Automating Storage Arbitrage in German Electricity Market

Mariia Bilousova[1,2,3](✉), Anton Motornenko[1], and Fabian Hofmann[1]

[1] Frankfurt Institute for Advanced Studies, Ruth-Moufang-Straße 1, 60438 Frankfurt am Main, Germany
[2] efl - the Data Science Institute, House of Finance Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main, Germany
[3] Goethe-Universität Frankfurt Fachbereich Wirtschaftswissenschaften, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, Germany

Abstract. This study examines the potential of energy arbitrage in the German electricity market as a way to increase the return on investment of battery storage technologies. The main goal is to develop and estimate the performance of automated arbitrage strategies for households using Tesla Powerwall energy storage. Based on historical prices of the German intraday electricity market, artificial intelligence algorithm is developed to find feasible charging and discharging strategies for battery storage. This is done by employing a Deep Q-Learning approach of Reinforcement Learning. As a baseline, a simple Expert System algorithm is suggested, that is based on buy/sell at fixed price approach. The maximal possible return from the arbitrage is explored by a linear optimization of the system under perfect price foresight. The Reinforcement Learning algorithm is found to achieve only $\sim 35\%$ of the maximal return which is only 5% more than the simplistic Expert system. Finally, the performance of both algorithms is compared to the already available results at other electricity markets.

Keywords: Energy storage arbitrage · Artificial intelligence · Reinforcement Learning · Expert System · Pypsa

1 Introduction

In an effort to mitigate climate change, the German government has established plans to increase the share of renewable energy in gross electricity consumption to at least 80% by 2030 and to reduce greenhouse gas emissions by 95% by 2050 compared to 1990 levels. At the beginning of 2020, 43% of the German power has been produced from renewable sources, e.g., wind, sun, water, and biomass. While conventional power plants traditionally provide the backup infrastructure to balance out mismatches between renewable power supply and demand, other balancing schemes have to step in place in order to reach these long term goals. Most importantly, an expansion of power transmission and energy storage is planned to shift energy in space and time more flexibly. However, in contrast

P. Schossig et al. (Eds.): IRES 2022, AHE 16, pp. 63–75, 2023.
https://doi.org/10.2991/978-94-6463-156-2_6

to transmission grid expansion, the expansion of storage technologies requires shorter planning phases as well as less inter-regional coordination. Storage technologies also come along with a list of system-beneficial qualities such as support for frequency regulation, voltage support as well as grid stabilization [Eyer and Corey, 2011]. In particular, fast reacting storage technologies as batteries have noiseless operation, low maintenance, high efficiency, and few installation constraints [Leadbetter and Swan, 2012].

The main concern about battery energy storage is its cost. The current design of the power market does not offer sufficient financial incentive for the essential expansion of the storage. To attract more investments in energy storage units, special tools and policies that stimulate sufficient return on investment are required. Primary economic advantage of the storage is that it can make use of the energy price fluctuations by charging at times when there is an excess of energy on the market and therefore the prices are low, and selling the energy at peak load times when the prices are high, known as energy arbitrage. Under current market setup, arbitrage is one of the major revenue sources for storage in electricity markets [Krishnamurthy et al., 2017]. Development of storage technologies in the recent decade and increase of its deployment among commercial consumers as well as households, prompted multiple studies to find an optimal strategy for storage behavior, especially, examining possibilities of arbitrage as a way to maximize the returns. Various optimization approaches have been applied to the problem: Linear Programming models, Mixed Integer Linear Programming [Krishnamurthy et al., 2017], Stochastic Optimization model [Shang and Sun, 2016], as well as more sophisticated techniques, such as Reinforcement Learning [Cao et al., 2020],[Harrold et al., 2021]. Even though, existing literature recognizes that arbitrage operations did not break-even in the recent past [Metz and Saraiva, 2018], mainly, due to high investment cost, it is expected that price arbitrage could play an important role in the competitiveness of energy storage in the nearest future [Campana et al., 2021]. Positive developments in arbitrage are the wide-spread occurrences of negative electricity prices in the recent years in some electricity markets, including Germany [Metz and Saraiva, 2018]. Some research [Wankmüller et al., 2017] suggests that storage facilities that pursue arbitrage purposes should use only the most profitable opportunities to increase the Net Present Value, considering the life span of the battery. Currently, one of the most promising approaches to solve the problem of profit maximisation in arbitrage is Reinforcement Learning (RL), specifically, Deep Reinforcement Learning (DRL). DRL combines the advantages of Reinforcement Learning and Deep Learning and it has already achieved convincing results in the field of Artificial Intelligence (AI) [Lee and Choi, 2019].

