Vrije Universiteit Amsterdam

Bachelor Thesis

Can the use of a home battery help reduce a household's CO2 emissions derived from electricity consumption? An analysis using Reinforcement Learning optimization models of the charge-discharge strategy.

Author: Matei Dogariu | *2724342*

1st supervisor: Thilo Kielmann

2nd reader: insert name here

A thesis submitted in fulfilment of the requirements for

the VU Bachelor of Science degree in Computer Science

July 4th , 2024

Abstract

This study investigates the performance of standalone battery storage systems (BSSs) in reducing CO_2 emissions derived from household electricity consumption. We found that overall CO_2 emissions could be reduced by 2.04% annually, and that a 5 kWh battery is most effective when considering the energy intensity of the household. The study was performed on a medium-sized Dutch household with a mean hourly energy consumption of 1.2 kWh. The results were obtained by simulating a household's energy consumption with a BSS using real-world carbon intensity and energy consumption data. In this simulated environment, two reinforcement learning algorithms (PPO and DQN) were trained to find an effective charge-discharge strategy, with the goal of maximizing CO_2 emission savings. Despite the, arguably, small percentage of saved greenhouse gas emissions, a standalone BSS can be installed in any household-type, is cheaper compared to photovoltaic (PV) systems, and can also serve as a backup power system. This study aims to fill the research gap regarding non-PV BSSs' effectiveness in the European region. Lastly, the method of determining the charge discharge strategy using Reinforcement Learning algorithms is novel in this field, as we believe this approach could maximize overall performance.

Table of Contents

1 Introduction

Research on post-industrial age human driven climate change has been well understood for the past few decades. Early on, the concept of energy decarbonization was proposed to be the long-term solution to reducing greenhouse gas emissions (Nakicenovic, 1993). One of the more challenging areas in need of decarbonization is household energy consumption. Beyond transitioning to a ‘greener’ energy mix in electricity grids, a solution needs to be found to the issue of mismatched consumption and production cycles (Munné-Collado et al., 2019). Household energy consumption peaks early in the morning and late-afternoon, with this pattern matching the daily peaks in energy grid carbon intensity (grams of CO2 emitted per kWh), as Zhang (2019) determined that high electricity demand has a direct negative effect on an electricity grid’s carbon intensity. Thus, shifting or spreading the peak demand in the household energy sector is an important step in grid decarbonization.

Literature on the topic mostly explores fitting detached houses with photovoltaic (PV) and battery storage systems (BSS), which show promising reductions in CO2 emissions which vary around 21% on a yearly basis (Khowaja et al., 2022). Although this approach shows promising results, it fails to account for multiple factors which hold it back from being universally viable. Such complex systems are expensive to build or to retrofit to existing housing (Hoppmann et al., 2014), and can only be installed on single family homes (detached houses). This shortcoming of PV systems leads to a large percentage of homes like apartments and multi-family units, which account for 46% of European housing (Eurostat, 2021), to not have a reliable method of energy use decarbonization. Additionally, PV-BSS can only use self-produced solar power, while failing to exploit the variability of grid carbon intensity.

This paper aims to address the above-mentioned drawbacks of residential PV systems by exploring the viability and performance of standalone grid connected BSSs which can efficiently exploit grid carbon intensity variability. This approach would address both the high cost of implementation and reduced use case of PV systems, while spreading the peak household energy demand which would, in itself, decrease a grid's carbon intensity. To investigate the viability of a standalone BSS, two reinforced learning (RL) agents are trained in a simulated environment of a Dutch household's hourly energy consumption and available carbon intensity data. Multiple battery size and carbon intensity scenarios are simulated and will provide a vast-enough learning and testing environment for the RL methods. The goal of the agents is to maximize saved CO2 emissions, when compared to the baseline household emissions without the BSS. The agents' overall performance across the different environment variables will be a good indicator of the general viability, efficiency, and performance of a standalone BSS in household energy decarbonization.

2 Related work

Since their introduction in 1991, the adjusted price of lithium-ion batteries (LIBs) has decreased steeply at around 13% per year (Ziegler & Trancik, 2021). The LIB became the forefront of decarbonization strategies in both the transportation sector (Stampatori et al., 2020) and in electricity grid systems (Tian et al., 2021). The economic and ecological viability of LIB powered electric vehicles (EVs) has already been achieved, according to Nimesh et al., 2021. On the contrary, using LIBs as grid energy storage devices to decrease carbon consumption or to improve load demand response has been shown to be simply impractical due to high implementation costs (Rajeevkumar Urs et al., 2024).

