Utilisation of grid-connected energy storage systems for supporting a biomass generator through upcoming market reforms

Katy Honour

Submitted in accordance with the requirements for the degree of

Integrated Doctor of Philosophy and Master of Science

The University of Leeds

School of Chemical and Process Engineering

December 2023

I confirm that the work submitted is my own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to express my gratitude to my supervisors for their expert guidance and unwavering support throughout the course of my research. I am thankful to Professor Timothy Cockerill, Professor Peter Taylor and Dr. Jan Palczewski for his rigorous academic support and valuable advice.

Special thanks are due to Peter Sage at Net Zero Research. I am also thankful to Scott Hume, who facilitated my visit to North Carolina during my PhD, an experience that significantly broadened my research ambitions and provided a unique perspective that has enriched my work.

I acknowledge the support of the Bioenergy CDT, whose resources and intellectual environment have been incredibly beneficial throughout my time in Leeds.

On a personal note, I want to thank my family and Kayode Sonaike for his support and encouragement, which was vital during the most challenging phases of my research.

Abstract

Great Britain has committed to achieving net zero greenhouse gas emissions by 2050. The sector supplying energy is the second-largest source of greenhouse gas emissions. To meet statuary requirements, National Grid ESO is exploring the increased use of non-dispatchable renewable energy sources and the phasing out of inertia-providing synchronous generators.

The lack of synchronous machinery could destabilise the grid, thus preserving some capacity is crucial. Utilising synchronous generators powered by biomass could offer a remedy; providing necessary inertia without the heavy greenhouse gas emissions associated with fossil fuel synchronous generators like coal, oil, and natural gas.

Energy storage systems can time-shift electricity generation and demand to balance production and consumption. Integrating biomass generators with ES enhances efficiency and reliability. The proliferation of renewables increases the frequency of generator cycling, which can double operational costs. Onsite energy storage mitigates this need, reducing expenses, wear and tear, and additional CO_2 emissions. Using energy storage for ancillary services or market arbitrage also reduces the need for biomass generator ramping.

This thesis explores arbitrage and Firm Frequency Response products that biomass and energy storage combined can buy and sell. It confirms the benefits of investing in an energy storage, highlighting profitability from Firm Frequency Response and resilience against changes in the electricity system. Novel methodologies for optimal energy storage operation without foresight are introduced using linear programming and reinforcement learning for decision-making.

Contents

List of Tables

List of Figures

Nomenclature

B - Set of actions for the biomass RL agent

BM - Balancing Mechanism

BMRS - Balancing Mechanism reporting service

b - Biomass dispatch

c - Energy storage charging

CAES - Compressed air energy storage

CBESS - Combined biomass and energy storage system

CCGT - Combined cycle gas turbine

d - Energy storage discharging

E - Set of actions for the ES RL agent

e - Energy storage dispatch

EFA - Electricity forward agreement

ES - Energy storage

$EScap$ - Energy storage capacity

fd - Frequency deviation

FES - Future energy scenarios

FFR - Firm Frequency Response

LCOS - Levelised cost of energy storage

MDES - Medium-duration energy storage

$Opex$ - Operational cost of operating a biomass plant

$price$ - Remuneration for power delivered to the National Grid

p_x - Power capacity of asset x

PFA - Perfect foresight agent

$Q(S, A)$ - State action value function

R - reward

rc - Cost of ramping a biomass plant

res - Vector of booleans (1 for timesteps maintaining response)

RL - Reinforcement learning

rr - Maximum ramping capability of biomass generator

NDR - Non-dispatchable renewable

SCES - Supercapacitor energy storage

SMES - Superconducting magnetic energy storage

SoC - State of charge

tr - Vector of booleans (1 for timesteps triggering FFR)

tv - Tendered volume in a given EFA block

w_o - Optimisation window

w_t - Training window

ws - Aggregated wind and solar factor

TSO - Transmission system operator

α - Step-size parameter

ϵ - Probability of the agent picking an action at random

η_c - Efficiency of charging energy storage

η_d - Efficiency of discharging energy storage

σ_x - Variable x determined by PFA

τ - Length of a timestep

γ - Discount factor

1 Introduction

Great Britain (GB) has pledged to reach net zero emissions relative to 1990 by 2050. The energy supply sector currently ranks as the second highest emitter of greenhouse gasses (just after the transport sector) and will need to change drastically to meet the legal target. This transition to net zero creates particular problems for the electricity system as part of the energy supply sector. Figure 1 demonstrates the impact of the transition. The two problems identified are a proliferation of intermittent non-dispatchable renewables (NDRs); energy sources that will lead to a mismatch between supply and demand in the grid and a decommissioning of large synchronous generators, leading to a lack of inertia. To demonstrate the first point, Grubb and Newbery (2018) documented that the share of renewables providing UK electricity in 2017 was 22%, up from 4% in 2008. Furthermore, GB's transmission system operator (TSO), National Grid ESO, has published a set of 'future energy scenarios' (*Future Energy Scenarios* (2021)) which project an increase in NDRs up to the year 2050. In the FES scenario with the most conservative estimate for installing new NDR capacity there is still 100 GW by 2030 and 200 GW by 2050.

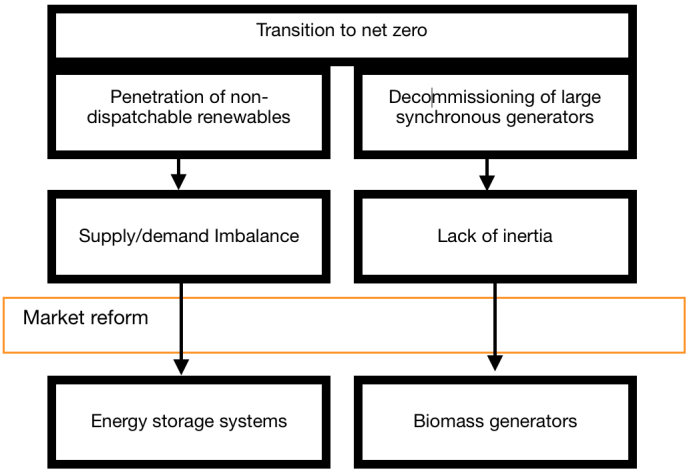

Figure 1: Problems and possible solutions for the electricity system driven by the transition to net zero

	Installed capacity (GW)
Biomass	1.58
Coal	12.30
Hydro	6.36
Gas	29.92
Nuclear	8.92
Oil	0.37
Solar	8.67
Waste	1.89
Wave/Tidal	0.04
Wind	23.20

Table 1: Capacity of generation types in the UK electricity system according to the UK publication *UK electricity production* (2022)

Meanwhile the UK government has published that the number of coal synchronous generators providing electricity has dropped from 70% in 1990 to less than 3% in 2020 (*End of coal power to be brought forward in drive towards net zero* (2020)). The political and financial will to decommission gas synchronous generators is less evident though literature such as McGlade et al. (2018) is clear that, without carbon capture and storage, gas for electricity, heating and hydrogen combined would need to be reduced to 10% of its 2010 value by 2050 to meet the net zero target. For further reference, table 1 shows the installed capacity of coal, gas, oil and biomass in 2022. Gas plants are hard to phase out of the system because in addition to inertia they provide flexibility.

A 2016 study of conventional combined cycle gas turbines (CCGTs) from Imperial College (Strbac et al. (2016)) projected that the flexibility from CCGT would need to increase

to account for the rise in NDRs before newer, greener technologies would be available to provide that flexibility and phase out gas synchronous generators. They projected a peak of CCGT in 2025. Additionally, cycling CCGT helps to avoid the unnecessary curtailment of renewable power. Biomass plants would be hard pressed to provide this flexibility due to high ramping constraints and increased wear and tear from constant cycling.

The shut down of synchronous generators contributes to a reduction in system inertia. When a deviation in system frequency occurs, synchronous generators automatically respond by absorbing or injecting kinetic energy. This corrects the deviation. System inertia refers to all the inertia in the rotating masses of synchronous generators that are online at a given time. The less system inertia, the quicker the system operator has to respond to maintain frequency. Periods of low inertia make the frequency more sensitive to an imbalance in supply and demand; an imbalance during low inertia results in a greater frequency deviation than the same imbalance during high inertia. Compounding the problem, renewables do not contribute to system inertia.

A solution explored by Ackermann et al. (2017) may lie in switching to biomass fuelled synchronous generators which can provide the same level of inertia while emitting few greenhouse gas emissions relative to coal, oil and gas. Ackermann et al. (2017) have asserted in their paper that the grid cannot be supported without any synchronous machines even with innovative designs for synthetic inertia, hence the need to maintain some biomass capacity.

Figure 1 shows biomass generators as a provider of inertia during the transition to net zero. The second solution displayed energy storage systems (ESS) are discussed subsequently.

Energy storage (ES) can be used to time-shift the output from generators, thereby compensating for the mismatch in supply and demand created by NDRs. ESS can also provide additional load when the net load drops below the must run output of nuclear by absorbing extra energy from the system. ESS technologies can shift energy on different time scales. ESS can be used by households with time-of-use tariffs, wind and solar farms who want to smooth their profiles and used by themselves to contract with National Grid ESO for ancillary services.

Possible downsides of investing in storage include upfront investment costs, loss of energy in round-trip inefficiencies and subsequently a higher carbon intensity of the elec-

tricity produced.

Biomass generators and ESSs can be combined in operation, capitalising on the advantages of both systems. ESSs have been used in a similar capacity for supporting wind turbines in Azad et al. (2015).

A key argument for using an ESS to support a biomass generator presented by Van den Bergh and Delarue (2015) is that compared to oil and gas, biomass burners are particularly susceptible to damage from repeated heat cycling. In general, an increase in NDRs will lead to an increase in the need for synchronous generator cycling. Van den Bergh and Delarue (2015) found costs for the synchronous generator operator can double operating in a system with 50% penetration of NDRs. Furthermore, it has been found by Eser et al. (2016) that a 35% penetration leads to a 63-181% increase in ramping and thus a decrease in average load factor and lower operating efficiency. An on-site ESS may relive the need for ramping, cutting costs, wear and tear and the additional CO_2 production associated with an increase in ramping. Particularly when providing additional ancillary services or performing arbitrage, the ESS can be employed to offset biomass ramping. The following thesis addresses this question of operating a combined biomass and energy storage system (CBESS) to provide Firm Frequency Response (FFR) and bid on the Balancing Mechanism (BM), which are both aspects of GB's electricity market.

An overview of both markets will be presented here briefly and discussed in significantly more detail later.

FFR is offered by National Grid ESO to ensure the reliability of GB's electricity supply as do all ancillary services. FFR covers rapid response measures to supply and demand imbalance. FFR works by contracting with providers to respond to deviations in system frequency with (in the case of the non-dynamic response) a fixed volume of power; a reduction in power for a high system frequency and an increase in power for a low system frequency. Each provider must respond within 30s and maintain their output for a 30 minutes. Relative to other worldwide regions this is a demanding response time but a short maintenance period (Raineri et al. (2006)).

At any time, CBESS operators seek to maximise the system performance against current market constraints. A benefit of ES to the CBESS is provide ancillary services to National Grid ESO. As such FFR provision is examined in some detail.

Deregulation of the electricity market began in GB in 1989 and was finalised in 1991.

In 2001 direct bilateral contracts between supply of energy and demand were introduced. There are three wholesale electricity markets operating in the UK; two of which are day-ahead (Epex Spot UK Power Auction and N2EX Day-Ahead Auction) while the third is the Balancing Mechanism (BM) which exists to account for discrepancies in supply and demand on a 30 minute interval. Of the three the BM is the most volatile. In the future the centralised control of the national grid is likely to be replaced by more decentralised mechanisms. To illustrate, the increasing mismatches in supply and demand discussed previously are resolved with centralised scheduling of assets on the BM. This approach treats demand as an uncontrollable feature created by the consumers. Decentralised iterations of the BM accommodate many more 'smart' technologies such as load control and ESSs that can time-shift demand. Therefore there is strong reason to believe a CBESS would benefit performing arbitrage on the BM.

Chapter 2 is a literature review that examines the composition of electricity systems, highlighting grid-enhancing techniques and technologies that can reduce the need for expensive transmission expansion. Chapters 3 and 4 set up the thesis to establish a methodology for operating a CBESS in the two markets previously mentioned. The motivation for such is to design a strategy that operates without foresight, allowing a picture to be gained of how a real CBESS owner could make operational decisions in real time and without the constrains used in other studies that require some form of foresight (Pimm et al. (2020)). In chapter 5 the methodology of mimicking such a perfect foresight agent was extended further to the CBESS, tested against the National Grid's Firm Frequency Response market, through a system of reinforcement learning, effectively recreating perfect foresight behaviour. The chapter also demonstrates that biomass operational costs significantly impact profits. And before an overall summary of the thesis is given in chapter 7, in chapter 6 there is a discussion of installed capacities of medium-duration energy storage (MDES) needed to scale with NDR capacity, projecting a need for 11 TWh by 2050 for NDR penetration. There is special emphasis on the role of MDES in intra-day balancing due to the diurnal cycle of solar power. The chapter examines MDES operation from 2022 to 2035.

In conclusion, the aim of this thesis is to explore the role of a CBESS working for National Grid ESO up to 2035, for its relevance to a transmission grid attempting to decarbonise.

Additionally, the research question can be stated as such;

With the combined operation of a biomass generator and energy storage system, to what extent can the system provide the services required by the national grid in Great Britain up to 2035?

1.1 Objectives

Now presented are the objectives of the thesis, with a explanation and justification for each one outlined beneath. The methods for achieving each objective follow.

- **O1** To do a literature review that covers the role of bioenergy and ES in the electricity system.

  To write a sufficient thesis, knowledge is required of the role of bioenergy and ES in the electricity system. The literature review will establish whether current biomass plants can provide flexibility and ES is currently used in co-operation with other forms of generation such as wind.

- **O2** To review historical frequency data and price signals from the UK electricity system.

  The review will situate the thesis in the UK and establish the time period being covered.

- **O3** To design a methodology for making the operational decisions for a biomass and ES combined system.

  Running a biomass and ES combined system requires some reasoning behind operational decisions. The methodology developed should make operational decisions that maximise profits and take full advantage of the ESS. The methodology should assume no knowledge of future events.

- **O4** To design a system for testing operational decisions.

  The operational decisions taken by the methodology developed in **O3** need to be assessed to determine if they are good ones.

- **O5** To develop case studies that where the methodology developed in **O3** can be applied.

The case studies should be chosen to reflect a biomass and ES combined system that could realistically fit into the current UK electricity system. Ideally the case studies should demonstrate a range of technologies and markets.

- **O6** Draw conclusions about the operation of a CBESS for arbitrage and providing FFR.

1.2 Previous conference publications

- D. Chernick, K. Graves, K. Honour, T. Penney, S.Wiseman, 'Valorising Indian Municipal Solid Waste: A Feasibility Study', University of Leeds Bioenergy Conference, 2019.

- K. Honour, 'Optimal control of energy storage systems for large thermal plant operators', The 1st FERIA Conference, the European Conference on Fuel and Energy Research and Its Applications, 2021.

The following chapter is a literature review that will place the thesis in its general context.

2 Literature review

This literature review discusses the make-up of electricity systems and the various markets that have and may evolve to ensure stability. These can be grid enhancing techniques in any form. Mirzapour et al. (2024) did a full analysis of these such techniques including hardware and software and, importantly, suggested how these technologies could reduce the need for transmission expansion, finding it to be one of the most expensive options for increasing stability. Furthermore, in Mirzapour et al. (2024), every technology analysed was put in its place considering environmental costs of the particular expansion, which instillation sites would improve regional efficiency and locational prices. Wu et al. (2024) improved on the work of Mirzapour et al. (2024) by including international trading routes that allow for the building of new capacity in each jurisdiction.

This literature review is split into several different parts and sets up the reader to understand further which technology types and jurisdictions will be focused on. The first introduces ESS and bioenergy generators, both of which can be shown to fulfill parts of

the necessary roles identified by Mirzapour et al. (2024). Section 2.1 focuses exclusively on ESS; the cost of operation and reviews comparative analyses of different technologies. Section 2.3 presents the operation of bioneregy in GB, exploring policy instruments that have propped up bioenergy expansion and summarising what researchers have concluded on the role of bioenergy in the future. Sections 2.4 through 2.5 introduce GB electricity markets and considers how they may evolve to 2050. A summary of the literature up to that point is presented in section 2.7 and the GB electricity system is placed in the context of the EU. This first part of the literature review sets the scene for the entire thesis, clearly describing the literature landscape that has preceded it.

Section 2.8 analyses the ways in which decision mathematics and, particularly, Markov decision processes have been used to develop bidding strategies in electricity markets. The goal of such an analysis is to establish how researchers have maximised profit from ESS.

The review finishes by pointing out objectives that can be drawn from the literature to direct the rest of the thesis (section 1.1) and the methods that can be used to achieve these (section 2.9). Finally, section 2.10 reiterates the gaps in the literature identified and places the scope of the remaining research in its place with regard to current literature.

2.1 Energy storage technologies

ESS have the potential to provide the flexibility the grid needs to support renewable technologies. This has led to a rich research field analysing both the ability of ESS to achieve this goal and their economic viability.

Energy storage can be defined by a set of parameters such as round-trip efficiency which are, furthermore, abstract features of the underlying technology. It is therefore worthwhile spending the opening salvo of this literature review discussing the technologies that have been developed and the parameters arising from their operation. Several ESS technologies (gravitational energy storage, superconducting magnetic energy storage (SMES), and supercapacitor energy storage (SCES)) are hard to gather reliable economic information for due to yet being deployed on grid-wide scale as was noted in Brandon et al. (2016). Even mature technologies are little understood leaving many potential investors unsure of a stable business model but as was further discussed in Brandon et al. (2016) research is being studied increasingly. A summary follows of available literature for the most popular ESS.

It should be noted that, throughout the literature there is an important distinction relating to the units used to report operational costs. Some use £/MWh as in Schmidt et al. (2019) while Zakeri and Syri (2015) prefer the use of £/kW-yr. £/kw-yr are a unit expressing the cost of capacity provided for 8760 hours. Converting between the two units can be done as so

$$\text{Opex (£/kW-yr)} = \frac{\text{Opex (£/kWh)}}{8760 \times \text{Capacity factor}} \tag{1}$$

The significance of the conversion comes from the inclusion of the capacity factor, allowing for the consideration of how the ES is used operationally.

For example, as in Zakeri and Syri (2015), the stages of developing operational strategies that vary with time and their changing capacity factors for the entire year were put into equation 1 and meant Zakeri and Syri (2015) could 1) decide on initial installation capacities and 2) compare different technology types. It will be demonstrated later that the unit of £/kW-yr can be more useful in calculations of levelised cost of energy storage (LCOS) and furthermore justified for use in this thesis for continuity with published literature.

2.1.1 Batteries

Batteries are a well developed technology for storing energy. Components include battery cells which are the fundamental units that store electrical energy. They consist of positive and negative electrodes, electrolytes, and separators. Different battery chemistries, such as lithium ion, lead-acid, or flow batteries, exhibit distinct characteristics and performance metrics. Lead-acid was the first chemistry to penetrate the market. However, Salvini and Giovannelli (2022) have documented that as of 2022 lithium ion batteries are the most developed in terms of market penetration with an efficiency of 0.8 and a lifetime of around 10 years using a methodology of examining LCOS. Rahman et al. (2021) developed a bottom-up techno-economic model for comparing battery technologies for the applications of energy trading, frequency regulation and voltage support finding lithium ion batteries to consistently achieve low operating costs.

Meanwhile, Park et al. (2016) posit that, for a battery with a long discharge time, money is most wisely spent on flow batteries. Flow batteries can be scaled up or down

in size relatively easily by adjusting the size of the electrolyte tanks. Furthermore, flow batteries have a relatively long lifespans, and they have the ability to be frequently cycled without degradation in capacity. More advanced technologies including models such as high energy density batteries have vastly higher capital costs compared to their low energy density counterparts.

The UK government aims to install 40 GW of battery ES by 2030 as was noted by *Guidance for generators: Co-location of electricity storage facilities with renewable generation supported under the Renewables Obligation or Feed-in Tariff schemes* (2020) who concurrently wrote that the largest capacity grid-scale battery is 1.3 GW located in the North Sea.

2.1.2 Supercapacitor energy storage

SCES is a technology whereby energy is stored in an electic field between electrodes. The electrodes are typically made of activated carbon or other conductive materials with a large surface area, allowing for efficient ion adsorption and desorption. The electrolyte facilitates the movement of ions between the electrodes, enabling the storage and release of electrical energy. Commonly used electrolytes include aqueous or organic-based solutions with suitable ionic conductivity. SCES have a low energy density (0.5-5 Wh/kg) compared to other ESS as noted by Brandon et al. (2016). SCES' fast-acting response and ability to perform a large number of cycles throughout its lifetime (up to 1 million (Akram et al. (2020))) makes it profitable for correcting high frequency mismatches in supply and demand. Further, Mostafa et al. (2020) hypothesised it could also be used for voltage support. Other advantages include a long lifetime (depending on operation) and >0.85 energy efficiency.

Akram et al. (2020) lists SCES projects that have been deployed for grid-scale applications, including a 4 MW installation in Spain providing frequency regulation.

2.1.3 Superconducting magnetic energy storage

SMES technologies store energy in their magnetic fields. SMES systems consist of superconducting coils made of materials with zero electrical resistance at low temperatures. These coils are typically made of niobium-titanium or niobium-tin compounds and are cooled to extremely low temperatures using cryogenic systems. SMES can play the same

functional role as SCES, delivering very fast response times while having a lower capital cost. Adetokun et al. (2022) wrote a definitive review of SMES and the literature on SMES from the last 10 years, including 1240 papers and finding a reported lifetime of 20 - 40 years and an efficiency of 0.95 - 0.98 although SMES has only, as of yet, reached the demonstration stage of deployment. Performance enhancement can be achieved by up to a factor of 5 by 2030 according to Adetokun et al. (2022) with the standardisation of the technology. By their reckoning, commercialisation could begin in 2025. The work of Adetokun et al. (2022) is particularly worth paying attention to in the context of this thesis as they have described various operational strategies for the SMES including model predictive control, neural network controllers and fuzzy logic controllers.

Disadvantages include a low energy density and tendency to self-discharge. Additionally, the SMES must be kept at consistently low temperatures (-269°C) to maintain their superconducting properties with adds considerably to the cost of operation. Targets for reduction in operating costs in the near future are ambitious, Zhu et al. (2013) projecting an 85% drop. That paper was published in 2013 and its ambitious projections proved fairly plausible as the 2021 paper by Khosravi et al. (2021) saw a drop over the preceding decade of 67%.

The US department of energy released a report as early as 2003 (Schoenung and Hassenzahl (2003)) on the use of SMES for voltage support and found that cost of installation scaled with energy capacity to the power of 2/3.

2.1.4 Hydrogen

Energy can be stored in the form of hydrogen fuel, its key advantage being that it can be stored for months, mitigating seasonal variations in supply and demand (Baumann et al. (2021)). Efficiencies of hydrogen storage have been placed at 0.2 - 0.5 depending on the exact technology making it unsuitable for correcting short-term mismatches. Guerra et al. (2020) ran a techno-economic analysis of seasonal hydrogen storage for a US electricity system with 61% penetration of non-dispatchanble renewables (NDRs), finding 1 week of discharge financially viable by 2025 and 2 weeks by 2050, assuming reductions in costs with increasing innovation.

Andersson and Grönkvist (2019) identified electrolysis equipment and storage facilities and two factors that contribute most to the cost of storing hydrogen. Equipment to

perform electrolysis of hydrogen has capital costs in the range just under 200 £/kW to a maximum of 1600 £/kW according to Chatenet et al. (2022) depending on the electrolyser technology. And further, Steward et al. (2009) have published that to store the hydrogen, costs depend on the nature of the sight with three orders of magnitude in the price differences between above and underground caverns for gas storage.

2.1.5 Pumped hydroelectric

Pumped hydroelectric relies on the gravitational potential that can be stored in water. Pumped hydroelectric systems consist of two reservoirs located at different elevations. The upper reservoir stores water at a higher elevation, while the lower reservoir is positioned at a lower elevation. The height difference between the two reservoirs determines the potential energy that can be stored. Pérez-Díaz et al. (2015) documented the commercialisation of pumped hydroelectric and found it to be one of the most prolific technologies for large-scale storage, worldwide installation was 130 GW in 2015. Guerra et al. (2020) expect pumped hydroelectric to increase capacity 5 fold to 2050. Despite their potential to expand in capacity, pumped hydroelectric is considered a fully developed technology as in Klumpp (2016) and thus Klumpp (2016) finds no potential for capital or operational cost reduction.

Pumped hydroelectric can achieve a discharge time at rated capacity over 200 hrs, putting it into the group of long-duration technologies (see table 2). Smaller capacity pumped hydroelectric can also be used to to provide inter and intra-day storage. Pumped hydroelectric has an efficiency approaching 0.8 (Klumpp (2016)).

Pérez-Díaz et al. (2015) presented a similar role as Adetokun et al. (2022) did for SMES, researching the operational strategies of pumped hydroelectric for the New York Independent System Operator with a presumed deterministic price profile. Mixed integer linear programming proved to generate the most profitable operational strategy and made the best use of water resources.

2.1.6 Compressed air

Compressed air ES (CAES) uses electricity to mechanically compress air and store it in a pressurised location. During discharge, the air can operate turbines to recover the electricity. To improve system efficiency, heat exchangers are often used to recover the heat generated during compression and store it for later use. The recovered heat can be

Duration	Service	Discharge time at rated capacity (h)	Cycles per year	ESS system
Long	Seasonal energy storage, emergency capacity	>200	250-300	Pumped hydroelectric, CAES, hydrogen
Medium	Capacity payments, inter-day storage, Balancing mechanism	4-200	300-400	Hydrogen, batteries
Short	Frequency regulation, supporting renewables, voltage support	<4	1000+	Batteries, SMES, SCES

Table 2: The roles of various energy storage systems in supporting the electricity system, divided by long, medium and short duration types

utilised during the expansion process to increase the efficiency of electricity generation.

Drury et al. (2011) wrote a comprehensive review of the costs associated with CAES providing multiple ancillary services in a range of electricity markets across the US to pick out the costs of operating a CAES. Similar to hydrogen, the costs of CAES include storage cavities. It is far cheaper to use natural geological formations for storing hydrogen. Klumpp (2016) published that there is room for around 20% reduction in capital and operational costs as CAES are further commercialised. Another review of CAES, (Bazdar et al. (2022)) identified its key drawbacks as a low maximum depth of discharge and long response time. The long response time places CAES in the long duration category (see table 2). Reported efficiencies of CAES range from 0.5-0.8.

A summary of the literature has been carried out, identifying key characteristics of the various technologies. The technologies can be used for a range of applications that are summarised in table 2.

Figure 2 shows the average operational and capital costs (averages are calculated from a range of 15 papers: Aburub et al. (2019), Andersson and Grönkvist (2019), Argyrou et al. (2018), Augustine and Blair (2021), Denholm et al. (2023), Ho et al. (2012), Khosravi et al. (2021), Kintner-Meyer et al. (2012), Ming et al. (2013), Pérez-Díaz et al. (2015), Rancilio et al. (2020), Schmidt and Staffell (2023), Sneezer (2021), Sundararagavan and Baker (2012), Zakeri and Syri (2015), Zhu et al. (2013)). It should be noted that the

operational costs reported do cover a range of operational strategies, including a range of grid-scale operations and ancillary services. Furthermore the methodologies used to arrive at the figures vary greatly. Importantly all the papers chosen exclusively focus on grid-scale applications, leaving out microgrids etc.

