Equation 4.9}$$

$$C_{MH2}(t) = c_{MH2} * P_{MH2FC}(t) \quad \text{Equation 4.10}$$

$$\Pi(t) + V_H(t + 168) = R_{V2G}(t) - C_{el}(t) - C_{MH2}(t) + H(t + 168) * p_H(t + 168) \quad \text{Equation 4.11}$$

### 4.3.3 Simulation and Evaluation Setup

To evaluate the performance of the trained supervised learning policy, we utilize an FCEV2G simulation environment. This simulation models the FCEV2G station's operations on an hour-by-hour basis. While the policy is trained offline on a static dataset of expert decisions, this simulation environment serves two critical roles in the overall process. First, during data preparation, it is used to generate a realistic sequence of state observations that form the basis of our training data. Second, after training, it is used to evaluate the final policy by executing its decisions and calculating the resulting profit or loss based on real market conditions.

The FCEV2G simulation operates on an hourly time step, where each step corresponds to one hour in the historical dataset. To evaluate the agent's performance, the dataset is divided into training and testing data. Each phase can use a different time range to assess how well the agent reacts to unseen data. In this

study, electricity and traffic data from Alberta between 2015 and 2025 are available. The training data consists of the start of 2015 to the end of 2024, whereas the testing data started from 2025 to July 7<sup>th</sup>, 2025. However, the test set's end date is not fixed. It extends to the most recent hour for which data was available from AESO at the time of the experiments, which creates a realistic deployment scenario.

The constants of the FCEV2G simulation environment and MILP can be seen in Table 4.4.

Table 4.4 SL and MILP Constants

Variable Name	Value	Unit	Type	Description
$H_{max}$	10	MWh	Integer	Maximum Hydrogen Storage Capacity
$n_{max}$	8	-	Integer	Number of Available Spots
$\eta_{el}$	0.6 [41]	$\frac{MWh \text{ of } H_2 \text{ Produced}}{MWh \text{ of Electricity}}$	Float	Electrolyzer Efficiency
$c_{MH2}$	120 [59]	$\frac{CAD}{MWh}$	Integer	Market Hydrogen Cost
$V_H(t + 168)$	[348.18, 150]	$\frac{CAD}{MWh}$	Float	Value of the Remaining Inventory
$P_{el,max}$	1	MW	Integer	Maximum Electrolyzer Capacity

One SL simulation hourly step consists of seven different parts. The first part is to load the actions and make sure they are in the given intervals by clipping them into their low and high values. The second part is to define the meanings of the decisions, which is the first action from the action space. If the decision is higher than 0, the V2G operation mode is active, as shown in Equation 4.12. The decision is scaled to the number of trucks participating ( $n_{V2G}(t)$ ) for that hour. The scaling is rounded up because the number of trucks should be an integer value. On the other hand, if the decision is lower than 0, the hydrogen production mode is active, which is in Equation 4.13. The absolute value of the decision is multiplied by the maximum electrolyzer capacity for the scaling to the actual electrolyzer capacity ( $P_{E2HT}(t)$ ). Finally, if the decision is 0, the station will remain idle, which can be seen in Equation 4.14.

$$\text{if } D(t) > 0.01 \text{ then } n_{V2G}(t) = \text{ceil}(D(t) * n_{max}) \text{ and } P_{E2HT}(t) = 0 \quad \text{Equation 4.12}$$

$$\text{if } D(t) < -0.01 \text{ then } n_{V2G}(t) = 0 \text{ and } P_{E2HT}(t) = |D(t)| * P_{E2HT,max} \quad \text{Equation 4.13}$$

$$\text{if } D(t) = 0 \text{ then } n_{V2G}(t) = 0 \text{ and } P_{E2HT}(t) = 0 \quad \text{Equation 4.14}$$

The third and fourth parts are to initialize variables and load the market conditions, respectively. The variables in the simulation are the total power generated ( $P_{FC2G}(t)$ ), total hydrogen consumption of the fuel cell ( $P_{FC}(t)$ ), which is the sum of market hydrogen used ( $P_{MH2FC}(t)$ ), and storage hydrogen used ( $P_{HT2FC}(t)$ ). Finally, the adjusted number of trucks that can be fully served with the given conditions ( $n_{V2G,adj}(t)$ ). The market conditions loaded are hourly electricity price ( $p(t)$ ), number of trucks in the station at that hour ( $n(t)$ ), and excess supply flag ( $f_{excess}$ ). AESO defines supply surplus as when the electricity price equals 0, which is the same definition applied here [62].

Starting with the V2G logic, which is the fifth part, the first condition is to check if the number of trucks participating is positive, and it's not a supply-surplus hour. If both conditions are met, the V2G operation mode starts. This condition can be shows as Equation 4.15.

$$\text{if } n_{V2G}(t) > 0 \text{ and } f_{excess} = 0 \quad \text{Equation 4.15}$$

The next step for the V2G operation mode is to calculate the needed power and hydrogen to satisfy the V2G at that specific hour. Before the calculation, we need to cap the agent's decision of the total number of participating trucks with the number of trucks in the station (Equation 4.16).

$$n_{V2G}(t) = \min(n_{V2G}(t), n(t)) \quad \text{Equation 4.16}$$

The truck profiles are defined by the same methodology in Sections 2.2 and 2.3. Then, the truck profiles are sorted from highest power output to lowest. The participating trucks are selected starting from the highest power to the lowest power to obtain higher electricity output. The total power output ( $P_{FC2G}(t)$ ) is the sum of the powers of the amount of selected truck profiles ( $n_{V2G}(t)$ ). The needed hydrogen ( $P_{FC}(t)$ ) is calculated from the sum of hydrogen consumption of the same amount of selected truck profiles.

Now that we've calculated the total hydrogen required for the hour, as we already decided how much to draw from storage, we need to check whether the currently stored hydrogen is sufficient to meet this decision. The second action, alpha ( $\alpha(t)$ ), specifies the fraction of the total hydrogen demand that is fulfilled using market hydrogen, and  $(1-\alpha(t))$  is the proportion fulfilled by stored hydrogen on-site. The sufficiency of stored hydrogen is determined as follows. Since the agent's action may request more hydrogen from storage than what is available, the storage hydrogen consumption is capped at the current storage level ( $H(t)$ ), as shown in Equation 4.17. This ensures that consumption does not exceed what is

physically stored. The total available hydrogen that can be dispatched ( $P_{FC,actual}(t)$ ) is the sum of the available storage hydrogen ( $P_{HT2FC,actual}(t)$ ) and the decided total market hydrogen ( $P_{MH2FC}(t)$ ), which is Equation 4.18.

$$P_{HT2FC,actual}(t) = \min\left(\left(1 - \alpha(t)\right) * P_{FC}(t), H(t)\right) \quad \text{Equation 4.17}$$

$$P_{FC,actual}(t) = P_{HT2FC,actual}(t) + \left(\alpha(t) * P_{FC}(t)\right) \quad \text{Equation 4.18}$$

If the total available storage hydrogen is sufficient for the participating trucks, there is no need for an adjustment in the total power generated ( $P_{FC2G}(t)$ ), the amount of storage hydrogen used ( $P_{HT2FC}(t)$ ), and the amount of market hydrogen used ( $P_{MH2FC}(t)$ ). The amount of storage hydrogen used is deducted from the current storage for the next hour. The condition, and the resulting equations are provided in Equation 4.19.

$$\text{If } P_{FC,actual}(t) \geq P_{FC}(t) \text{ then } P_{HT2FC}(t) = P_{HT2FC,actual}(t) \quad \text{Equation 4.19}$$

$$\text{then } P_{MH2FC}(t) = P_{FC}(t) - P_{HT2FC}(t)$$

$$\text{then } H(t + 1) = H(t) - P_{HT2FC}(t)$$

If the total available stored hydrogen is insufficient to fuel all the selected participating trucks, the simulation determines the maximum number of trucks that can be supported. It begins by reducing the number of participating trucks to  $n_{V2G}(t) - 1$  and recalculates the total hydrogen needed ( $P_{FC}(t)$ ). This loop continues, decrementing the truck count by one each time, until either a feasible number of trucks is found, or the count reaches zero. If a valid number is found, the V2G operation proceeds with this reduced number of trucks. The amount of market hydrogen usage is recalculated according to Equation 4.19. If the count reaches zero, the hour is classified as idling mode instead of V2G operation.

The sixth part, which is the hydrogen production mode, starts with checking if there is available storage capacity ( $H_{avail}(t)$ ), as provided in Equation 4.20. The logic is to cap the agent's decision with the available storage capacity, and then use the adjusted decision ( $P_{E2HT,actual}(t)$ ) for the hydrogen production mode. The equation series (Equation 4.21, Equation 4.22, and Equation 4.23) for the hydrogen production mode can be seen below.

$$H_{avail}(t) = H_{max} - H(t) \quad \text{Equation 4.20}$$

$$P_{FC,actual}(t) = P_{HT2FC,actual}(t) + (\alpha(t) * P_{FC}(t)) \quad \text{Equation 4.21}$$

$$P_{E2HT,max}(t) = \frac{H_{avail}(t)}{\eta_{el}} \quad \text{Equation 4.22}$$

$$P_{E2HT,actual}(t) = \min(P_{E2HT}(t), P_{E2HT,max}(t)) \quad \text{Equation 4.23}$$

If adjusted electrolyzer input ( $P_{E2HT,actual}(t)$ ) is positive, then storage will be updated with the adjusted input. This condition is depicted in Equation 4.24. Or else, the operation mode will be idling instead.

$$\text{If } P_{E2HT,actual}(t) > 0 \text{ then } H(t+1) = H(t) + (P_{E2HT,actual}(t) * 1h * \eta_{el}) \quad \text{Equation 4.24}$$

The final part consists of computing the profit ( $\Pi(t)$ ) for the hour. From the V2G operation mode, the revenue ( $R_{V2G}(t)$ ) is generated by selling electricity to the grid. Otherwise, the costs are from the electricity consumed by the electrolyzer ( $C_{el}(t)$ ) and the market hydrogen usage ( $C_{MH2}(t)$ ). The financial components for the final part are indicated in Equation 4.25, Equation 4.26, Equation 4.27, and Equation 4.28.

$$R_{V2G}(t) = P_{FC2G}(t) * p(t) \quad \text{Equation 4.25}$$

$$C_{el}(t) = P_{E2HT,actual}(t) * p(t) \quad \text{Equation 4.26}$$

$$C_{MH2}(t) = c_{MH2} * P_{MH2FC}(t) \quad \text{Equation 4.27}$$

$$\Pi(t) = R_{V2G}(t) - C_{el}(t) - C_{MH2}(t) \quad \text{Equation 4.28}$$

## 4.4 Training Procedure for the SL Agent

The supervised learning agent was trained using behavioral cloning. This imitation learning technique treats the task as a supervised learning problem, where the goal is to train a policy that minimizes the difference between its predicted actions and the best actions from the expert dataset. We utilized the Python imitation library's implementation of behavioral cloning, which trains an Actor-Critic Policy by minimizing the mean squared error between the policy's action outputs and the expert's actions for a given set of state observations.