In this paper, the possibility of energy storage arbitrage as a way to increase the return on investment in the German intraday electricity market is studied. First, the maximal possible return from the energy arbitrage is estimated using historical data by performing linear optimization of the problem given perfect price foresight. This is then used as an idealistic upper bound for different arbitrage strategies to be used under uncertainty. The main objective of the study is to develop and estimate the performance of automated arbitrage strategies. First, an Expert System is introduced as a baseline algorithm to estimate the return using a simplistic arbitrage strategy. Then, an advanced Deep Reinforcement Learning approach for arbitrage is presented. Finally, the performance of all algorithms with respect to each other is estimated and compared to the already available results at other electricity markets.

2 Modeling

2.1 Theoretical Background

To make use of energy arbitrage, a storage device buys power at times when the prices are low and sells it at times when the prices are high. Therefore the profit f over T discrete time steps with an assigned duration of Δt is given by

$$f(d) = \sum_{t=1}^{T} \eta\, p(t)\, d(t)\, \Delta t \tag{1}$$

where $d(t)$ is the power dispatch of the storage device in kW and $p(t)$ is the electricity price in €/kWh at time t. The objective of an operating strategy is to maximize the profit, i.e.

$$\max_{d} f(d) \tag{2}$$

The optimal dispatch $d(t)$ defines at which rate the storage should be charged ($d(t) < 0$), discharged ($d(t) > 0$), or kept idle ($d(t) = 0$) to gain profit from the energy trading. The state of charge of the battery $e(t)$, which corresponds to the net amount of energy stored in the storage at t, is defined as

$$e(t) = \sum_{t'=1}^{t} \eta\, \Delta t\, \max\left(d(t'), 0\right) + \frac{1}{\eta}\, \Delta t\, \min\left(d(t'), 0\right) \tag{3}$$

where η describes the charging and discharging efficiencies and η^2 is the round-trip efficiency. The energy storage is bound to a maximal storage capacity $e_{\max}$, i.e.

$$e(t) < e_{\max},\ \forall t \tag{4}$$

In the same way, the dispatch is bound to a maximal discharging and charging capacity, i.e.

$$d_{\min} \leq d(t) \leq d_{\max},\ \forall t \tag{5}$$

Unless otherwise stated, the discharge is assumed to have only three possible values $d = [d_{\min}, 0, d_{\max}]$, which means that charge/discharge occurs always at maximal rate. Moreover, the maximal charge and discharge rates are considered to have the same power P, $-d_{\min} = d_{\max} = P$.

2.2 Input Data

For this study, the European Power Exchange (EPEX SPOT) intraday historical market data for Germany for the years 2019 and 2020 were chosen as price data $p(t)$ [epexspot, 2020]. In the Day Ahead (DA) and Intraday (ID) electricity market, prices can fall below zero. This happens when highly inflexible power generation (including weather-dependent renewables such as wind and solar) coincides with low demand. With the rising share of renewable energy in the market, occurrence of negative prices has become more frequent. Secondly, the trade on the ID market goes on continuously. ID market serves to balance out the real-time demand and allows last minute adjustments. Electricity can be traded up to 5 min before delivery through hourly, half-hourly or quarter-hourly contracts. Therefore, it is the most volatile market with the highest range of price fluctuations, see Fig. 1 where the two market prices are compared. These two factors make ID market the most appealing for the arbitrage. To align the model with the input price data, the time-dependence of all quantities is set to a temporal resolution of $\Delta t = 15\,\text{min}$.