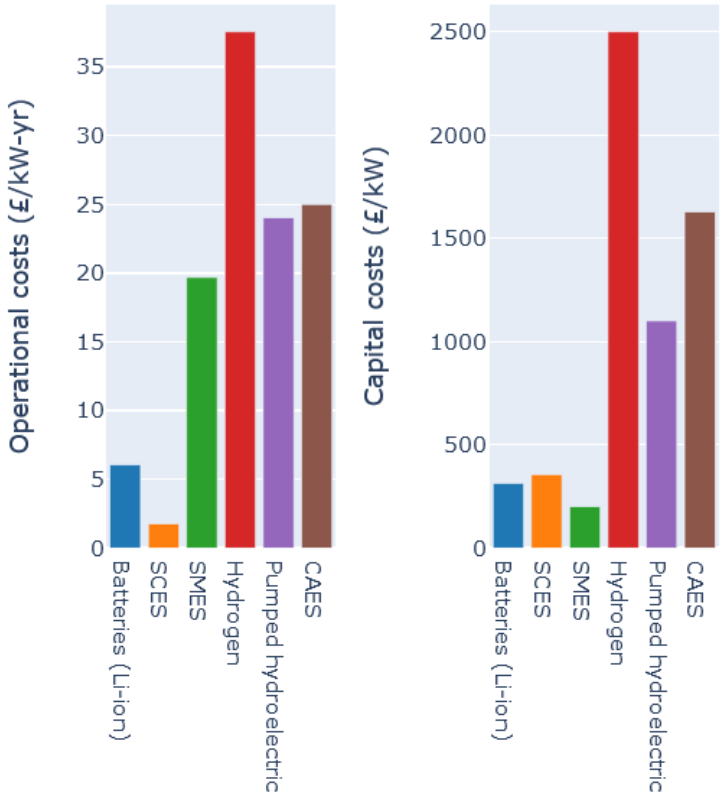

Figure 2: Results of a literature review displaying the average operational (£/kW-yr) and capital (£/kW) costs for several of the most popular energy storage technologies.

Few papers in the literature address the uncertainties inherent in their reported costs with the exception of Zakeri and Syri (2015) (which finds considerable uncertainty) and Schmidt et al. (2019). Zakeri and Syri (2015) uses a Monte Carlo method which takes values at random from a given range and repeats the methodology 10,000 times. Their paper reports a -17% to $+39\%$ change in costs for various ESS and that hydrogen storage entailed the biggest uncertainty. Furthermore, that ESS in general with low efficiencies are subject to the most variation in cost using the Monte Carlo method.

Schmidt et al. (2019) reported a similar finding in regard to round-trip efficiencies, estimating it to be the most sensitive parameter in predicting ESS levelised cost of storage across all technologies studied, based on a Monte Carlo simulation. Notably, of the

literature reviewed, Schmidt et al. (2019) has the most transparent methodology and a free, interactive version of their model is available (Schmidt and Staffell (2023)).

The levelised cost of energy storage (LCOS) is a concept that constantly recurs throughout the literature. Figure 3 shows the increase of interest in this subject between 2000 and 2020 and the next section is dedicated to examining this topic.

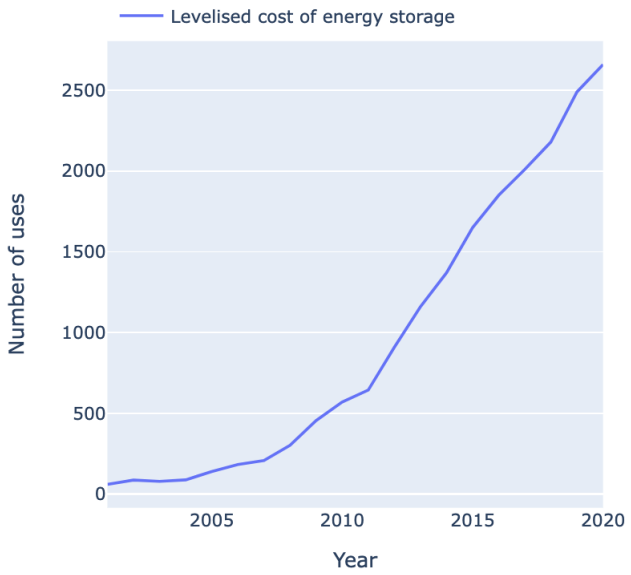

Figure 3: Appearance of 'levelised cost of energy storage' between 2000 and 2020.

2.2 Levelised cost of energy storage

The LCOS is a concept taken from the levelised cost of electricity, a metric designed to evaluate the profitability of an investment venture. LCOS calculates the cost of storing energy over the full lifetime of an installed technology divided by the electricity delivered from the ES during that lifetime. The popularity of LCOS increased when researchers discovered that relying on the levelised cost of electricity tended to overestimate the optimal reliance on ES self-consumption for solar farms (Luerssen et al. (2021)). For example, variations in electricity price may encourage a solar farm to self-consume, charging a connected ESS, and selling to the grid when prices have increased. The calculations relied on the arbitrage effect of time shifting the electricity but underestimated the additional cost of storing energy in the given ESS.

Papers calculating common ES technologies' LCOS are numerous as well as using LCOS to determine the viability of novel ES technologies such as gravitational ES (Botha

and Kamper (2019)) and chilled water ES (CWES) (Luerssen et al. (2021)). The LCOS can be calculated thusly

$$LCOS = \frac{I + \sum_{n=1}^{N} \frac{\text{Operational cost}}{(1+r)^n} + \sum_{n=1}^{N} \frac{\text{Charging cost}}{(1+r)^n} + \frac{\text{End of life cost}}{(1+r)^{N+1}}}{\sum_{n=1}^{N} \frac{\text{Electricity delivered}}{(1+r)^n}} \tag{2}$$

where I are the investment costs, r is the discount rate and N is the lifetime of the project. Both operational costs, charging costs and electricity delivered are summed in each time period n, giving LCOS the unit of 'unit of money' per 'per unit of energy delivered'. This is most often quoted as £/kW-yr or £/MWh depending on the requirements of the paper. The methodology for calculating LCOS differs throughout the literature, mainly on the point of whether to vary charging costs when charging from an electricity grid. Additionally, the last term in the numerator of equation 2 is ignored in a some analyses such as in Jülch (2016).

As can be seen from equation 2 the calculated LCOS will depend greatly on the operational strategy of the ES and therefore is a good metric for comparing technologies for operating in different markets.

Figure 4 demonstrates the varying LCOS as technologies are used for in a range of markets. It shows a collection of data published by Schmidt et al. (2019), namely the LCOS for 9 ES technologies operating in 3 different markets; the Balancing Mechanism, providing frequency storage and providing seasonal storage. It should be noted that in the column displaying seasonal storage, technologies operating over 100,000 £/MWh discharged are excluded as they cannot fit on the scale of the graph (for example SCES has a an LCOS for seasonal ES of 1,158,000 £/MWh discharged). The results show that the range of batteries and pumped hydroelectric have the lowest LCOS at intra-day storage, lithium ion batteries are cheapest when providing frequency regulation while CAES, pumped hydroelectric and hydrogen are fairly evenly matched at providing seasonal storage.

Furthermore, the factors effecting LCOS can be generally fit into 3 categories:

1. **Electricity market economics**: This category includes market electricity prices, important when considering when to charge and discharge the ESS and which market would be most suitable.

2. **Physical features of the ES**: This category entails the physical features of the ESS system that might limit the ES stored and efficiency of storage, such as self-

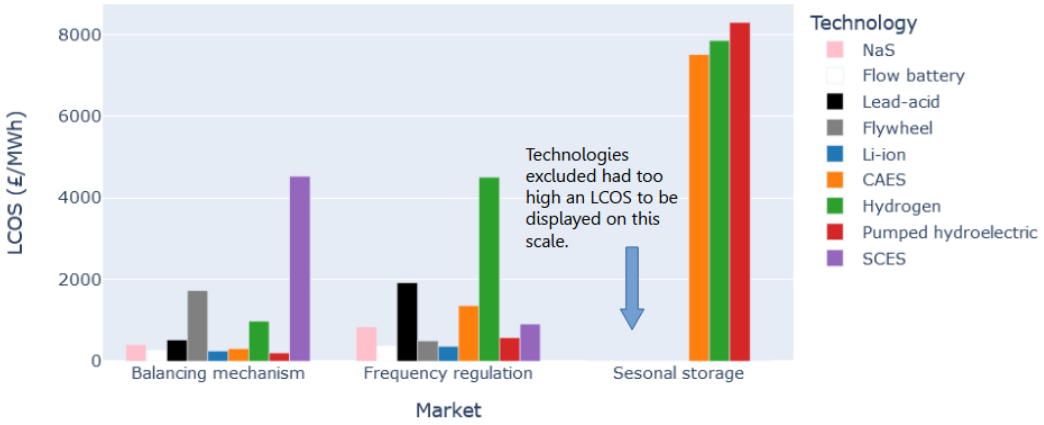

Figure 4: A bar chart showing the levelised cost of energy storage for technologies performing seasonal storage, operating on the Balancing Mechanism and performing frequency regulation.

discharge, depth of discharge and storage capacity.

3. **Costs of storage**: Capex, discount rates, costs of repair and any other costs that are accrued in operating the ESS components.

In summary, the literature suggests that, over the next decade, lithium ion batteries will continue to provide the majority of grid-scale ES for short-medium duration applications due to increased production of batteries worldwide and economies of scale. The head start of this technology is predicted to make it hard for other, currently less developed technologies, to enter the market, a economic phenomenon known as the lock-in effect. Schmidt et al. (2019) finds batteries retaining the most competitive LCOS up to 2050. Furthermore, Luerssen et al. (2021) (which examined CWES) conducted a comparison of CWES to a lithium ion battery and found the CWES to have a lower LCOS but still predicted improvements in lithium ion batteries over the coming decades and so concluded it to be a favourable technology to the more novice CWES. For long duration applications pumped hydroelectric, CAES and hydrogen are equally competitive, with any one technology yet to pull ahead in the marketplace.

2.3 The role of biomass in the UK

Bioenergy has been a key part of GB's plan to reach net zero by 2050 under the assumption that burning biomass does, indeed, remove a net amount of CO_2 from the atmosphere.

This has been disputed by some researchers including Pierrehumbert (2022). Pierrehumbert (2022) suggested that some stringent carbon accounting needs to be used especially when evaluating biomass supply lines. Politically, Pierrehumbert (2022) accuses both anti-bioenergy activists and bioenergy operators of publishing biased research.

The biggest biomass to electricity plant currently operational in the UK is Drax Power plant with 4 650 MW synchronous generators. Pierrehumbert (2022) follows the history of Drax. Since 2020 Drax used 8% coal, 2% hydrogen, 15% natural gas and 75% biomass. Drax imports its wood from the US and Brazil. The Brazllian wood is waste wood from the leather industry which would otherwise be left on the forest floor to decompose releasing CO_2, or burnt. The wood from the US is grown specifically for Drax supply lines, such managed forests have been criticised for their lack of biodiversity.

In Cross et al. (2021), the authors investigated renewable energy technology policies and their respective impacts across the UK, Sweden, Denmark, and Finland, nations characterised by high economic development and vigorous climate and energy objectives. They scrutinised the advancement and challenges encountered in bioenergy policy and deployment from 1990 to 2019. Whereas all four countries are subject to the EU's legally binding renewable energy and emission targets, divergent paths and ambitions materialised at the national level. They provided a synthesis of both qualitative (from stakeholder events and in-depth interviews) and quantitative analyses (of bioenergy generation performance). Although Denmark was not part of the qualitative study, it was included in quantitative analyses to widen perspectives and serve as an additional comparison for the UK, especially considering the similarities in reliance on agricultural resources and imported biomass. it was found that bioenergy, particularly in heat and power generation, stands as a pivotal sector in each country's renewable energy framework. Their interview of stakeholders in bioenergy for electricity (Cross et al. (2021)) stated that their main concern was political questions over whether bioenergy would continue to be seen as a green alternative though they supported CfD schemes.

Alterations in government attitudes and policies generate a multifaceted impact, notably in bewildering investment landscapes. A stable policy environment, the stakeholders stated, is paramount for sustaining and augmenting investment flows into bioenergy capacity, given the protracted and capital-intensive nature of energy projects. Subtle or abrupt shifts in policy, such as alterations in subsidy schemes, tax incentives, or reg-

ulatory frameworks, can notably perplex investor confidence and decision-making. For example, a stakeholder referenced a 2010 change in Norway's feed-in tariffs, altering the prerequisites for qualifying for incentives. In pursuit of risk mitigation, the stakeholder deterred investments in bioenergy expansion.

Addressing this concern raised in the 2021 study, Welfle and Röder (2022) developed the Bioeconomy Sustainability Indicator Model (BSIM) to evaluate carbon emissions and potential neutrality of burning biomass using 126 indicators. The BSIM was used to analyse energy crops, agricultural residues and municipal waste. For municipal waste, the biggest sustainability risk came from poor efficiency of burning municipal waste and the greatest strength was the strong emissions regulation of burning municipal waste already in place. The same results were found for agricultural residues. For energy crops, again, the biggest risk was concern over energy inefficiency and the greatest strength was that ecosystems potentially affected by growing new energy crops are already protected by law. Further schemes that have encouraged new biomass installation can be found in figure 5 with references from Bruce and Ruff (2018), *Enhanced Frequency Response: Frequently asked questions* (2016), *Enhanced Frequency Response Market Information Report* (2018), *Firm Frequency Response Interactive Guidance* (2017), *Frequency Response Market Information Report: Monthly Report January 2022* (2022), *GSR023: Clarification of the applicability of the N-1-1 criterion* (2018), *Operability Strategy Report* (2020), *System Needs and Product Strategy* (2017) and Cross et al. (2021).

Figure 6 shows the projections for biomass capacity being built in GB in the future as produced by National Grid ESO, first of all for direct sale of electricity to the grid and second for all other competing biomass uses including biomethane blending, hydrogen production, aviation, road transport and industrial processes combined. All FES show biomass capacity rising and falling throughout the first half of the century, with Leading the way and Falling short both plateauing and the capacity in System transformation and Consumer transformation rising sharply between 2030 and 2035. Both System transformation and Consumer transformation rely on a medium level of decarbonisation making biomass (a technology with a high carbon density relative to wind and solar) a viable option. Leading the way uses other less carbon intensive electricity sources and Falling short fails to find any investment in more biomass capacity for electricity production. Falling short and System transformation utilise biomass in the rest of the economy and

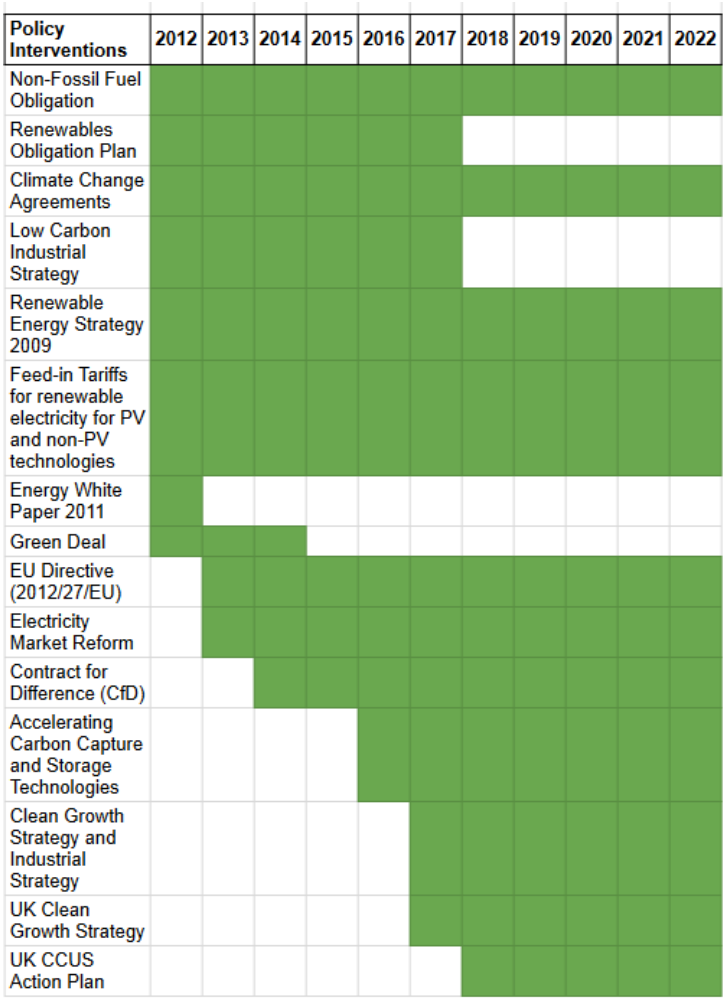

Figure 5: Policy incentives introduced since 2012 affecting new capacity of bioenergy and energy storage systems.

potentially pushing up competition for biomass supplies. This comes from a high degree of biofuels in road transport for both, biofuel for hydrogen production in System transformation and blending biomethane into the gas system in Falling short. All in all, the scenarios require that up to 61% of GB's agricultural land would need to be committed to biomass production.

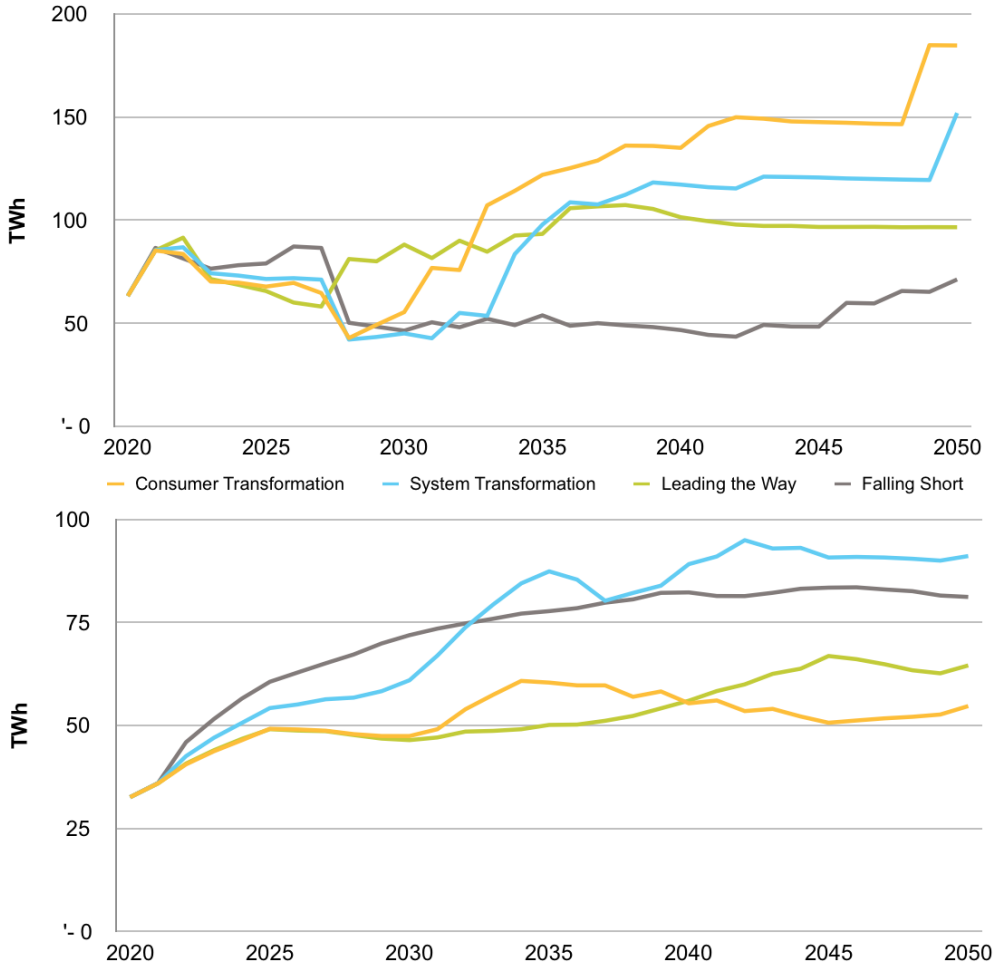

Figure 6: Biomass projections 2020-2050 (*Future Energy Scenarios* (2021)). Above: projected overall biomass capacity providing electricity to the grid. Below: Use of biomass in the UK for biomethane blending, hydrogen production, aviation, road transport and industrial processes combined.

2.4 Electricity markets

From the technologies themselves there is now a presentation of the markets available for them to operate in. The motivation behind having an electricity market is to facilitate the efficient production, distribution, and consumption of electricity. Electricity markets provide a platform for buyers and sellers to trade electricity as a commodity, allowing for the matching of electricity supply and demand in an economically efficient manner. Producers can determine the optimal level of electricity generation based on market prices, ensuring that resources are utilised effectively. Consumers, on the other hand, have the freedom to choose their electricity suppliers and consumption patterns based on their

preferences and needs.

Electricity markets provide a mechanism for determining the price of electricity based on supply and demand fundamentals. Through the interaction of market participants, prices are established in real-time or through forward contracts, reflecting the current or future scarcity of electricity. Price discovery promotes transparency and facilitates informed decision-making by market participants. By introducing competition into the electricity sector, markets foster cost reductions. Market participants, including generators, retailers, and service providers, are incentivised to improve their operations, invest in new technologies, and deliver electricity at competitive prices. This competitive environment encourages innovation in renewable energy, energy storage, demand response, and other areas of the electricity industry (Hogan (2022)).

Electricity markets operate under regulatory frameworks that establish rules, standards, and safeguards to protect consumers and ensure fair market competition. Regulatory authorities monitor market behavior, enforce compliance, and oversee the functioning of market mechanisms. This oversight helps maintain market integrity, prevent market abuse, and protect consumer interests.

All this is mentioned to provide a framework for how ES can be thought of in relation to the markets. Will it further facilitate the listed goals of the market itself?

2.4.1 Energy markets

The perceived ability of an ESS to provide good economic returns depends vastly on the market environment. There has been a common trend towards decentralisation in recent years documented by Gutierrez-Alcaraz and Sheble (2004), which is characterised by price signals adjusted dynamically based on the balancing of supply and demand. Literature refers to markets that pay for power exchange between grid and generator as energy markets. Another set of markets exist, referred to as ancillary markets, which will include a market that works to support the transmission system. They are examined in later sections.

Despite the range of countries and operators explored in the literature, a common feature is the presence of a spot market, where the temporal variations in electricity price are greatest. The practise of taking advantage of variations by buying electricity when it is cheap and selling when expensive is known as arbitrage.

Another common feature is a day-ahead market (a market that is closed 24 hours in advance). The day-ahead market tends to be less volatile and, as such, less risky strategies look into bidding in day-ahead markets or a combination of spot and day-ahead. While electricity prices hit their highest peak in the spot market, a larger volume of energy is traded on the combined day-ahead market and bilateral contracts, with Walawalkar et al. (2007) finding a ratio of 1:9 on the New York system operator.

In Britain the Balancing Mechanism (BM) trades in half hour periods for balancing supply and demand. The BM has a nationwide settling price making it redundant where the operator sits geographically. The prices on the BM need to be set high enough to encourage generators to part load, so that they can increase production.

The exact capacity and sizing of ESS required for a given system has been studied and the power to energy ratio is crucial to determining the suitability of an ESS for a given system according to Zhang et al. (2014). They find that, for a 50% renewables penetration, the required capacity of ES is at least 5% of the overall installed generation. The paper provides a comprehensive exploration into the escalating demand for ES capacity. Engaging with 9 years of demand and generation data, and emphasising a 70% round-trip storage efficiency and 15% curtailment of wind, the paper suggests that an optimal mix of 84% wind and 16% solar PV will necessitate approximately 43 TWh of storage capacity for the UK, equating to an investment of around £165.3 billion or roughly 7% of the UK's GDP in 2021. Variations in renewable energy mix profoundly impact the requisite storage capacity, evidenced by a surge to 159.7 TWh for 100% wind and a decrease to 74 TWh for 100% solar. The optimal mix offers a storage duration of approximately 29.5 days, considering the capacity and maximum rated discharge power. Storage capacity, it's suggested, would ideally be decentralised and proximate to generation sites, employing a diverse technological arsenal for efficacy. Lastly, the paper underscores the noteworthy impact of over-generation on storage capacity and overall electricity costs, revealing potential annual savings of approximately £4.8 billion with a 15% over-generation.

Although the above research has shown ESS to be profitable, counter arguments should not be ignored. A number of papers find ESS unattractive in energy markets. An examination of the profitability of a battery ESS (BESS) carried out by Cho and Kleit (2015) concludes that the cash flows generated by a BESS are unlikely to exceed the initial investments even when ancillary markets are also accounted for. The difference in results

may be attributed to the specification of the ESS being a battery and the inclusion in Cho and Kleit (2015) of a limited number of charging cycles and operating years. Or else that Cho and Kleit (2015) addresses the time-valued nature of money, discounting cash flows that occur in the future. These additions make direct comparisons between models difficult.

An interesting case for the future of ESS is presented in Sioshansi et al. (2009). It discusses arbitrage but also notes that as ESS proliferate, there will be a flattening of daily peaks and opportunities for arbitrage will decrease. Biggins et al. (2022) confirmed this in their study, where profits from arbitrage are on average overestimated by 28%.

Storage is particularly beneficial in the presence of negative prices. Negative prices on a national grid can occur due to an oversupply of electricity relative to demand, particularly during periods of low consumption and high generation from renewable sources like wind and solar. When production exceeds demand, and the grid's storage and export capacities are insufficient to absorb the excess, electricity prices can fall below zero. This phenomenon incentivises producers to reduce output, thus maintaining grid stability and preventing overloading. For a storage operator, this is a win-win situation, in which the operator makes a profit and charges their storage device, which can be discharged for further profit at a later stage. In a study of a wind power plant coupled with grid-scale storage by Zhou et al. (2011) profit is made by storing energy when the wind farm produces more than the transmission line connecting it to the grid can handle, preventing the need for wind curtailment. The economical advantage of storage grows when there are more negative prices. However the environmental advantage decreases as there is, on average, less storage space available and consequently more wind curtailment. McKone and Wolfs (2019) warns against the pursuit of negative prices because relying on them in business planning may be overly optimistic.

As well as selling to the grid the owner of an electricity hungry asset such as a desalination plant or ammonia plant can make their process 'green' by using renewables. In such a case the owners will need ESS to flatten the output of renewables such as wind or solar.

A gap in the literature is identified here as, while there is research into all three of the markets described above, there is little examination of how an ESS could optimally be

operated in multiple markets. For example, an ESS could be dispatched in the frequency market one day and the spot market on another depending on the characteristics of the day and which market is most profitable. Alternatively, the ESS could be dispatched in different markets in different years, depending on how the dynamics of the electricity system change and the ESS ages. In the literature this different market at different times approach is referred to as sequential stacking, while parallel stacking refers to providing many services simultaneously and has been recommend previously as a promising research field.

The time scales for some ancillary and energy markets overlap, implying that ESS could be capable of operating in markets simultaneously. In fact, both Wogrin and Gayme (2015) and Drury et al. (2011) explore this option with both concluding this is a profitable venture. A better understanding of ESS playing in multiple markets, simultaneously or otherwise, would benefit the restructuring of electricity markets to encourage investment in stability and three ancillary markets are explored in detail in the following sections.

2.4.2 Voltage support

An analysis of GB's voltage support is presented here as maintaining voltage within limits is a mandatory service in any electricity grid. The most common technology for offering voltage support is currently not ES but generators, capacitors and inductors and static VAR compensators however ES can provide this important service and ESS, particularly battery ES have been increasing in popularity in recent years. Providing voltage support in GB relies on the sale of reactive power.

A provider of voltage support to the national grid must respond to signals from National Grid ESO and keep their voltage between $\pm$ 10% of the requested nominal voltage. Heavily loaded overhead lines absorb reactive power and lower voltage. Low loads (typically occurring at the weekend and in summer months) lead to a necessary injection of reactive power and an increase in voltage. National Grid ESO has stated that over the next 5 years more reactive power absorption will be necessary as NDRs proliferate.