A solution to decarbonizing and improving load demand in electricity grids are household BSSs (Hesse et al., 2017; Leadbetter & Swan, 2012). Leadbetter and Swan assessed that battery sizes of 5kWh fit the needs of a low electricity intensity household, while 22kWh are more appropriate for high intensity residences (those that utilize electric heating systems). These BSSs can either be installed with PV systems, or as standalone systems just connected to the electricity grid.

When considering solar PV systems alongside a residential BSS, research has shown that such integrated systems are remarkably effective at better utilizing self-overproduction of PV energy (Barchi et al., 2018; Quoilin et al., 2016). On the other hand, quantifying the economic viability of household PV-BSS systems has been challenging. According to Hoppmann et al. (2014), the economic viability of PV-BSS systems remains unclear, due to a multitude of factors that can affect the viability computation. These factors include the high price of PVs installation, local weather conditions, grid prices and carbon intensity, household energy intensity, local regulations, and other geographically dependent factors.

With this paper's goal in mind, little research has been done on the effectiveness of decarbonization through standalone residential BSSs, when compared to the amount of literature on more 'ideal' household PV-BSSs. Research by Vejdan et al. (2019) shows that the effectiveness of household BSS depends heavily on the energy mix in local grids. Power grids with a predominant non-renewable energy mix led to no, or negative, impact of a household with only a BSS. In addition, the study found that such systems can also increase total emissions, when connected to grids which have a low carbon intensity variation.

Another paper which supports these findings is the research of Bistline & Young (2020). This study considers the generalization of viability of standalone residential BSSs, by including impact on emissions across multiple regions in the US. In addition, the emission model used is more complex, as it accounts for additional inefficiencies like roundtrip efficiency losses (Bobanac et al., 2022), dispatch effects, and investment effects. Bistline & Young (2020) also conclude that BSSs may or may not lead to increases or decreases of greenhouse gas emissions based on a multitude of influencing factors.

The most recent direct study on the viability of household energy decarbonization through grid-connected BSSs furthers research in the space by investigating whether used electric vehicle batteries can be repurposed as BSSs (Khowaja et al., 2022). The author's approach aims to increase an EVs battery's useful lifespan, to avoid premature obsoletion and to decrease initial costs and effective emissions when compared to using a completely new BSS. The study looks at the effectiveness of both BSS systems with and without photovoltaic energy generation. In the case of grid-connected BSS, their method of charge/discharge optimisation is a simple, mathematical, approach summarised as 'charge when emission factors are low' and 'discharge when emission factors are high' ('high' and 'low' when compared to the daily, cyclical, distribution of emission factors). Khowaja et al. (2022) found through this approach that around 0.12 tons of CO_2 were saved in 2018 when averaging six homes. Additionally, the authors found that the net carbon savings fluctuated throughout the year, with maximums in the spring season, and minimum in the fall season. Thus, this paper also supports the statement that the largest impact on CO_2 savings is the carbon intensity variation throughout days and seasons.

2.1 Motivation

Considering Hoppmann et al. (2014) show that the economic viability of household PV systems remains uncertain, it is important to recognize that a simpler to implement, and universally adoptable, solution to household energy usage must be researched. Another important factor missed in the papers by Vejdan et al. (2019) and Bistline & Young (2020) is the inability to install a PV system on any household except for the detached house. According to the most recent study on population residence (Eurostat, 2021), 46% of Europeans live in flats. These types of homes can never be retrofitted with PV systems, so other ways of decarbonizing their energy consumption must be explored. With this goal in mind, this paper tries to fill the research gap concerning the efficiency in reducing household CO2 emissions with a grid connected BSS. Advantages of such systems include lesser cost of implementation (when compared to traditional PV systems), less space requirement, less complexity and universal compatibility with any grid connected home type. Additionally, these systems can be deployed in multi-unit homes, where no other decarbonization methods can be retrofitted and used. The overall cost and effective environmental impact of a BSS can be improved further by upcycling out of warranty EV battery stacks, as outlined by Khowaja et al. (2022). The objective of a grid connected BSS is to store low carbon intensive energy from the grid to better address high energy demand moments in a day which coincide with high carbon intensity, thus reducing overall household CO2 emissions. More specifically, we investigate how well such battery systems can perform across multiple environmental characteristics.