The Actor-Critic Policy network is a common architecture in decision-making tasks. This structure consists of two primary components that share a common feature extraction network. The first component, which is the actor, is the part of the network responsible for outputting the action. The second component,

which is the critic, is designed to estimate the value of a given state. Although we are using a supervised learning approach, where the critic's output is not directly used for training the final policy, we retain this standard architecture. The behavioral cloning algorithm focuses exclusively on training the actor head to mimic the expert's actions, while the critic head remains largely unused in this phase.

To prevent overfitting and improve the model's ability to generalize to unseen data, we employed a custom neural network architecture featuring dropout layers. Our feature extractor consists of multiple dense layers, with the output of each layer being processed by a Rectified Linear Unit (ReLU) activation function. The feature extractor learns to identify and combine important signals, filter out noise, and discover non-linear relationships within the raw data. This resulting feature vector, which is a compressed yet highly informative summary of the original state, is then passed to the final layers of the policy network (the actor) to make the ultimate decision.

The ReLU function is a nonlinear transformation. It allows positive values to pass through unchanged while converting all negative values to zero. This process introduces non-linearity into the model, which enables it to learn more complex patterns than a simple linear model could. Additionally, a dropout layer with a probability of 0.2 randomly sets a fraction of the active neuron outputs to zero during each training update. This combination of ReLU and dropout forces the network to learn robust representations. It prevents it from becoming overly reliant on any single feature and enhances its performance on the out-of-sample test set.

The training process is further enhanced using a decaying learning rate. Instead of using a fixed learning rate throughout training, we implement a linear decay schedule. The learning rate is initialized at a higher value ( $3e-4$ ) to facilitate rapid learning in the initial epochs and is gradually reduced to a lower final value ( $1e-5$ ) as training progresses over 5000 epochs. An epoch represents one complete pass through the entire training dataset. This approach allows the model to make exploratory updates early on and then fine-tune its parameters with more precise adjustments in the later stages.

To ensure that the final selected model represents the peak performance achieved during the entire training run, we implemented a checkpointing strategy. The full training process was divided into chunks of 100 epochs. After each chunk, the current policy was evaluated on the full in-sample training dataset to calculate its total profit. If this profit exceeded the best score recorded so far, a checkpoint was created by saving the current policy's network weights and the associated normalization statistics. This methodology guarantees that the final model used for analysis is the one that achieved the highest performance during training, which protects against any performance degradation in the late stages of training.

## 4.5 Chapter Summary

This chapter detailed the complete methodological blueprint for developing and validating a behavioral cloning agent for FCEV2G operational control. The foundation of this methodology is the creation of a high-quality expert dataset. This was generated using a rolling horizon MILP optimization, which was itself enhanced with a lookahead mechanism to produce a decision policy for the agent to imitate. Following this, the agent's architecture was defined, featuring a 48-dimensional observation space engineered with multi-horizon price forecasts to provide critical foresight, and a two-dimensional continuous action space for nuanced operational control.

The agent's policy, a deep neural network, was then trained using behavioral cloning. To ensure the resulting model was robust and high performing, the training procedure incorporated several key techniques: dropout for regularization, a decaying learning rate for stable convergence, and a checkpointing system to select the model from the epoch of peak performance. Finally, a rigorous evaluation framework was established, utilizing a custom-built simulation to test the policy on distinct training (2015-2024) and out-of-sample test (2025) datasets. This structured approach allows for a comprehensive assessment of both the agent's imitation accuracy and its ability to generalize to new market conditions.

## Chapter 5 Supervised Learning-based FCEV2G Operation

### 5.1 Comparing SL Agent and MILP Expert: Training Dataset

The initial phase of our analysis evaluates the supervised learning agent's performance on the training dataset (2015-2024). This comparison against the MILP expert on the data it was trained on serves two key purposes. First, to validate the effectiveness of the behavioral cloning training process itself, and second, to quantify the baseline performance gap that exists even under familiar conditions.

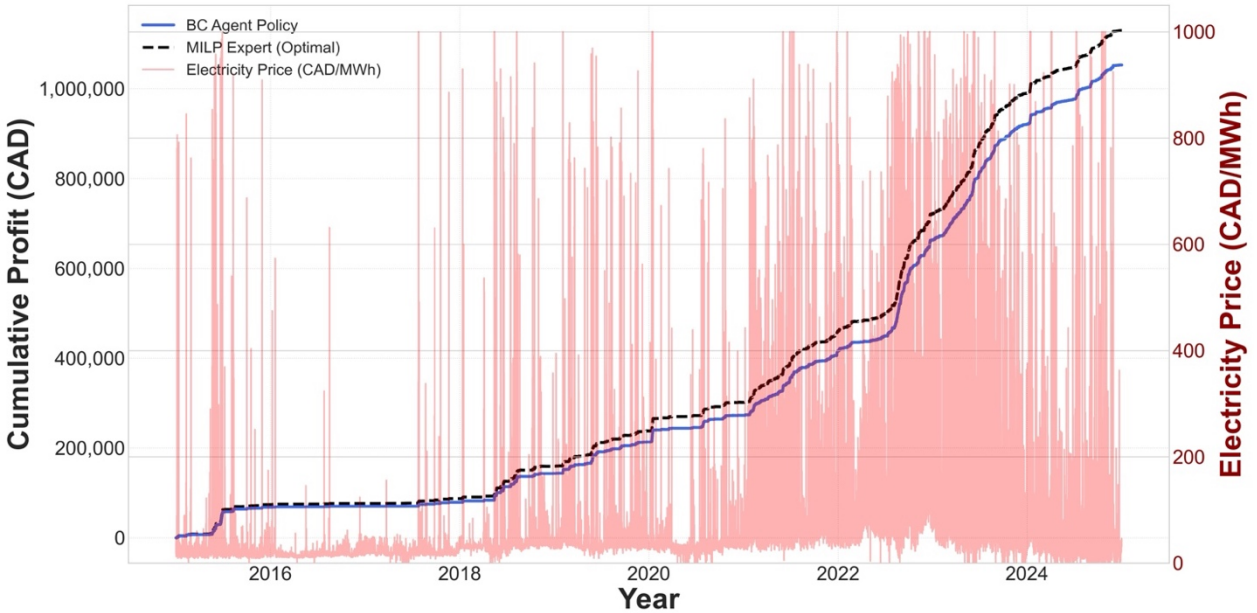


Figure 5.1 The Cumulative Profit Difference between SL Agent and MILP Expert on Training data

The profit difference between the FCEV2G operations by the supervised learning (SL, behavioral cloning) agent and MILP expert can be seen in Figure 5.1. The total profit of the operation by the expert is 1,129,486.04 CAD, whereas the total profit of the agent is 1,052,803.22 CAD. This means the behavioral cloning agent successfully captured 93.2% of the total possible profit identified by the expert optimizer. Crucially, while the MILP expert required over six minutes (380 seconds) of computation time to generate this optimal plan, the trained SL agent can produce a decision for any given hour in milliseconds, making it suitable for real-time deployment. To understand the source of this 76,682.82 CAD performance gap, we analyzed the three operational modes of the system at any given hour, i.e., hydrogen production mode (G2H), V2G operation mode (V2G), or remaining idle. Furthermore, V2G actions were categorized based on their dominant hydrogen source. An action was labeled V2G (from Storage) if most of the hydrogen consumed in that hour came from the onsite hydrogen storage. Otherwise, the action is labeled as V2G

(from Market). This allows for a comprehensive analysis in the primary strategic intent of each V2G operation decision. In general, a rational agent should perform hydrogen production (G2H) at low electricity prices, V2G at high electricity prices with storage hydrogen, and remain idle at unprofitable electricity prices. It should also minimize the use of market-purchased hydrogen because of its high price.