Zhou et al. (2011) noted that compared to other ancillary services, voltage support requires a fast response time, especially in GB where National Grid ESO requires the one of the shortest response times internationally.

Konidena (2020) proposes that, in order to incentivise the further proliferation of

ES providing voltage support National Grid ESO should offer location based prices for voltage support and that those prices themselves be set via reverse pricing. The first proposition follows from the locational variations in voltage on the transmission grid. Locational pricing enables dynamic resource allocation by providing pricing signals that reflect the real-time costs and needs of voltage support across different grid locations, fostering economically efficient deployment of resources. This way, congestion within the grid, particularly in areas with high renewable integration, can be mitigated. They consider that locational pricing could be implemented through a systematically structured Locational Marginal Pricing (LMP) model, wherein prices are set based on the marginal cost of serving the next unit of demand.

The second proposition allows providers to have more control over setting their pricing models.

Jay et. al. (2019) supported the first proposition for a hypothersised IEEE 24-Bus system with a high penetration of NDRs. The objective of the research was to show that the cost of reactive power to the system operator would be lower should the system operator adopt the locational pricing model. They also found that the resultant prices were less volatile.

In conclusion, the analysis underscores the critical role of ESS in providing voltage support within the GB electricity grid, a service traditionally dominated by generators, capacitors, inductors, and static VAR compensators.

Furthermore, the integration of biomass systems with ESS highlights a novel approach to voltage support. Biomass dispatch, informed by the SoC of the ESS, can complement the capabilities of energy storage providing voltage support.

2.4.3 Frequency regulation and inertia

Storage can be useful in a system with a developed market for frequency regulation. With increasing penetration of NDRs, there is likely to be an increase in over and under-frequency events in power grids. The rapid response of ESS can balance supply and demand and correct frequency deviations. Table 2 shows the technologies designed for frequency response (batteries, SMES and SCES). Frequency response can be sold on the day-ahead market (DAM) based on the TSOs predictions for the next day. Frequency response can also be sold in a real time market (RTM) which makes up for any miscalcu-

lations of the DAM.

Designing the ESS to have sufficient charge when needed to respond to a frequency event is complicated and involves careful planning of when an ESS should be charging and discharging. This is referred to as SoC restoration and is covered extensively in section 2.6. Zhang et. al. (2019) explored the optimal bidding strategy of an ESS providing frequency response and found that a good SoC restoration strategy is paramount. They found that their ESS could meet 100% of frequency response requests from the TSO using their algorithm, while the ESS could only meet 47.7% of requests when simply bidding the ESS full capacity. This is a key finding as failing to meet requests in the market can lead to penalties or disqualification from the market.

Frequency markets have been identified by Kintner-Meyer et al. (2012) as potentially the most profitable option for ESS operators. Markets in some counties are better suited to ESS participation than others. For example, the Dutch system (studied by Strassheim et al. (2014)) allows an ESS operator to withdraw their bid after being accepted in the DAM if they find they cannot meet their commitment. This lowers risk to the ESS operator and encourages their participation in opportunistic trading strategies.

The most developed frequency regulation market exists in the Irish transmission system, where there is a high penetration of wind generation (see Maldet et al. (2022) and Do et al. (2020)). GB splits its own frequency regulation into primary and secondary products. More specifically, primary products must be able to respond within 10 seconds and maintain their output for 30 seconds. Secondary response must be delivered within 30 seconds and maintained for 30 minutes. National Grid ESO allows bids to be placed up to 3 months in advance (*BMRS API and Data Push User Guide* (2019)).

Junyent-Ferr et al. (2015) explored the intersection of frequency regulation with inertia. They presented their findings on a low inertia future of GB's electricity market. In their model the power represented by dP (MW) resulting from a frequency deviation of df (s^{-1}) from the nominal frequency of f^N (s^{-1}) is given by:

$$dP \approx -\frac{2HS}{f^N}\frac{df}{dt} \tag{3}$$

where H and S are the inertia time (s) and power on the grid (MW) respectively. In this context inertia time is the time taken for the system to respond to a deviation in

frequency. Equation 3 suggests that there will be a bigger inertial response depending on $\frac{df}{dt}$, the RoCoF (rate of change of frequency). Rearranging equation 3 gives

$$\frac{dP}{S} \approx -\frac{2H}{f^N}\frac{df}{dt} \tag{4}$$

$\frac{dP}{S}$ is referred to as the headroom. It is the unitless fraction of the overall power that the system must be able to respond with during frequency deviations and has been used by Junyent-Ferr et al. (2015) to inform, for example, the level of ES that would be needed to balance the electricity grid.

Junyent-Ferr et al. (2015) present three scenarios for the response to the increase in NDRs on the grid. The first models the grid with high inertia (low inertial response time), the second with reduced inertia (higher inertial response time) and the third also with a reduction in inertia but a substantial fleet of ES to replace the loss. As expected, the second scenario shows a rather larger sensitivity to a change in frequency. The RoCoF is 10% higher in the second scenario than the former. With the addition of ES the RoCoF is restored to a lower rate. Also, the scenario including ES results in a transient drop in frequency that lasts only on the order of seconds before recovering to steady-state.

Synchronous generators can provide inertia to the grid, which protects it from over and under-frequency events. In fact, a key role of bioenergy on the grid is to provide inertia, especially as NDR proliferation decreases inertia. As will be described further in section 2.5.1 the clearing order of the market favours NDRs due to the low marginal cost of increasing their output. In short, when energy prices are low, NDRs are still willing to increase their output as their is little additional cost to them. Synchronous generators that provide inertia have fuel costs, meaning they will most likely be priced out of the market when energy prices fall.

Missing from the literature is a full analysis of whether bioenergy could realistically provide the inertia the grid requires on a fossil fuel free grid, especially in the case of open cycle gas turbines (OCGTs) which provide fast ramping load following to the grid. However, advancements in biomass preprocessing, like torrefaction, and innovations in combustion technologies are aspects under research to improve the overall responsiveness and efficiency of biomass power plants (Ho et al. (2012)).

It has been argued by Junyent-Ferr et al. (2015) that power electronics attached to

NDRs can effectively provide a form of synthetic inertia. The major drawback of this approach is that inertia provided by synchronous generators is capable of self-correcting. The physical mechanics providing inertia mean that their is no need for a form of central control. Synthetic inertia requires central control and thus is more prone to mistakes as documented by Ackermann et al. (2017). This aspect also requires a more careful design of how National Grid ESO will manage frequency events.

National Grid ESO is going through a period of reform in which it is developing products more suitable to ESS as described in *The Electricity Trading Arrangements* (2019). The electricity market reform of National Grid ESO for this purpose began in 2013 and figure 7 shows the effects of the reform in its historical context.

In conclusion, ES has the potential to support the national grid and help it balance electricity frequency. Parameters of ES suggest it could profit from current frequency products offered by National Grid ESO and could profit further as National Grid ESO reforms. Inertia from synchronous generators is valuable as it does not require management. Bioenergy may be offer the inertia at relatively low emissions but further research is needed.

2.5 Future of electricity markets

The future of the electricity markets in the UK will be very relevant to chapter 6. Electricity futures are an unpredictable investment. According to Grubb and Newbery (2018), average time from initial investment in a grid-scale capacity project to the start of the project is six years.

National Grid ESO has published projections to encourage confidence among potential investors. This section of the literature review focuses on the *Future Energy Scenarios* (2022) (FES) (as published by National Grid ESO) which outline scenarios for the future of the electricity grid. The FES are updated each year. The 4 scenarios reach to 2050, 3 of which achieve net zero by 2050. The scenarios can be plotted on axes of level of social change and speed of decarbonisation. In brief, they are

- **Falling short**: Low level of social change and slow speed of decarbonisation. There is little change to societies consumption patterns combined with little decarbonisation Compared to the other three scenarios. By 2050, this scenario sees a 80% drop in emissions relative to 1990.

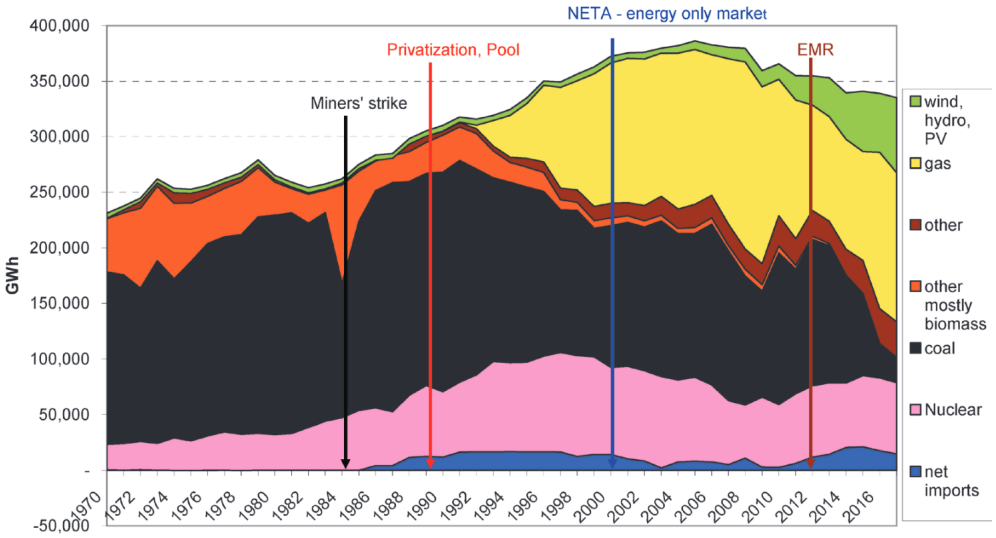

Figure 7: The history of capacity in GB since 1972. Reproduced from Grubb and Newbery (2018)

- **System transformation**: Mid level of social change and medium speed of decarbonisation. Hydrogen is used to partially decarbonise heating and supply side flexibility expands to incorporate renewable proliferation.

- **Consumer transformation**: High level of social change and medium speed of decarbonisation. Consumer changes mean that the electricity system is characterised by demand side flexibility rather than supply side.

- **Leading the way**: High level of social change and high speed of decarbonisation. The speed of decarbonisation is swift due to the fact that the price of commissioning wind and solar projects is lower than gas by 2025.

Its unclear in the future energy scenarios what policies inform the movement to increased renewables and reduction in demand. Fan et al. (2018) have documented the inability of the liberalised energy market to generate the price signals necessary for the requisite investment.

Ofgem have leaned towards fixed feed-in tariffs for offering incentive to invest in their paper *Guidance for generators: Co-location of electricity storage facilities with renewable generation supported under the Renewables Obligation or Feed-in Tariff schemes* (2020). The argument for fixed feed-in tariffs (FiTs) was popular around the early 2010s during the early days of renewable penetration. Accreditation decisions for such energy instal-

lations are diligently made on a case-by-case basis. For installations no matter the size, it's imperative for a FiT licensee to validate compliance. Owners of FiT installations must continually comply with stringent metering prerequisites to qualify for support payments, with the FiT scheme disallowing the use of multiple meters to record any electricity imported or to deduce net generation. It is essential that FiT licensees guarantee that generation payments, especially in scenarios where storage is co-located with an accredited FiT installation, are precisely made only for electricity that is verifiably generated by the FiT installation to ensure regulatory and financial integrity of the system.

National Grid ESO suggest that green certificate trading is more effective than fixed feed-in tariffs for incentivising investment (*Future Energy Scenarios* (2021)). The Renewable Obligation Certificates (ROCs) were a type of green certificate introduced in 2002 and terminated in 2017 and required suppliers to cover some percentage of their volume with RoC certified energy. To be in adherence, electricity produced by an accredited station must also be utilised in a 'permitted way', encompassing several scenarios such as being consumed by the operator, supplied to customers in GB or Northern Ireland via a private wire network, exported to a distribution or transmission system where customer supply cannot be evidenced, or any combination thereof. Particularly crucial is the scenario involving generating stations with co-located storage facilities, wherein operators must decisively illustrate compliance with the 'permitted ways' to qualify for the issuance of RoCs. Further nuances dictate that only the net renewable electricity, defined rigorously as

$$\begin{aligned}\text{Net renewable generation} = (\text{Gross output} - \text{input} \\ \text{electricity}) \times \text{Renewable qualifying percentage}\end{aligned} \tag{5}$$

The feed-in tariff with contract for difference (CfDs) have also been explored in Fan et al. (2018) and found to be a favourable incentive scheme. A CfD guarantees a price to generators, expecting them to pay the difference if the market price goes above the guaranteed price and making up the difference in the case of a market price shortfall. Results suggested the guaranteed price should be set high (10 p/kWh) before coming down as new renewable installations bring down the cost of energy.

In short, there is no consensus on whether feed-in tariffs, CfDs or green certificate

trading encourages more investment. This may come from the range of measurement metrics used, including 'fitness' (a sum of subsidy cost, energy cost and the wholesale price on the grid) in Fan et al. (2018) and risk in Bunn and Yusupov (2015).

2.5.1 Future of ancillary services

The national lockdowns of 2020 and 2021 during the covid-19 pandemic can be informative when considering the future of ancillary markets. During this time demand across GB was decreased considerably without a proportionate decrease in renewable output (up to 25% decline in electricity usage (Kirli et al. (2021))), effectively increasing the percentage of electricity coming from renewables. Despite the lower energy prices caused by the lack of demand, NDRs are always incentivised to sell as much electricity as possible due to their low marginal cost of increasing their output and the financial instruments that protect them from low energy prices. Synchronous generators have a higher marginal cost of increasing their output as they rely on fuel costs. The declined reliance on synchronous generators lead to a fall in inertia. Additionally, National Grid ESO has a legal responsibility to maintain as much ancillary services as can cover the unexpected shutdown of the largest provider of capacity at any one time as documented in the *GSR023: Clarification of the applicability of the N-1-1 criterion* (2018). As nuclear plants ramping and shutdown costs make them relatively impervious to energy price fluctuations, the largest provider at any time is likely to be a nuclear plant.

While the decline in energy prices has been documented, National Grid ESO paid 3 times the 2019 average for ancillary services during 2020 (National Grid ESO (2022)) as downward flexibility was paramount. The main expense went towards a new product Optional Downward Flexibility Management. The new market attracted 4.5 GW (National Grid ESO (2022)). This increase in price of flexibility suggests the growth of the ancillary market as NDRs proliferate in the coming decades.

There have been studies into increasing the efficiency of these ancillary markets up to 2050. Vivero-Serrano et al. (2019) examines the idea of linked charge/discharge bids on profitability of ESS operating in ancillary markets. Currently National Grid ESO runs its BM on linked price/quantity bids, where a bid is submitted as a £/MWh-MWh pairs for charging and discharging separately. This setup is common in Europe and for several of the US TSOs. For several of these TSOs the ESS is considered a load or a generator

depending on whether it is offering to charge or discharge, while in GB the ESS belongs to a third class of bidder that can provide both services.

Vivero-Serrano et al. (2019) uses a bilevel model; one minimising the expenses for the storage operator and the other simulating the operation of the grid and find that the linked charge/discharge bids are most profitable for ESS entering the market. The results also indicate that a lack of transparency from the transmission operator about how the market clears lowers ESS profits but also makes the market less efficient.

Huang et al. (2018) explores the degree of centralisation optimising the participation of ESS in ancillary markets. They analyse 3 degrees of centralisation. In exploring the implications of centralised and decentralised markets for ESS operators, certain hypothetical scenarios emerged that underline divergent opportunities and challenges across these market structures. Huang et al. (2018) found that a centralised market might ostensibly present a structured and stable environment for ESS operators. In this scenario, a centrally regulated price mechanism minimises market volatility, potentially providing ESS operators with predictable and stable revenue streams. The possibility of large-scale grid stabilisation projects could also be higher, given the centralised procurement of ancillary services. However, ESS operators face stringent regulations and are susceptible to policy changes that are orchestrated to balance the larger grid's needs rather than optimising for localised solutions (National Grid ESO (2022)). The ESS operator's contribution to grid stability was dampened by rigid tariff structures and potential bureaucracy in adapting to technological innovations and dynamically changing market conditions.

Contrastingly, a decentralised market was found to pave the way for a more flexible, localised, and agile energy landscape. Innovative ES solutions were found for providing ancillary services. ESS operators were exposed to enhanced risks due to price volatility and inconsistent regulatory frameworks across different regions.

In conclusion Huang et al. (2018) found that the 2 most decentralised options provided the same operational ESS dispatch, leading the authors to conclude that the semi-centralised market structure would be most advantageous, allowing the operator to impose socially beneficial constraints.

Pandžić et al. (2018) looked into developing ESS using private investment or relying on TSOs to make decisions about new capacity of ES or re-enforcing transmission lines to reduce ancillary service costs. The results showed that it is always better for the system

operator to invest in strengthening their transmission lines rather than building ES, due to their long lifetime. In a case study where where the TSOs were restricted to only investing in ES, it was more advantageous for new ES capacity to come from private investment. Pandžić et al. (2018) is among a subset of papers that consider social welfare, and find that dependency on private investment is also best for social welfare.

2.5.2 Future energy scenario implications for medium-duration energy storage

This section focuses on medium-duration ES (MDES). The justification for this is partly due to projections from the FES which will become relevant in chapter 6. MDES, typically defined by an energy-to-power ratio ranging from 4 to 200 hours, plays a crucial role in balancing supply and demand over several hours to days. Unlike short duration systems, which handle rapid fluctuations and provide grid stability for short periods, MDES addresses longer-term discrepancies in energy availability, such as those caused by the diurnal cycle of solar power or extended periods of low wind generation. These systems are essential for integrating renewable energy sources into the grid, as they can store excess energy generated during peak production times and discharge it when production is low. By doing so, MDES helps to smooth out variability on the grid.

Transitioning to a net zero energy system: Smart Systems and Flexibility Plan 2021 (2022) illustrates an approach towards fortifying the UK's electrical storage and interconnection capacities through the mid-2020s and beyond to 2030. A chief aspiration is to establish a premier regulatory framework that propels the deployment of electricity storage.

Figure 8 shows projections for a typical load in Leading the way. In Leading the way there is a high projected capacity of NDRs and ES. From 00:00 - 05:00 demand is low but NDR capacity is so low as for there to be a net requirement of further capacity to be dispatched. Later in the day, load ramps up, as does NDR generation, creating a surplus. Load then continues to increase outstripping NDR generation and this continues into the evening as both load and NDR generation decrease. The ES projected capacity can charge during the hours of the day with an NDR surplus (as shown in yellow) and discharge during periods of high load relative to NDR delivery (as shown in blue). This analysis suggests that an MDES could be appropriate for managing the fluctuations of

NDRs over a 24 hour period.

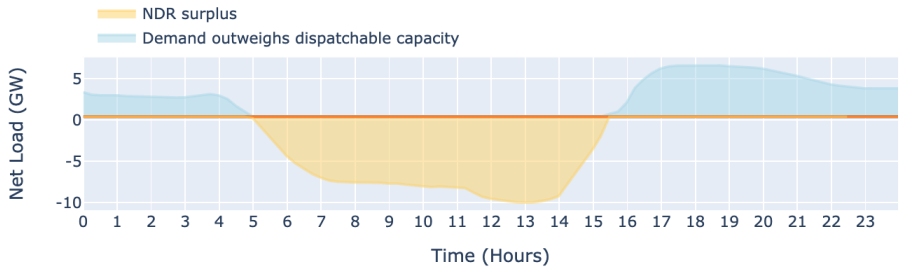

Figure 8: The net load in 2050 through an average day in Leading the way. The periods of non-dispatchable renewables surplus are shown when the predicted capacity of energy storage can shift load to periods when demand outweighs dispatchable capacity.

The length of time an ES can deliver its rated capacity is the parameter that defines the duration as listed for various technologies in table 2. An MDES would be needed to fill the intra-day storage gap projected.

It was assumed that each day the battery charged during periods when the NDR capacity outstripped demand and discharged when the NDR capacity could not match demand. The operational costs per unit energy charged and discharged were fixed at the average calculated in section 2.1 and presented in figure 2 which is a simplification of truly running a lithium ion battery but demonstrative of the range of total operational costs accrued throughout the day. Figure 9 shows the frequency of operational costs for operating a lithium ion battery in Leading the way. The results were found by running a simulation with the capacities listed in Leading the way against wind and solar data from 2019. From figure 9, the two plateaus at the highest and lowest daily expenses represent days illustrative of the seasonal effects of demand. Low operational expenses occur on days where NDR production generally match demand and high operational expenses occur on days as depicted in figure 8 where there is high NDR production stored followed by a deficit in the evening - nightime hours. The majority of days (80%) sit in the middle range.

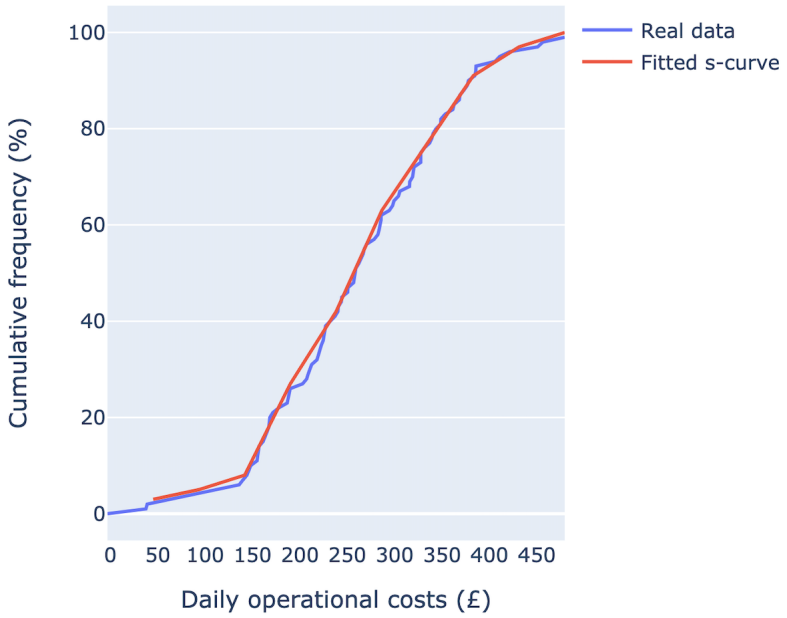

Figure 9: Cumulative frequency of operational costs for a lithium ion battery in Leading the way. Based on the operational costs shown in figure 2.

With the MDES established, there are two further points to consider from the literature that are addressed in this thesis:

1. **Levelised cost of electricity:** There is a blindspot in the literature that overestimates the levelised cost of electricity (LCOE) when calculating the LCOE for renewables in papers such as Schoenung and Hassenzahl (2003), Sundararagavan and Baker (2012) and Khosravi et al. (2021). Reported LCOEs fail to incorporate the rising LCOE when some percentage of renewable energy is curtailed. For example, an x factor decrease in output due to curtailment increases the LCOE by a factor of $\frac{1}{1-x}$. The non-dispatchability of renewables adds significant cost to their effective LCOE. This additional cost of lack of flexibility provides another argument for investing in MDES.

2. **Demand side markets:** Demand side markets (DSM) were under-studied in the literature in the early 2000s according to due to their being relatively immature. The BM was historically reliant on load-following power plants (generally carbon intensive gas) and there is very little inclusion of DSM on the BM. However interest in it has increased and the last comprehensive report on DSM from National Grid ESO was published for 2016/17. In that year they purchased a total of 2.71 GW of

demand side response (DSR) from a combination of frequency response, short term operating reserve and fast response. 0 MW of this came from ES in 2016/17 and 10 MW from load response. The remainder came from generation balancing. National Grid ESO is planning to change this in favour of more ES and load response through financial incentives for MDES as documented in their paper *Transitioning to a net zero energy system: Smart Systems and Flexibility Plan 2021* (2022).

In summary, the FES point to a need for further investment in MDES and MDES affords the grid more flexibility which National Grid ESO is bound to invest in further according to their Smart Systems and Flexibility Plan.

2.6 Combined biomass and energy storage

This section describes an ESS supporting a biomass generator, referred to throughout this thesis as a combined biomass and energy storage system (CBESS).

Generation Integrated Energy Storage (GIES) refers to the integration of energy storage systems directly into power generation facilities or plants. Instead of having standalone energy storage systems, GIES combines the capabilities of ES with the power generation assets themselves. In general, GEIS enable power plants to respond rapidly to grid demand fluctuations and support grid operations. Lai and Locatelli (2021) explored this concept for thermal energy storage demonstrating the flexibility that improved efficiency by reducing the need for cycling of the generators. They demonstrated black start capability and reduced the downtime associated with generator ramp-up, enabling faster power restoration after disruptions. National Grid ESO have explored the benefits of GEIS to the transmission grid (*Future Energy Scenarios* (2021)) in overcoming congestion issues. By absorbing excess generation during periods of congestion or high grid demand and releasing it during off-peak periods, GIES helps alleviate grid bottlenecks, reducing transmission constraints The CBESS is a specific example of such a GIES.

Often discussed in the literature are the 'degrees of freedom' ancillary service markets offer and how these affect SoC restoration (see Rancilio et al. (2020)).

To elaborate, various products contain different degrees of freedom. Purchasing the old enhanced frequency response product from National Grid ESO requires a response 'within a deadband' i.e. within a fixed range of volume response ($\pm$ 9%). Purchasing

other products requires the delivery of a fixed volume, tightly regulated by National grid ESO. The degrees of freedom affect the SoC restoration of the ESS (how the ESS chooses to manage its SoC). Higher degrees of freedom allow the ESS operator to manage the SoC in such a way as to maintain the desired SoC with more accuracy and avoid cycling if that is the wish of the operator. For ESS and particularly battery ESS a higher degree of freedom is preferred.

The CBESS offers a unique opportunity for SoC restoration and management with charging coming directly from the biomass generator at some times instead of always relying on charging from the grid, an SoC restoration pattern under-researched in the literature. The CBESS is therefore an interesting system paired with a low degree of freedom ancillary product. The CBESS will be particularly useful considering the skew towards over-frequency events on GB's national grid.

There is a large gap in the literature concerning CBESS but there a few studies, particularly on other technology types being supported by ES. For example, Razmi et al. (2021) studied the operation of a CHP plant supported by compressed air ES and Azad et al. (2015) looked into supporting a windfarm using an ESS. Both papers find that the ES reduced operational costs. Azad et al. (2015) is particularly detailed, addressing voltage control by the combined system, illustrating that the GIES plays an integral role in enhancing the dynamics of voltage profiles, particularly through the management of both active and reactive power. For wind power it is important to address transmission curtailment, and Azad et al. (2015) show that ESS can store and controlledly inject energy, ensuring stability and avoiding the disconnection of wind turbines.

Also, present in the literature, are detailed descriptions of biomass ramping rates and start-up costs can be found in Van den Bergh and Delarue (2015) which are paramount in understanding how a CBESS would potentially operate. For example, in a unit commitment problem, a methodology often used for developing operational strategies, 40% of of the costs can be saved by fully accounting for cycling. Cycling costs are particularly important to consider for biomass plants, which, more so than other fuels, suffer wear and tear when cycling. They are bound to go up as NDRs proliferate and a bioenergy plant may be called upon to cycle its output. Cycling costs can be split into 4 categories

- **Efficiency costs**: Wear and tear harm mechanical parts and the equipment becomes less efficient.

- **Maintenance costs**: Costs of replacing equipment broken due to wear and tear.
- **Direct start-up costs**: Additional fuel costs incurred during start up or shut down.
- **Ramping costs**: Additional fuel costs incurred during ramping.

The National Renewable Energy Lab (NREL) has reported that a largely underestimated contributor to the cycling costs are unplanned outages that can cause maintenance costs and direct start-up costs.

Finally in this section is a discussion of how ESS can be sized when co-located with various forms of generation, including a biomass generator. There are a range of papers dedicated to sizing ESS for synchronous generators including Lian et al. (2014). A common technique for sizing ESS is to use linear programming, maximising a certain metric subject to constraints. Khosravi et al. (2021) chooses to maximise the net present value of ES in their analysis and in addition use a min-max method to propagate uncertainty.