3 Methodology

To investigate whether a home battery can be used to aid in the decarbonization of household energy consumption, a computational modeling approach was used. The method employed in this paper involves the design of a household-like environment following the standard created by OpenAI's Gymnasium framework with real-world data on energy consumption and carbon intensity. In this standardized environment, two reinforcement learning algorithms were trained with the goal to minimize the effective CO2 emissions using the battery's charge/discharge cycle. The following section will provide, in addition to an in-depth description of the proposed environment, a high-level overview of the RL agents used in training, as deep understanding of the two RL algorithms is not necessary for this study. Note that StableBaseline3's implementation of the DQN and PPO agents was used, as both agents have interfaces directly compatible with Gymnasium-compliant environments

3.1 Carbon Intensity and Electricity Consumption Datasets

The study is performed using data specific to the Netherlands region. Carbon intensity live and historical datasets, from 2021 onwards, are freely available for research purposes from the Electricity Maps (2024) data portal. The metric used as an environmental parameter is 'LCA Carbon Intensity', which takes a holistic approach at assessing the effective carbon intensity. The carbon intensity dataset contains values from the year 2021 and resolution is hourly.

In contrast, residence-specific energy consumption data is difficult to access, as it is protected by EU privacy regulations. Through the research of Uttama Nambi et al. (2015), a high-resolution (1 second) dataset (Dutch Residential Energy Dataset) of the energy consumption of a single household is publicly available. Ultimately, the chosen resolution was one hour, as that is the maximum resolution in which electricity grid carbon intensity is available. The processed DRE dataset contains datapoints from 5th of July 2015 to 5th of December 2015.

With respect to having the same year for both datasets, we assess that the results presented are still scientifically accurate as the environment always picks the same hour of the same day from both datasets. This is the only way to account for seasonal differences and keep scientific relevance for the area of the Netherlands, even though older carbon intensity data or newer household energy consumption data that match were not available.

3.2 Gymnasium Environment

The basics of Reinforcement Learning place an agent inside an environment, who is tasked with making decisions while being given observations and rewards. The Gymnasium toolkit was selected in the creation of the custom environment because it has well-written documentation and is the most used Python environment-building library (Brockman et al., 2016; Körber et al., 2021). The framework's goal is to have a multitude of environments with the same interface, such that it acts both as a collection of benchmarks and as an accessible springboard for RL applications. To expand on the latter point, the framework also supports the building of custom environments using its default classes as building blocks, through which we constructed this household carbon emission environment.

All Gymnasium environments, including this one, interact with the agent as follows: the agent makes an action for each step, after which it receives an observation and a reward. Reinforcement Learning theory states that environments must follow the Markov Property, where each of their states summarize completely all relevant information (Sutton & Barto, 2015). I.e., the system should provide concise information about the future of the system in a way that is independent of previous states. The response of such Markovian environment at time $t + 1$ must only depend on the action taken and the environment's response at time t. Sutton and Barto provide a mathematical representation of a Markovian environment:

$$p(s', r \mid s, a) = \Pr\{ R_{t+1} = r,\ S_{t+1} = s' \mid R_t, S_t \} \qquad \textbf{(3.1)}$$

Where p is the probability of transitioning to state s' with reward r, from state s with action a, and Pr is the complete probability distribution. Note how Pr at $t + 1$ only depends on the state and action taken at t. In real-world applications the Markovian property is maintained only as an assumption. In essence, the state needs to be as good as possible of a basis to predict future actions and rewards (Sutton & Barto, 2015).

The environment has two collections, namely the action space and the observation space. In the case of our custom environment implementation the action space is discrete, having the values 0, 1 and 2. Each is mapped to the possible actions 'Discharge', 'Nothing' and 'Charge'. The observation space is continuous, and contains data about carbon intensity, energy consumption, battery charge value and emissions delta.

Lastly, the simulated environment rewards agents when the maximum absolute value of the difference between emissions emitted with the battery and baseline emissions increases. A formalized mathematical description of the proposed reward function:

$$r = \begin{cases} 1, & \Delta_{t+1} > \max(\Delta_t) \\ -0.01, & otherwise \end{cases} \tag{3.2}$$

$$\Delta_t = E_{baseline} - E_{battery} \tag{3.3}$$

In other words, the goal of the agent is to maximize saved CO2 emissions in the household environment, when compared to the same system without a battery.