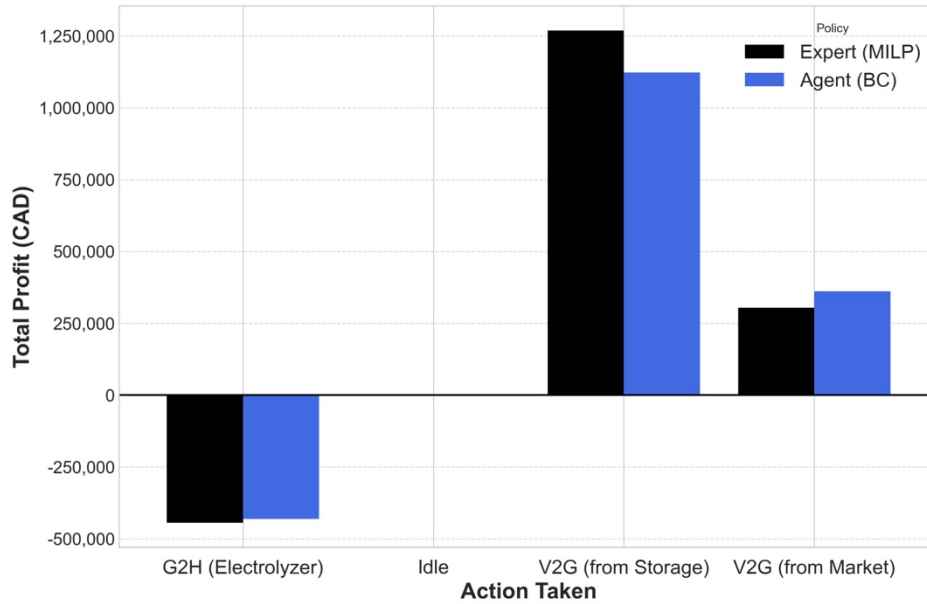


Figure 5.2 Profit Distribution of the Actions Taken by Agent vs Expert

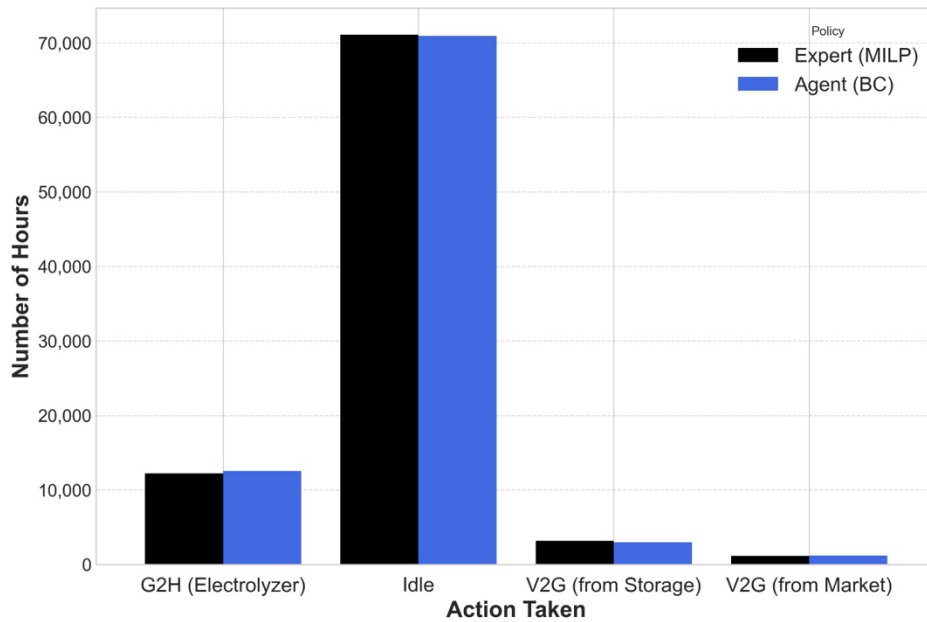


Figure 5.3 Frequency Distribution of the Actions Taken by Agent vs Expert

Figure 5.2 and Figure 5.3 shows the distributions of profits and the frequencies of the actions by the agent and expert, which confirm that the agent learned the fundamental principles of the task. Both the agent and the expert identified using storage hydrogen for V2G as the primary profit tool and running the electrolyzer as a necessary step for building low-cost hydrogen reserves. Furthermore, the number of hours spent idling is similar between the two policies, which demonstrates that the agent learned from the training data when it is best to take no action. Despite these similarities in the strategies, the figure reveals an important difference. The difference is in the V2G (from Storage) category, where the expert generated approximately 78,072 CAD more profit than the agent. This suggests the expert was more effective at capitalizing on the opportunities to conduct V2G.

Another key observation from Figure 5.3 is that the 'Idle' mode is the most frequent action for both the expert and the agent, which accounts for over 70,000 hours of the simulation period. This high frequency of inaction is a core feature of an optimal profit-maximizing strategy in a volatile energy market. As profitable opportunities are dictated by the electricity price, hydrogen production is only profitable during periods of low electricity prices, while V2G operation is only profitable during high-price peaks. Yet most hours in a typical year do not fall into these extreme categories. During these intermediate price periods, the optimal decision is to remain idle. This strategy conserves capital by avoiding unprofitable hydrogen production and preserves the valuable stored hydrogen for future high-price discharge events.

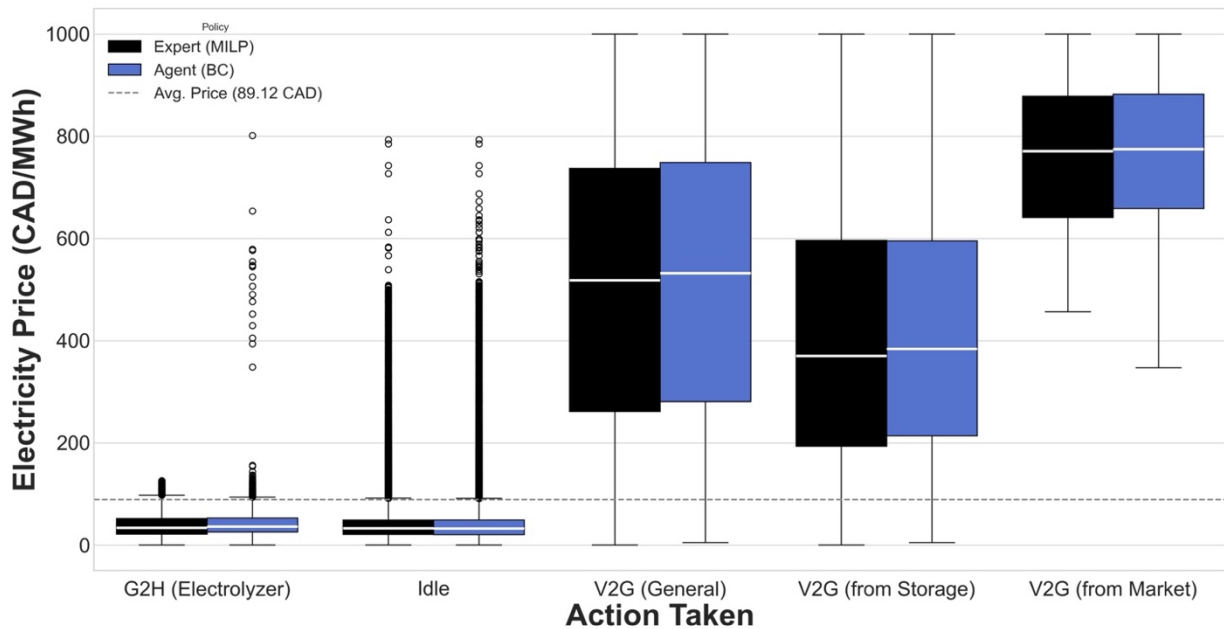


Figure 5.4 Comparison of Agent vs. Expert Strategy by Price

The strategic behavior of the behavioral cloning agent vs MILP expert is analyzed in Figure 5.4 using a box and whiskers plot. This visualization displays the distribution of electricity prices for each action, where the box represents the interquartile range (IQR), the white line marks the median, and the whiskers extend to 1.5 times the IQR. Any prices falling outside this range are classified as outliers. The plot reveals that the agent has learned the core economic principles of the task. The operations from both the agent and the expert exhibit the same patterns: low median prices for G2H actions and high median prices for V2G actions.

We looked further into the electricity prices when a certain action takes place. Starting with the G2H action, the agent's median price for initiating G2H is 36.10 CAD/MWh, slightly higher than the expert's 33.51 CAD/MWh. Additionally, the concerning part is that the expert has 261 G2H actions classified as outliers, with a maximum price of 125.89 CAD/MWh. In contrast, the agent has 342 outliers and made a G2H decision at a higher maximum price of 801.27 CAD/MWh. However, an investigation into these high-priced G2H events reveals the agent's strategy. We conducted a targeted analysis of the 15 instances where the agent initiated G2H at prices above 400 CAD/MWh. Thirteen of these decisions were strategically justified by the imminent arrival of a higher price (above 600 CAD/MWh) within the next 72 hours. For example, a G2H action at 405.26 CAD/MWh was followed by a V2G opportunity at 750.19 CAD/MWh. This shows that the agent not only learns to buy low but also proactively builds hydrogen reserves in anticipation of a future selling opportunity.

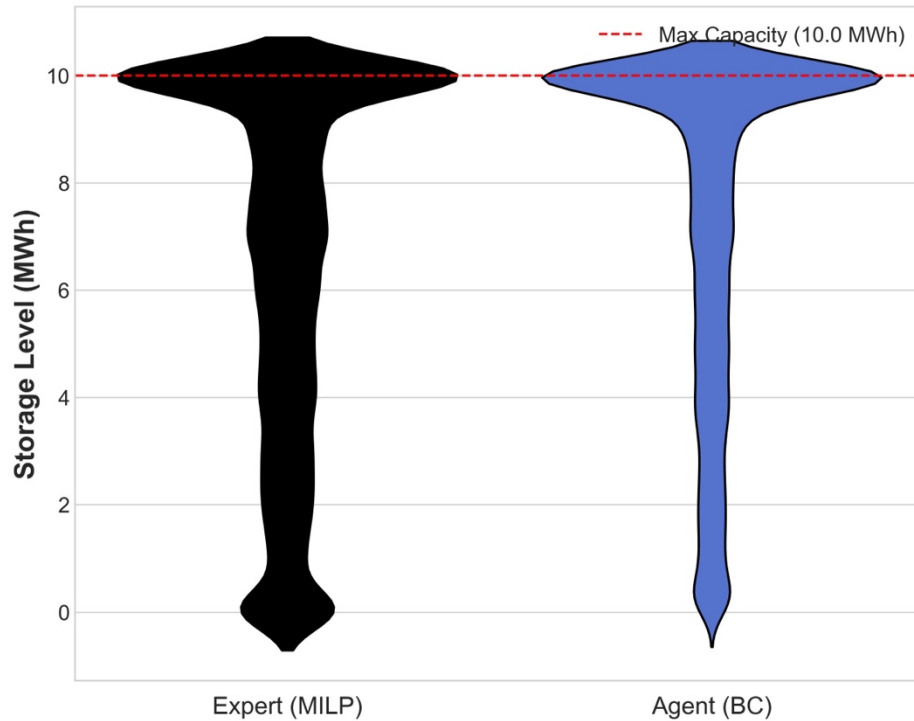


Figure 5.5 Overall Distribution of Hydrogen Storage Levels

Another action to analyze is the V2G operation with onsite stored hydrogen. Counter-intuitively, the agent's median price for using storage hydrogen (383.75 CAD/MWh) is higher than the expert's (369.82 CAD/MWh). Despite selling V2G electricity at a higher price, the agent's total profit is less. This contradictory result can be explained by examining the agent's overall storage management strategy, which can be seen in Figure 5.5. Statistics on the hydrogen tank levels show that the agent maintains a fuller tank on average than the expert (mean of 7.38 MWh vs. 6.50 MWh and median of 9.09 MWh vs. 7.34 MWh). This indicates that the agent has learned a more conservative storage policy. It is less willing to deplete its stored hydrogen reserves. Therefore, the agent achieves a higher median V2G price because it is more selective. It only discharges when prices are exceptionally high to justify using its stored energy. However, this selectiveness means it omits moderately profitable V2G opportunities that the expert acts on. This explains how selling at a better median price can lead to lower overall profit, as the agent prioritizes energy security over maximizing transaction volume.