2.7 An EU comparison

There are a range of regulatory frameworks across the EU. National Grid ESO has favourable attitudes towards innovation according to Biancardi et al. (2021), though it has an extremely poor track record with regard to energy storage.

An examination of patent applications across various countries by Baumann et al. (2021) can be informative when looking at how technologies have developed. Patent applications for a technology have been shown to reach their peak when that technology is first penetrating the market. GB has a high number of patents on bioenergy innovation compared to other EU countries, while Germany and France lead the way on lithium ion batteries (also documented in Baumann et al. (2021)).

Cross et al. (2021) studied the role of bioenergy in several EU countries. Bioenergy for electricity in the UK was comparable to the Nordic countries; Denmark, Sweeden and Finland. The development of bioenergy in all four countries was driven by policies focused on energy security.

The regulators and TSOs play a key role in ensuring the security and affordability of a grids electricity supply. They have increasingly been put under pressure to innovate during international efforts to decarbonise. Biancardi et al. (2021) examined the proposition that the framework adopted by a TSO would encourage a higher level of innovation. They

found that TSOs are generally unwilling to innovate at all and require incentives from regulators to make any changes. The average European TSO spent less then 0.5% of the their profits on innovation according to a 2016 report carried out by the European Network of Transmission System Operators for Electricity (ENTSOE) (*Research, development and innovation roadmap* (2016)). Their SWOT (strengths, weaknesses, opportunities and threats) analysis identified the a major threat to innovation to be a tendency towards short-term over long-term projects. More long-term goals include enhanced transmission components, power electronics converters, digitalisation, amalgamating intelligent controls with hardware infrastructure and superconductivity applications. The exploration and incorporation of new storage technologies and management systems might augment TSO capabilities (*Research, development and innovation roadmap* (2016)). Pandžić et al. (2018) supported this with their conclusion that the best way for the TSOs to invest was to strengthen transmission lines, a relatively high-cost investment. A key recommendation of the ENTSOE report is that national governments enforce TSOs to commit more of their budget to research and development.

According to ENTSOE and backed up by Pandžić et al. (2018) half of all investment for the integration of ES should come half from EU contributions and half from private investment. However, the majority of investment for innovating market design should come from the TSO, meaning that that the TSO is indirectly responsible for encouraging new ES capacity.

2.8 Markov decision process models

Any storage device faces the question of when to charge and when to discharge. The task of developing a bidding strategy (both MW dispatch and price bid) can be split into two parts; firstly creating a model that captures the relevant features of the universe being considered, secondly choosing a strategy to examine the universe and decide upon the bidding strategy that optimises reward. A simple solution from from Durante et al. (2017) is to set a price that the marginal price of the market must exceed, at which point the storage device will discharge. The storage device will charge when the marginal price drops below some other lower threshold. The threshold is set based on previous prices and in the simplest cases is not adjusted as the year progresses. In other cases the threshold can be adjusted as time progresses and data is added to the historical data set (for an

example and further analysis of this see Martins and Miles (2021)).

A more sophisticated set of solutions are found by having a rolling window, that allows decision makers to see into a window of future events, that is updated as time moves forward (demonstrated by Pimm et al. (2020)). This method is called model predictive control, and its complexity is dependant on the length of the rolling window chosen and the objective function. Arnold and Andersson (2011) expressed the disadvantages of such approaches that include overestimating returns during times of consistently low prices.

This review will focus on using a Markov decision process (MDP) to design the universe as MDPs are well-suited to considering future events without assuming knowledge of what will occur in the future with 100% certainty.

A Markov decision process is a stochastic process for which the probability distribution for moving into another state is only dependant on the current state and independent of any past decisions. To develop an MDP, there must be a state space, a list of possible actions that can be taken and a reward associated with each state and action. The future rewards are compared to establish a policy function: a list of the most rewarding actions that can be taken in a given state and a value function: a representation of rewards recouped by taking a certain action. The following is a description of the features of various MDPs used in the literature to study ES.

2.8.1 State space

First of all, the MDP state space can be continuous as in Löhndorf and Minner (2010) or discrete, each possessing distinct characteristics and applicable use cases, with inherent advantages and disadvantages. Sutton and Barto (2020) and Tang and Agrawal (2020) can be read for further comparisons of the two approaches but for this analysis it is important that, owing to their simplicity, models with discrete states are generally easier to implement and debug, making them more accessible in various practical scenarios, especially where states naturally fall into distinct categories (e.g., inventory levels, task completion stages). Discretised methods are subject to the curse of dimensionality at increasing resolution. In the continuous case the value function is replaced with an approximated value function and Löhndorf and Minner (2010) finds the approximated policies to be more rewarding than optimal policies of a discretised space. Continuous state spaces can represent environments with high fidelity and precision since they are not constrained to

discrete values or intervals. Many real-world phenomena, especially in physics and engineering, naturally exist in a continuous state (e.g., temperature, speed, and position), making continuous state spaces relevant for modeling such scenarios. This format can handle complex, high-dimensional problems by capturing more nuanced information and detailed variations within the environment. However, ensuring convergence to an optimal policy or value function is often more mathematically intricate and computationally demanding in continuous spaces. Modeling and learning in continuous spaces requires larger datasets to adequately capture and generalise across the state space.

2.8.2 The transition matrix

The transition matrix determines how the system evolves between states and is made up of an array of probabilities for moving from one state to another.

Mitkovska-Trendova et al. (2014) describes determining a transition matrix that is truly representative of the system being described is vital. Accurate results can only be obtained with accurate probabilities in the transition matrix and confidence in results of an MDP are affected by accuracy of transition probabilities. In the case of modelling electricity markets, uncertainties which must be captured include; uncertainties about supply and demand in future time periods and uncertainties about system constraints such as transmission capacity.

A common technique for constructing a transition matrix demonstrated by Haji Bashi et al. (2018) is to obtain historical data and assign probabilities of state transitions based upon historical frequency. In this case the reliability of the historical data is important. Haji Bashi et al. (2018) found the use of historical data to be a straightforward and intuitive approach that is based directly on observed system behavior. Despite the apparent advantages Haji Bashi et al. (2018) also noted that, in real-world scenarios, where the system dynamics may evolve, the transition matrix may age quickly. Mitkovska-Trendova et al. (2014) uses the methodology of historical frequency augmented by expert opinion, leveraging domain-specific knowledge which might be inaccessible to data-driven methods. Using expert opinion can be useful when empirical data is scarce or unavailable though prone to bias and might lack the rigor of data-driven approaches.

Further, Mitkovska-Trendova et al. (2014) examined theoretical modeling where transition probabilities are derived from theoretical models and equations, often based on

scientific principles that describe system dynamics. Typically grounded in scientific or operational research principles and, given a solid foundational basis, theoretical modeling can predict transitions in scenarios even without historical data. This method may incorporate assumptions that might not always align with real-world complexities. With respect to ESSs, Mitkovska-Trendova et al. (2014) suggest it would be possible to create a transition matrix by employing physical and engineering principles related to battery chemistry, degradation, and energy management.

MDPs with a time dependant transition matrix are referred to as non-time-homogenous Markov models. As electricity markets are clearly periodical, it follows that the probabilities of certain transitions will be higher or lower depending on the time of day. This can be taken into consideration by segmenting historical data into hour long time periods and using the historical frequency method to develop a transition matrix for each hour.

2.8.3 Multi-player models

MDPs can be developed into Makov games, MDPs that include several players. For simplicity, studies often adopt the assumption that the actions of an individual generator have no effect on the marginal price of the market. However, the MDP can be expanded to model the actions of several generators and therefore takes into account how the actions of a single bid may influence decisions taken by other generators as each agent aims to maximise its own objective or utility function, which may or may not align with the objectives of other agents.

Ye et al. (2019) demonstrated the use of multi-player MDPs analysing imperfect electricity markets. They do this to reiterate the point made above that it is unreasonable in such circumstances to assume price taker status. Ye et al. (2019) highlights issues in previous papers such as severe modeling and computational complexities and the need for complete knowledge of techno-economical characteristics. In response, Ye et al. (2019) proposes a novel methodology. This innovative approach allows the multiple players to gain precise feedback on the impact of their offering strategies on the market clearing outcome and enabling them to formulate more profitable bidding decisions by thoroughly exploring the entire action domain. The effectiveness of this methodology was validated through case studies on a test market operated over the IEEE Reliability Test System, demonstrating that the proposed method promises significantly higher profits than existing

RL methods.

2.8.4 Solving strategies

There are several methods for determining the optimal bidding strategy. An MDP can be solved with dynamic programming i.e. policy iteration as in Zhou et al. (2011) or value iteration as in Song et al. (2000). These solving strategies suffer from the curse of dimensionality with a large state space. Although policy iteration is faster at converging, it requires an initial guess to begin. With reference to section 2.8.2, these solving strategies are only appropriate when there is a high level of confidence in the accuracy of the transition matrix. They require a full knowledge of the environment and never learn from experience.

The following set of solutions can be applied to problems with a continuous state space. Approximated value functions allow a model to explore the state space and represent the value of any state. In Luo et al. (2019), the authors develop a method of reinforcement learning that finds a balance between the advantages and disadvantages of value and policy iteration. Their methodology doesn't need an initial guess and converges reasonably quickly. Reinforcement learning has been implemented for energy storage operating in parallel with a wind farm in Oh and Wang (2020) and to model a battery in Cao et al. (2020).

In summary, the economics look promising as capital costs of ESS are projected to continue to fall over the next decade (Way et al. (2022)). Worldwide investment and international pressure to keep global emissions low will consistently drive down ES costs as they benefit from increased capacity. Despite this, the market still needs to offer the right incentives if capacity is going to keep up with projected targets. This literature review has detailed how markets are changing and the National Grid ESO projections to 2050.

However, the FES are projections of the best capacity investments, in the case of Leading the way, already assuming that the climate goals will be met. Less studied is which policy structures could realistically lead to achieving the climate goals. The little literature addressing this problem show no consensus.

Finally, MDP models show a promising route to planning generation dispatch in upcoming electricity markets. Crucially, the transition matrices developed to understand the

markets can act as an important feed-back loop. The transition matrices can inform market designers about how market players are interpreting the market and the probabilities being considered making operational decisions.

2.9 Methods

The listed methods are meticulously structured to align directly with the stated objectives, ensuring a cohesive approach to addressing each research goal. Every methodological step has been itemised to correspond to an individual objective, with **M1** relating to **O1** and **M2** to **O2** etc. The intention being that there be a clear and systematic exploration of the research that enhances the comprehensibility and traceability of the results.

- **M1** Read and compile most relevant literature. Identify gaps in the literature.
- **M2** Start by identifying a trustworthy source for frequency data and price signals. Establish statistical trends in the data sets. Connect trends to changes in the electricity system such as new markets and changing electricity mix.
- **M3** Start with the simple test case of a ES system and explore methodologies for making operational decisions. Apply the methodology to the biomass and ES combined system.
- **M4** Make operational decisions using perfect foresight. Compare the perfect foresight agent to methodology designed with **M3**.
- **M5** Carry out calculations of net present value for the addition of co-located ES to identify the most informative case study.
- **M6** Analyse results generated during **M5** and compare profits across the range of markets and years. Extract useful information and conclude with the most beneficial to potential CBESS operators.

Operational decisions for alternatively biomass generators and ESS have been generated by solving a linear programming problem such as is common throughout the literature for various generation setups (see Bracale et al. (2015), Pereira et al. (2014) and Pimm et al. (2020)). Such cases present a best case scenario for operating a generation system. Currently perfect or rolling foresight models rely on their knowledge of upcoming events

to make operational decisions. These papers, very common in dispatch planning, predict out-sized profits and vast carbon reductions that cannot be achieved by an agent acting myopically. Arnold and Andersson (2011) demonstrated how model predictive control (which employs a rolling window of foresight) can overestimate the efficiency of ES compared to a model with an uncertain forecast. This common mistake, has informed the decision in this thesis to focus on generating operational decisions without foresight. A reinforcement learning (RL) agent can be trained to make operational decisions with no foresight. Such an agent delivers a more realistic picture of how a CBESS operates to deliver electricity to National Grid ESO. The best case scenario can be used to train a RL agent in how to behave. After training, the RL agent can make no foresight decisions that will replicate the behaviour of a perfect foresight agent (PFA). The RL training is the methodology explored in this thesis described by **M3** and fulfilling **O3**. To illustrate the difference figure 10 shows a PFA and an RLA operating with and without foresight respectively. For a problem with a horizon of N time periods the PFA optimises over all the observations $(o_1 - o_N)$ producing a set of N actions. These actions can be used as training data to teach the RLA. Once trained, the RLA acts as depicted in figure 10 producing an action in each time period from a single observation.

2.10 Scope of research

Figure 11 shows the original contributions to the literature the thesis can provide. Turquoise circles identify topics commonly covered in the existing literature. Short duration ES (SDES) and medium duration ES (MDES) have been studied previously as have various electricity generating technologies, optimised to operate on both the BM and providing FFR with perfect foresight. The original contributions (shown in dark blue) include modelling a biomass generator supported by an ESS and examining the operation of a biomass generator providing FFR. A significant contribution lies in using a reinforcement learning agent to mimic the behaviour of the perfect foresight agent.

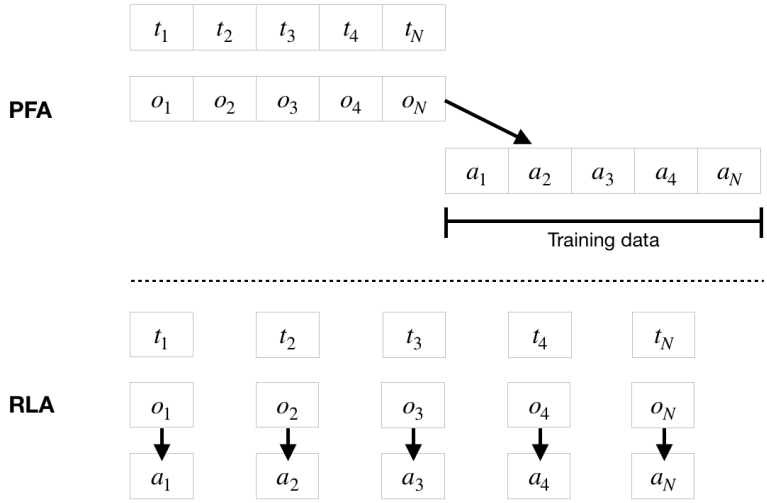

Figure 10: A depiction of the training regime for a perfect foresight agent and a reinforcement learning agent. The perfect foresight agent takes the observations (o_n) for N time periods and uses them all in a single optimisation to produce a set of actions. This set of actions is optimal and used as the training data for the reinforcement learning agent. Once trained the reinforcement learning agent can generate actions taken from one observation from one time period i.e. without foresight.

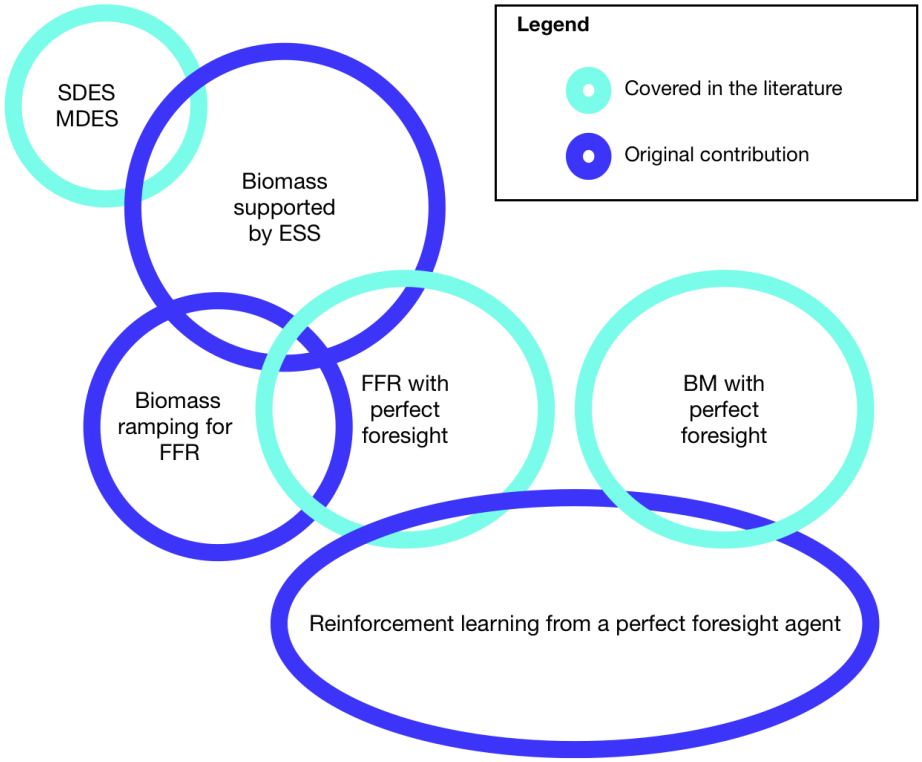

Figure 11: Original contribution thesis has to research field.

3 Methodology

This section presents the methodology repeated throughout the remainder of the thesis used to establish the best operational performance of the system being considered. It describes the setting up of a linear programming problem for operation of an agent with perfect foresight. It then goes on to introduce the reinforcement learning (RL) techniques for learning from the perfect foresight agent (PFA). Finally, this methodology's application to other chapters is outlined.

The key feature of the PFA is that it has perfect knowledge of the environment, what will occur in the future and can plan scheduling to maximise profits. The actions of the PFA will be used subsequently to teach the RLA.

The problem of solving the PFA is constructed as a linear programming problem. The objective function for the PFA maximises profits subject to constraints introduced by the technical limits of the ESS and biomass generators specific to each chapter.

The linear program problems were solved using Matlab's `linprog` and was undertaken on ARC3, part of the High Performance Computing facilities at the University of Leeds, UK.

During RL, an agent learns by updating a state-action value function ($Q(S, A)$). The state-action value function represents the value of the agent, when finding itself in state S, choosing action A.

In each learning step the agent moves from state S to S' via action A. The information learnt during this step is added to the state-action value function. $Q(S, A)$ was updated using temporal difference learning in the following way:

$$Q(S, A) = Q(S, A) + \alpha[\ R + \gamma Q(S', A') - Q(S, A)] \tag{6}$$

where R is the reward generated during the transition from state S to S' via action A, α is a step-size parameter and γ is the discount factor. Looking at the form of equation 6 the value function updates with the reward (R), a high reward increases the value of the of the state and action and encourages the RLA to repeat the action. $Q(S', A')$ is also present in equation 6 and used to update the value function. $Q(S', A')$ is an estimation of the value of going into the next state and the subsequent action. Together $R + \gamma Q(S', A')$ represent the value of the current state and action because the immediate reward (R) is present

and all future rewards (represented by $\gamma Q(S', A')$). The discount factor (γ) controls how the agent measures distant rewards as opposed to those in the immediate future. γ needs to be in the range [0,1). A discount factor of 0 implies the agent is myopic. A discount factor approaching 1 suggests the agent will base its actions on the outcome of all future rewards.

The square brackets contain the difference between $R + \gamma Q(S', A')$ and $Q(S, A)$. This difference gives the name to the methodology temporal difference learning. It is the difference between the estimation of $Q(S, A)$ and the new estimation of the value function $R + \gamma Q(S', A')$. For example if the old estimation $Q(S, A)$ is 1 and the new estimation $R + \gamma Q(S', A')$ is 1.5, equation 6 becomes $Q(S, A) = 1 + \alpha(1.5 - 1)$. If α were 1 the new value of $Q(S, A)$ would be 1.5.

In reality α is rarely set to 1. The purpose of the step-size parameter is to moderate the learning, preventing a single update having too great an influence on the state-action value function as a whole. If in the example α were set to 0.5, the new value of $Q(S, A)$ would be 1.25. In that case the update still moved $Q(S, A)$ from 1 towards 1.5, but the influence was reduced.

After some number of iterations the 'best' action can be found for a given state by maximising over the set of actions.

$$a = max_A(Q(S, A)) \tag{7}$$

3.1 Reward function

The reward function is used to evaluate the action taken, its purpose is to take information from the environment and return a single number that can be used in equation 6. The reward function can be a function of any kind, so long as it maps many variables to one. The agent uses the reward function to maximise its reward, the actions taken should maximise cumulative reward, rather than immediate reward. A well designed reward function will be vital to ensuring that the agent converges on the global optimal rather than a local optimal. There are no specific rules across the literature for choosing a reward function though there are many papers published in the field. Zakaria et al. (2021) have shown that correct selection of the reward function can discover the global maximum and increase the speed of their agent's learning.

The research field uses certain metrics, which will also be used here, to evaluate the merits of different reward functions, namely reward average.

Reward average references the mean reward over every time-step as learning continues. Reward average should increase as the learning continues, however it will naturally fluctuate as the agent finds new areas of the state space to exploit. Reward average is not the most sophisticated metric developed throughout the literature but has been shown by Sutton and Barto (2020) to work for problems (such as the ones presented in this thesis) that employ temporal difference learning. Further justification will be given in this thesis for the use of a PFA but an important one is outlined here. The reward average can constantly be compared to the optimal action during learning; the action taken by the PFA. This will be vital for assessing the usefulness of the reward function being used, enable correct shaping of the reward function and parameter tuning of the reward function.

In subsequent chapters, there will be descriptions of the particular reward function developed for those RLAs.

3.2 ϵ-greedy policy

To choose actions during learning an ϵ-greedy algorithm is used. The ϵ-greedy algorithm is renowned across the field. An in depth description of its development can be found in Sutton and Barto (2020). In this algorithm the agent chooses the action that maximises its current value function. However, with probability ϵ the agent will pick an action at random. This gives the agent the opportunity to take actions that have never been tried before and potentially discovering new actions have great benefit. The equation for exploitation of the value function is the same as that referenced in equation 7.

3.3 Energy storage system lifetime

After the operational strategy has been established it is useful to calculate the lifetime of the ESS and ensure that the operational strategy is not degrading it too severely. Battery lifetime is the period during which the battery can provide services before its energy capacity degrades to a predefined level (typically 80% of its full capacity). A widely used term to count the battery life is the battery cycle-life, i.e. the number of cycles at a certain depth of discharge that the battery can provide before it reaches the end of its life.

The relationship between battery cycles and lifetime is non-linear. Many parameters

contribute to battery aging including complex charge/discharge profiles in different applications of the BESS. High energy/power capacity ESSs, such as pumped hydro, CAES and large batteries farms that aim to achieve energy time shift may only undergo one to two cycles in a day. ESSs used in applications like wind power smoothing and frequency regulation may experience tens of irregular half cycles every day. The calendar-life ageing components and cycling related ageing components of the battery jointly contribute to the degradation of battery life. The calendar-life ageing components of BESSs greatly depend on the battery type, the chemistry of cell and the working environmental conditions, such as ambient temperature and humidity. Cycling related ageing is the battery degradation resulting from amp hour (Ah) throughput the battery in charge/discharge cycles. There are a variety of stress factors which have varying degrees of impacts on battery degradation during each cycle. Stress factors include, but are not limited to, Ah throughput, maximum discharge rate, interval between full charges, duration at high DOD and partial cycling of the BESS (see Van den Bergh and Delarue (2015))

A cycle counting scheme named 'rainflow counting' has been used in this thesis to extract the number of irregular, overlapping cycles at different depths of discharge. This technique is particularly vital as it provides a more nuanced understanding of how battery discharge cycles influence the overall degradation and eventual end-of-life of the battery. Rainflow counting is among 3 methods most commonly employed in the literature (Amzallag et al. (1994)); the other two being range-pair counting (which does not consider the complete loading history like rainflow counting) and peak-valley counting (which does not consider the intermediate cycles). In rainflow counting first the depths of discharge are discretised into bins. The smaller cycles are defined by the turning points in the ES charging/discharging time series. When finished the algorithm produces a histogram showing the cycles at each depth of discharge. The analysis was carried out using Matlab's `rainflow` function.

Depending on the battery decided upon, there will be an exponential relationship between 'cycles to death' $c(d)$, where the number of cycles c is a function of the depth of discharge d. The depth of discharge is discretised into N boxes (with each box denoted by n). If NOC is the number of cycles, the rainflow counting algorithm will return a number of cycles at each discretised depth of discharge (NOC_n) then the lifetime consumed (X)

can be given by

$$X = \sum_{n=1}^{n=N} \frac{NOC_n}{c(d)_n} \tag{8}$$

This equation crucially quantifies the fraction of the battery's lifetime that has been utilised, effectively representing the proportion of total available cycles that have been expended prior to the battery reaching a state of significant degradation. X is thus a critical metric, indicating the extent to which a battery has been 'used up' and how close it is to becoming inefficient or failing altogether. The significance of X extends beyond mere numerical value; it provides essential insights into the battery's operational efficacy and health over time. By understanding X, engineers and battery operators can make informed decisions about battery usage patterns, maintenance schedules, and replacement timings, all of which are vital.

In conclusion, presented are 6 steps to applying this methodology to a problem.

1. Identify the action space of the system. In the context of this thesis these will be dispatch volumes and price bids.

2. Identify the state space of the RLA being considered by finding information in the environment that will inform the actions being taken.

3. Conceptualise the problem as a linear programming problem, what needs to maximised and the technical constraints on the system.

4. Perform test runs with various optimisation and training windows. Balance computational work with how quickly the PFA and RLA diverge.

5. Run the training algorithm.

6. Perform test runs operating the PFA and RLA with the same information. Scatter the results from the test runs to show how effectively the RLA is reproducing the PFA's behaviour.

4 Modelling the Balancing Mechanism

This chapter of the PhD is designed to demonstrate the work that has been done on agents that can make control decisions for an ESS operating in a given energy trading market.

The agents designed and presented here are taught after examining the behaviour of the PFA i.e. an agent operating with perfect foresight. Once the agent has adopted the behaviour of the PFA it can then dictate behavior in real time and match the outcomes of the PFA without needing any foresight whatsoever. The agent being able to make decisions without foresight is vital as decisions must be made to really run a ESS.

The example presented shows that the designed agent can reproduce the PFA's behavior when operating a battery ESS on the Balancing Mechanism (BM) for arbitrage. Later in the thesis, the same techniques will be adapted to other ESS setups and markets. For now however, the BM is focused on because the data is publicly available, high-quality data utilising reliable measurement methods. It should be free from errors, biases, or inconsistencies.

The task at hand was to train an agent that could choose price and volume bids for submitting to the BM. The problem has been approached using the methodology of training reinforcement learning agents. Before describing the training and testing of the reinforcement learning (RL) agents, two sections are presented outlining relevant information to setting up the BM and working with a PFA. The first is a description of how the BM works and an analysis of price dynamics on the BM. The second section introduces the PFA and how it behaves when operating an ESS on the BM.

4.1 Historical Balancing Mechanism data

The historical data was taken from the Balancing Mechanism Reporting Service's (BMRS) API, which hosts historical data including the marginal prices on the BM. The agents were trained on a year's worth of data from 01-01-2019 to 01-01-2020. With settlement periods lasting half an hour, this amounts to 17,568 data points. This is the year's data chosen to be presented and analysed here due to the abnormalities created in the system during the Covid-19 pandemic affecting subsequent years' data.

The BM operates under a pay-as-bid payment structure by which it is meant that, after receiving all bids and offers, the system operator (SO) grants bids and offers in order of cheapest to most expensive until it has fulfilled its volume requirements. The maximum the SO will pay for power is the marginal price. Every offer below and bid above the marginal price is paid at face value. Before the SO has settled this price they can forecast the imbalance volume (the difference between the actual demand for electricity and the

amount of electricity generated). Unsurprisingly then, there is a correlation between the imbalance volume on the grid and the marginal price as shown in figure 12. The correlation is high, the gradient on figure 12 being 9.4.