The technical parameters of the battery storage is aligned to reported values of the Tesla Powerwall 2 which can be considered representative for anal-

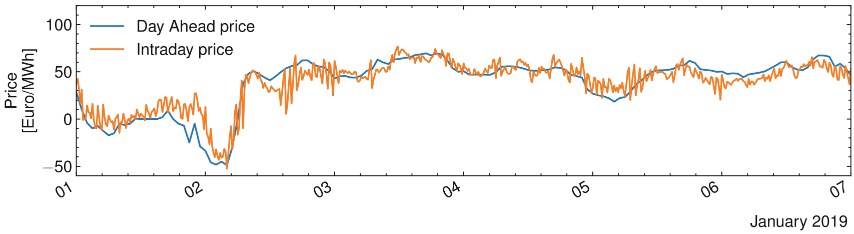

Fig. 1. Comparison of Day Ahead and Intraday electricity prices for the first week of January 2019.

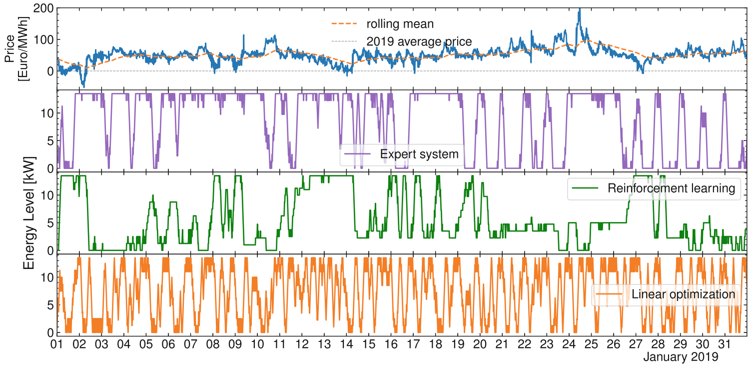

Fig. 2. Storage state of charge with the linear optimization for January 2019. The results correspond to the optimization shown in Fig. 3.

ogous home energy storage solutions. It has a storage capacity of 13.5 kWh and a dispatch capacity of 5 kW. The battery has 100% depth of discharge and 90% round trip efficiency [Tesla, 2022]. The warranty for non-solar self-consumption is 37.8 MWh of aggregate throughput, which equals to 2800 full charging/discharging cycles. At the time of writing, the price of the battery including software and installation is about 12200 EUR. In the following chapters, unless stated otherwise, the parameters of the Tesla Powerwall 2 will be used for the storage, i.e. $P = 5$ kW, $E_{\max} = 13.5$ kWh and $\eta = 0.9$.

3 Results

3.1 Linear Optimization

The maximal possible profit reachable in a setup presented in Sect. 2.1 can be calculated using linear optimization approach. The computational implementation uses the PyPSA package [Brown et al., 2018] as an interface to create the linear problem which is solved using the Gurobi solver.

The results of the linear optimization are presented in Figs. 3, 2. The optimization suggests that the maximal possible profit in 2019 is 337 € for energy arbitrage with the Tesla Powerwall 2. The profit is monotonously and almost linearly accumulated through the whole year as can be seen from Fig. 3. This can be achieved with rather active and rapid response to the electricity price dynamics. Figure 2 illustrates the response of the storage to the electricity price dynamics. One can observe that in order to obtain the maximal profit the storage is almost never idle and reacts to almost all electricity price changes by charging or discharging.

3.2 Expert System

An Expert System (ES) is a computer program that reproduces the behavior of a human expert in a distinct domain of knowledge [Liebowitz, 1995]. ES have been successfully applied in various spheres since 1980-s and have emerged as useful,

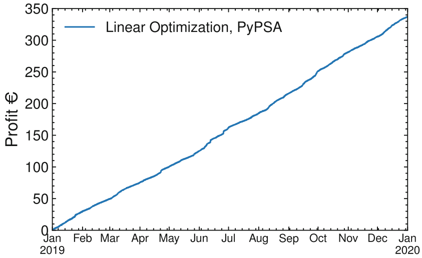

Fig. 3. Maximal possible profit gained in one year from energy arbitrage with the Tesla Powerwall 2 storage. Electricity prices are from the EPEX intraday German market data for year 2019. The maximal profit from the energy arbitrage with given storage is 327 €.