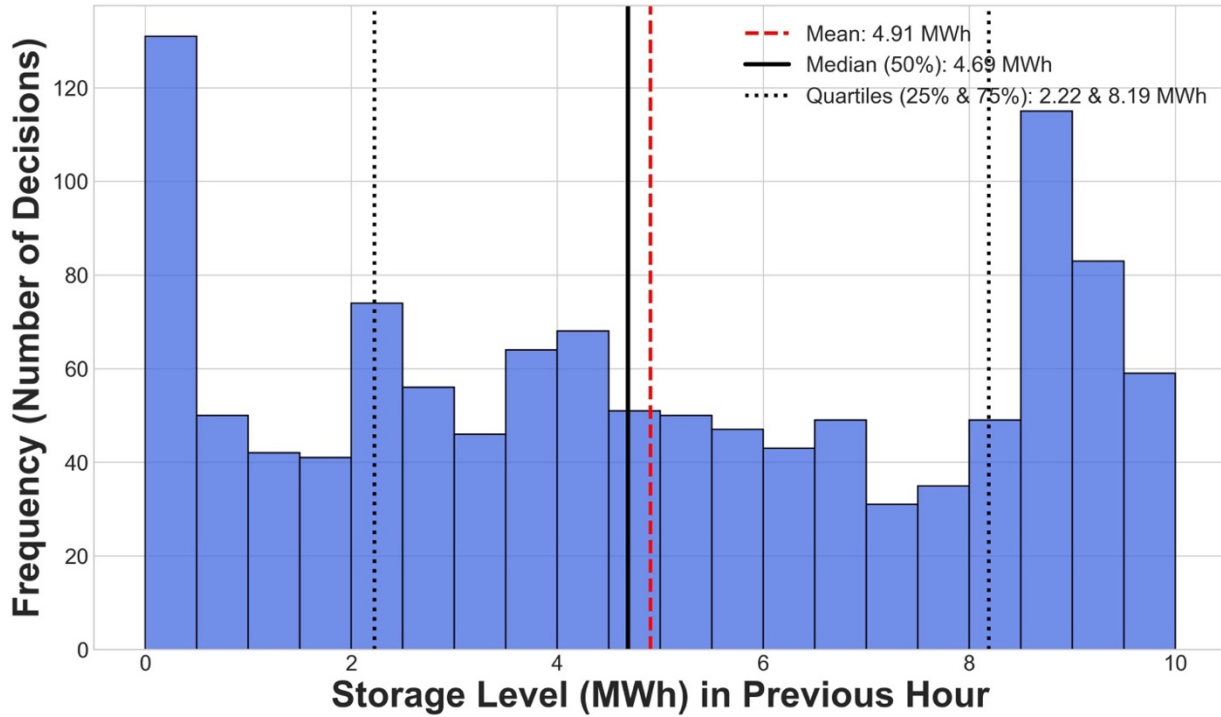


Figure 5.6 Agent's Storage Level One Hour before V2G (from Market) Decision

The V2G (from Market) category also highlights the agent's inconsistent strategic behavior. This action represents a high-risk, high-reward strategy. The expert's policy has a minimum price of 456.40 CAD/MWh for doing V2G with market hydrogen. Whereas the agent is more aggressive, initiating this strategy at a minimum price of 347.08 CAD/MWh. This willingness to engage in V2G at a lower electricity price is one of the reasons why the agent's total profit from this category was lower than the expert's. To understand the underlying strategy, an analysis was performed on the agent's hydrogen storage level in the moments immediately preceding these 1,184 decisions. Figure 5.6 shows the agent's storage level at the hour right before a V2G action takes place. Distribution varies with a mean of 4.91 MWh and a median of 4.69 MWh. This appears that there is no single generalized strategy for this action. Furthermore, in approximately 25% of cases, the agent's storage was below 2.22 MWh. As a full V2G dispatch requires over 3 MWh of hydrogen with eight participating trucks, the agent is forced to buy from the market to avoid missing a good V2G opportunity.

However, in the remaining 75% of cases, the agent's behavior is explained by the same strategy explained previously. In these instances, the agent has sufficient hydrogen stored but chooses not to use it. The agent has learned to give high value to its stored hydrogen, so that it reserves it for the most profitable

events. And it is willing to pay the high price of market hydrogen to service high-intermediate V2G opportunities. In summary, the expert's perfect foresight allows it to calculate precisely when to use stored hydrogen, whereas the agent's learned policy applies its conservative rule too broadly, which leads to lower overall profitability.

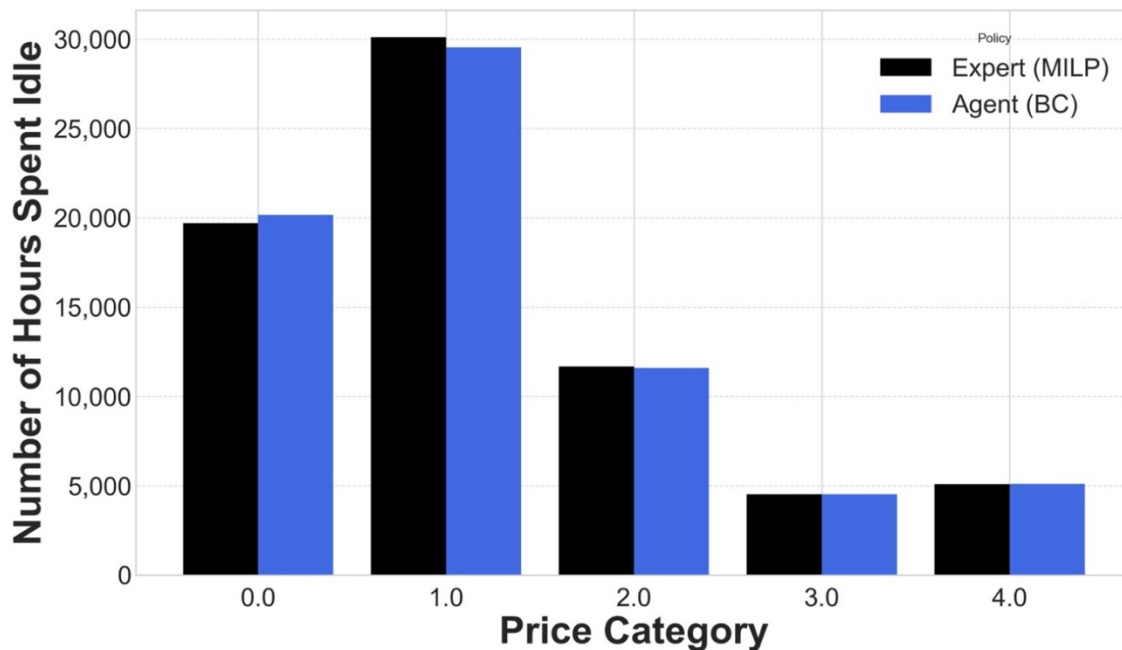


Figure 5.7 Distribution of Agent's Idle Hours by Price Category

As idling is the most frequent action taken by both the agent and the expert, we investigated the idling action patterns. The frequency analysis of these decisions, broken down by price category, can be seen in Figure 5.7. In the lowest price category 0, where hydrogen production is most profitable, the agent chose to idle over 20,000 times. That's because during these low-price idling periods, the agent's hydrogen tank had a median storage level of 10.0 MWh, its maximum capacity, and a mean storage level of 8.97 MWh. The agent learned that when the tank is full, it must be idle. Even the price of electricity is low, as it cannot store any more hydrogen because of physical constraints.

The other aspect of the agent's policy is the decision to idle during the highest-price category 4. Among these 5,093 instances, 96.39% of the cases where the agent idled at a high price, even though a higher price will appear within the next 72 hours. This means that the agent is not ignoring the current profitable high prices. Instead, it has learned to make a trade-off: the agent acts to neglect a good selling

opportunity now to preserve its finite resources for an even better opportunity that it anticipates is coming soon.

By analyzing the agent's learned behavior to explain the 6.8% profit gap, we can clarify the differences between the agent's strategies and those of the expert. We found that the agent can make forward-looking decisions with a limited view of the entire dataset. It correctly learned to act on more profitable prices by capturing the expert's long-term planning logic. However, this sometimes caused the agent to make suboptimal decisions, such as producing hydrogen and idling at high prices to gain future profits. Another issue is the agent's overly conservative inventory policy. For V2G (from Storage), the agent is very selective, waiting for higher prices before using its stored hydrogen. While this results in a higher median V2G price, it also causes the agent to miss out on larger volumes of intermediate profitable trades that the expert captures, leading to lower overall profit. Additionally, the value placed on storage influences the agent's V2G (from Market) strategy. It often pays higher prices for market hydrogen to support intermediate V2G opportunities, aiming to preserve cheaper reserves for a potential future peak. In summary, the behavioral cloning agent is not an exact clone but a different version of the expert. It has learned long-term anticipation, but its performance is limited by its learned strategy.

## **5.2 Comparing SL Agent and MILP Expert: Test Dataset**

Having established the agent's baseline performance and behavior on familiar data, the second phase of our analysis evaluates how well its learned operational strategy generalizes to unseen market conditions. This section therefore presents a comprehensive comparison of its operational decisions and resulting profitability against the expert on the out-of-sample test dataset (2025). This evaluation is the critical test of the policy's robustness, providing the most realistic measure of its effectiveness in a real-world deployment scenario.

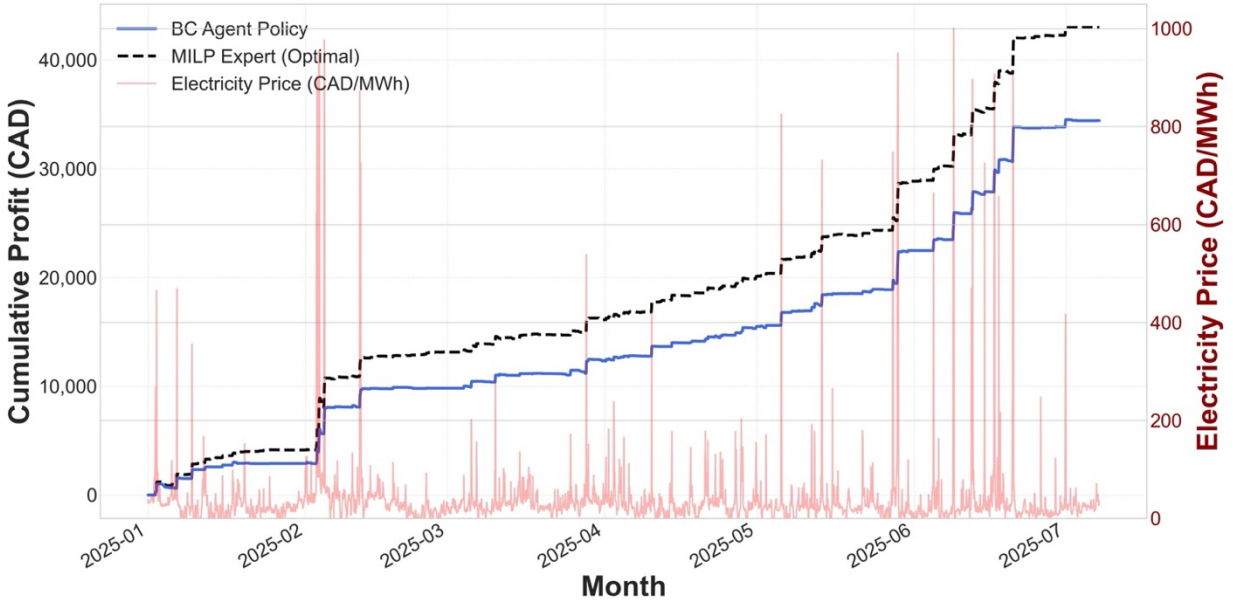


Figure 5.8 The Cumulative Profit Difference between Behavioral Cloning Agent and MILP Expert on Testing data

To evaluate the generalization capability, the trained agent was deployed on an unseen test dataset spanning from January 2025 to early July 2025. This out-of-sample test is to measure the model's effectiveness. The profit difference between the behavioral cloning agent and MILP expert for the testing dataset can be seen in Figure 5.8. The total profit of the expert is 42,815.80 CAD, whereas the total profit of the agent is 34,404.60 CAD. This means the behavioral cloning agent successfully captured 80.4% of the total possible profit identified by the expert optimizer, which is lower than in the training dataset. The same analysis for the actions can be done to understand the profit difference to determine whether the lower profit by the agent is due to an overfitting problem or the limited long-term anticipation.