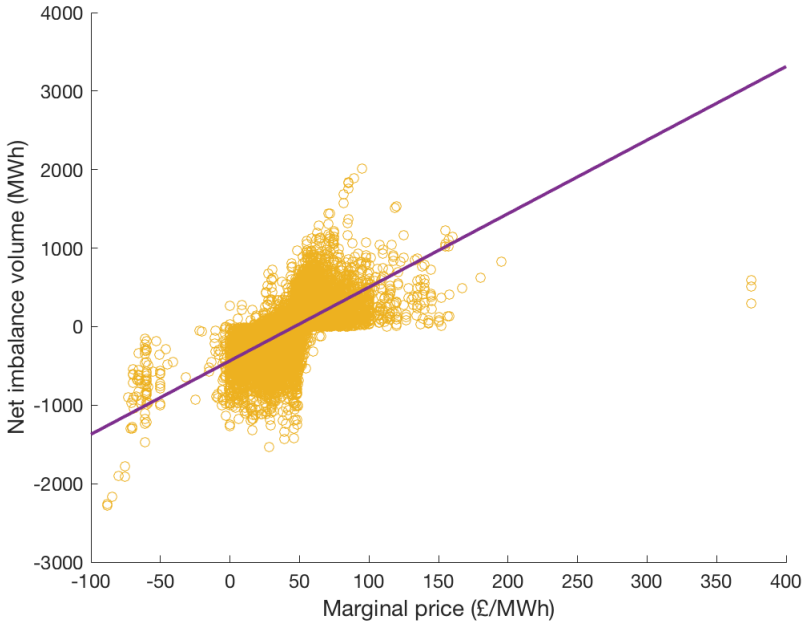

Figure 12: Scatter plot of net imbalance volume against marginal price.

In using the marginal price from the BMRS API, the assumption has to be made that the bids and offers placed by the RL agent do not affect the calculation of the marginal price. The assumption of being a price taker is explored further in section 4.2.

The following is a summary of the statistical elements of the marginal price on the BM for the training data specified. The marginal price has a mean of 42.63 £/MWh and a standard deviation of 20.03 £/MWh, making it a relatively volatile BM (for further discussion see Klæboe et al. (2015)).

Key to characterising marginal price dynamics is identifying price spikes as they add the most opportunity for arbitrage. Price spikes are 'uncharacteristically large prices' according to Sandhu et al. (2016). What constitutes uncharacteristically large varies is rather subjective and across the literature. A simple methodology for labelling a price spike would be with a fixed thresholds i.e. a price spike could be identified as a price that breaches a certain price threshold for a certain period of time. Setting the threshold and time period introduces the subjectivity and a level of arbitrariness. For this report the

largest 1% of reported marginal prices are considered spikes.

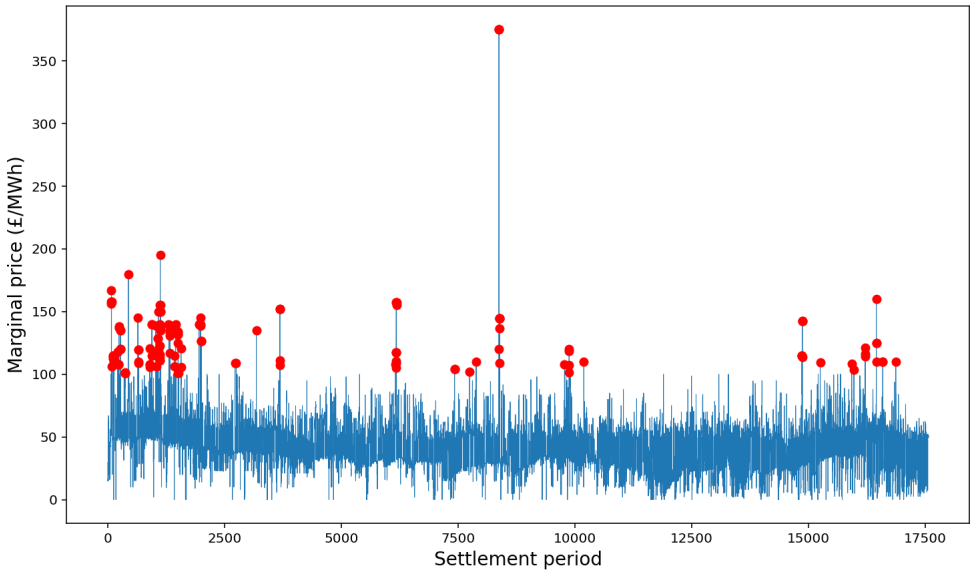

Figure 13: Marginal prices for the specified training period. Price spikes highlighted by a red dot.

With this definition of price spikes, the smallest spike has a marginal price of 100 £/MWh, more than double the average marginal price previously calculated. Figure 13 shows price spikes highlighted by a red dot. To perform arbitrage on the BM it could be helpful to predict price spikes. Although it would appear at first that price spikes are impossible to predict, an indication that a price spike is about to occur can come when the average marginal price over the last five timesteps increases. Figure 14 shows boxplots of the average marginal price over the last five timesteps for the full dataset compared to the average marginal price over the last five timesteps immediately before a price spike.

Furthermore, price spikes are affected by many factors; they are more likely to occur during periods of high demand and a low level of demand. This can be confirmed by looking at the Net Imbalance Volume data quoted above. They are more likely to occur during extreme weather events, when the transmission lines are particularly overloaded and periods of high oil and gas prices. Linear regressions to confirm this have been taken using data from `renewables.ninja` (Staffell and Pfenninger (2016) and Pfenninger and Staffell (2016)), the BMRS API and *Crude oil and natural gas database from Trading Economics* (2022). Specifically, frequently used was the ROLSYSDEM dataset from the BMRS API, which provides rolling system demand forecasts.

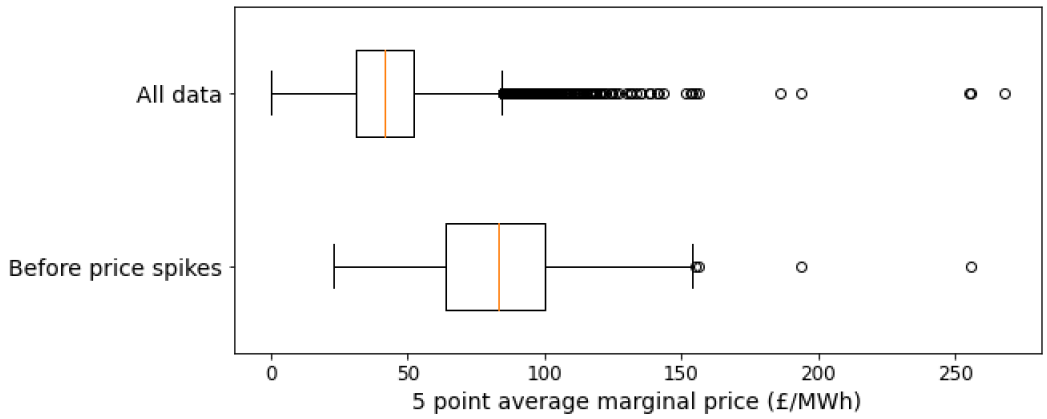

Figure 14: Boxplots of 5 point averages for all data and 5 point averages before price spikes occur.

The correlation between overloaded transmission lines and the standard deviation in spike prices can play a pivotal role in market dynamics and grid management. Overloaded transmission lines are a common issue in power systems, often resulting from peak demand periods, generation unavailability, or unexpected outages. When power lines reach their capacity limit, they cannot transmit additional electricity needed to meet demand in certain areas, leading to a condition known as congestion. This congestion is a primary driver for variations in electricity prices across different regions of the grid, fundamentally influencing the marginal price of electricity. In this case, as marginal prices increase, the range of prices across different areas also widens. This spread is notably reflected in the standard deviation of prices, which measures the amount of variability or dispersion from the average price. When transmission lines are overloaded, the frequency and magnitude of price spikes generally increase. This increase in volatility attracts the attention of both regulators and market participants, as it indicates underlying inefficiencies or imbalances in the system. An increasing standard deviation in the context of rising marginal prices can be indicative of several systemic issues. First, it suggests that the power system is operating under stress, with supply struggling to meet peak demand effectively. Second, it points to potential inadequacies in transmission infrastructure or its management, where the existing capacity is insufficient to handle peak loads without resorting to costly alternatives. Third, it highlights the impact of geographical disparities in electricity generation and consumption, which can exacerbate the effects of congestion.

Going forward, the effect of seasonality on the price spikes is examined.

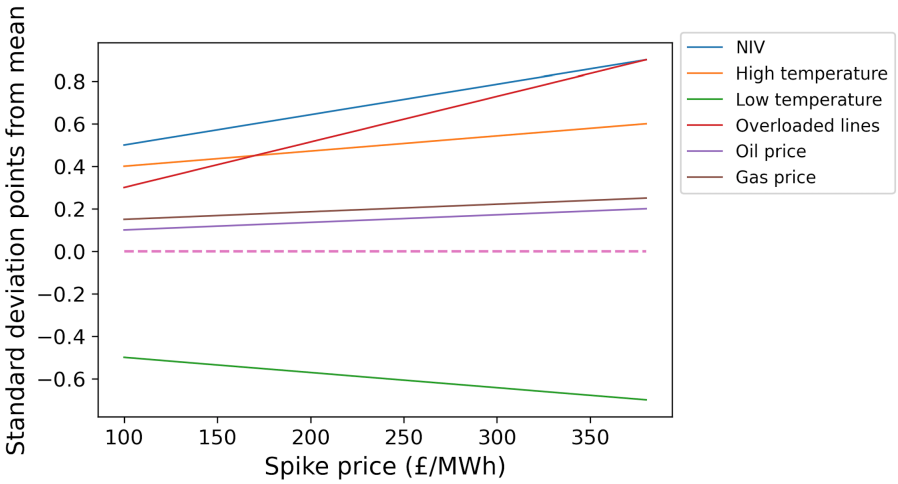

Figure 15: Linear regressions of spike price to Net Imbalance Volume, high temperature, low temperature, overloaded lines, oil price and gas price.

It would seem from the original data that there is a slight negative trend in the data (see figure 13). This may lead to misrepresentation of the number of price spikes at the beginning of the year. However, applying linear least squares regression revealed a gradient of -0.00087 which is small enough to be ignored. The trend means that on average prices at the end of the year will be decreased by 0.00087 £/MWh per settlement period.

Also, the BM is a mechanism for making up errors in forecasting and so it was originally supposed there would be no seasonality that may need to be corrected for.

To attempt to confirm the data is stationary (i.e. not seasonal) an augmented Dickey-Fuller test (Fuller (1976)) was applied. The null hypothesis of the test was that the data is non-stationary and the alternate hypothesis was that it is stationary. The threshold for rejecting the null hypothesis was set at 1%. The test was applied to the raw data, a profile of the marginal price averaged over a day of settlement periods, a profile of the marginal price averaged over a week of settlement periods and a profile of the marginal price averaged over a month of settlement periods. Table 3 shows the resultant test statistics and whether the statistic passes the augmented Dickey-Fuller test. In each case the null hypothesis can be rejected, with the exception of the test on the profile of the marginal price averaged over a week of settlement periods.

Subsequently, the weekly variation was examined in closer detail using moving average-based model. This has been applied across the literature including for the European Energy Exchange electricity spot market (Stefan and Rodney (2007)). The data was

	Marginal price profile	Daily average profile	Weekly average profile	Monthly average profile
Test statistic	-12.06	-8.96	-3.21	-5.67
Null hypothesises rejected?	Y	Y	N	Y

Table 3: Table of augmented Dickey-Fuller tests results.

deconstructed into three separate components:

$$p_t = T_t + S_t + e_t \tag{9}$$

where p_t is the marginal price profile, T_t is the trend over the full year, S_t is the weekly component and e_t is the residuals of randomly fluctuating elements of the profile.

First the data was passed through a linear convolution filter to extract the trend. The trend being the aspect of the data that changed gradually through the year. The trend was removed from the price profile and the average for each weekly period was assumed to be the oscillatory trend. From the oscillatory trend it is apparent why the augmented Dickey-Fuller test could not reject the null hypothesis for the profile of the marginal price averaged over a week of settlement periods. The residual results show the data after both the trend and weekly oscillations were removed and reflected the price spikes.

In conclusion, the findings from the augmented Dickey-Fuller test and the subsequent detailed analysis of weekly price variations provide a robust foundation for future explorations, particularly in designing and refining observation spaces for reinforcement learning agents. By establishing that the marginal price data, with the exception of the weekly profile, is predominantly stationary, a more predictable baseline from which machine learning models can learn and make predictions can be built. The detailed decomposition of the price data is all information that a reinforcement learning agent would need to handle in subsequent sections.

4.2 Price taker status

To use the assumption that our agent can act as a price taker it is necessary to determine if our generated bids and offers would significantly alter the marginal. The following methodology has been designed to simulate the decision process of the National Grid when setting the marginal cost of an increase/decrease in power on the grid. Depending

on the needs of the system the marginal price is referred to as the system buy price (SBP) (in the event of under-generation) or the system sell price (SSP) (in the event of over-generation).

The majority of trading arrangements are agreed between independent parties (producers and consumers of electricity) Facchini et al. (2019). However, when there is an imbalance between generation and demand on the grid, the National Grid will take 'balancing actions' which involve instructing parties to increase or decrease generation or demand.

When settling on an SSP/SBP, the National Grid considers which balancing actions are taken purely for balancing generation and demand and which actions are taken to manage 'system needs' such as voltage levels *BMRS API and Data Push User Guide* (2019). Actions taken for system balancing are 'flagged' and not included in generating the marginal price. There are two stages of flagging; first stage and second stage flagging. Flagged actions are potentially system balancing and may be removed from the marginal price calculation (*The Electricity Trading Arrangements* (2019)).

The processes involved in determining the SSP/SBP are outlined in the following sections.

4.2.1 Continuous acceptance duration limit

Actions that last for less than 10 minutes are first stage flagged (*BMRS API and Data Push User Guide* (2019)). They are considered to be system balancing and are not included in calculation of SSP/SBP.

4.2.2 De Minimis tagging

Actions offering volumes under the De Minimis acceptance threshold first stage flagged and are considered to be system balancing. As of 1 April 2019 the De Minimis acceptance threshold was set at 0.1 MWh (*BMRS API and Data Push User Guide* (2019), *Imbalance Pricing Guidance* (2019)).

4.2.3 Arbitrage tagging

When a bid is submitted to the National Grid that is simultaneously more than an offer for the same volume the National Grid will buy the volume and sell it straight on at a

profit. This mechanism is considered to be system balancing (*BMRS API and Data Push User Guide* (2019)).

4.2.4 Classification

If a first stage flagged action is more expensive than the most expensive first stage unflagged action it is deemed to be system balancing. If it is less expensive it is considered energy balancing and stays part of setting the marginal price.

At the end of the classification process all of the first stage flagged actions become second stage flagged.

4.2.5 Net imbalance volume tagging

In this process the offers and bids cancel each other out to leave a surplus of either bids or offers. In cases where there is less generation than demand there will be a net volume of offers. In cases where there is more generation than demand there will be a net volume of bids.

The SSP/SBP is then set at the most expensive offer/bid that remains after carrying out the preceding processes.

The BMRS API contains records for every bid and offer submitted to the BM at each settlement period (see section 5.2.52 of *BMRS API and Data Push User Guide* (2019)). The code written to simulate the BM takes this data and replicates the decision making process outlined in sections 4.2.1 to 4.2.5. The code has been validated against historical data from the BMRS API. The simulation was run over the entire year of 2018. The majority of errors in simulating the BM fall close to 0. In fact the median error is our simulation is just 0.1786 £/MWh. Once the reinforcement learning agent has been developed its bids will be added to the simulation of finding the marginal price to see if its price taker status can be confirmed.

With an understanding of the BM, there is now presented a case study that will show the process of designing the RLA.

4.3 Case study

In this case study a 4 MWh battery was used with a maximum power flow of ± 2 MW. The specific parameters used here are irrelevant to the context, as the case study is only presented to demonstrate that the RLA can be successful in mimicking the PFA and the parameters do not capture the essential characteristics of the system being portrayed. With these parameters the PFA's actions were generated using open-source software which utilises mixed interger linear programming and can be accessed on github (see Green (2024)). The software assumes perfect round-trip efficiency of the battery. The software allows for setting the initial SoC of the battery and aims to completely empty the battery at the final timestep. In this study the initial SoC was set at 1.

Analysis of the perfect foresight agent's behaviour shows no correlation with time of day. Also, the mean and standard deviation of the perfect foresight agent's actions are broken down by month in table 4. Both statistics are stationary across the year, a reflection of the original price profile showing no seasonality.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	-0.04	-0.05	-0.05	-0.05	-0.04	-0.05	-0.05	-0.06	-0.05	-0.05	-0.05	-0.05
Standard Deviation	1.39	1.39	1.39	1.37	1.36	1.35	1.41	1.43	1.51	1.44	1.45	1.59

Table 4: Mean and standard deviation of the perfect foresight agent's power output for 2019.

4.4 Choosing an observation space

To generate the optimal price and volume bids two separate agents were trained with their own environments. Arguments for why this setup was chosen are now presented.

4.4.1 Bayesian lasso regression

Bayesian lasso regression, in this context, is used to aid in linear regression. The marginal price timeseries can be regressed with respect to several different variables and bayesian lasso regression should help to decide which variables to choose in the final model.

Bayesian lasso regression is a regression analysis technique that can help to identify the variables that are most useful in characterising a time series. It can be helpful in choosing the observation dimensions as it is a methodology for suggesting which variables are best

correlated. In this scenario, if a variable is correlated with the marginal price timeseries it may suggest that it will make a useful observation for the RLA.

An advantage of this technique is that many variables can be suggested and with the Matlab econometrics toolbox, the lasso function will suggest which variables are most useful. This is important because, with our discretised value function, it is important to keep the dimension size of the value function relatively low to reduce computing time and avoid over-fitting.

The following variables were used as indicators for predicting how the marginal price will evolve:

- The marginal price in the previous timestep
- The marginal price in the previous but one timestep
- The marginal price in the same settlement period of the previous day
- The average marginal price over the last 5 timesteps
- The maximum marginal price over the last 5 timesteps
- The minimum marginal price over the last 5 timesteps
- The average marginal price over the last 10 timesteps
- The maximum marginal price over the last 10 timesteps
- The minimum marginal price over the last 10 timesteps
- The average marginal price over the last day
- The maximum marginal price over the last day
- The minimum marginal price over the last day

With the maximum number of indicators set to 4, the results showed the best indicators were the average marginal price over the last 5 timesteps, the maximum marginal price over the last 5 timesteps, the minimum marginal price over the last 5 timesteps and the marginal price in the previous timestep. Restricted to 3 indicators, the minimum marginal price over the last 5 timestep was no longer included and with 5 indicators the marginal price the same settlement period of the previous day was to be included.

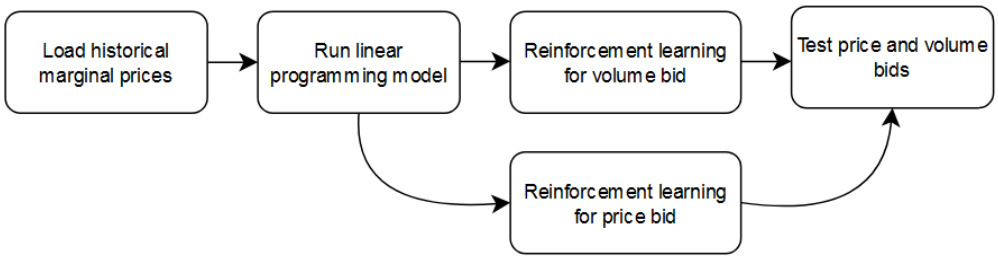

Figure 16: Flowchart of training and testing whole system.

The reward function is used to evaluate the value of the action taken.

To illustrate the process of constructing a reward function an RLA was developed that could generate price and volume bids for a generic ESS. If the RLA agent is successful it should learn to bid volumes with a bid price below the market price when discharging and above when charging. In $Q(S, A)$ a 2-D action space (A) made up of

1. a_1 - Price bid (£/MWh)
2. a_2 - Volume bid (MW)

The reward function was designed to maximise profit as it would be if the bids were really placed. When discharging the reward was given by:

$$R = \begin{cases} -a_1 \times a_2 & \text{if } \ 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } a_1 < p \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where Cap is the capacity of the ESS and p is the marginal price on the BM. While charging the reward was given by:

$$R = \begin{cases} -a_1 \times a_2 & \text{if } \ 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } p < a_1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

While discharging, a positive reward is generated by multiplying the volume and price bids (a discharge is defined as a negative number). The first constraint on that reward ensures the SoC stays between 0 and 1. The second ensures that the system only receives a reward if the price bid is below the marginal price. Breaking either constraint results in

a reward of 0. Likewise, the system will receive a negative reward while paying to charge, subject to the physical limitations of the ESS and that the price bid exceeds the marginal price.

Before continuing, it is helpful to define a metric which will be refereed to as average success rate. The success rate will be defined as the fraction of PFA's cumulative profit obtained by the RLA during a 200 settlement period test.

$$\text{Success rate} = \frac{\text{RLA's cumulative profit}}{\text{PFA's cumulative profit}} \tag{12}$$

In the scenario that the RLA generates a loss while operating, the success rate will be negative. The quoted success rates will be averaged over 1000 tests to capture a full range of behaviour.

This section presents the cumulative profits for a combination of observation dimensions shown in table 5. The numbers in the first column refer to:

1. The marginal price in the previous timestep
2. The average marginal price over the last 5 timesteps
3. The maximum marginal price over the last 5 timesteps
4. The minimum marginal price over the last 5 timesteps
5. PFA's state of charge

In choosing an observation space, the dimensions need be designed to deliver information that will help the reinforcement learning agent about PFA's behaviour. Each point in the observation space will be visited a handful of times during learning and mapped to

Observation dimensions	Average range of the PFA's actions	Average success rate for 1000 trials
1, 2	0.37	0.90
1, 3	1.02	0.56
1, 5	0.32	0.91
2, 5	0.30	0.83
3, 5	0.59	0.78

Table 5: Average range and average success rate for training the reinforcement learning agent with a range of observation spaces. (Only the 5 best of the combinations are displayed).

a recommendation from the PFA about how to behave. The range of behaviour from the PFA at a given point in the observation space indicates how well the chosen observation dimensions can be used to predict PFA's behaviour. A high range may suggest that the observation space is not useful for learning this particular behaviour. A low range may suggest consistency in the PFA's behavior whenever this instance of observations occurs or may suggest that the agent is over-fitting. The 'average range of the PFA's actions' in table 5 quotes the average range for every point in the specified observation space.

The way this is calculated can be explained thusly; take an imaginary but illustrative point in the state action value function ($Q(s_{example-1}, a_{example-1})$) which in this instance is visited twice during the learning process. It should be noted that this is an unrealistically low number of times to visit but is written here for comprehension. In the first visit to $Q(s_{example-1}, a_{example-1})$ say the reward is 0. In the second visit to $Q(s_{example-1}, a_{example-1})$ say the reward is 2. Then the range of rewards given by this point in the state action value function is 2, In a contrasting example: in the first visit to $Q(s_{example-2}, a_{example-2})$ say the reward is 0. In the second visit to $Q(s_{example-2}, a_{example-2})$ say the reward is 0. Then the range is 0. There may be many reasons for this to be but, importantly it can give a sense of how that point in the state action value function is changing as learning progresses.

Table 5 shows that the average range of PFA's actions are generally low, with the exception of observation space 1,3 which does not contain a dimension for the average price over the last 5 timesteps. This indicates that the average price is important in pinning particular instances of the PFA's behaviour to a single point in the observation space. It is also interesting to note that including observation spaces 1 and 5 appear most commonly in the table of top 5 combinations. The analysis reveals a subtle negative correlation between the average success rate and average rage. Specifically, as average success rate experiences an increase, there is a corresponding slight decrease observed average range. The strength of this association is relatively weak, with a correlation coefficient of -0.19.

Combination 1,5 has the highest average success rate and it is interesting that it is combined with a relatively low average range suggesting that it may be the best choice for observation space in this instance.

Not listed in the table is the combination 3,4 which has the lowest average success rate, recouping a loss on average. On examination of the results, this loss occurs because

in this dataset there was a higher level of bids that needed to be truncated to maintain SoC, leading to a loss in profits.

The observation space (S) was 2-D made up of

1. The marginal price in the previous timestep
2. ESS state of charge (SoC)

Dimensions of the observation space were discretised into bins of width 5 £/MWh except for the final bin which acounted for all prices above 150 £/MWh. The SoC discretisation had a width of 0.1 a lower limit of 0 and an upper limit of 1.

The best price bids from this learning agent are presented in figure 17 for an SoC of 0.5 and proposing a power output of -1MW. This reward function is referred to as square drop-off because the reward is either given in full or the system gets a reward of 0. The results show that below p_{t-1} = 20 £/MWh the agent learns very little about the right behaviour and the resultant policy is nonsensical. Above p_{t-1} = 20 £/MWh this approach generates a crude threshold rule, that for p_{t-1} = 20-60 £/MWh bid 20 £/MWh. For p_{t-1} = 60-100 £/MWh bid 40 £/MWh, for p_{t-1} = 100-150 £/MWh bid 60 £/MWh and bid 100 £/MWh for p_{t-1} above 150 £/MWh. Referring to figure 18, the scatter plot has been limited to showcase only data points where the marginal prices are above 0 £/MWh. This decision was taken to eliminate any potential noise caused by zero or negative pricing, which can occur under specific market conditions but are not the focus of this analysis. The primary objective of this visual representation is to highlight the distinct

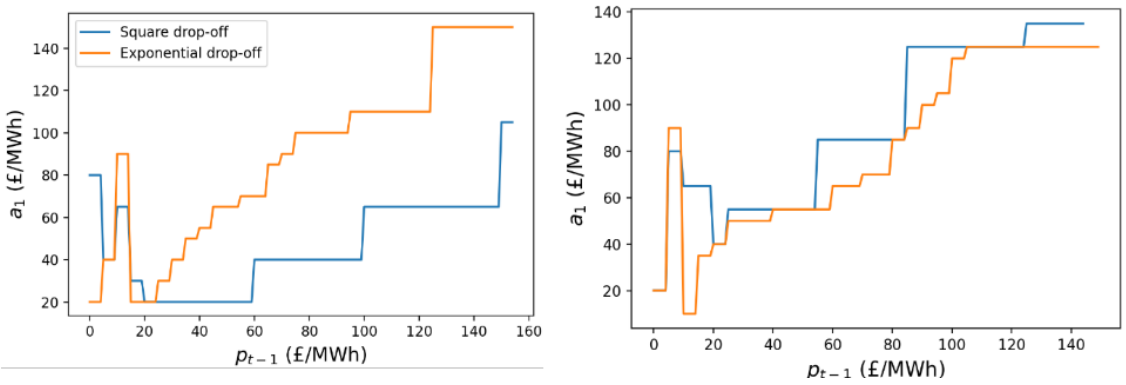

Figure 17: Best price bid during discharging (left) and charging (right) according to a single reinforcement learning agent that generates both price and volume bids with reward shaping.

correlation that emerges when marginal prices exceed £75/MWh. This methodological choice is substantiated by the hypothesis that significant operational insights and economic implications are predominantly observable within this higher price bracket. The exclusion of lower values, therefore, enhances the clarity of the data presentation, ensuring that the substantive aspects of the correlation are neither obscured nor diluted by less relevant data points. Subsequently, the produced policy may be a consequence of the correlation between p_t and p_{t-1}; that the agent learns that the marginal price in the current timestep will most likely be the same as the last. However because of the sharp drop-off in reward (as the reward function returns 0 above the marginal price while discharging), the agent is cautious and recommends a bid firmly below the most likely marginal price.