deployable systems worldwide [Liebowitz, 1991]. For the given energy arbitrage problem, a model that sells and buys energy at a fixed threshold was set up. In such approach, the energy dispatch solely depends on price dynamics and has a completely determined behavior. The used Expert System can be formulated with the following rules for energy dispatch:

$$d^{\mathrm{ES}}(t) = \begin{cases} P \Leftarrow\; p(t) > p_{\mathrm{tr}} \;\&\; E < E_{\max} & \text{(discharge/sell)} \\ 0 \Leftarrow\; p(t) > p_{\mathrm{tr}} \;\&\; E = E_{\max} & \text{(idle)} \\ 0 \Leftarrow\; p(t) = p_{\mathrm{tr}} & \text{(idle)} \\ 0 \Leftarrow\; p(t) < p_{\mathrm{tr}} \;\&\; E = 0 & \text{(idle)} \\ -P \Leftarrow\; p(t) < p_{\mathrm{tr}} \;\&\; E > 0 & \text{(charge/buy)} \end{cases} \tag{6}$$

The Expert System has only one parameter, the threshold price p_{tr}, which defines its performance. The sensitivity of the system to the value of p_{tr} was estimated by performing a scan in a range $p_{\mathrm{tr}} \in \{\min p(t), \max p(t)\}$. The results of the scan are presented in Fig. 5 where the total annual profit of arbitrage for the EPEX 2019 data is presented. The mean value has shown to be the optimal threshold p_{tr} for the ES 5. That is, when the storage is not full and the price of electricity is below mean the storage always charges (buys electricity); when the price of electricity is above mean the storage discharges (sells electricity). With the given strategy of an ES algorithm, energy storage managed to earn approximately 100 € for the year 2019, as well as for the year 2020. That is only ∼30% of the maximal possible value obtained from the linear optimization. During one year trading period the battery went through 700 full charge/discharge cycles, while the maximal possible number of cycles per year is ∼1500. As a result, ES strategy lets the storage stay idle in total for almost half of the time.

Figure 4 presents state of charge of the battery following the Expert System strategy. The storage responds only to significant changes in electricity price. During relatively long periods of time when the price is continuously higher or smaller than the threshold, the storage doesn't perform any actions and hence loses the ability to profit.

3.3 Reinforcement Learning

Reinforcement Learning (RL) is a computational approach to learning from interaction. It provides a mathematical framework for an agent (algorithm) to learn various strategies and find the actions that maximize numerical rewards. RL has been applied to various problems and has shown good results in the field of autonomous driving, industrial control etc [Guan et al., 2015a]. In the recent years, there has been a rising number of research in the field of electricity markets, energy storage, management in microgrids, electrical vehicle charging, energy arbitrage etc. In [Guan et al.,2015b], RL is applied to optimize the charging and discharging strategy of battery energy storage systems in industrial parks. In [Cao et al.,2020], authors apply DRL to optimize the battery energy arbitrage considering an accurate battery degradation mode. In [Guan et al., 2015a] RL was applied to minimize electricity bills of residential energy storage systems [Xu et al., 2019a]. RL was also used to learn a market clearing strategy for the local event-driven market in [Chen and Su, 2019] by using a Q-learning accounting for different rewards across multiple episodes. In [Xu et al., 2019b] DRL approach was enhanced by incorporating the proximal policy optimization (PPO) algorithm with a recurrent neural network to represent the price time series. In [Wang and Zhang, 2018] authors derive a temporal arbitrage policy for storage via Reinforcement Learning for the New England real-time electricity market data. They show that their RL algorithm performs significantly better than online modified greedy algorithm offered in [Qin et al., 2016].

One of the main problems of successful energy arbitrage is finding the patterns of the electricity price. Despite the data being quite stochastic, it is still possible to distinguish certain trends. Figure 7 illustrates the Fourier modes of the electricity price, it can be seen that German electricity market has pronounced daily, weekly, monthly, seasonal and yearly modes. Figure 6 shows that the prices are often lower on the weekend, which goes in line with the fact that demand is higher during the working days. To consider these temporal periodic and systematic trends in the energy arbitrage, we train a deep Reinforcement Learning algorithm.