3.3 Deep Q-Network Algorithm

Q-Learning is a primitive, off-policy, RL method which has as goal to converge towards Q, which is an approximation for the optimal action-value function q (Sutton & Barto, 2015; Watkins & Dayan, 1992). This strategy matches the Markov Property of environments that each state only depends on the previous one, and thus all have an optimal action which can be derived from the q function. Unlike most RL algorithms, Q-Learning does not have the maximization of the reward as goal. This algorithm approaches the Q function if and only if all state-action pairs are updated throughout, which is a trivial requirement (Sutton & Barto, 2015; Watkins & Dayan, 1992).

The modern DQN algorithm uses a neural network, instead of a lookup table, to approximate the Q function. Additionally, it assumes future rewards are 'worth' less than immediate ones, with future rewards being discounted by a factor γ per step (Mnih et al., 2013). One limitation of DQN that impacts its performance on the presented task is its relative sample inefficiency due to its inability to perform enough exploration, which leads to suboptimal policies (Jin et al., 2018). This limitation may be especially impactful to the task at hand, as charging the battery inherently leads to negative immediate rewards, and at timestep 0 the battery is fully discharged. An optimal agent for our modelled environment must learn to balance late future discounted return and effective exploration of the domain.

3.4 Proximal Policy Optimization Algorithm

PPO is a modern implementation of vanilla policy gradient methods, introduced by Schulman et al. (2017). Alike Q-Learning, this family of algorithms also relies on the Markov Decision Process assumption. All policy gradient algorithms alter their parametrized policy to optimize for long-term cumulative reward. How this algorithm improves is by simplifying the computational complexity of the objective optimization; from a second-order derivative matrix which is computationally intensive, to first order minibatch Stochastic Gradient Descent which is easier to compute. This simplification can be achieved by changing a hardline constraint to a penalty in its policy optimization (soft constraint). This change achieves two objectives, increased exploration rate through sub-optimal decisions, and considerably faster convergence rate. In simpler terms, PPO approximates the second-order derivative using SGD, which can lead to 'wrong steps', but is considerably faster to compute. To decrease the likelihood of these 'wrong steps', a penalty is added to the objective function. This was shown, empirically, to increase viability and accuracy on both large- and small-scale problems. Schulman et al. (2017) found their PPO algorithm improves significantly in sample complexity and overall performance in ATARI games, when compared to stable versions of ACER(Wang et al., 2016) and A2C (Mnih et al., 2016). As reduced sample complexity and comprehensive exploration are advantages of PPO, the algorithm is likely well-suited for the task explored in this paper.

4 Results

The viability of a grid-connected BSS can be determined by calculating the percentage of effective CO2 emissions saved over different time frames. The analysis of results across seasons and battery sizes can lead to a more comprehensive answer to whether the systems in question are effective at household decarbonization in the Netherlands. Important to note that 'N/A' values correspond to negative CO2 gains, i.e., the BSS emits more CO2 in the set time frame compared to the baseline CO2 emissions. The reason for having 'N/A' entries instead of the resulting values is that negative CO2 gains are predominantly caused by inadequacies in the RL models.

4.1 Seasonality Comparison

Table 1 showcases the percentage of saved CO2 emissions distributed across seasons. For each season, five models are trained for each model type. Among those, the model with the highest mean episodic reward is selected. The results are obtained by averaging the result of 30 episodes run by the two selected PPO and DQN models. Battery size is 5kWh and the hourly charge/discharge rate is $\frac{battery\ size}{4}$.

	Spring	*Summer*	*Fall*	*Winter*	*Overall*
PPO	3.02%	2.14%	1.20%	1.73%	2.04 %
DQN	0.06%	N/A	N/A	N/A	N/A

Table 1 - Percentage of CO2 savings across RL methods and seasons

4.2 Battery Size Comparison

Table 2 showcases the percentage of saved CO2 emissions across multiple battery sizes. For each battery size five instances of each model are trained; the one with highest mean episodic reward is selected. The presented results are obtained by averaging the results across 30 episodes for each RL algorithm and battery size. The charge/discharge rate remains $\frac{battery\ size}{4}$.