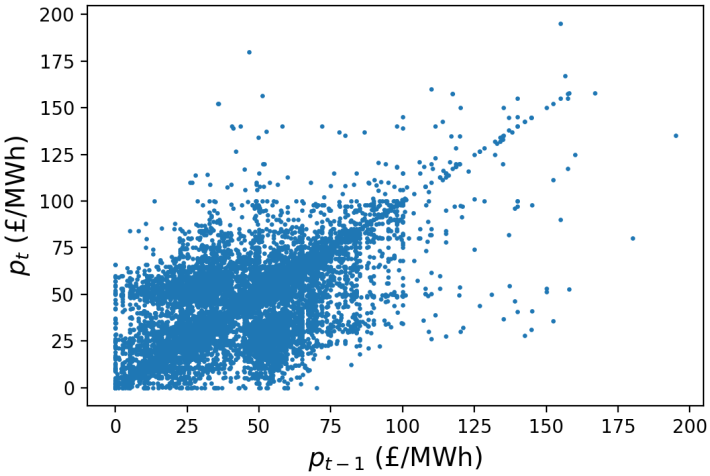

Figure 18: Scatter plot of p_t against p_{t-1}. Importantly this small section of the data is displayed for clarity.

After that attempt a new reward function was tested that gave a reward even when the agent exceeded the marginal price during discharging or offered under the marginal price during charging. The function was thus during discharging:

$$R = \begin{cases} -a_1 \times a_2 & \text{if } \ 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } a_1 < p \\ -(a_1 \times a_2) exp(-0.1(a_1 - p)) & \text{if } \ 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } p < a_1 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

and thus during charging:

$$R = \begin{cases} -a_1 \times a_2 & \text{if } 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } p < a_1 \\ -(a_1 \times a_2)exp(-0.1(a_1 - p)) & \text{if } 0 < SoC + \frac{a_2}{Cap} < 1 \text{ and } a_1 < p \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The extra part of the reward function means that the system can still gain a reward getting the price wrong though there is an exponential drop off as the bid diverges from the marginal. This allows the agent to take more risks in bidding.

4.4.2 Scaling the exponential reward function

Using this reward function the resultant policy is shown in figure 17 called exponential drop-off. The policy shows more variation than the policy developed by the agent with a square drop-off and more closely follows the linear regression between p_t and p_{t-1}.

Recreating the linear regression may arise from the exponential part of the reward function that encourages the RLA to match the marginal price as closely as possible. Instead of matching the marginal price it may be better to match some optimal behavior. This is where the PFA could come in useful, using the PFA's action in the exponent of the reward function.

Futhermore, the results from this study suggested it would be best to decouple the generation of price of volume bids into two separate agents.

In this conceptualisation the reward for the volume bid agent was given by:

$$R = exp(-|\sigma_{volume} - a_2|) \tag{15}$$

where σ_{volume} is the volume bid of the PFA.

The reward for the price bid agent was given by:

$$R = exp(-|\sigma_{price} - a_1|) \tag{16}$$

where σ_{price} (like σ_{volume}) is the price bid of the PFA.

This is a very specific example; the important point being that the reward function

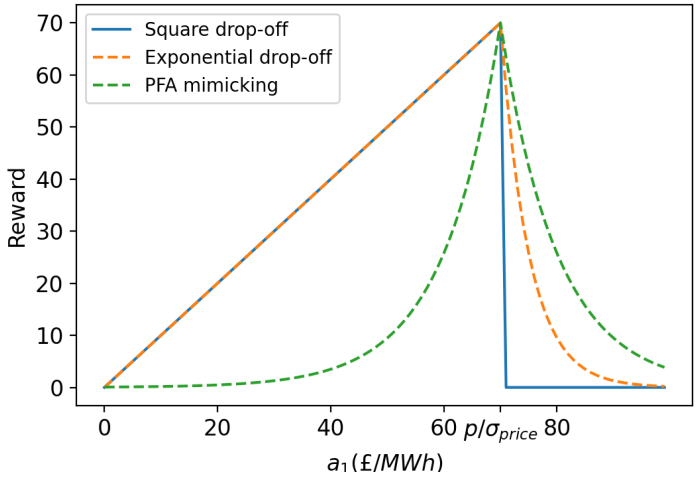

Figure 19: Reward function shapes for discharging.

uses an exponential and (using this exponential) compares the actions of the PFA to the subsequent actions taken by the RLA. in the scenario of equation 16 the price bid of the PFA is being directly compared to the price bid of the RLA while being trained. In subsequent cases it may not be the price bid specifically being examined and such, the reward function can be generalised to cover more examples that will be presented later in the thesis;

$$R = exp(-|\sigma - a|) \tag{17}$$

i.e. the reward is also generated from the difference between the PFA's actions and RLA's. This recreates the form of the case study that initiated this practise but allows for enough generalisation such that it can be and will be applied to the action and observation spaces of other scenarios.

4.5 Maintaining state of charge

Once both agents have been trained, they can generate price and volume bids imitating PFA's behaviour based entirely on historical data. However, the agent has no awareness of the battery ESS current SoC.

In order to maintain a reasonable SoC during testing volume bids that would exceed

SoC are truncated:

$$b'_v = \begin{cases} 1 - (1/2) \times SoC \times C & \text{if } \; SoC + b_v > 1 \\ -(1/2) \times SoC \times C & \text{if } \; SoC + b_v < 0 \end{cases} \tag{18}$$

4.6 Parameter tuning

In this section the parameters in the reward functions are examined, to find the values that produce the best trained agents. The first parameter explored is the 'steepness' (s) of the reward function of the price bidding agent:

$$R = \begin{cases} 2exp(-s(p - b_p)) & \text{if } \; \sigma_2^{price} = 0 \text{ and } b_p < p \\ 2exp(-s(b_p - p)) & \text{if } \; \sigma_2^{price} = 1 \text{ and } p < b_p \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

This reward function generates a reward of 2 if the reinforcement learning agent correctly mimics the PFA with an exponential drop-off as the RLA's behaviour diverges from the PFA's. The parameter s determines how forgiving the reward function is of this divergence. A high s indicates that differences between the RLA's behaviour and the PFA's will be penalised harshly. The results of adjusting this parameter are displayed in figure 20 and shows a peak of average success rate at $s = 0.5$. In figure 20, an additional set of data points was meticulously gathered and integrated to refine the understanding and identification of the local maxima. Shown in red, the targeted area was surrounding the suspected peak values.

Likewise the steepness of the exponential drop-off has been varied for the volume bidding agent, with results displayed on the right in figure 20 for the equation.

$$R = exp(-s|b_v^{PFA} - b_v|) \tag{20}$$

The figure suggests a value of s being 1 was most successful.

At the beginning of learning, the agent is quite ignorant, being initialised to value every state and action equally. Therefore the need for exploration (discovering new strategies) is high. Towards the end of learning exploitation is more useful. The ad hoc rule was

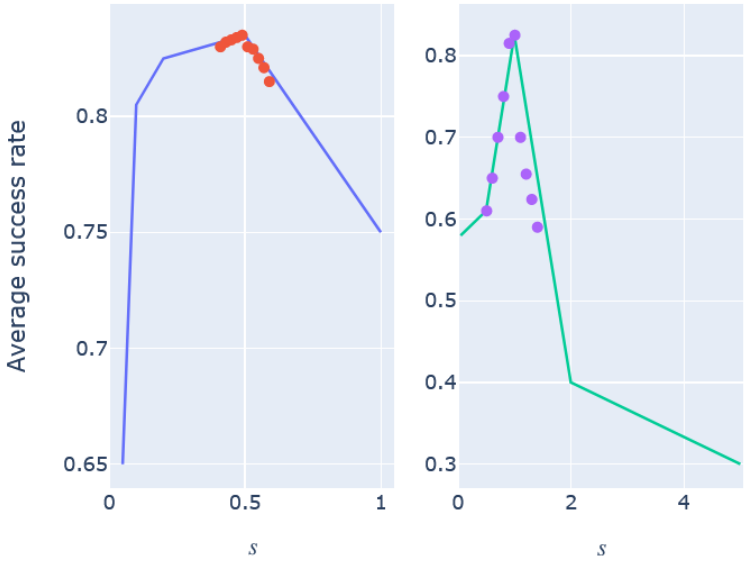

Figure 20: Average success rate while varying parameter s in the reward function of the price bid agent (left) and volume bid agent (right).

applied that after each learning step ϵ was decreased by 1/(Maximum number of learning steps), ensuring that ϵ decreased uniformly, reaching 0 by the end of the learning process.

To demonstrate how the policy improves over time figure 21 is presented. To create this figure, each agent was trained for 1000 episodes, after which the extracted polices were compared to PFA for a set of test data. The success rate was calculated and plotted and then learning continued for another 1000 episodes. This process was repeated for a total of 500,000 episodes and the results for 'descending ϵ' (as the method of slowly decreasing ϵ has been called) is shown in figure 21 in green.

Also presented is the same process for fixed values of ϵ. With a high ϵ, the average success rate gradually increases but takes a long time to do so as time is used inefficiently exploring parts of the action space that are not useful. It can be speculated that if the agent was left to train for a much longer period, its average success rate would continue to rise. By contrast, with a low ϵ the agent finds rewarding actions quickly continues to repeat them, not discovering any new information or significantly improving its success rate during learning.

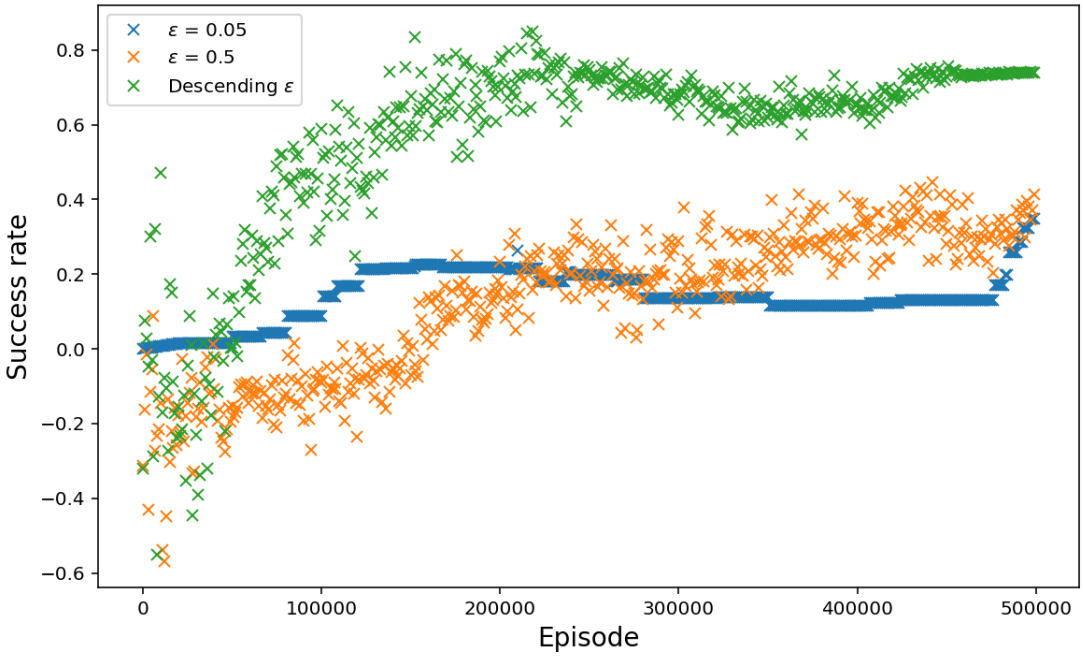

Figure 21: Success rate as learning develops for various values of ϵ.

4.7 Results

A scatter plot of RLA's vs PFA's profits is presented in figure 22, demonstrating the range in profits achieved in imitating PFA's behaviour. Also shown is the line y=x i.e. the hypothetical maximum profit achievable if the RLA were to perfectly replicate the PFA. At the upper end of the profit scale, the RLA is less successful at imitating the PFA, achieving a lower percentage of PFA's profit. This can be explained by the presence, during these testing periods, of price spikes i.e. sudden jumps in price that are uncommon. The PFA could see these and discharge during those settlement periods. However, these price spikes were not well learnt by the reinforcement learning agent due to the small reward generated when the RLA attempts to discharge SoC it does not possess. All this is clearly visible in the divergence of the results from y=x in figure 22. It points to a need to further study these edge cases in the state action value function.

The results can be put in context by comparing them to the work of Cho and Kleit (2015) who compared a backcasting technique to maximum possible revenues as has been done here. Backcasting is how they refer to a technique of optimising over the last 2 weeks and then applying that recommendation to the upcoming 2 week period, under the assumption the next 2 weeks will be more or less identical to the last 2. They claimed to

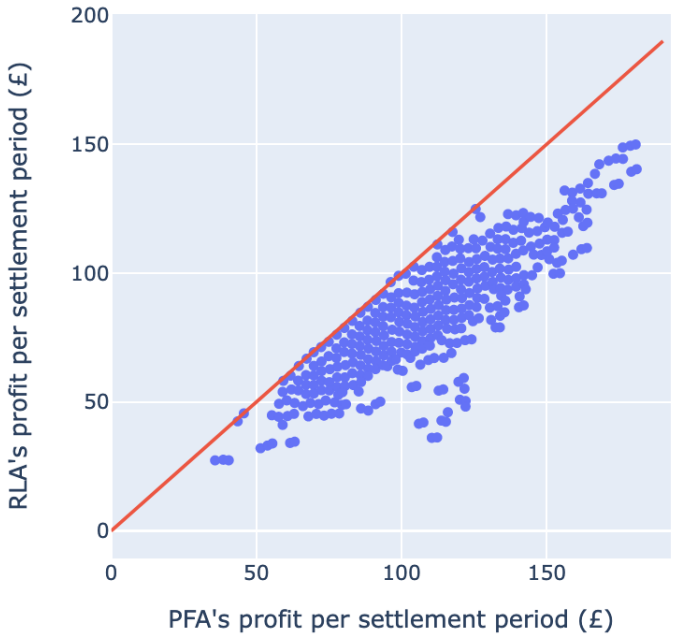

Figure 22: Cumulative profit for the perfect foresight agent and the reinforcement learning agent during 200 settlement periods run for 1000 tests. Also shown is the line y=x.

reach 45-85% of maximum possible revenues with a slight skew towards underachieving, suggesting a greater level of robustness in the methodology presented here. Furthermore, in Madahi et al. (2024) (which also used reinforcement learning to try and develop an agent without foresight) a method was used that did not use a PFA but a system based directly on the statistics. They found a 12.6% increase from using a method of rule-based control strategies (discussed in section 2.8) which does not compare to the increase in profits shown as the reward function based on the PFA was developed here and so suggests the use of the PFA can add to profits generated.

Finally as described in section 4.2, once the code was developed, the price and volume bids generated by the RLA were added to the list of bids and offers recorded by BMRS for each settlement period. Then the same decision making process for setting the marginal price was applied. It was observed that the influence of the RLA's bids on the marginal price was notably minimal, impacting it in only 7% of instances. Furthermore, within this small fraction, the average modification to the marginal price was a modest 2.05 £/MWh, underscoring how slight the changes typically were. Even though there was an instance where the adjustment reached as high as 23 £/MWh, such occurrences are exceptionally rare. This data strongly supports the rationale for considering the RLA as predominantly

a price taker, given the very limited scope and scale of its impact on market prices.

4.8 Conclusion

In conclusion, this chapter has made significant strides in understanding the dynamics of marginal pricing in 2019, highlighting not only the variability and trends within the market but also the potential for strategic engagement through energy storage systems. By meticulously analysing and interacting with the provided data, there was a successful demonstration that the capacity of such systems to can operate effectively. Crucially, the exploration of such data have laid a solid foundation for the development of a reinforcement learning agent. This agent, with its ability to navigate and learn from a rich observation space, represents a significant step forward in predictive accuracy and operational efficiency. Furthermore, the implementation of a series of flagging and tagging processes has underscored the feasibility of modelling the Balancing Mechanism.

Future work will aim to expand the observation spaces to encompass additional variables such as frequency and others arbitrage opportunities within the energy market. This evolution of the observation space is anticipated to deepen understanding of the multifaceted nature of energy systems. The exploration of frequency data, in particular.

5 Combined biomass and energy storage system providing Firm Frequency Response

The methodology of mimicking a PFA is extended for the combined operation of an energy storage (ES) system and biomass generator. The combined system was tested against the National Grid's 'Firm Frequency Response' (FFR) market and in particular there is an examination of providing non-dynamic frequency response. The methodology was tested for robustness, against changes in market contracts including increased sensitivity to frequency fluctuations, biomass Opex and increasing tendered volume. It was found that the reinforcement learning (RL) agent can recreate the behaviour of PFA fairly well. It was also found that biomass Opex has the biggest impact on profits.

In this chapter a study is presented of operating a combined biomass generator and energy storage system for contracting to the National Grid's FFR market. It is worth starting this report with a description of FFR. That is followed a description of historical

frequency data in the UK (section 5.1) and a view of the costs associated with biomass operation (section 5.2). This is followed by the presentation of a case study used throughout the report (section 5.3). The perfect foresight model for operating the system or 'PFA' can be found in section 5.4 and the RL technique for mimicking PFA in section 5.5. Finally the results are presented section 5.7.

FFR works by contracting with providers to respond to deviations in system frequency with (in the case of the non-dynamic response) a fixed volume of power; a reduction in power for a high system frequency and an increase in power for a low system frequency. The provider must monitor frequency deviations, be able to respond in 30 seconds and maintain their response for 30 minutes and further rules can be found in National Grid ESO's *Firm Frequency Response Interactive Guidance* (2017). At the time of writing (February 2022) 250 MW of Non-dynamic FFR was projected to be purchased for March 2022 as published in the *Frequency Response Market Information Report: Monthly Report January 2022* (2022).

The fixed volume required is negotiated beforehand with National Grid, as is the sensitivity to frequency deviation required by the provider i.e. the provider needs to fix the frequency deviation ±Hz at which they will respond. It is the responsibility of the provider to monitor and respond to frequency deviations and FFR regulation states that 'The standard deviation of response error over a 30 minute period must not exceed 2.5% of the contracted response'.

The key differences found in operating on the FFR market as opposed to the Balancing Mechanism presented in the previous chapter are as follows. Instead of the 30 minute settlement periods tendered on the BM, the FFR splits each day into 6 4-hour 'EFA blocks' (EFA block boundaries are presented in table 6). Procurement of tendered volume in any given EFA block can be attained well in advance (up to a month beforehand). The prices for tendered volume are fairly steady and, therefore, it is assumed that the tendered volume has been awarded in advance and the price setting element of the previous chapter has been removed. Instead focus is given entirely to the biomass and ES volume dispatch.

EFA block	1	2	3	4	5	6
Time boundaries	23:00-03:00	03:00-07:00	07:00-11:00	11:00-15:00	15:00-19:00	19:00-23:00

Table 6: EFA block boundaries.

The service generates two revenue streams; an availability price (£/hr) a response energy payment (£/MWh). The availability price is paid for every hour the provider commits to monitoring frequency and delivering response. The response energy payment is given depending in the volume the provider is called upon to deliver.

FFR pricing for 2019 is available for public access on the National Grid website. Analysis of this data shows that the lowest settling price for tendering of power is 3.31 £/MWh. The lowest price is taken in this study in order to ensure our system is awarded the contract and it can be assumed the operator is called upon to respond to frequency deviations in every EFA block. Likewise the availability price was set at 126.94 £/hr.

Next is presented an analysis of the the National Grid's system frequency according to data obtained from the BMRS API for the year 2019. 2019 was chosen for being the most recent year of data available unaffected by the SARS-CoV-2 pandemic.

5.1 Historical frequency data

First of all, it should be addressed that the BMRS API frequency data contains some gaps for the year 2019 as can be seen in table 7.

The frequency data fluctuates around 50 Hz as would be expected from national statutory requirements which are that the National Grid must maintain a system frequency of 50 ± 0.5 Hz. Figure 23 shows the frequency of various frequency deviations for 2017-2020. Although the shape of the curves look consistent, the times with deviations away from 50 Hz by some small amount (up to $\pm$ 0.02) are becoming more common as inertia goes down it becomes harder to maintain a steady 50 Hz. Additionally the tendency for an excess of

Date	Time
02/09/2019	00:53:45-07:54:00
25/04/2019	09:25:45-14:46:00
22/05/2019	09:25:45-18:14:00
31/05/2019	19:57:45-21:06:00
06/06/2019	09:25:45-12:46:00
09/07/2019	10:47:45-11:04:00
02/09/2019	21:45:45-21:56:00
26/09/2019	09:31:45-13:48:00
24/10/2019	09:31:45-15:08:00

Table 7: Dates and times where data was unavailable from the Balancing Mechanism Reporting Service's API and hence left from the data set.

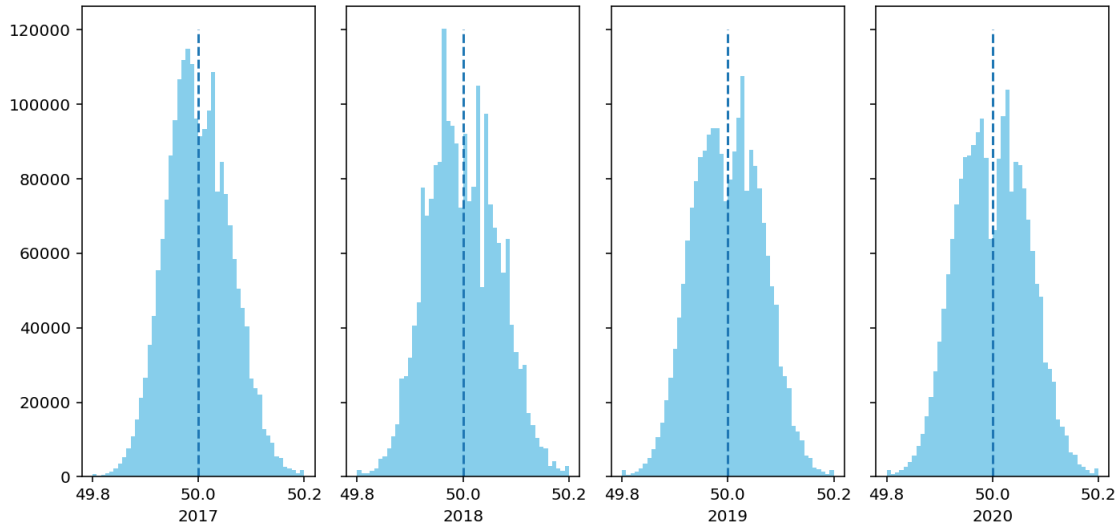

Figure 23: Histogram of system frequency from 2017-2020.

under frequency events is changing, with a more symmetrical curve in later years. Both effects may indicate less predictability in the future as there is a wider range of system frequencies that regularly occur.

It also appears that there is a lower number of reports of the system frequency being exactly 50 Hz. It may be a consequence of the introduction of the highly sensitive frequency regulation described leading to regular over-corrections.

Frequency deviations are characteristically different across EFA blocks. As shown in figure 24, there is generally less frequency deviation in EFA block 1. Furthermore, figure 25 shows the mean frequency per settlement period for the summer months (March-August). A trend towards large frequency drops in the morning is apparent. Also plotted is the system load per settlement period. The periods of low load appear correlated with these larger frequency deviations.

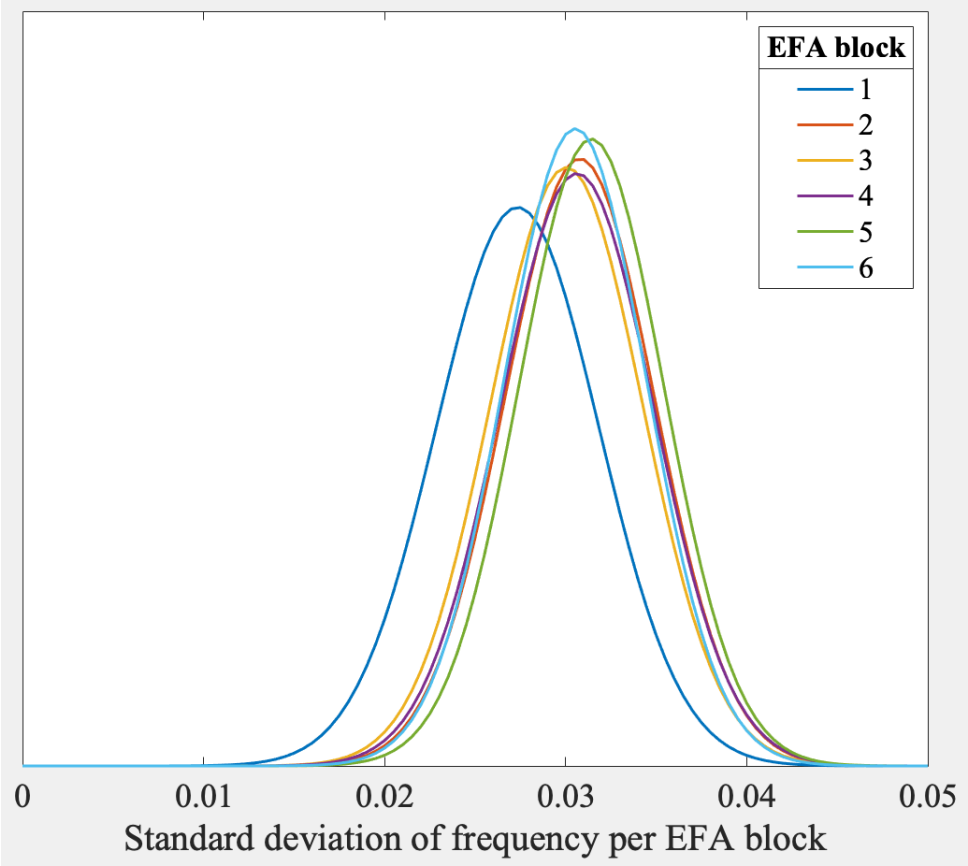

Figure 24: Standard deviation of frequency for EFA blocks 1-6.

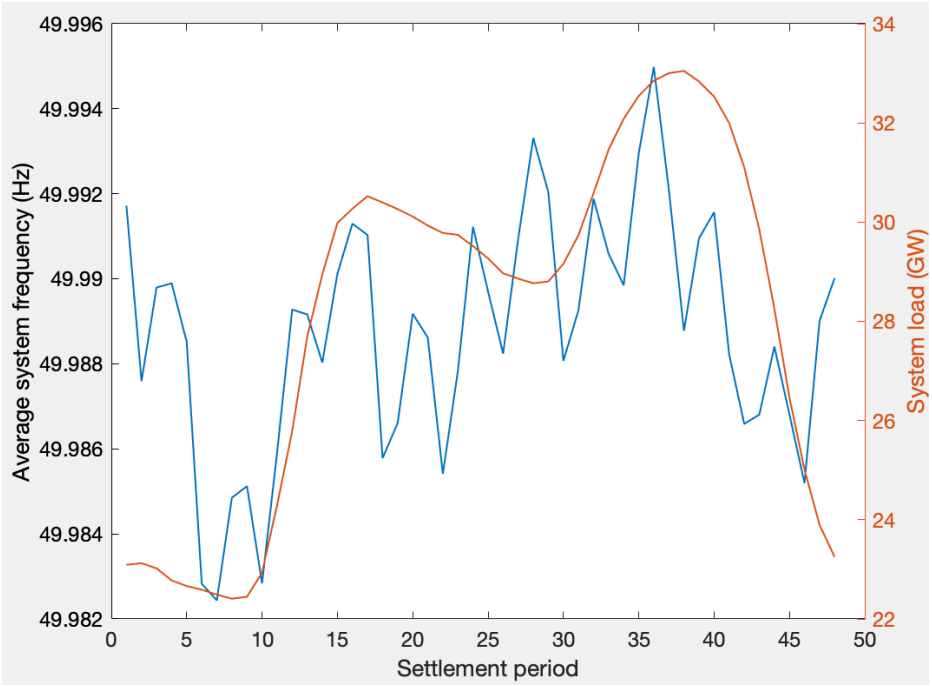

Figure 25: System frequency and system load per settlement period averaged over the summer months (March-August) 2019.