This is done by creating an environment (storage and energy market) and an agent (arbitrage algorithm) that interact with each other through selling and buying electricity, hence, performing the arbitrage. Then the optimal arbitrage strategy is obtained by performing a Deep Q-learning [Mnih et al., 2013]. The objective of the Deep Q-learning is to obtain a neutral network that is capable of mapping the state of the environment to the most effective action (Fig. 8).

In the following, the environment is described by its state s_i which is then used by the agent to decide the best action at the given timestep t_i. The state reads as a 7-dimensional vector:

$$s_i = \Big\{ H(t_i),\, d_w(t_i),\, d_m(t_i),\, m(t_i),\, p(t_i),\, p(t_{i-1}),\, e(t_i) \Big\}, \tag{7}$$

where each component of the vector correspond to hour, day of the week, day of the month, month, state of charge, electricity price, previous electricity price,

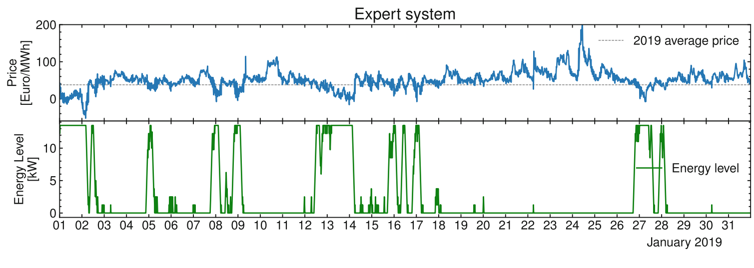

Fig. 4. Storage state of charge it January 2019 following the Expert System strategy. The results correspond to the Expert System arbitrage shown in Fig. 5.

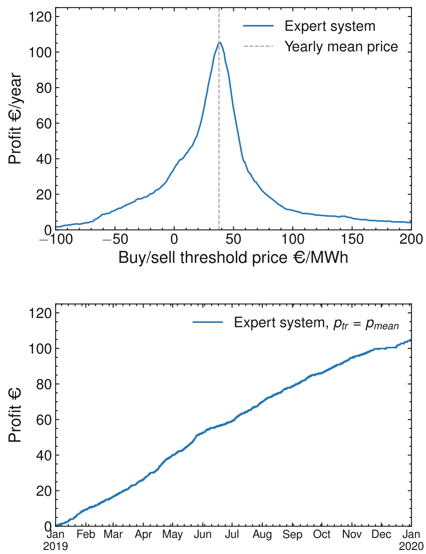

Fig. 5. Profit gained in one year by the Expert System from energy arbitrage with the Tesla Powerwall 2 storage. Electricity prices are from the EPEX intraday German market data for the year 2019. Upper: Performance (Profit) of the Expert System as a function of the threshold price p_{tr} Lower: Dependence of the profit and time for ES with 2019 mean electricity price as p_{tr}.

respectively, at a timestep t_i. The agent is capable of three actions a_i:

$$a_i = \begin{cases} 1 & \text{(discharge/sell)} \\ 0 & \text{(idle)} \\ -1 & \text{(charge/buy)} \end{cases} \tag{8}$$

Executing an action a_i in a specific state $s(t_i)$ provides a reward r_i to the agent which allows him to estimate the success of the action. Through its learning the agent explores outcomes of different actions at different states and thus "memorizes" the actions that lead to the highest rewards. Here we follow the reward function r_i suggested in [Wang and Zhang, 2018]:

$$r_i = \begin{cases} \Big(p(t_i) - \overline{p}(t_i)\Big)P & \text{(discharge/sell)} \\ 0 & \text{(idle)} \\ \Big(\overline{p}(t_i) - p(t_i)\Big)P & \text{(charge/buy)} \end{cases} \tag{9}$$

where the average price $\overline{p}_i$ is calculated by

$$\overline{p}(t_i) = (1-\eta)\,\overline{p}(t_{i-1}) + \eta\, p(t_i), \tag{10}$$

with $\eta = 0.01$.

The agent is trained by performing deep Q-learning, which we implement using a well-developed machine learning framework `PyTorch` [Paszke et al., 2022]. In the implementation, we follow instructions from [Paszke, 2022] and adapt to the current problem, i.e., state size and number of actions. The training is performed on the 2019 EPEX intraday energy price data.