	1 kWh	*3 kWh*	*5 kWh*	*7 kWh*	*10 kWh*
PPO	0.61%	1.81%	1.93%	0.97%	1.17%
DQN	N/A	N/A	N/A	N/A	N/A

Table 2 - Percentage of CO2 savings across RL methods and battery sizes

4.3 One Run PPO

To showcase in more detail how a successful simulation looks like, an above average (in performance) run is displayed in Figures 1 and 2. In the 720-hour timeframe (30 days) simulated, the total baseline emissions were 312.68 kg of CO2. With the BSS, 9.68 kg of CO2 were emitted less than the baseline. This leads to a reduction of 3.09% in effective carbon dioxide emissions.

In Figure 1 the reward pattern of the algorithm can be seen. With each new maximum in saved CO2 emissions, the algorithm receives a positive reward. An upward shape with consistent slope is desired, as it means the algorithm consistently found CO2 emissions gains. Similarly, Figure 2 represents the hourly difference between the two running emission values. Here, the actual charge/discharge strategy can be observed. Positive slope means battery was discharged; null slope means no action taken; negative slope means battery was charged.

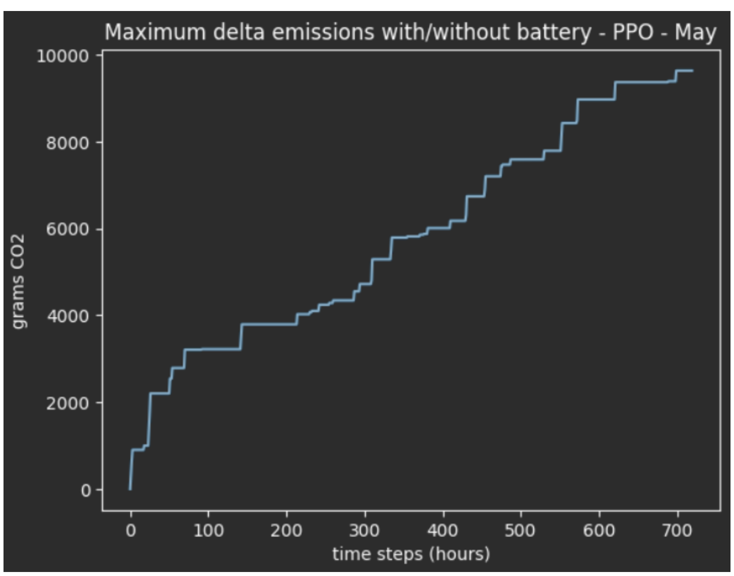

Figure 1

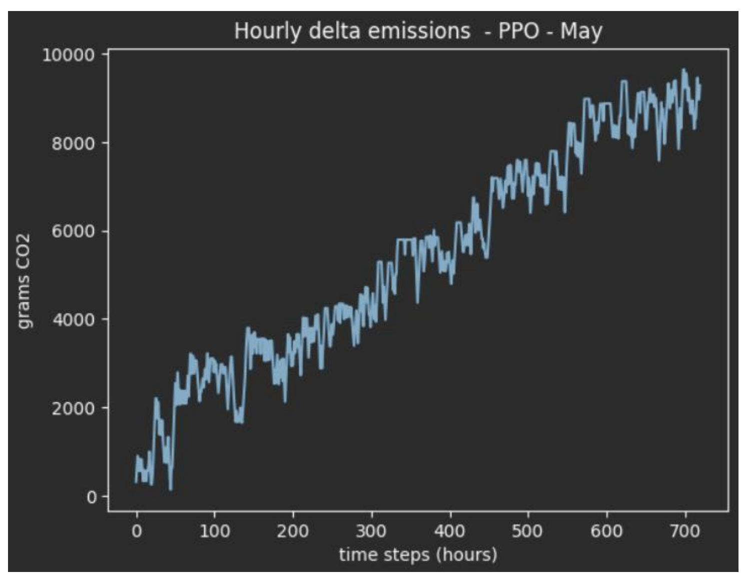

Figure 2

5 Discussion

To address the multiple ‘N/A’ entries in Tables 1 and 2, the Deep Q-Network algorithm simply could not explore enough of the domain space to find adequate strategies. The DQN algorithm, like all RL methods, has stochastic exploration built in. However, it can only perform well if most of the exploration is done by the first 30% of timesteps. Beyond the exploration fraction threshold, the algorithm has between 1-5% chance of making a random step, thus adjusting such complex strategy, as is needed here, almost never happens. Because of this, almost all DQN models trained were inadequate at performing the task at hand. In most cases, the models’ strategy would cause more CO_2 emissions than the baseline. This behaviour kept repeating no matter the hyperparameters nor the amount of time steps for which the model was trained. Thus, the ‘N/A’ results do not represent the inability of a BSS to save CO_2 emissions, but rather the inability of the DQN algorithm to find a strategy which can save emissions.