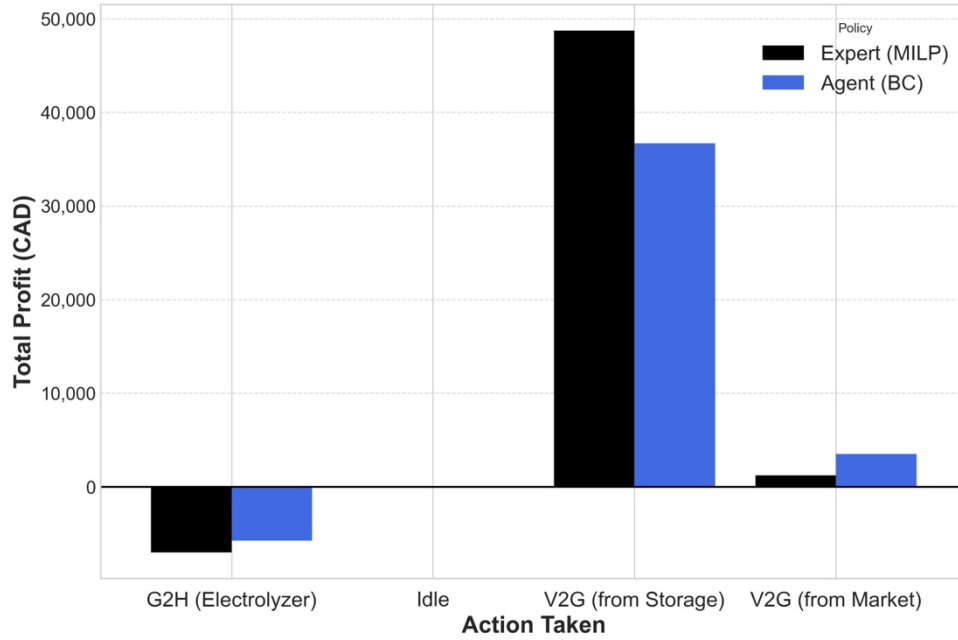


Figure 5.9 Comparison of Agent vs Expert on Profit Contribution

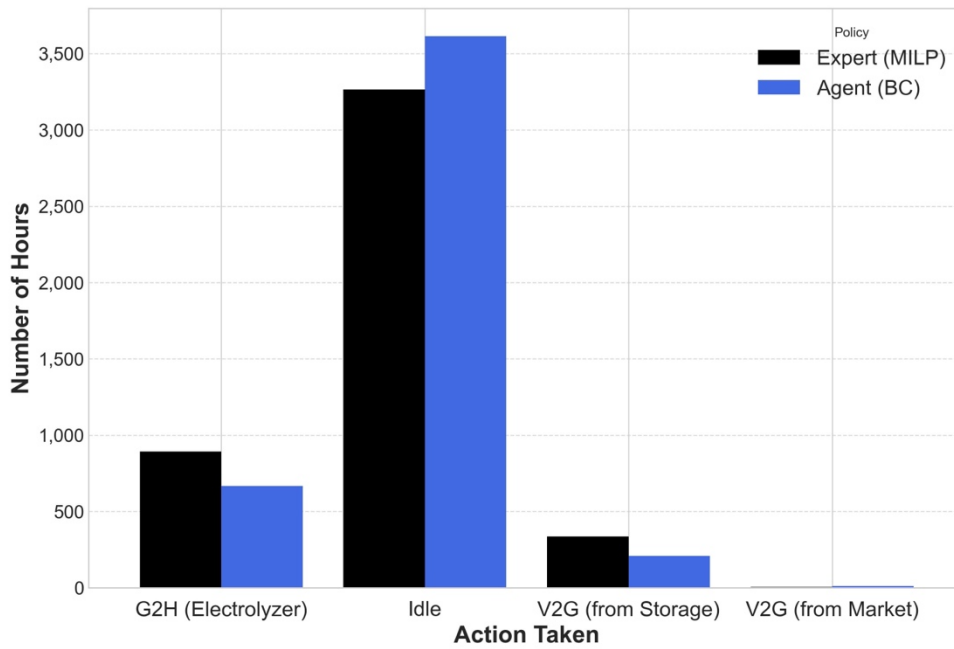


Figure 5.10 Comparison of Agent vs Expert on Action Frequency

The agent's strategic behavior on the test set is consistent with its training set performance. The agent's conservative inventory policy can be seen in Figure 5.9 and Figure 5.10. In the V2G (from Storage) category, it shows that the agent executed the action 206 times compared to the expert's 335 times. This is consistent with the agent's strategy of waiting for higher prices. And this risk-averse strategy is the largest single contributor to the performance gap.

Furthermore, the agent's V2G (from Market) strategy is also analyzed. From the limited dataset for this action is limited, we can see the agent did it more often than the expert (11 vs. 5 times). This resulted in higher profits in this action, but this sub-optimal decision lowers the total profit. This increased reliance on market hydrogen is not due to overfitting. Instead, it is a direct result of the agent's conservative strategy of its own stored reserves, forcing it to make a sub-optimal choice to avoid missing V2G opportunities altogether.

The same conservative storage strategy applies to the hydrogen production. As the V2G (from Storage) action is chosen less, the stored hydrogen is used less, so that the hydrogen production is less necessary.

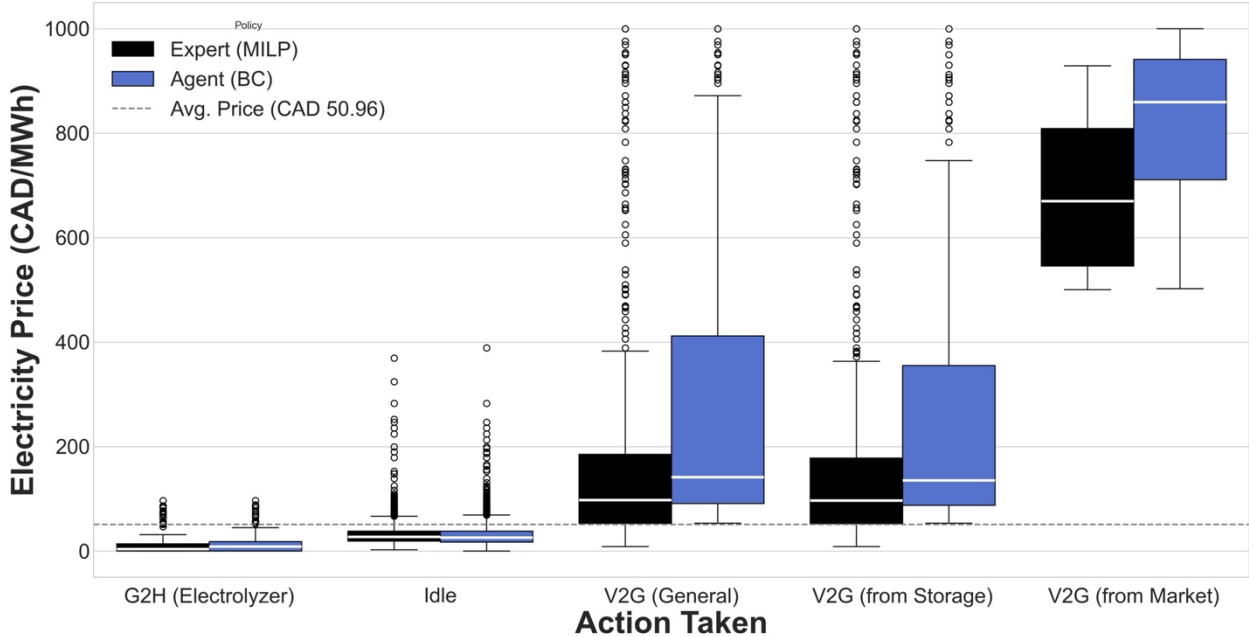


Figure 5.11 Comparison of Agent vs. Expert Strategy by Price in the Test Set

A comparison of agent vs. expert strategy by electricity price can be used to identify the agent’s strategy, which can be seen in Figure 5.11. Starting with the hydrogen production, the agent's median price for initiating G2H was 8.20 CAD/MWh, which is higher than the expert's 3.49 CAD/MWh. To investigate these high-priced G2H events, we conducted a targeted analysis of the 26 instances where the agent initiated G2H at prices above 38.17 CAD/MWh, which is the 75% quartile. It was found that 25 instances were followed by the imminent arrival of a higher electricity price (above 91.39 CAD/MWh) within the next 72 hours. This again confirms the agent’s limited long-term strategy, which was also seen in the training dataset.

On the V2G (from Storage) action, the agent's median selling price from storage was 135.55 CAD/MWh, which is higher than the expert's 96.83 CAD/MWh. Like what we found during the training, the agent waits for better prices, but this wait also causes the agent to act less frequently (206 times vs. the expert's 335 times), which is the primary driver of the profit gap. For the idling action, the agent's median price (25.68 CAD/MWh) remained close to the expert's (27.00 CAD/MWh). And finally, because of the conservative storage strategy, the V2G (from Market) actions are done at a wider range of electricity prices than in the expert’s operation, which was also seen in the training. The agent's less selective approach to this high-risk action shows that the similar behavior during its training.