The best performance of the RL optimal arbitrage policy gains only $\sim$ 115 € which is only 15% improvement over the simple Expert System. The performance of the RL approach is presented in Fig. 9. The RL arbitrage strategy is prone to the drawback of the ES which doesn't react to small changes in electricity prices and as a result the storage stays idle for most of the time.

A possible explanation for the poor performance of the RL approach is the definition of the reward function r_i. The agent obtains the highest reward when the electricity prices change largely, and, hence, it favors to stay idle and wait until large fluctuations of the electricity price to maximize its reward. However, finding a different rewarding function which will encourage rapid reactions on fine price dynamics is not straightforward and requires a separate study.

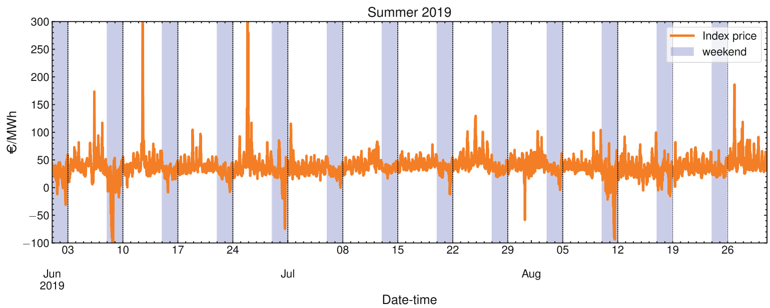

Fig. 6. Electricity price dynamics in summer 2019. Blue bands correspond to weekends when the prices are systematically lower.

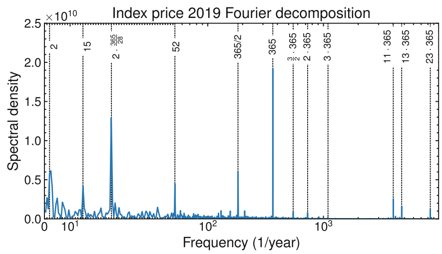

Fig. 7. Fourier modes of the electricity price. German electricity market has pronounced modes (vertical lines) which can be associated with hourly, daily, weekly, monthly, etc., periodic trends.

3.4 Performance Comparison

Result comparison of linear optimization, Expert System and Reinforcement Learning algorithms is demonstrated in Fig. 10. The performance of the ES algorithm with a fixed threshold (profit of $\sim$ 100 €) is similar to the performance of a more complex RL approach with a dynamic threshold value (profit of $\sim$ 115 €). Reinforcement Learning offers only about 15% more returns. However, the returns of both algorithms are more than three times lower than the maximal possible returns obtained from the linear optimization performed with PyPsa ($\sim$ 337 €). Taking into account the initial investment of $\sim$ 12200 EUR, the approximate breakeven time would be more than 100 years in the scenario where battery degradation is not considered and average electricity does not change significantly on yearly average basis.

To compare performance of ES with the results of the RL algorithm applied in similar research, the ES developed within this study was applied to the data for the New England US energy market used in [Wang and Zhang, 2018]. For this comparison, the capacity and parameters of the battery were set up as in [Wang and Zhang, 2018]. As shown in the Fig. 11 results of RL algorithm developed in the corresponding paper are analogous to the results of the Expert System with the buy/sell threshold set at the average price of electricity annually. RL yearly yields 40000 $ [Wang and Zhang, 2018] and our simplistic ES returns for the same period approx. 39000 $. The maximal possible return was examined with linear

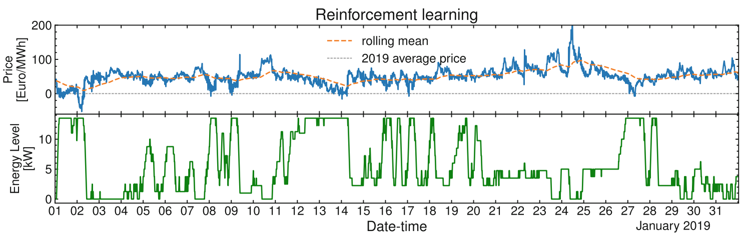

Fig. 8. Storage state of charge in January 2019 using the Reinforcement Learning approach. The results correspond to the RL arbitrage shown in Fig. 9.