When considering what results would validate/invalidate the goals and assumptions of this paper, it is important to note that the environment which the agents are trying to solve is ultimately a real-world simulation. This aspect of the investigation warrants some consideration in two ways. On the one hand, having above expectation results can be caused either by omitting some real-world inefficiencies or the instance of the RL tool performed above average. On the other hand, having poor results does not entail that the system cannot be ‘solved’, but rather that the tools (RL methods) used were inadequate for the task. To better grasp the conditions for the interpretation of the results, we can substitute our BSS environment with a theoretical racetrack. The racetrack is, obviously, a real-world entity that behaves predictably, just like the BSS simulation. In this theoretical case, this paper would aim to firstly investigate whether it is feasible to complete laps around the racetrack, and if so, to then find out how fast different cars could complete the track. Alike our results, not being able to finish a lap around the racetrack does not mean the track is faulty in some way, but rather that the cars were poorly equipped for the task to begin with. In the same vein, having surprisingly fast lap-times does not entail the track is easier to drive around than expected, but rather that corners were possibly cut.

When investigating the performance of the models in saving CO2 across seasons, the PPO results align with similar results from established literature on the subject matter (Khowaja et al., 2022). The model consistently performed the best in the months of March through May, while performing the worst in September through November. On average, our PPO model was able to save around 70 kg, or 2.04%, out of yearly CO2 emissions of 3534 kg. We attribute the fluctuations in performance to the changing variance of electricity grid energy mix throughout the year, which is directly correlated with the carbon intensity data used in this study. Comparing a standalone BSS with another which uses on-site renewables (PV or wind-capture energy) to charge, the BSS performs considerably worse solely at offsetting CO2 emissions. At the same time, such system does not need to only be used for CO2 reduction purposes. First of all, a household battery system can also serve as backup power in case of power outages, or for energy arbitrage purposes (monetary profit from fluctuating energy rates). Second of all, when considering a more holistic approach to reducing CO2, repurposed automotive Lithium-Ion Batteries have been shown to decrease the overall emissions impact, and to significantly increase the household viability of a BSS (Khowaja et al., 2022).

Regarding the different battery size investigation, the same case as above applies for the Deep Q-Networks methods; it remains unsuitable for the task at hand. Regardless, the results obtained for the PPO algorithm in this section can be broken down further. For battery sizes from 1 to 5-kWh the PPO algorithm performed in a regular fashion, with consistent upward slopes in the resulting emission difference graphs. The mismatch between the 2.04% overall value resulted from the first part of the investigation and the 1.93% result from the 5-kWh battery size test is due to the stochastic nature of the PPO algorithm. Even though the same model was used for both investigations with identical parameters, the difference in overall outcome stands at 5.5%. Considering that the household from which electricity consumption data was derived is considered a low energy intensity household (Uttama Nambi et al., 2015), our results coincide with those of Leadbetter & Swan (2012) that a 5-kWh BSS fits the needs of a low electricity intensity household best. However, the results for the 7 and 10-kWh battery sizes uncover an issue with the training of the RL agents. With battery sizes 7 to 10 times that of the hourly consumption of the household, the PPO algorithm struggled to understand the observation space because of how much charging the battery would affect emissions in the short term. Thus, these surprisingly low results are not representative of the actual possible real-world reduction in CO2.

Due to their relative inefficiency, it is unlikely for a household to install a new standalone BSS just for 2% CO2 savings annually. Although, if certain governmental incentives are set-up for homeowners, and considering both the extra duty as electricity backup and the possibility of upcycling out of commission EV LIBs, then the question of the real-world use case of household BSSs remains an open one.

5.1 Limitations

While this research showcases the theoretical potential of grid-connected BSSs to reduce household CO2 emissions, there are some limitations that would hinder such system in a practical environment. Despite using real-world data on carbon intensity and energy consumption, energy transfer inefficiencies of LIBs were not taken into account. Currently, Lithium-Ion batteries have around 90% roundtrip efficiency rates (Bistline & Young; 2020).