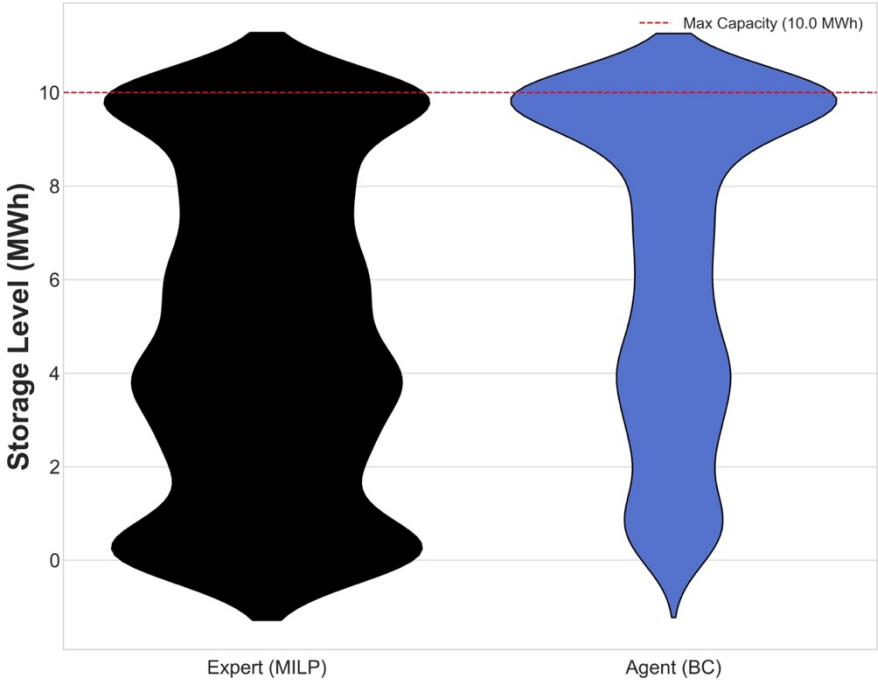


Figure 5.12 Overall Distribution of Hydrogen Storage Levels on Test Dataset

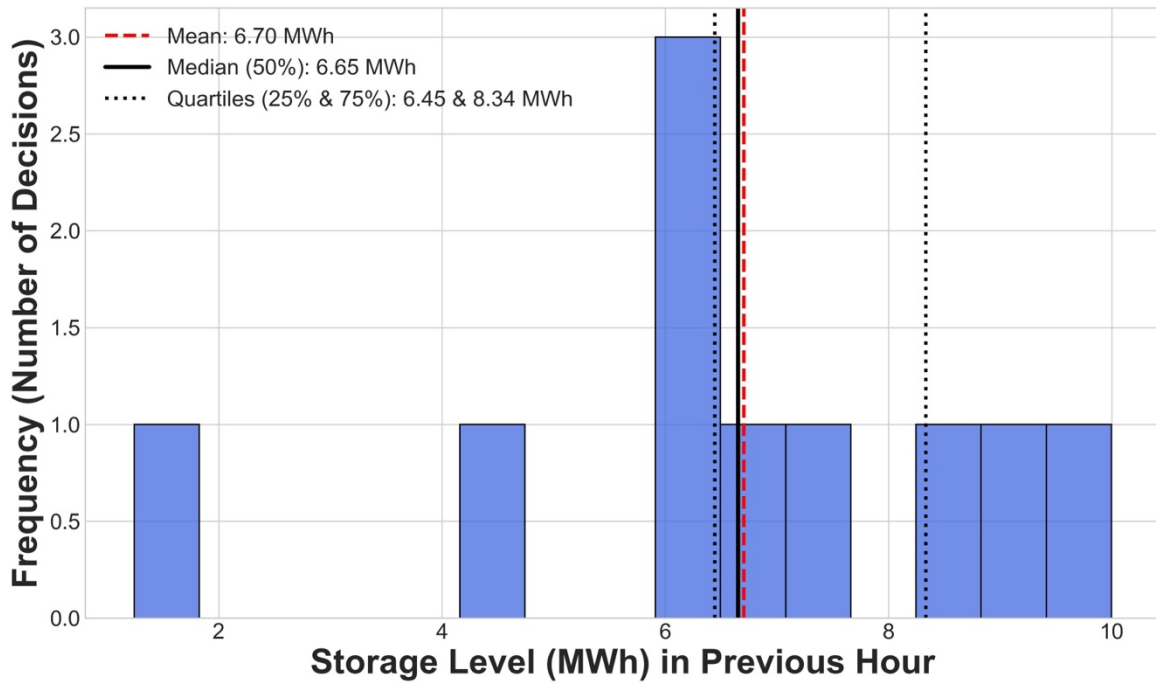


Figure 5.13 Agent's Storage Level when Deciding to Use Market Hydrogen

Figure 5.12 shows the overall distribution of hydrogen storage levels, which demonstrates the agent's conservative inventory policy. The agent maintained a fuller tank than the expert, with a mean storage of 6.47 MWh compared to the expert's 5.00 MWh, and a median of 7.39 MWh versus the expert's 4.72 MWh. This tendency to hoard hydrogen provides the strongest evidence of generalization, which is demonstrated by the agent consistently reproducing the same sub-optimal storage strategy on the unseen test data that it exhibited during the in-sample training analysis. Furthermore, the strategy behind the agent's V2G (from Market) decisions also remained consistent with the training. The analysis of the agent's storage level just before the V2G (from Market) action reveals a similar pattern to the training (Figure 5.13). This confirms that the agent chooses to pay a premium for market hydrogen to avoid depleting its cheap reserves.

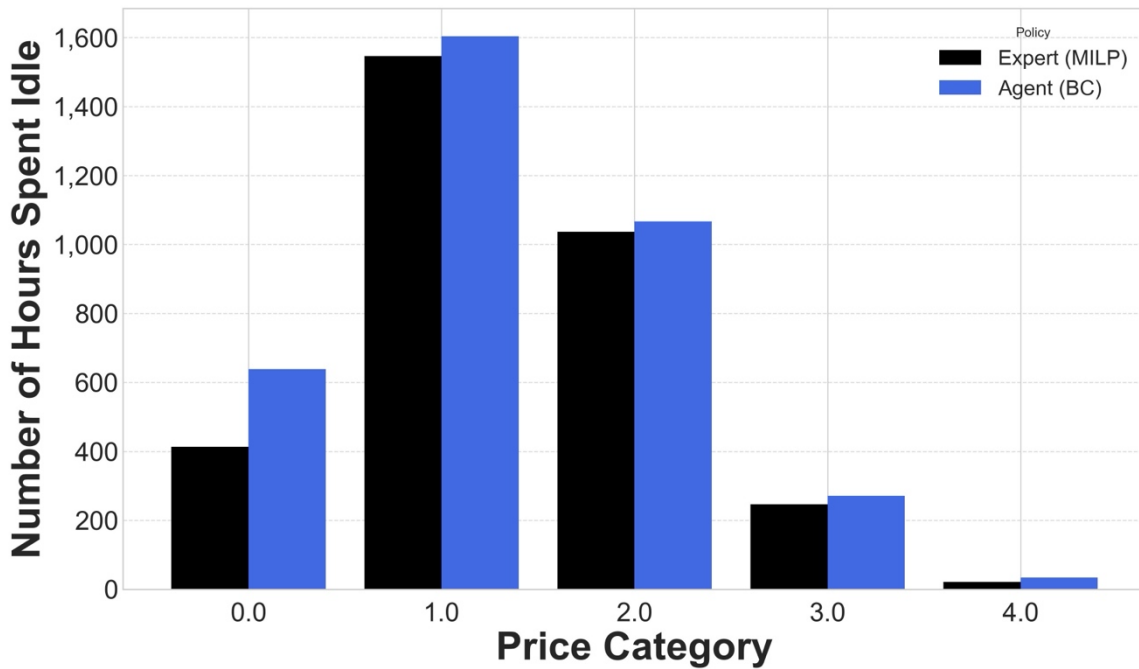


Figure 5.14 Comparison of Idle Hours by Price Category for Test Data

Figure 5.14 shows the comparison of the idle hours by price category for the testing data. The agent's response to physical constraints remained the same with the training. In the hours when the agent takes the idling action during low electricity price hours (price category 0), the median hydrogen storage level was at its maximum capacity of 10.0 MWh. This is consistent with the training set behavior. Additionally, when faced with high-price V2G opportunities on the test set, the agent still chose to idle in certain situations. A forward-looking analysis revealed that 88.24% of these decisions were followed by a higher electricity price within the next 72 hours. While this is a slight decrease from the 96% accuracy observed on the training set, this high success rate on unseen data provides strong proof of the agent's ability to successfully use its forecasts to wait for future price peaks.

The evaluation on the unseen test dataset provides evidence that the behavioral cloning model did not overfit and successfully generalized its learned policy. The agent's strategy is defined by a conservative inventory strategy. This was most evident in its higher average storage levels on the test dataset. This strategic hesitation is the primary driver of the profit gap. It exhibits as a frequent failure to commit to V2G (from Storage) actions during profitable periods, a cautious tendency that was consistently observed in both training and testing. Its strategy of idling during high electricity price periods was due to projected future price peaks (88.24% of cases on the test set) and idling during low electricity price periods was consistently due to a full storage tank. However, this also demonstrates the agent's weaknesses in making better decisions to achieve higher profits. The agent has a less selective approach to the V2G (from Market) action.

Therefore, the agent has a final 80.4% performance on the test set as a result of a generalization of a sub-optimal policy interacting with new market conditions.

### **5.3 Chapter Summary**

This chapter presented a comprehensive performance analysis of the supervised learning agent, comparing its operational strategy against the MILP expert on both in-sample training data and an out-of-sample test set. The evaluation revealed that the behavioral cloning agent was highly effective, capturing 93.2% of the expert's profit on the training data. Deeper analysis into the performance gap identified the agent's core learned strategy: a conservative storage policy characterized by a tendency to hoard stored hydrogen. While this led to the agent missing some moderately profitable V2G opportunities, this behavior was often a forecast-driven choice to wait for better future prices.

The final evaluation on the unseen 2025 test data confirmed that this entire strategic identity generalized successfully without overfitting, with the agent achieving a robust 80.4% of the expert's profit. The agent's performance on the test set was defined by this same strategic patience; its decisions to idle were consistently justified by its forecasts or physical constraints. Ultimately, this chapter concludes that while the behavioral cloning agent is not a perfect mimic, it successfully distills the expert's long-term foresight into a rapid and robust data-driven policy, which proves its capability for near-optimal performance in unseen market conditions.