5.2 Biomass cost of operation

Including a biomass generator in the system requires an analysis of the costs associated with running, ramping, starting and shutting down a biomass generator. These operations are collectively known as cycling and introduce wear and tear to the biomass generator caused by thermal gradients in the machinery. Costs also include any loss in efficiency caused by cycling. Some papers (Eser et al. (2016)) have also concluded that, compounding financial costs, increased cycling can cost more in CO_2/MWh.

Thermal generators more generally have been studied in future scenarios with increased renewables in the energy mix. It has been suggested by Pereira et al. (2014) that there will be an increase in frequency of cycling in these scenarios and so it becomes increasingly important to include cycling costs in modelling biomass generators.

Duggal and Venkatesh (2015) presented a study of thermal generators operated in tandem with a battery ESS. They used a linear programming approach to solve a 24-hr scheduling problem for several combinations of generator and battery ESS capacity.

Assumptions for the operational costs (£/MWh) of a biomass generator are taken from Duggal and Venkatesh (2015) and presented in table 8 due to the similarity in setup between this study and that one. Likewise the ramping costs (£/δMW) are taken from Van den Bergh and Delarue (2015).

The maximum ramp rate achievable by a generator variesacross studies. The National Renewable Energy Laboratory has produced a study which found that 'A typical large fossil-fired thermal generator may be able to ramp 1% of its capacity in 1 minute'

	Value	Unit	Reference (where applicable)
ES capacity	20	MWh	
ES power output	5	MW	
ES charge and discharge efficiency (η_c,η_d)	0.9		*Cost Projections for Utility-Scale Battery Storage: 2021 Update*
Biomass capacity	450	MW	
Biomass ramp rate	4.5	MW/min	
Biomass Opex	2.79	£/MWh	Duggal and Venkatesh (2015)
Biomass ramping cost	1.53	£/δMW	Van den Bergh and Delarue (2015)

Table 8: Parameters for operating biomass and energy storage systems.

(National Renewable Energy Laboratory (2005)). Furthermore, an assessment of ramping capabilities was carried out using data from the BMRS API, which lists the generation per BM unit. Draxx operates a biomass generator with BM unit name 'DRAXX-1' which has a capacity of 645 MW according to Elexon BM unit details. The generation data was available per settlement period. The maximum difference in generation between settlement periods for the year 2019 was 243.7 MW. This translates to a ramping capability of 8.12 MW/min (1.26% of capacity per minute). The true capability may be higher as this estimate assumes constant ramping throughout the settlement period but the biomass generator is at least capable of this rate. The results reflect the NREL paper and the NREL rule is used in this study (see table 8).

5.3 Biomass with energy storage system case study

In this section a case study is developed as a lens through which to view the technical aspects of operating a biomass and ES together to provide FFR. What follows is a justification for the sizing of the ES picked, foresight windows and other parameters that characterise the case study. First of all, the ES technology and sizing is addressed. Sizing of the ES is decided by assessing the net present value of several power capacities. Net present value (NPV) can be used in order to size the ESS technology suitable for this case study.

The NPV of an investment can be calculated using

$$NPV = -I_0 + \sum_{n=1}^{T} \frac{C_n}{(1+r)^n} \tag{21}$$

where I_0 is the initial investment cost, n is the year after the initial investment, T is the total lifetime of the investment C_n is the cash flow in year n and r is the discount rate.

Lithium ion batteries would be a suitable technology to provide FFR, having a response time of less than 40 ms (Meng et al. (2020)). Lithium ion batteries are popular for their high efficiency and high energy and power density. The initial investment costs have been studied by the National Renewable Energy Laboratory (NREL) who have collated information on lithium ion grid-scale batteries. They researched 6, 4 and 2hr systems. The 4hr system is a battery with a power to capacity ratio such that is can sustain its rated power output for 4hrs. For example, a 20 MWh battery would need a power rating of 5

MW to be a 4hr system. In selecting the 20 MWh battery capacity for frequency response applications in this study, the decision was underpinned by a strategic alignment with the project's scope. This capacity was chosen so that the whole system could be compared in subsequent chapters when the combined system will be used for forms of arbitrage and ultimately think of the operation of the system over its entire lifetime. The NREL report put the capital expense of the lithium ion batteries at 263, 287 and 423 £/kWh for 6, 4 and 2hr systems respectively see table 9.

20 MWh Battery	Storage duration (hrs)	Power capacity (MW)	Cost (£/kWh)
	2	10	423
	4	5	287
	6	3.33	263

Table 9: The cost per kW of installing a lithium ion battery according to *Cost Projections for Utility-Scale Battery Storage: 2020 Update* (2020)

The cashflow (C_n) is hard to predict before performing the optimisation and teaching the RL agent, but optimistic projections can be used initially. In that case, it was assumed that the availability fee was paid for each EFA block and that frequency response was required 100% of the time. In each case it was assumed that the battery tendered for FFR at its full rated power capacity i.e. 3.33, 5 and 10 MW for 6, 4 and 2hr systems respectively. Using this approach the income is the same in each n^{th} year after the initial investment.

Literature reviews on the NPVs of lithium ion batteries use a range of discount rates but the most relevant is a 2021 study that specifically looks at FFR in the UK and and recommends a 10% discount rate (Biggins et al. (2021)).

Figure 26 shows the NPVs for the three different batteries for a 20 year period. The figure shows that a 4hr system crosses 0 first and gets the best NPV past the 3 year mark (hence the decision to use a 4hr system in this thesis).

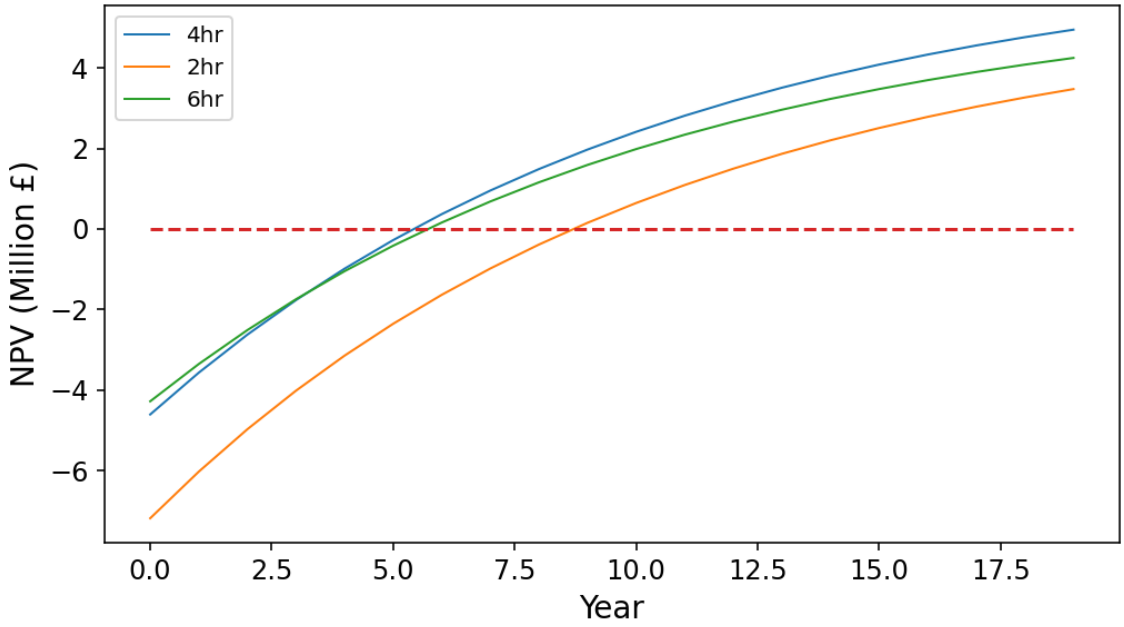

Figure 26: Net present value for 2, 4 and 6hr lithium ion battery over 20 years.

Initially the following case study was used. A 450 MW biomass generator was contracted to provide 5 MW of non-dynamic FFR for the year 2019 at 3.31 £/MWh / 126.94 £/hr. The minimum part load was set at 440 MW and other operational parameters were set as described in section 5.2 and presented in table 8.

The case 'Biomass no FFR' was run first, followed by 'Biomass with FFR'. Both are presented in the results to illustrate the comparison of a biomass generator and ES combined system. In the case of Biomass with FFR the ramping rate constraint of 4.5 MW/min is capable of tendering for 1 MW for FFR, meeting the requirement of responding within 30 seconds. Finally, biomass with FFR is retrained for increasing frequency sensitivity.

It was contrasted with the addition of a 20 MWh capacity ESS with a 5 MW power capacity and efficiency rating of 0.9. A summary of the cases explored and their parameters is presented in table 10

The case study assumed that there were no shutdowns and the combined system was contracted to operate in all EFA blocks.

The biomass and ES dispatch was then calculated with and without perfect foresight (sections 5.4 and 5.5 respectively).

	Biomass no FFR	Biomass with FFR	Combined system with FFR
Contracted EFA blocks	0	1-6	1-6
Tendered volume (MW)	N/A	1	5
Frequency trigger ($\pm Hz$)	N/A	0.2	0.2
Availability price (£/hr)	N/A	126.94	126.94
Response energy payment (£/MWh)	N/A	3.31	3.31

Table 10: Parameters for contracting firm frequemcy response in the three cases of Biomass no firm frequemcy response, Biomass with firm frequemcy response and Combined system with firm frequemcy response.

5.3.1 Exploring parameters

ES and biomass generator combined operation may be capable of delivering frequency regulation but the operational costs and profits are deeply dependant on supply and transportation of biomass, degradation effects in the biomass generator and future trends in frequency regulation by system operators.

The first point (supply of biomass) is particularly hard to estimate and very dependant on feedstock. Production and harvesting of a certain biomass feedstock can be estimated but the availability for supplying thermal generators can be complicated by demand for food, extracting sugars or carbon offsets in the form of biomass (documented by Zhang et al. (2010) and van Kooten and Johnston (2016)). Given the potential disruptions to biomass production discussed, it is prudent to explore further. Conducting a sensitivity analysis on the forms of contracts with National Grid ESO, operational expenditures of biomass facilities, and the newly anticipated ramping costs could provide crucial insights. This approach mitigates and highlights potential risks. Therefore, after the initial case study was analysed some parameters were varied to examine their impact on biomass and ESS operation and subsequent profits.

The parameters explored were

- **Frequency sensitivity** - Tested at 0.2, 0.15 and 0.1 Hz
- **Tendered volume** - Tested at 1 MW intervals from 1-10 MW
- **Biomass Opex** - Varied from -50% to 50% of its original value

- **Ramping cost** - Varied from -50% to 50% of its original value

5.3.2 Foresight

Following from this section, the biomass and ES dispatch were determined first with perfect foresight using a linear programming approach (section 5.4). It is unrealistic to determine the profitability of a combined system offering FFR by only presenting the perfect foresight case. Therefore a reinforcement learning method was used to imitate the actions of the perfect foresight agent or 'PFA'. The perfect foresight agent was optimised over a day, then the RL agent was trained from that data. The perfect foresight agent was then optimised for the next day, followed by that day's RL training and so forth. The justification for choosing a day-long horizon is presented in section 5.6.1 after elements of methodology have been explained.

5.4 Perfect foresight agent

In this section the perfect foresight agent is presented. The aim of this section is to schedule the biomass and ES dispatch to provide FFR. The key feature of this agent is that it has perfect knowledge of the frequency triggers that will occur in the future and can plan scheduling to maximise profits. The actions of the perfect foresight agent will be used subsequently as the PFA to teach the reinforcement learning (RL) agent.

The problem of solving the perfect foresight agent is constructed as a linear programming problem. The objective function for the perfect foresight agent is

$$\mathbf{max}(b_t - c_t + d_t) \times price_t - b_t \times Opex - |\delta b_t| \times rc + tv_t \times fd_t \tag{22}$$

where b_t is the biomass dispatch at time t, c_t is the energy storage charging at time t, d_t is the energy storage discharging at time t, $Opex$ is the operational cost of operating a biomass plant, $price_t$ is the remuneration for power delivered to the National Grid, rc is the cost of ramping a biomass plant, tv is the tendered volume in a given EFA block and fd_t is the frequency deviation at time t.
subject to

$$0 \leq \sum_t (\eta_c c_t + \frac{d_t}{\eta_d})\tau \leq EScap \tag{23}$$

where $EScap$ is the energy storage capacity, η_c it the efficiency of charging energy storage, η_d the efficiency of discharging energy storage and τ is the length of a timestep. Constraint 23 ensures the ES system maintains a state of charge between 0 and 1.

$$0 \leq c_t \leq p_c \tag{24}$$

$$0 \leq d_t \leq p_d \tag{25}$$

$$mpl \leq b_t \leq p_b \tag{26}$$

where mpl is the minimum part load and p_x is the power capacity of asset x. Constraints 24, 25 and 26 limit the power output of each asset to its power capacity.

$$-rr \leq \delta b_t \leq rr \tag{27}$$

where rr is the maximum ramping capability of biomass generator. Constraint 27 limits the ramping of the biomass generator.

$$\delta b_t - \delta c_t + \delta d_t = tv_t \times fd_t \text{ when } tr_t = 1 \tag{28}$$

where tr contains the initial frequency trigger, with a value of 1 where the frequency first passes 50.2 or 49.8 and 0 elsewhere.

$$\delta b_t - \delta c_t + \delta d_t = 0 \text{ when } res_t = 1 \tag{29}$$

where res has a value of 1 for the 120 timesteps after the frequency trigger and 0 elsewhere. Constraints 28 and 29 ensure the system meets the demand of providing its tendered FFR. The frequency data gathered is sampled at 15s periods. In the case study the contracted frequency trip was $\pm$ 0.2 Hz. The frequency profile was translated into a profile of required volume by monitoring for deviations of this size and for 120 timesteps (30 minutes) excluding original trigger the tendered volume was required. Further frequency trips within those 120 timesteps did not trigger a further action. The vector **tr** contains the initial frequency trigger, with a value of 1 where the frequency first passes 50.2 or

49.8 and 0 elsewhere. At the trigger point constraint 28 ensures the output changes by the tendered volume. The vector **res** has a value of 1 for the 120 timesteps after the frequency trigger and 0 elsewhere. Constraint 29 ensures the output of the combined system maintains constant for the 120 timesteps after the frequency trigger. Figure 27 shows the purpose **tr**, **tv**, **fd** and **res**.

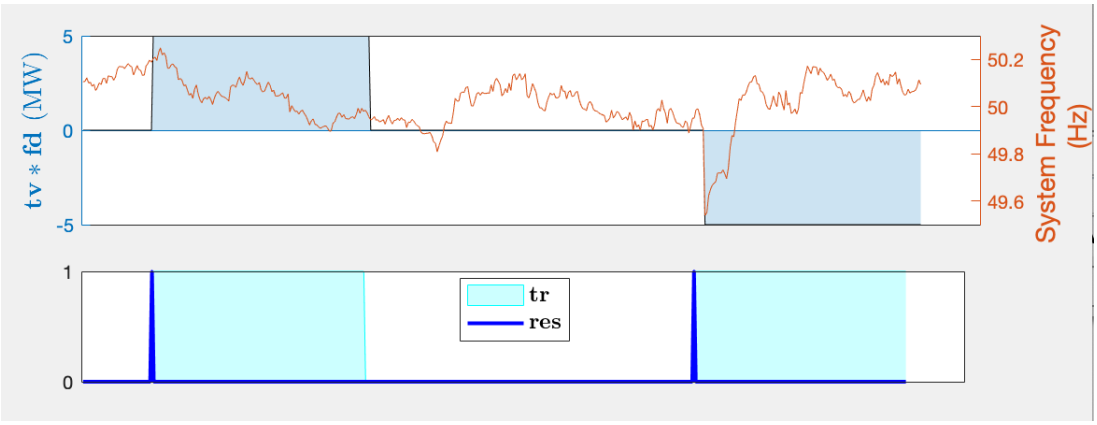

Figure 27: Above: The volume required by firm frequemcy response for the sample of historical frequency data. Below: **tr** (which contains the initial frequency trigger, with a value of 1 where the frequency first passes 50.2 or 49.8 and 0 elsewhere) and **res** (which has a value of 1 for the 120 timesteps after the frequency trigger and 0 elsewhere). When **tr** is 1 constraint 28 applies and the combined system changes its volume. When **res** is 1 constraint 29 applies and the combined system maintains its response. **tv** is the tendered volume in a given EFA block and **fd** is the frequency deviation.

The base case was solved by setting c_t and d_t to 0 in equation 22 28 and 29 and removing constraints 23-25.

The linear program was solved using Matlab's `linprog` and was undertaken on ARC3, part of the High Performance Computing facilities at the University of Leeds, UK.

5.5 Reinforcement learning agent

In this section, two state-action value functions are presented which were trained separately meaning that at the end of training there were two value functions ($Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$). The first had an action space of biomass dispatch and the second ES dispatch.

The remainder of this section is split into 5 subsections that explain the observation spaces for agents $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$, the reward functions for $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$, the policy for exploring the best actions, the interactions with PFA's actions,

and the training procedure in that order.

5.5.1 Observation space

The observation space of $Q_{Bio}(S, A)$ was made up of

- ES state of charge in the previous timestep
- Biomass dispatch in the previous timestep
- Frequency in the current timestep

The observation space of $Q_{ES}(S, A)$ was made up of

- Frequency in the current timestep
- Biomass dispatch in the current timestep

The observation spaces of the biomass and ESS intersect significantly. For instance, the biomass dispatch constitutes a component of the ESS observation space, providing critical input for its operational decisions. While the ESS dispatch itself does not directly influence the biomass observation space, the SoC of the ESS, a crucial determinant of its operational capacity, is an integral part of the biomass observation space. The SoC can only be accurately determined after the ESS dispatch has been finalised, creating a dependency that necessitates synchronous learning and decision-making. This interdependence means that actions taken by one agent inevitably influence the observation space and subsequent decisions of the other. Specifically, for the biomass system to make optimal dispatch decisions, it must account for the current SoC of the ESS. This dependency ensures that the biomass system can adjust its operations based on the available energy storage capacity, leading to more efficient and coordinated overall system performance. Therefore, the biomass observation space must dynamically incorporate the ESS SoC to make informed and effective decisions.

Expressing the variables as vectors they are ES state of charge (**SoC**), system frequency (**f**), biomass dispatch (**b**), and ES dispatch (**e**). ES dispatch comprises the combination of charging (c) and discharging d from section 5.4 $\mathbf{e} = c - d$. Each vector contains T elements, each element representing the ES state of charge, system frequency, biomass dispatch and ES dispatch at time t. The vector **f** contains the historical frequency data. **SoC**, **b** and **e**

will be filled as learning proceeds. Figure 28 demonstrates how this occurs. Starting with the biomass observation space at $t = 1$, the observations generate the biomass dispatch in the first timestep (b_1). The biomass dispatch then makes up a part of the ES observation space. From the ES observations the ES dispatch (e_1) is generated. The ES dispatch is used to update the SoC for the second timestep.

$$SoC_t = \begin{cases} SoC_{t-1} + \frac{\eta_c e_t \tau}{EScap} & \text{if } e_t \geq 0 \\ SoC_{t-1} + \frac{e_t \tau}{\eta_d EScap} & \text{otherwise} \end{cases} \tag{30}$$

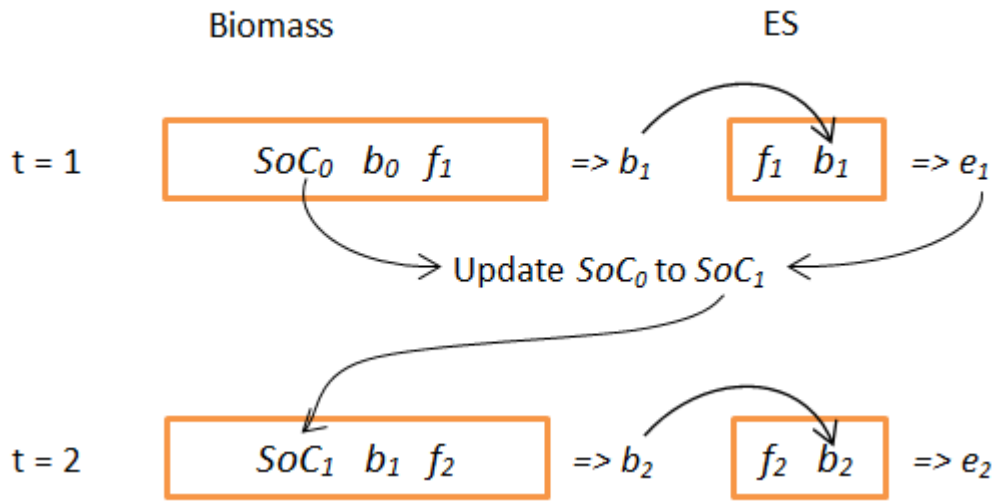

Figure 28: An illustration of the connected learning of the Biomass and energy storage agents. The left and right hand sides show the biomass and energy storage systems respectively. First the biomass observation space contains SoC_0, b_0 and f_1. The observation space is used to generate the biomass dispatch (b_1). The biomass dispatch is then used in the energy storage observation space. The energy storage dispatch (e_1) is generated and then combined with SoC_0 to calculate SoC_1. In the next timestep SoC_1 is used in the observation space of the biomass. The resultant biomass dispatch (b_2) is used in the energy storage observation space and so on for t = 1,2,3,...

Subsequently the new SoC (SoC_1) is used in the biomass observations in the second timestep 2 to generate biomass dispatch b_2. b_2 is used in the ES observation space in timestep to generate e_2 and so on until the vectors are full. An algorithmic representation can be found on page 106.

Both $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$ were represented by a matrix, meaning that each observation and action dimension was discredited. Discretisation of ES dispatch, biomass

dispatch and state of charge for observation and action spaces are presented in table 11.

Observation dimension	Bin width	Lower limit	Upper limit
State of charge	0.1	0	1
ES dispatch (MW)	1	-5	5
Biomass dispatch (MW)	1	440	450

Table 11: Discretisation of variables for action and state spaces.

The frequency was discretised into three boxes; any frequency above 50.2 Hz, any frequency between 50.2 and 49.8 Hz and any frequency below 49.8 Hz. Likewise, when the frequency triggers were adjusted to examine the effects on profits the frequency discretisation was adjusted accordingly so that, for example, if the frequency trigger were ± 0.15 Hz the boxes would be 0-49.85, 49.85-50.15, 50.15+ and so forth for other frequency triggers.

The rewards for actions b and e are presented in equations 31 and 32.

$$R = exp(-|b - \sigma_b|) \tag{31}$$

$$R = \begin{cases} exp(-|e - \sigma_e|) & \text{if } \ 0 < SoC \times EScap + \eta_c c\tau + \frac{d\tau}{\eta_d} < EScap \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

where σ_x is PFA's dispatch decision for asset x.

Notably, there is nothing in the reward function that explicitly references the maximum ramp rate of the biomass generator. However the form of the observation space and reward should teach the RL agent to keep within the ramping rate. The biomass ramp rate in the previous timestep is part of the observation space. The PFA will only recommend a subsequent dispatch that is constrained by the ramping rate. Equation 31 shows that the RL agent will be penalised for choosing an action other then PFA's and so will be discouraged form choosing actions that violate the ramping constraint. In this way the RL agent will naturally learn the ramping constraint.

5.6 Interactions with the perfect foresight agent

In this subsection there is a description of the training process for developing the state action value functions.

It was established at the start of this section that the rewards will hinge on the difference between the PFA and the RLA's actions, with a high reward encouraging the RLA to repeat a given action and a low reward discouraging a given action.

However consider the following example; the PFA chooses an action and updates its environment. The RLA chooses a different action and updates its own environment. Using equation 17 the resultant reward is small or 0. This is an important learning step for the RLA, which is now less likely to repeat the same mistake. However, if the RLA were to make several large mistaken actions the environments of the RLA and PFA would start to diverge significantly. This could be a problem as the optimal PFA action will be dependant on its an environment and this could lead to the RLA being given the wrong advice from the PFA for its current environment.

Looking at this problem, it helps to consider two parameters.

1. Optimisation window (w_o): How far into the future the PFA looks when scheduling dispatch

2. Training window (w_t): How much training is done before the environments are realigned

Figure 29 illustrates the relationship between the optimisation window and training window. The PFA schedules actions over the optimisation window with its state represented by σ_t at timestep t. The RLAs are trained on data from the optimisation. At the end of the training window the states of the two agents can be compared. To account for any divergence the optimisation is repeated from the end of the training window with states of the agents set to equal.

In an ideal world, the optimisation window would be long and the training window short. For example, if the optimisation window were a year long and the training window 15s there would be no state divergence. The optimisation would schedule dispatch perfectly and the RLA would attempt to learn the PFA's action in that timestep. The RLA would pass information about its state to the PFA and a new year-long optimisation would begin and so on with year long optimisations for every timestep in the year. Such a methodology would be fairly time-consuming and computationally intensive.

Take the example of training a biomass generator supported by an ESS. As was established in section 3.1, the best training regime uses two RLA; one learning the dispatch of

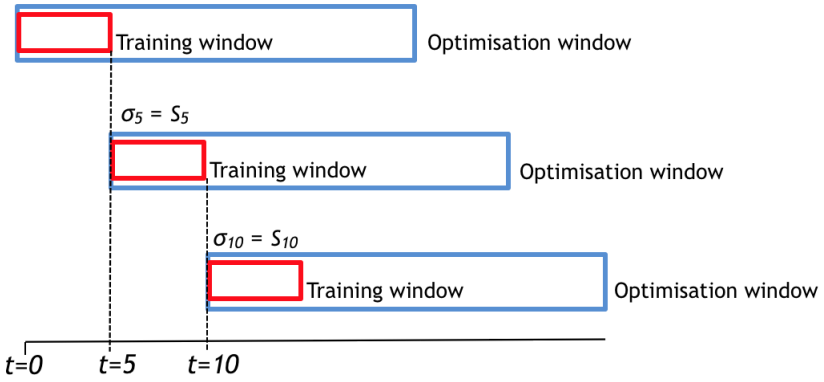

Figure 29: A diagram to demonstrate the optimisation and training windows. The PFA schedules dispatch over the optimisation window (blue boxes). The RLAs are updated over the training window (red boxes). At the end of training the RLA's state is passed to the PFA, which is used to reset the PFA's environment for the start of the next optimisation window.

Algorithm 1 Algorithm for learning Q_{Bio} and Q_{ES}

1: Initialise Q_{Bio}, Q_{ES}, $\sigma_{Bio,0}$, $\sigma_{ES,0}$
2: n = 1
3: **while** $n < N$ **do**
4: Get PFA's actions for $n : n + w_o$
5: Train Q_{Bio} and Q_{ES} (as in algorithm 2) for $n : n + w_t$
6: $\sigma_{Bio,n+w_t} = S_{Bio,n+w_t}$, $\sigma_{ES,n+w_t} = S_{ES,n+w_t}$
7: $n = n + w_t$
8: **end while**

the biomass generator ($Q_{Bio}(S, A)$) and one for the ESS ($Q_{ES}(S, A)$).

Algorithm 1 shows the order of gaining the PFA's actions and training the RLA, moving between optimisation window and training window as was described previously.

Algorithm 2 shows how $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$ were developed during the learning step referenced in algorithm 1. It summarises all the information presented so far in this chapter. It shows the actions for the biomass dispatch at time t (b_t) and ES dispatch at time t (e_t) being generated from their respective state-action value functions via the ϵ-greedy policy. In lines 8 and 9 the rewards are generated from equation 17. The actions for $t + 1$ are generated by the ϵ-greedy policy in lines 11-17. And finally the update steps occur with temporal difference learning, as described in equation 6.