Another limitation of this study is the unavailability of newer and more varied energy consumption data. Because electricity data from a single household was used, consumption patterns should not be generalized from the findings of this study. Additionally, life cycle analysis emissions of the BSS was not considered, i.e., maintenance periods and early degradation due to intensive use. Lastly, the RL agents selected in this study were not fully appropriate in solving the task at hand (mostly the DQN agent), as a custom PPO implementation would probably be more efficient at learning a more optimal charge/discharge strategy.

All of the mentioned factors would impact the overall effectiveness of the BSS in reducing CO2 emissions; efficiency rates and lifetime degradation of the BSS reducing performance, while better underlying data and RL agents would increase performance.

5.2 Future Research

Future work should use a purpose-built implementation of the Proximal Policy Optimization that would allow for better training times and longer time intervals than what was possible in this study. Additionally, considering battery roundtrip efficiency and expanding the study to take into account electricity consumption data from multiple households would considerably increase the generalizability of the obtained results. An alternative area that is in need of further exploration could be in quantifying the effective reduction in greenhouse gas emissions from repurposing EV batteries as household BSSs.

6 Conclusion

This study has quantified the potential reduction in household CO2 emissions by using a grid-connected battery storage system (BSS). The results of this study can suggest the real-world viability of BSSs in areas where electricity grids have high variability in carbon intensity.

However, our estimations could be improved by also taking into account battery inefficiencies, newer and more varied data, and LCA costs associated with using a BSS. When comparing a standalone BSS with one connected to PV, its performance in reducing CO2 emissions is certainly inferior. On the other hand, a standalone BSS has less up-front cost of installation, is compatible with all household types (including apartments) and can have additional purposes as backup power storage. We expect this study to aid in providing emission-conscious homeowners a better-defined outlook at the properties of standalone BSSs, which we assess is an area which is in need of further, in-depth, research.

References

Barchi, G., Miori, G., Moser, D., & Papantoniou, S. (2018). A Small-Scale Prototype for the Optimization of PV Generation and Battery Storage through the Use of a Building Energy Management System. *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 1–5. https://doi.org/10.1109/EEEIC.2018.8494012

Bistline, J. E. T., & Young, D. T. (2020). Emissions impacts of future battery storage deployment on regional power systems. *Applied Energy*, *264*, 114678. https://doi.org/10.1016/j.apenergy.2020.114678

Bobanac, V., Bašić, H., & Pandžić, H. (2022). One-way voltaic and energy efficiency analysis for lithium-ion batteries. *13th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2022)*, 261–266. https://doi.org/10.1049/icp.2023.0003

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). *OpenAI Gym*. https://arxiv.org/abs/1606.01540

Electricity Maps. (2024). Netherlands 2021 Hourly Carbon Intensity Data (Version January 17, 2024). In *Electricity Maps Data Portal*.

Eurostat. (2021). *House or flat: where do you live?* https://doi.org/10.2908/ILC_LVHO01

Hesse, H., Martins, R., Musilek, P., Naumann, M., Truong, C., & Jossen, A. (2017). Economic Optimization of Component Sizing for Residential Battery Storage Systems. *Energies*, *10*(7), 835. https://doi.org/10.3390/en10070835

Hoppmann, J., Volland, J., Schmidt, T. S., & Hoffmann, V. H. (2014). The economic viability of battery storage for residential solar photovoltaic systems – A review and a simulation model. *Renewable and Sustainable Energy Reviews*, *39*, 1101–1118. https://doi.org/10.1016/j.rser.2014.07.068

Jin, C., Allen-Zhu, Z., Bubeck, S., & Jordan, M. I. (2018). *Is Q-learning Provably Efficient?* https://arxiv.org/abs/1807.03765

Khowaja, A., Dean, M. D., & Kockelman, K. M. (2022). Quantifying the emissions impact of repurposed electric vehicle battery packs in residential settings. *Journal of Energy Storage*, *47*, 103628. https://doi.org/10.1016/j.est.2021.103628

Körber, M., Lange, J., Rediske, S., Steinmann, S., & Glück, R. (2021). *Comparing Popular Simulation Environments in the Scope of Robotics and Reinforcement Learning*. https://arxiv.org/abs/2103.04616