## Chapter 6 Conclusion and Recommendations for Future Work

### 6.1 Conclusion

This thesis addressed the need for intelligent energy management strategies at the intersection of renewable energy and transportation. Specifically, it set out to determine the economic viability and optimal control strategy for a FCEV2G station operating within the volatile electricity markets of Ontario and Alberta. The research progressed through two distinct methodological stages to first establish a baseline and then propose an advanced solution.

The initial phase involved the development of an operational simulation based on fixed constraints and historical data. The results of this baseline model were expected. A simple, and non-optimized FCEV2G station is not currently profitable. The analysis identified barriers to viability, including significant round-trip efficiency losses in the hydrogen cycle, the prohibitive cost of market hydrogen, and the absence of a dynamic dispatch strategy to prioritize actions during peak price periods. This analysis proved that to profit the implementation of an intelligent control policy is necessary.

To overcome the limitations of a rule-based approach, this research transitioned to a machine learning method. An intelligent agent was developed using a behavioral cloning framework, where a deep neural network was trained to mimic the decisions of a perfect-foresight MILP expert. The comprehensive evaluation of this agent demonstrated definitive success. The agent learned a highly profitable policy, achieving 93.2% of the expert's optimal profit on the training set and a robust 80.4% on an unseen test set. This high performance confirms the practical feasibility of the FCEV2G operation, proving that near-optimal results are achievable in real-time without perfect foresight.

The results revealed that the agent's success stemmed from learning a conservative inventory policy. This allowed it to act on long-horizon strategies by interpreting its forecast features, thereby solving the myopic dispatch problem that plagued the baseline model. However, this same risk aversion defined its primary limitation, manifesting as a hesitation to use stored hydrogen during moderately profitable periods where the expert was more aggressive. The analysis confirmed that the agent successfully generalized this entire strategic identity to the test set, demonstrating its robustness. The performance gap is therefore not due to overfitting but is a direct result of this learned cautiousness interacting with new market conditions.

In conclusion, this research delivers a successful proof-of-concept for an intelligent FCEV2G operational controller. We have demonstrated that a data-driven agent can manage the station's complex operations not just profitably, but feasibly in real-time. The primary contribution is the validation of a method to distill the long-term operational strategy of a slow optimizer into a rapid and robust AI policy.

## 6.2 Recommendations for Future Work

The success and limitations of the behavioral cloning agent highlight several promising directions for future research. The next step could be to surpass the performance ceiling of the MILP expert is by transitioning to a Reinforcement Learning (RL) framework. However, preliminary investigations revealed that the FCEV2G environment is economically brittle. Standard RL exploration, where the agent tries a wide range of actions, often leads to financially catastrophic decisions. This results in large negative rewards that destabilize the entire learning process. The agent could learn what not to do but may struggle to discover the narrow path of profitability. Therefore, the recommended future work is to conduct a cautious fine-tuning process. The behavioral cloning agent's policy can be used to warm start RL agent, which ensures the agent begins in a safe region. The agent should then be fine-tuned using an algorithm like Proximal Policy Optimization (PPO). The key is to find the perfect parameters that encourage safe exploration. This pre-training and cautiously fine-tuning methodology would allow the agent to safely explore the policy space to find a better strategy. It could potentially discover superior responses to market conditions that even the expert did not exhibit.

The current work relies on two key simplifications regarding the FCEV technology itself. Firstly, the PEMFC polarization curve model, which dictates the vehicle's power output and efficiency, was scaled up from a smaller experimental stack. Future research should focus on integrating a higher-fidelity model that is more representative of a heavy-duty FCEV truck. Secondly, the simulation assumes that there is no degradation during V2G. Degradation is highly dependent on operational stressors like high-power operation. An area for future work is to implement a dynamic degradation model. By incorporating a model that links V2G actions to a degradation rate, the power output from FCEV can change during V2G. This could add a new strategy to the agent, where partial V2G could be a new option instead of a full hour.

The performance of the agent is influenced by the quality of its forecasted inputs, which are the Alberta traffic data and the 24-hour and 72-hour price signals. The current study utilized simplified forecasts. To better assess the agent's real-world viability, future work should focus on improving the realism of these signals. This would involve replacing the current methods with state-of-the-art time-series forecasting models (e.g., LSTMs, Transformers) for both traffic and electricity prices. Crucially, the agent should then be trained and evaluated using the imperfect and noisy outputs of these models.

## References

- [1] Iea, “Electricity 2025,” 2025. Accessed: May 15, 2025. [Online]. Available: <https://www.iea.org/reports/electricity-2025#overview>
- [2] “What Is Climate Change? | United Nations.” Accessed: May 20, 2025. [Online]. Available: <https://www.un.org/en/climatechange/what-is-climate-change>
- [3] “The Paris Agreement | UNFCCC.” Accessed: May 20, 2025. [Online]. Available: <https://unfccc.int/process-and-meetings/the-paris-agreement>
- [4] “The hydrogen colour spectrum | National Grid.” Accessed: May 22, 2025. [Online]. Available: <https://www.nationalgrid.com/stories/energy-explained/hydrogen-colour-spectrum>
- [5] “How clean is green hydrogen? | MIT Climate Portal.” Accessed: May 22, 2025. [Online]. Available: <https://climate.mit.edu/ask-mit/how-clean-green-hydrogen>
- [6] “Analysing the future cost of green hydrogen | PwC.” Accessed: Jul. 14, 2025. [Online]. Available: <https://www.pwc.com/gx/en/issues/business-model-reinvention/how-we-fuel-and-power/analysing-future-cost-of-green-hydrogen.html>
- [7] “Green Review | The emergence of natural hydrogen — a gamechanger in energy - Green Review.” Accessed: Jul. 14, 2025. [Online]. Available: <https://greenreview.com.au/resources/the-emergence-of-natural-hydrogen-a-gamechanger-in-energy/>
- [8] “ADVANCED CLEAN ENERGY STORAGE | Department of Energy.” Accessed: Jul. 14, 2025. [Online]. Available: <https://www.energy.gov/lpo/advanced-clean-energy-storage>
- [9] “Hydrogen Storage | Department of Energy.” Accessed: May 22, 2025. [Online]. Available: <https://www.energy.gov/eere/fuelcells/hydrogen-storage>
- [10] “TRE FCEV - Nikola Hydrogen-Electric Semi-Truck.” Accessed: Jul. 14, 2025. [Online]. Available: <https://www.nikolamotor.com/tre-fcev>
- [11] “2024 Toyota Mirai - Toyota USA Newsroom.” Accessed: Jul. 14, 2025. [Online]. Available: <https://pressroom.toyota.com/vehicle/2024-toyota-mirai/>
- [12] T. Zhang, J. Uratani, Y. Huang, L. Xu, S. Griffiths, and Y. Ding, “Hydrogen liquefaction and storage: Recent progress and perspectives,” *Renewable and Sustainable Energy Reviews*, vol. 176, p. 113204, Apr. 2023, doi: 10.1016/J.RSER.2023.113204.

- [13] “Hydrogen Storage and Transport: Technologies and Costs.” Accessed: Jul. 14, 2025. [Online]. Available: <https://escholarship.org/uc/item/83p5k54m>
- [14] A. Alobaid, M. Kamil, and K. Abdelrazek Khalil, “Metal hydrides for solid hydrogen storage: Experimental insights, suitability evaluation, and innovative technical considerations for stationary and mobile applications,” *Int J Hydrogen Energy*, vol. 128, pp. 432–456, May 2025, doi: 10.1016/J.IJHYDENE.2025.04.232.
- [15] “The Techno-Economics of Hydrogen Pipelines - The Transition Accelerator.” Accessed: Jul. 14, 2025. [Online]. Available: <https://transitionaccelerator.ca/reports/the-techno-economics-of-hydrogen-pipelines/>
- [16] “Gaseous Hydrogen Compression | Department of Energy.” Accessed: Jul. 14, 2025. [Online]. Available: <https://www.energy.gov/eere/fuelcells/gaseous-hydrogen-compression>
- [17] “Liquid Hydrogen Delivery | Department of Energy.” Accessed: Jul. 14, 2025. [Online]. Available: <https://www.energy.gov/eere/fuelcells/liquid-hydrogen-delivery>
- [18] “Use of hydrogen - U.S. Energy Information Administration (EIA).” Accessed: May 25, 2025. [Online]. Available: <https://www.eia.gov/energyexplained/hydrogen/use-of-hydrogen.php>
- [19] “Fuel Cell Benefits - Plug Power.” Accessed: Jul. 14, 2025. [Online]. Available: [https://www.plugpower.com/fuel-cell-power/fuel-cell-benefits/?utm\\_source=chatgpt.com](https://www.plugpower.com/fuel-cell-power/fuel-cell-benefits/?utm_source=chatgpt.com)
- [20] “How hydrogen combustion engines can contribute to zero emissions | McKinsey.” Accessed: Jul. 14, 2025. [Online]. Available: [https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-hydrogen-combustion-engines-can-contribute-to-zero-emissions?utm\\_source=chatgpt.com](https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-hydrogen-combustion-engines-can-contribute-to-zero-emissions?utm_source=chatgpt.com)
- [21] “Fuel Cell Basics | Department of Energy.” Accessed: May 25, 2025. [Online]. Available: <https://www.energy.gov/eere/fuelcells/fuel-cell-basics>
- [22] “PEM Fuel Cells 101: How They Work and Why They Matter - Plug Power.” Accessed: May 25, 2025. [Online]. Available: <https://www.plugpower.com/blog/pem-fuel-cells-101-how-they-work-and-why-they-matter/>
- [23] “Polarization Curves.” Accessed: May 26, 2025. [Online]. Available: <https://www.fuelcellstore.com/blog-section/fuel-cell-information/polarization-curves>