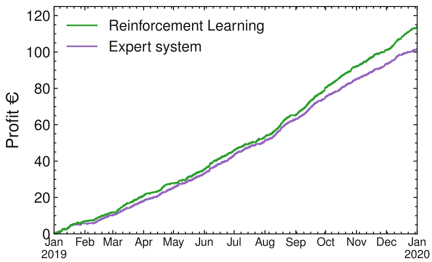

Fig. 9. Cumulative profit obtained from the Reinforcement Learning approach (green) compared to the Expert System results (purple).

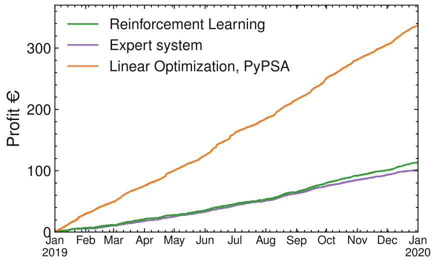

Fig. 10. Result comparison of linear optimization (orange), Expert System (purple) and Reinforcement Learning (green) algorithms. The results correspond to the Reinforcement learning in Fig. 9.

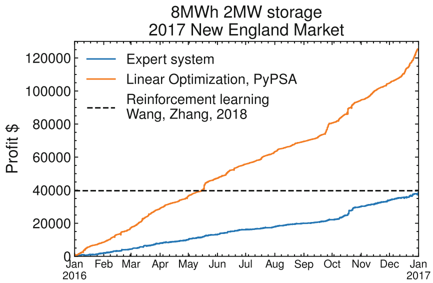

Fig. 11. Result comparison of linear optimization (orange), Expert System (blue) and Reinforcement Learning by [Wang and Zhang, 2018](dashed line) algorithms.

optimization the same as for the Tesla Powerwall for the German market. The linear optimization suggests that the maximal return at New England Market with 8MWh storage is 126000 $. This shows that both RL algorithms, the one developed within this study as well as the RL algorithm developed by [Wang and Zhang, 2018] perform similarly relative to the simple Expert System and explore only 1/3 of the maximal possible profit.

4 Conclusions

Different strategies for energy arbitrage with the Tesla Powerwall 2 storage are studied. The results of the study show that energy arbitrage in the German electricity market is in general possible. Intraday market is the most attractive market for the arbitrage operations due to its balancing function and frequent wide price spread over short periods of time. Nevertheless, as previously mentioned in [Metz and Saraiva, 2018], arbitrage does not break even within life span of the battery to justify the investment in the battery storage technologies for the purpose of monetary gains. However, if we consider battery storage as a necessary technology to balance supply and demand in a highly volatile renewable energy system, price arbitrage would be a benefit. This study demonstrated that use of a simple deterministic algorithm, expert system, with a fixed buy/sell threshold yields similar returns as application of a more sophisticated DRL approach. However, both approaches reach only up to $\sim 30\%$ of the maximal possible return which was estimated with linear optimization. The simple expert system was also applied to the New England US energy market data used in [Wang and Zhang,2018] and compared to the results of their RL algorithm. Similarly to the German electricity market, the profit of both algorithms was about $\sim 30\%$ of the maximal possible profit. Moreover, as mentioned in [Gao and Yu, 2021], even when the performance of DRL-based algorithms in distribution system applications is promising, the learned control policies are embedded in deep neural networks and therefore are hard to interpret as compared to the Expert System. This, in its turn, leads to practical difficulties and limited possibilities to check algorithm for safety properties.

Acknowledgments. The work is financially supported by the Federal Ministry of Economic Affairs and Energy (BMWi) under grant nr. FKZ03EI1028A (EnergiesysAI) and efl the Data Science Institute. A.M. acknowledges the Stern-Gerlach Postdoctoral fellowship of the Stiftung Polytechnische Gesellschaft.

References

Brown, T., Hörsch, J., and Schlachtberger, D. (2018). PyPSA: Python for Power System Analysis. *Journal of Open Research Software*, 6:4.