Leadbetter, J., & Swan, L. (2012). Battery storage system for residential electricity peak demand shaving. *Energy and Buildings*, *55*, 685–692. https://doi.org/10.1016/j.enbuild.2012.09.035

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T. P., Harley, T., Silver, D., & Kavukcuoglu, K. (2016). *Asynchronous Methods for Deep Reinforcement Learning*. https://arxiv.org/abs/1602.01783

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). *Playing Atari with Deep Reinforcement Learning*. https://arxiv.org/abs/1312.5602

Munné-Collado, I., Aprà, F. M., Olivella-Rosell, P., & Villafáfila-Robles, R. (2019). The Potential Role of Flexibility During Peak Hours on Greenhouse Gas Emissions: A Life Cycle Assessment of Five Targeted National Electricity Grid Mixes. *Energies*, *12*(23), 4443. https://doi.org/10.3390/en12234443

Nakicenovic, N. (1993). *Decarbonization as a long-term energy strategy.* https://pure.iiasa.ac.at/id/eprint/12727/1/decarbonization1994.pdf

Nimesh, V., Kumari, R., Soni, N., Goswami, A. K., & Mahendra Reddy, V. (2021). Implication viability assessment of electric vehicles for different regions: An approach of life cycle assessment considering exergy analysis and battery degradation. *Energy Conversion and Management*, *237*, 114104. https://doi.org/10.1016/j.enconman.2021.114104

Quoilin, S., Kavvadias, K., Mercier, A., Pappone, I., & Zucker, A. (2016). Quantifying self-consumption linked to solar home battery systems: Statistical analysis and economic assessment. *Applied Energy*, *182*, 58–67. https://doi.org/10.1016/j.apenergy.2016.08.077

Rajeevkumar Urs, R., Mussawar, O., Zaiter, I., Mezher, T., & Mayyas, A. (2024). Navigating the Cost-Efficiency Frontier: Exploring the viability of Grid-Connected energy storage systems in meeting district load demand. *Energy Conversion and Management*, *299*, 117828. https://doi.org/10.1016/j.enconman.2023.117828

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). *Proximal Policy Optimization Algorithms*. https://arxiv.org/abs/1707.06347

Stampatori, D., Raimondi, P. P., & Noussan, M. (2020). Li-Ion Batteries: A Review of a Key Technology for Transport Decarbonization. *Energies*, *13*(10), 2638. https://doi.org/10.3390/en13102638

Sutton, S. R., & Barto, G. A. (2015). *Reinforcement Learning: An Introduction* (2nd ed.). The MIT Press. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Tian, Y., Zeng, G., Rutt, A., Shi, T., Kim, H., Wang, J., Koettgen, J., Sun, Y., Ouyang, B., Chen, T., Lun, Z., Rong, Z., Persson, K., & Ceder, G. (2021). Promises and Challenges of Next-Generation "Beyond Li-ion" Batteries for Electric Vehicles and Grid Decarbonization. *Chemical Reviews*, *121*(3), 1623–1669. https://doi.org/10.1021/acs.chemrev.0c00767

Uttama Nambi, A. S. N., Reyes Lua, A., & Prasad, V. R. (2015). LocED. *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 45–54. https://doi.org/10.1145/2821650.2821659

Vejdan, S., Kline, A., Totri, M., Grijalva, S., & Simmons, R. (2019). Behind-the-Meter Energy Storage: Economic Assessment and System Impacts in Georgia. *2019 North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS46351.2019.9000287

Wang, Z., Bapst, V., Heess, N., Mnih, V., Munos, R., Kavukcuoglu, K., & de Freitas, N. (2016). *Sample Efficient Actor-Critic with Experience Replay*. https://arxiv.org/abs/1611.01224

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3–4), 279–292. https://doi.org/10.1007/BF00992698

Zhang, H. (2019). Effects of electricity consumption on carbon intensity across Chinese manufacturing sectors. *Environmental Science and Pollution Research*, *26*(26), 27414–27434. https://doi.org/10.1007/s11356-019-05955-9

Ziegler, M. S., & Trancik, J. E. (2021). Re-examining rates of lithium-ion battery technology improvement and cost decline. *Energy & Environmental Science*, *14*(4), 1635–1651. https://doi.org/10.1039/D0EE02681F