- [24] C. Zhang, Y. Zhang, L. Wang, X. Deng, Y. Liu, and J. Zhang, “A health management review of proton exchange membrane fuel cell for electric vehicles: Failure mechanisms, diagnosis techniques and mitigation measures,” *Renewable and Sustainable Energy Reviews*, vol. 182, Aug. 2023, doi: 10.1016/j.rser.2023.113369.
- [25] H. L. Nguyen, J. Han, X. L. Nguyen, S. Yu, Y. M. Goo, and D. D. Le, “Review of the Durability of Polymer Electrolyte Membrane Fuel Cell in Long-Term Operation: Main Influencing Parameters and Testing Protocols,” *Energies 2021, Vol. 14, Page 4048*, vol. 14, no. 13, p. 4048, Jul. 2021, doi: 10.3390/EN14134048.
- [26] V. Patil *et al.*, “Degradation mechanisms in PEM fuel cells: A brief review,” *Mater Today Proc*, Apr. 2023, doi: 10.1016/J.MATPR.2023.03.603.
- [27] P. C. Okonkwo, I. Ben Belgacem, W. Emori, and P. C. Uzoma, “Nafion degradation mechanisms in proton exchange membrane fuel cell (PEMFC) system: A review,” *Int J Hydrogen Energy*, vol. 46, no. 55, pp. 27956–27973, Aug. 2021, doi: 10.1016/J.IJHYDENE.2021.06.032.
- [28] D. Qiu, L. Peng, X. Lai, M. Ni, and W. Lehnert, “Mechanical failure and mitigation strategies for the membrane in a proton exchange membrane fuel cell,” *Renewable and Sustainable Energy Reviews*, vol. 113, p. 109289, Oct. 2019, doi: 10.1016/J.RSER.2019.109289.
- [29] S. Kundu, M. W. Fowler, L. C. Simon, and S. Grot, “Morphological features (defects) in fuel cell membrane electrode assemblies,” *J Power Sources*, vol. 157, no. 2, pp. 650–656, Jul. 2006, doi: 10.1016/J.JPOWSOUR.2005.12.027.
- [30] S. R. Dhanushkodi, S. Kundu, M. W. Fowler, and M. D. Pritzker, “Study of the effect of temperature on Pt dissolution in polymer electrolyte membrane fuel cells via accelerated stress tests,” *J Power Sources*, vol. 245, pp. 1035–1045, Jan. 2014, doi: 10.1016/J.JPOWSOUR.2013.07.016.
- [31] A. A. Topalov, S. Cherevko, A. R. Zeradjanin, J. C. Meier, I. Katsounaros, and K. J. J. Mayrhofer, “Towards a comprehensive understanding of platinum dissolution in acidic media,” *Chem Sci*, vol. 5, no. 2, pp. 631–638, Dec. 2013, doi: 10.1039/C3SC52411F.
- [32] E. Padgett *et al.*, “Mitigation of PEM Fuel Cell Catalyst Degradation with Porous Carbon Supports,” *J Electrochem Soc*, vol. 166, no. 4, pp. F198–F207, Feb. 2019, doi: 10.1149/2.0371904JES/XML.
- [33] R. Alink, D. Gerteisen, and M. Oszcipok, “Degradation effects in polymer electrolyte membrane fuel cell stacks by sub-zero operation—An in situ and ex situ analysis,” *J Power Sources*, vol. 182, no. 1, pp. 175–187, Jul. 2008, doi: 10.1016/J.JPOWSOUR.2008.03.074.

- [34] J. Liang, L. Fan, Q. Du, Y. Yin, and K. Jiao, “Ice Formation during PEM Fuel Cell Cold Start: Acceptable or Not?,” *Advanced Science*, vol. 10, no. 24, Aug. 2023, doi: 10.1002/ADVS.202302151,.
- [35] J. Zhao and X. Li, “A review of polymer electrolyte membrane fuel cell durability for vehicular applications: Degradation modes and experimental techniques,” *Energy Convers Manag*, vol. 199, p. 112022, Nov. 2019, doi: 10.1016/J.ENCONMAN.2019.112022.
- [36] E. Wallnöfer-Ogris, F. Poimer, R. Köll, M. G. Macherhammer, and A. Trattner, “Main degradation mechanisms of polymer electrolyte membrane fuel cell stacks – Mechanisms, influencing factors, consequences, and mitigation strategies,” *Int J Hydrogen Energy*, vol. 50, pp. 1159–1182, Jan. 2024, doi: 10.1016/J.IJHYDENE.2023.06.215.
- [37] V. F. Valdés-López, T. Mason, P. R. Shearing, and D. J. L. Brett, “Carbon monoxide poisoning and mitigation strategies for polymer electrolyte membrane fuel cells – A review,” *Prog Energy Combust Sci*, vol. 79, p. 100842, Jul. 2020, doi: 10.1016/J.PECS.2020.100842.
- [38] X. Li, “Principles of Fuel Cells,” *Principles of Fuel Cells*, Dec. 2005, doi: 10.1201/9780203942338.
- [39] M. Yilmaz and P. T. Krein, “Review of benefits and challenges of vehicle-to-grid technology,” *2012 IEEE Energy Conversion Congress and Exposition, ECCE 2012*, pp. 3082–3089, 2012, doi: 10.1109/ECCE.2012.6342356.
- [40] S. Sagaria, M. van der Kam, and T. Boström, “Vehicle-to-grid impact on battery degradation and estimation of V2G economic compensation,” *Appl Energy*, vol. 377, p. 124546, Jan. 2025, doi: 10.1016/J.APENERGY.2024.124546.
- [41] D. Ding and X. Y. Wu, “Evaluating the economic and carbon emission reduction potential of fuel cell electric vehicle-to-grid,” *DeCarbon*, vol. 7, p. 100096, Mar. 2025, doi: 10.1016/J.DECARB.2024.100096.
- [42] N. Naik and C. Vyjayanthi, “Optimization of Vehicle-to-Grid (V2G) Services for Development of Smart Electric Grid: A Review,” *2021 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2021*, 2021, doi: 10.1109/SMARTGENCON51891.2021.9645903.
- [43] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science (1979)*, vol. 349, no. 6245, pp. 253–255, Jul. 2015, doi: 10.1126/science.aac4520.

- [44] Z. M. Mosammam, P. Ahmadi, and E. Houshfar, "Multi-objective optimization-driven machine learning for charging and V2G pattern for plug-in hybrid vehicles: Balancing battery aging and power management," *J Power Sources*, vol. 608, p. 234639, Jul. 2024, doi: 10.1016/J.JPOWSOUR.2024.234639.
- [45] N. Mhaisen, N. Fetais, and A. Massoud, "Real-Time Scheduling for Electric Vehicles Charging/Discharging Using Reinforcement Learning," *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies, ICIoT 2020*, pp. 1–6, Feb. 2020, doi: 10.1109/ICIOT48696.2020.9089471.
- [46] O. ; Almughram *et al.*, "A Reinforcement Learning Approach for Integrating an Intelligent Home Energy Management System with a Vehicle-to-Home Unit," *Applied Sciences 2023, Vol. 13, Page 5539*, vol. 13, no. 9, p. 5539, Apr. 2023, doi: 10.3390/APP13095539.
- [47] A. Kumar, J. Hong, A. Singh, and S. Levine, "WHEN SHOULD WE PREFER OFFLINE REINFORCEMENT LEARNING OVER BEHAVIORAL CLONING?," 2022.
- [48] Z. Huang, J. Wang, X. Fan, R. Yue, C. Xiang, and S. Gao, "Economical Electric Vehicle Charging Scheduling via Deep Imitation Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 11, pp. 18196–18210, 2024, doi: 10.1109/TITS.2024.3434734.
- [49] D. Wang, J. Coignard, T. Zeng, C. Zhang, and S. Saxena, "Quantifying electric vehicle battery degradation from driving vs. vehicle-to-grid services," *J Power Sources*, vol. 332, pp. 193–203, Nov. 2016, doi: 10.1016/J.JPOWSOUR.2016.09.116.
- [50] C. Zhang, A. Kotz, K. Kelly, and L. Rippelmeyer, "Development of heavy-duty vehicle representative driving cycles via decision tree regression," *Transp Res D Transp Environ*, vol. 95, p. 102843, Jun. 2021, doi: 10.1016/J.TRD.2021.102843.
- [51] J. Marcinkoski, R. Vijayagopal, J. Adams, B. James, J. Kopasz, and R. Ahluwalia, "Technical Targets for Hydrogen-Fueled Long-Haul Tractor-Trailer Trucks," Oct. 2019. Accessed: Jun. 04, 2025. [Online]. Available: [https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/19006\\_hydrogen\\_class8\\_1ong\\_haul\\_truck\\_targets.pdf?Status=Master](https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/19006_hydrogen_class8_1ong_haul_truck_targets.pdf?Status=Master)
- [52] J. Huya-Kouadio and B. D. James, "Fuel Cell Cost and Performance Analysis," 2023. Accessed: Jun. 04, 2025. [Online]. Available:

- [https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/review23/fc353\\_james\\_2023\\_o-pdf.pdf](https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/review23/fc353_james_2023_o-pdf.pdf)
- [53] H. L. Nguyen, J. Han, H. N. Vu, and S. Yu, “Investigation of Multiple Degradation Mechanisms of a Proton Exchange Membrane Fuel Cell under Dynamic Operation,” *Energies* 2022, Vol. 15, Page 9574, vol. 15, no. 24, p. 9574, Dec. 2022, doi: 10.3390/EN15249574.
- [54] “MTO Traffic Volume Data.” Accessed: Jun. 05, 2025. [Online]. Available: <https://www.library.mto.gov.on.ca/SydneyPLUS/TechPubs/Portal/tp/tvSplash.aspx>
- [55] L. Hu, J. Huang, M. Gallo, and M. Marinelli, “The Impact of Fuel Cell Electric Freight Vehicles on Fuel Consumption and CO2 Emissions: The Case of Italy,” *Sustainability* 2022, Vol. 14, Page 13455, vol. 14, no. 20, p. 13455, Oct. 2022, doi: 10.3390/SU142013455.
- [56] “Waze Reported Traffic Patterns | MTO iCorridor.” Accessed: Jun. 05, 2025. [Online]. Available: <https://icorridor-mto-on-ca.hub.arcgis.com/pages/traffic-report>
- [57] “Technical Targets for Proton Exchange Membrane Electrolysis | Department of Energy.” Accessed: Jun. 05, 2025. [Online]. Available: <https://www.energy.gov/eere/fuelcells/technical-targets-proton-exchange-membrane-electrolysis>
- [58] “Global Hydrogen Review: Assumptions annex.” [Online]. Available: [www.iea.org](http://www.iea.org)
- [59] “Techno-Economics of a New Hydrogen Value Chain Supporting Heavy Duty Transport - The Transition Accelerator.” Accessed: Jun. 05, 2025. [Online]. Available: <https://transitionaccelerator.ca/reports/techno-economics-of-a-new-hydrogen-value-chain-supporting-heavy-duty-transport/>
- [60] “Data Directory.” Accessed: Jun. 05, 2025. [Online]. Available: <https://www.ieso.ca/power-data/data-directory>
- [61] “AESO Application Programming Interface (API) » AESO.” Accessed: Jun. 17, 2025. [Online]. Available: <https://www.aeso.ca/market/market-and-system-reporting/aeso-application-programming-interface-api>
- [62] “Information Document Supply Surplus ID #2011-010R,” Apr. 2024.