Campana, P., Cioccolanti, L., Francois, B., Jurasz, J., Zhang, Y., Varini, M., Stridh, B., and Yan, J. (2021). Li-ion batteries for peak shaving, price arbitrage, and photovoltaic self-consumption in commercial buildings: A monte carlo analysis. *Energy Conversion and Management*, 234:113889.

Cao, J., Harrold, D., Fan, Z., Morstyn, T., Healey, D., and Li, K. (2020). Deep reinforcement learning-based energy storage arbitrage with accurate lithium-ion battery degradation model. *IEEE Transactions on Smart Grid*, 11(5), 4513–4521.

Chen, T. and Su, W. (2019). Indirect customer-to-customer energy trading with reinforcement learning. *IEEE Transactions on Smart Grid*, 10(4), 4338–4348.

epexspot (2020). Epex market data. https://www.epexspot.com/en/market-data

Eyer, J. and Corey, G. (2011). Energy storage for the electricity grid: Benefits and market potential assessment guide. pages 1–232.

Gao, Y. and Yu, N. (2021). Deep reinforcement learning in power distribution systems: Overview, challenges, and opportunities. In *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5.

Guan, C., Wang, Y., Lin, X., Nazarian, S., and Pedram, M. (2015a). Reinforcement learning-based control of residential energy storage systems for electric bill minimization. pages 637–642.

Guan, C., Wang, Y., Lin, X., Nazarian, S., and Pedram, M. (2015b). Reinforcement learning-based control of residential energy storage systems for electric bill minimization. pages 637–642.

Harrold, D., Cao, J., and Fan, Z. (2021). Data-driven battery operation for energy arbitrage using rainbow deep reinforcement learning.

Krishnamurthy, D., Uckun, C., Zhou, Z., Thimmapuraam, P., and Botterud, A. (2017). Energy storage arbitrage under day-ahead and real-time price uncertainty. *IEEE Transactions on Power Systems*, PP:1–1.

Leadbetter, J. and Swan, L. (2012). Battery storage system for residential electricity peak demand shaving. *Energy and Buildings*, 55:685–692.
Lee, S. and Choi, D.-H. (2019). Reinforcement learning-based energy management of smart home with rooftop solar photovoltaic system, energy storage system, and home appliances. *Sensors*, 19:3937.
Liebowitz, J. (1991). *Proceedings of the World Congress on Expert Systems: Orlando, Florida, December 16–19, 1991....* Pergamon Press.
Liebowitz, J. (1995). Expert systems: A short introduction. *Engineering Fracture Mechanics*, 50(5), 601–607.
Metz, D. and Saraiva, J. (2018). Use of battery storage systems for price arbitrage operations in the15- and 60-min german intraday markets. *Electric Power Systems Research*, 160.
Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. A. (2013). Playing atari with deep reinforcement learning. *ArXiv*, abs/1312.5602.
Paszke, A. (2022). Pytorch: Reinforcement learning (dqn). https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html
Paszke, A., Gross, S., Chintala, S., Chanan, G. (2022). Pytorch. https://pytorch.org/
Qin, J., Chow, Y., Yang, J., and Rajagopal, R. (2016). Online modified greedy algorithm for storage control under uncertainty. *IEEE Transactions on Power Systems*, 31(3), 1729–1743.
Shang, D. and Sun, G. (2016). Electricity-price arbitrage with plug-in hybrid electric vehicle: Gain or loss? *Energy Policy*, 95:402–410.
Tesla (2022). Tesla powerwall. https://www.tesla.com/en_eu/powerwall
Wang, H. and Zhang, B. (2018). Energy storage arbitrage in real-time markets via reinforcement learning. pages 1–5.
Wankmüller, F., Thimmapuraam, P., Gallagher, K., and Botterud, A. (2017). Impact of battery degradation on energy arbitrage revenue of grid-level energy storage. *Journal of Energy Storage*, 10:56–66.
Xu, H., Li, X., Zhang, X., and Zhang, J. (2019a). Arbitrage of energy storage in electricity markets with deep reinforcement learning.
Xu, H., Li, X., Zhang, X., and Zhang, J. (2019b). Arbitrage of energy storage in electricity markets with deep reinforcement learning.