Algorithm 2 Pseudocode demonstrating Q_{Bio} and Q_{ES} are trained together, where T is the length of the training window. Rand(0,1) represents drawing a random number between 0 and 1. Randin(X) represents picking a random number in the set X.

for t=1:T **do**
  if rand(0,1) $< \epsilon$ **then**
    $b_t = \text{randin}(B)$
    $e_t = \text{randin}(E)$
  else
    $b_t = max_B(Q_{Bio}(S_{Bio,t}, B))$
    $e_t = max_E(Q_{ES}(S_{ES,t}, E))$
  end if
  Get R_{Bio}
  Get R_{ES}
  if rand(0,1) $< \epsilon$ **then**
    $b_{t+1} = \text{randin}(B)$
    $e_{t+1} = \text{randin}(E)$
  else
    $b_{t+1} = max_B(Q_{Bio,t+1}(S_{Bio,t+1}, B))$
    $e_{t+1} = max_E(Q_{ES,t+1}(S_{ES,t+1}, E))$
  end if
  $Q_{Bio}(S_{Bio,t}, b_t) += \alpha(R_{Bio} + \gamma Q_{Bio}(S_{t+1}, b_{t+1}) - Q_{Bio}(S_t, b_t))$
  $Q_{ES}(S_{ES,t}, e_t) += \alpha(R_{ES} + \gamma Q_{ES}(S_{ES,t+1}, e_{t+1}) - Q_{ES}(S_{ES,t}, e_t))$
  $\epsilon = \epsilon - 1/T$
end for

5.6.1 Training window

For the initial case study of combined system providing FFR (paramters in table 10) the following training windows were tested a full day, 3 EFA blocks, 1 EFA block and 1 hour.

In section 3.2 the concept of an ϵ-greedy policy was introduced, with ϵ gradually decreasing. At the start of learning epsilon was high and thus random actions likely to take place. At the end of learning ϵ is low and random actions very unlikely. This implies there will be a lot of divergence in SoCs at the start of the learning process and it would follow to use a shorter training window at the start of learning. To illustrate this the following process was carried out for combinations of ϵ and training window length.

A training window was taken for which it was assumed the optimal ES dispatch was 0 throughout the training window. For each timestep in the training window a random number r was drawn between 0 and 1. If $r < \epsilon$ the ES dispatch was drawn at random from the action space and the SoC updated. Otherwise it was assumed the ES dispatch is also 0. The setup of an optimisation window with optimal ES dispatch of 0 throughout

is a little contrived but now the SoC divergence can be plotted. Example paths for ϵ = 0.9 and 0.1 at training lengths 1 and 24 hrs are shown in figure 30. They illustrate that a longer training window increases divergence, as does a higher ϵ.

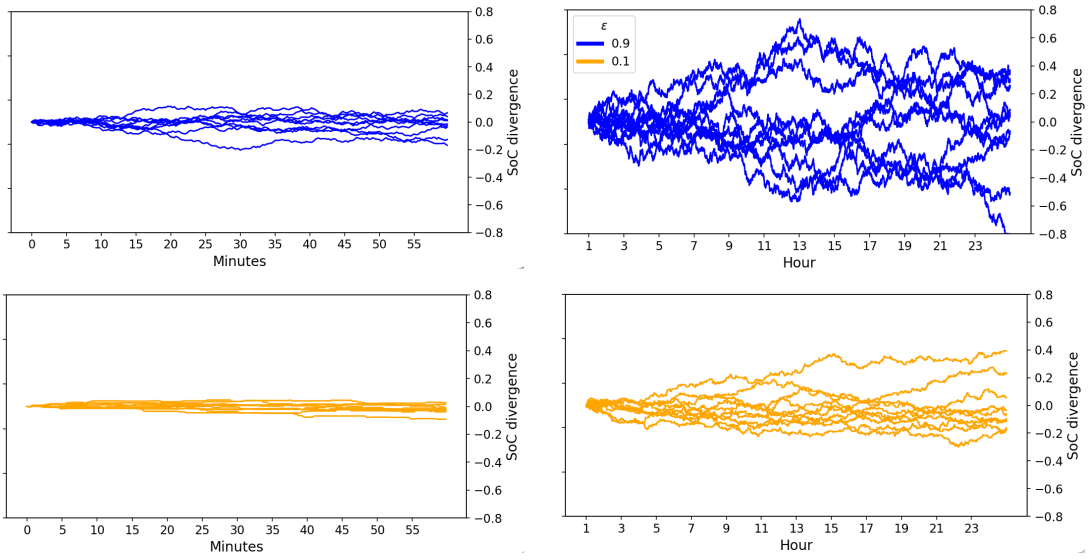

Figure 30: Example paths. Graphs on the left show a training window of 1 hr, on the right 24 hrs. Orange show paths taken with ϵ = 0.1 and blue ϵ = 0.9.

.

In conclusion at the start of learning the training window has to kept short, and towards the end of learning the training window can be relaxed to increase the speed of the methodology.

It was decided that for ϵ = 1-0.4 a training window of an hour would be used, 2 hrs at ϵ = 0.4-0.3 and 1 EFA block for the remainder of the time.

5.6.2 Optimisation window

Now that the suitable training windows have been established, an examination of the optimisation window is presented. The optimisation window needs to be at least as long as the training window and should be longer. Otherwise the PFA may make poor decisions, especially at the end of the optimisation window.

For a training window of 1 EFA block, several optimisation windows were tested. Figure 31 shows the difference in PFA's actions for different optimisation windows for a random EFA block. The EFA block (20th June 2019, EFA block 4) was drawn at random. It shows that when the optimisation window is reduced to 1 EFA block PFA's behaviour

changes to empty the ESS at the end of the optimisation window. The reason for such is as follows; the ESS managed under these conditions strategically seeks to empty its stored charge by selling energy back to the grid. Already knowing that there will not be a requirement to discharge in the EFA block due to a frequency trip, this ESS will sell all its available charge instead of 'playing safe' and keeping its SoC at a higher level.

500 random EFA blocks were looked at and it was found that the 1 day and 1 week optimisation windows always agreed, the 1 EFA block optimisation window always diverged from the others and in 16.4% of cases the 3 EFA block window also diverged from the others. In conclusion, the 1 day optimisation window seems sufficient to capture PFA's behaviour in the training window.

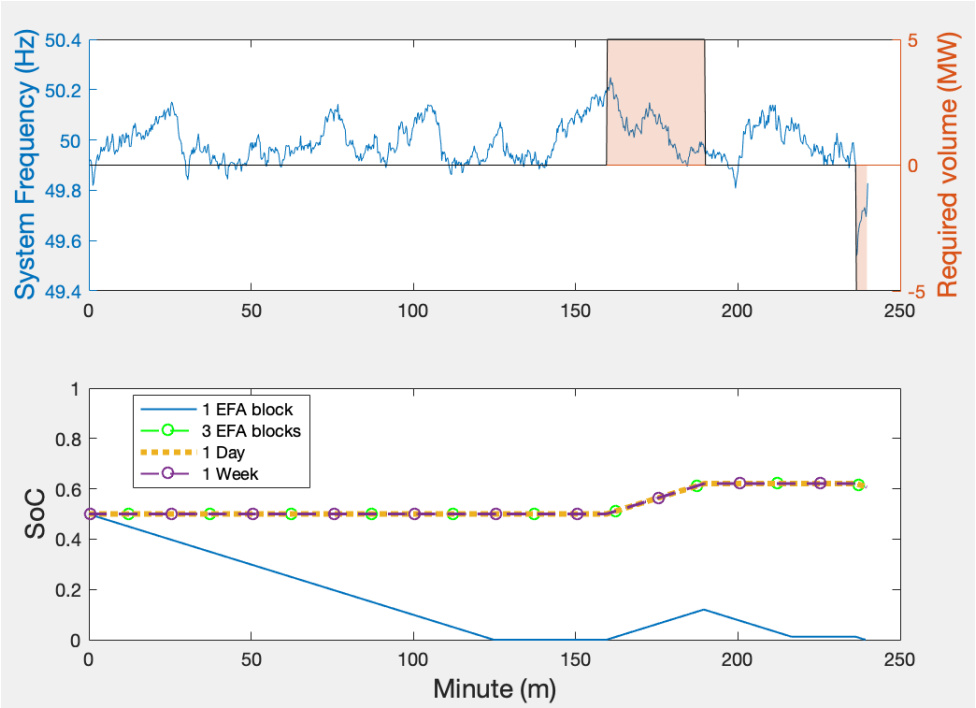

Figure 31: Training window for 4 different optimisation lengths.

5.6.3 Training

In this subsection there is a description of the training process for developing $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$. First of all, the frequency data was split into training and testing sets. Only the training data was used to develop $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$, leaving a set of testing data that the agent has never seen before and so can test the efficacy of the learnt policy. The training set contained 80% of the full data set, comprising 01-01-2019 - 19-10-

2019 and containing 1,681,920 data points. For testing EFA blocks were drawn at random from the remainder of the year.

Algorithm 2 shows how $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$ were developed during the learning step referenced in algorithm 1. In that algorithm (algorithm 2) The notation for $Q_{Bio}(S, A)$ would become $Q_{Bio}(SoC_t, b_t, f_{t+1}, b_{t+1})$ and $Q_{ES}(S, A)$ $Q_{ES}(f_t, b_t, e_t)$.

The order of generating actions and updating observations are important as they reflect the interacting action and observation spaces of the biomass and the ESS as described in section 5.5.1. At the start of 1 the initialisation sets $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$ to 0 for every state and action, $\sigma_{SoC,0}$ to 0.5 and $\sigma_{b,0}$ to 445 MW. Table 12 shows the parameters α and γ that were used in this study.

In section 5.3 the parameters that needed to be varied were listed. The actions of the perfect foresight agent are represented by σ. In retraining for new parameters, σ in the reward functions are updated with PFA's new actions, the remainder of the methodology is carried out in the same manner.

Learning parameters	
Step-size parameter (α)	0.1
Discount factor (γ)	0.99

Table 12: Parameters used in reinforcement learning when learning to mimic the perfect foresight agent.

5.7 Results

5.7.1 Baseline results

The initial case study with a 450 MW biomass generator etc is presented as described in section 5.3. Contracting with a minimum part load of 440 MW, the PFA maximises profits by operating at a constant 450 MW output, leading to a profit of £4752 for an EFA block.

Biomass no FFR			
Average profits per EFA block (£)	4752		
Biomass with FFR			
Frequency trigger (± Hz)	0.2	0.15	0.1
Average profits per EFA block (£)	5170.08	5158.24	5027.56

Table 13: Baseline results for a biomass generator with no energy storage systems. Average profits per EFA block are presented for a biomass generator without Firm Frequency Response, a biomass generator with Firm Frequency Response at varying frequency triggers and at various minimum part loads

Including FFR to the initial case study immediately increases profits see table 13. The availability price (126.94 £/hr) for providing across the EFA block should add £507.76 to profits. That the average profit has not increased by this degree from the 'Biomass no FFR' case suggests that the cost of ramping the biomass generator outweighs the income from providing response during frequency triggers. Therefore it is the availability price that keeps FFR profitable without an ESS.

Without an ESS, extra frequency sensitivity causes average profit per EFA block to decline which supports the conclusion that ramping the biomass for FFR is eating into profits. The cost of ramping outweighs any benefits from proving any extra volume in the EFA block. This suggests that should frequency deviations become more frequent, providing FFR will become less profitable.

5.7.2 Initial case study

In this case study, with an ES system capacity of 20 MWh, profits per EFA block are presented in figure 32. The average profit for the RL agent was £5248.73 per EFA block, just under the £5252.89 average profit for PFA. Discrepancies in the profit come from more regular cycling of the biomass plant from the RL agent.

Figure 33 shows PFA and RL behaviour for a representative EFA block. The EFA block in the diagram (EFA block 1 10-12-2019) was chosen at random from the testing data. It contains three frequency triggers. The PFA maintains a steady biomass output while the RL agent ramps the biomass to takeover the FFR requirement from the ESS.

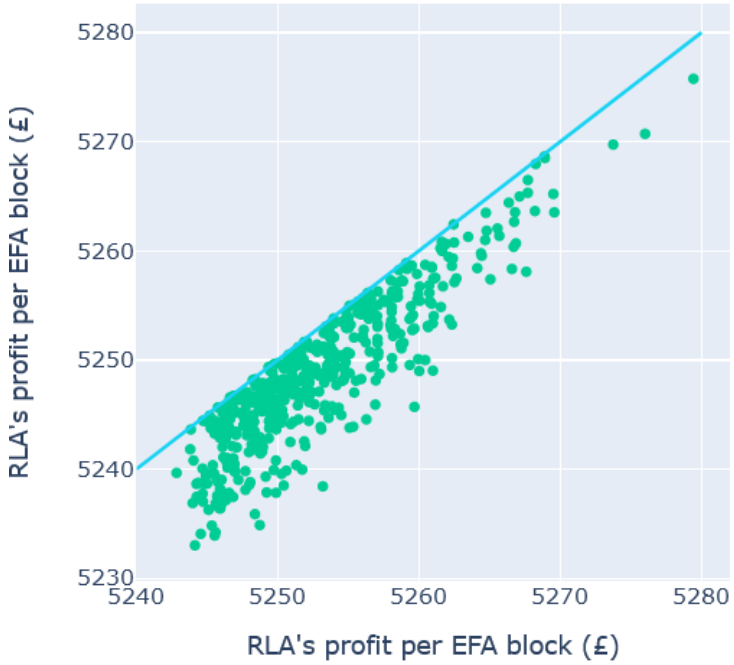

Figure 32: Scatter of reinforcement learning agent vs perfect foresight agent's profits per EFA block for 500 EFA blocks. Red line shows maximum where the reinforcement learning agent matches the perfect foresight agent's profits.

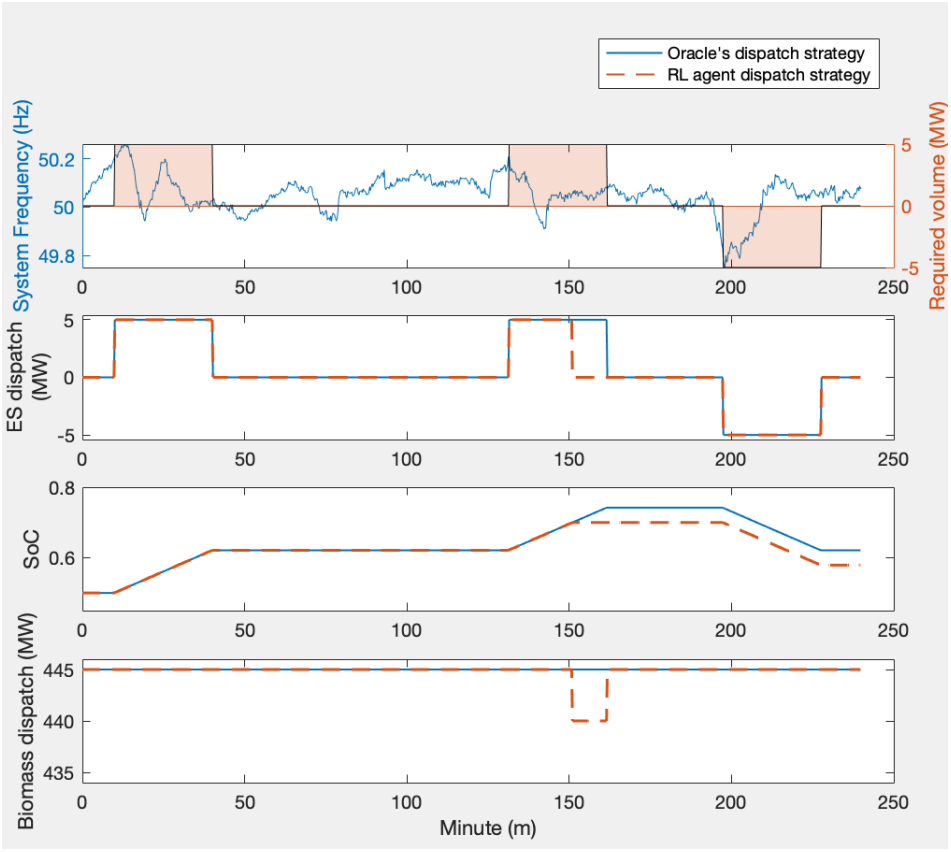

Figure 33: System frequency, required volume, energy storage and biomass dispatch and state of charge for a representative EFA block with 3 frequency trips

Figure 34 shows the policy for the biomass agent at an ES SoC of 0.7. It can be seen that at a low frequency the biomass generator maintains a constant output. At a high frequency the biomass ramps down and at $49.8 \leq f_t \leq 50.2$ the biomass ramps up to 445 MW before returning to a constant output.

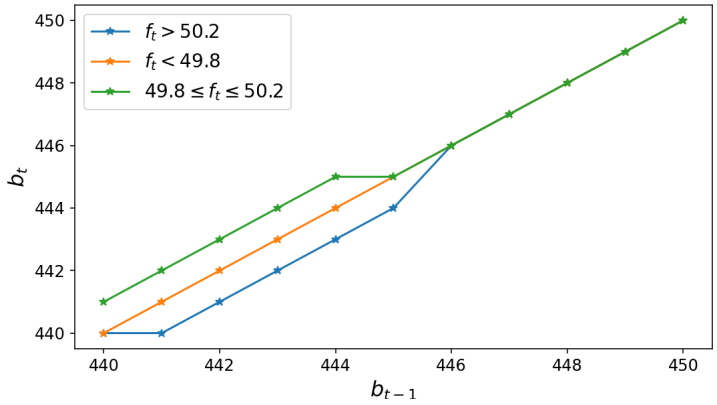

Figure 34: Biomass dispatch for $SoC_t = 0.7$ for combinations of b_{t-1} and f_t

5.7.3 Retraining

5.7.3.1 Varying frequency sensitivity The case studies with stricter frequency monitoring are presented in table 14, showing that profits directly gained from providing frequency response grow as more sensitive frequency triggers are contracted. The official guidance from National Grid is that 'The standard deviation of response error over a 30 minute period must not exceed 2.5% of the contracted response'. Measured this way, there is no 30 minute period over the year that breaches the contract terms even with the highest level of contracted frequency sensitivity.

5.7.3.2 Varying tender volume Figure 35 shows the results of varying the tendered volume. Between 1 and 5 MW there is a roughly linear increase in profits as the availability price remains fixed and the response energy payment accounts for increased profits. Profits

Frequency trigger (Hz)	Profit gained from providing FFR (£/EFA block)
± 0.2	5248.73
± 0.15	5255.05
± 0.1	5293.46

Table 14: Profits gained from providing Firm Frequency Response for increasing sensitivity to frequency deviations.

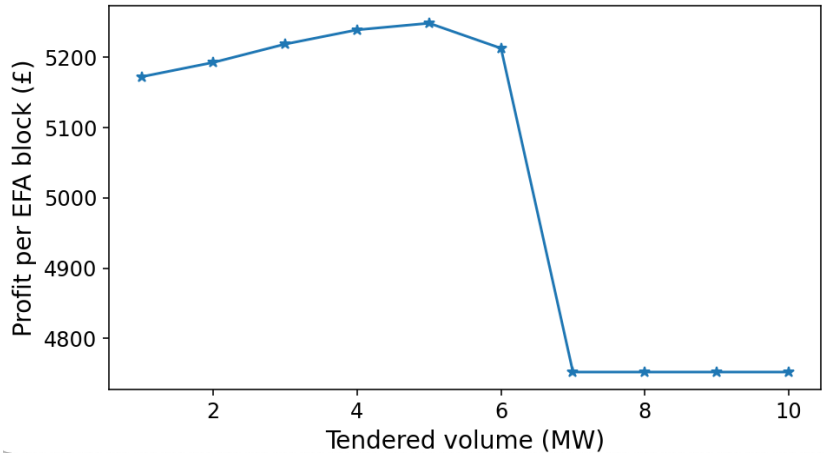

Figure 35: Profits while varying tendered volume for the biomass and energy storage combined system

peak at 5 MW and decrease over that volume. This can be explained by noting that the maximum power output of the ESS was 5 MW. Supplying volume over the ESS power capacity requires the biomass generator to ramp its output and leads to a decrease in overall profits. It can be concluded that optimally the operator can bid to provide the full capacity of their ESS.

Above 6 MW biomass ramping restrictions stop the combined system being able to provide the tendered volume. The system would then lose its response energy payment and availability price and revert to the operation of a biomass generator not providing FFR.

5.7.3.3 Varying operational biomass parameters The operational parameters effect the profits linearly as shown in figure 36 with changes in biomass opex having the biggest impact and maximum ramping cost the least. This follows from the vast volume generated by the biomass (all of which is subject to opex costs) compared to the relatively small volume being ramped.

It is informative to look at the operation of the ES to see the profit impacts of varying ramping cost and rate. As demonstrated by figure 33 at a certain SoC the biomass will ramp to take over provision of FFR. The system does this to provide a safety net of SoC. Then if there is a necessity to provide FFR the ESS is prepared. In the EFA block in figure 33 during an over-frequency event as SoC passes 0.7, the biomass generator ramps to take over FFR provision.

Table 15 presents the frequencies at which the biomass generator takes over FFR provision as the ramp rate and cost are varied. The upper SoC shows when biomass takes over in an over frequency event and the lower SoC indicates when biomass takes over in an under frequency event. Decreasing ramping cost allows for more ramping and less dependence on the ESS. Subsequently the ESS keeps its SoC within a smaller range while biomass provides more of the FFR. Increasing ramping cost makes ramping less desirable and ESS provides more FFR. Despite this, the overall loss in profits suggests the increased cost outweighs the money saved by avoiding ramping.

Ramping cost										
	-50%	-40%	-30%	-20%	-10%	+10%	+20%	+30%	+40%	+50%
Upper SoC	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8
Lower SoC	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.3	0.3

Table 15: State of charges during periods of providing volume

5.8 Energy storage system lifetime providing Firm Frequency Response

As was outlined in section 3.3, the lifetime of the ESS can be calculated, given the time-series of charging and discharging generated by the PFA and RLA. The cycling of the battery for the PFA over 2019 led to a lifetime loss of 9%. For the RLA the lifetime loss amounted to 7%, an interestingly smaller loss than that accrued by the PFA.

5.9 Conclusion

In conclusion, a biomass generator operating alone can make a profit from FFR provided the operator sets their availability price sufficiently high. With the realistic availability price presented here, the biomass operator is even able to generate profit with increased sensitivity to frequency deviations. However contracting for increased sensitivity does not generate additional profits due to the cost of ramping. Adding an ESS makes FFR provision more profitable. Profits are maximised when the operator sets the FFR tender at the power capacity of the ESS. Adding an ESS also makes it profitable to contract at increasing frequency sensitivities. This means that

1. The operator can choose to contract for smaller frequency triggers.
2. In the event that frequency deviations become more common the operator can maintain their contract and increase their income.

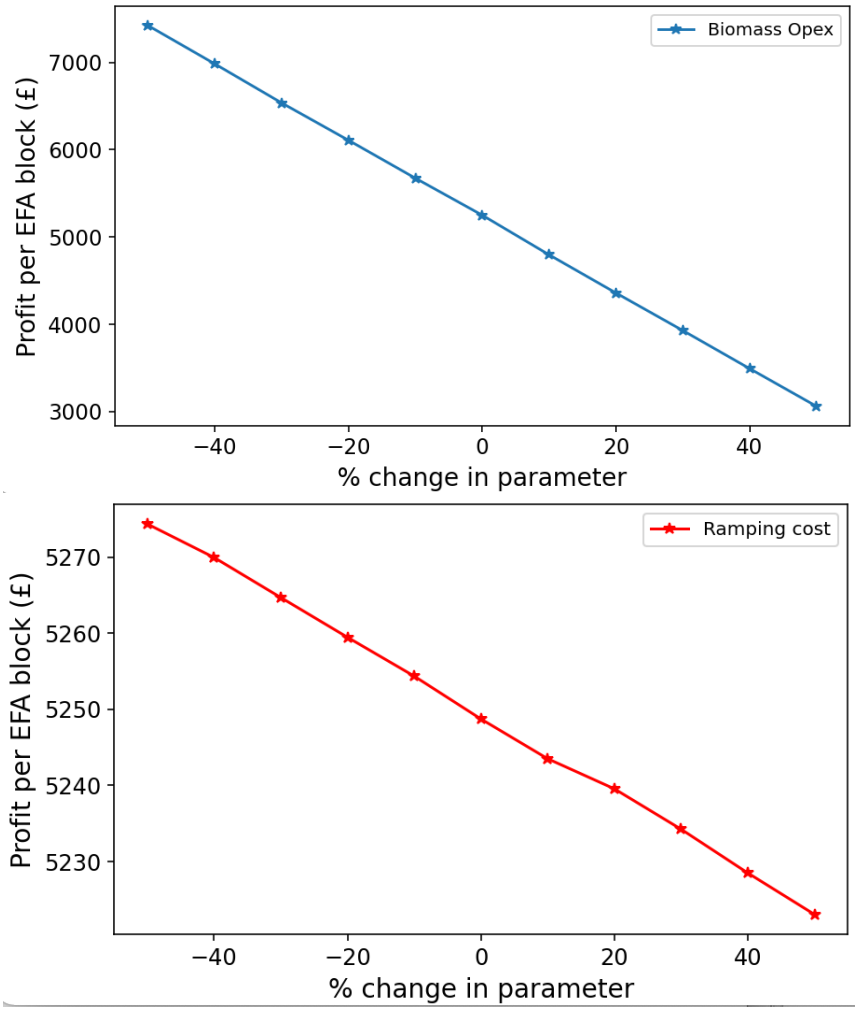

Figure 36: Profits while varying operational biomass parameters including biomass Opex, ramping costs and maximum ramping rate. Plotted on separate axes to show the smaller variations for ramping cost and rate.

Increased ramping costs reduce the dependence on biomass for providing FFR while still reducing overall profits. With even a 50% increase in ramping costs operating the biomass generator with ES remains profitable. On the other hand, an increase in biomass Opex of 20% results in a profit under £4752 which is the profit from operating the biomass without ES or providing FFR. Looking at figure 36 it seems the combined system offering FFR would pass the £4752 mark at an Opex increase of 17%. In the future further research into the Opex should be carried out as it is an important and sensitive parameter in determining system operation and FFR profitability.

In conclusion, though FFR is a critical service in the realm of electricity grid management, it's pivotal to recognise that the operational strategies and mechanisms in energy

markets are not static. These systems must adapt over time to harness new opportunities and address emerging challenges. Transitioning from the well-established concept of FFR, the next chapter delves into the nuanced world of arbitrage on the BM. The link between FFR and arbitrage is founded on the operational flexibility and market insight required to exploit economic opportunities within the grid's BM. Arbitrage involves capitalising on short-term price fluctuations within the balancing framework. The skills and insights gained from managing frequency can be pivoted to perform successful arbitrage. As energy markets grow more volatile and renewable penetration increases, the importance of adaptive operational strategies becomes apparent. This evolution necessitates a shift in focus from mere participation in grid stability to actively seeking profit-making opportunities through strategic energy sales and purchases. Thus, as there is a transition into discussing arbitrage, it's clear that the operational paradigms of energy storage and grid support services may need substantial adaptation throughout their lifetime to remain effective and profitable in a rapidly changing energy landscape.

6 Medium-duration storage

Pupo-Roncallo et al. (2021) studied the need for MDES in 2018, finding that the installed capacities of ES (duration >10 h, <100 h) need to scale linearly with NDR capacity. In the study (which projects a 100% penetration of NDRs to 2050) MDES should maintain 11 TWh of capacity. The study minimises the total cost of electricity when there is a mixture of all duration lengths. Short duration energy storage (SDES) has insufficient power-to-capacity ratios to perform inter-day arbitrage while long duration ES has a low turnaround efficiency for electricity. The objectives and main purpose of this chapter is to examine the operation of an MDES starting in the year 2022 and continuing to operate all the way to 2035. The aims are to consider if an MDES can be used over that 13 year time period, and how this compares to the FFR operation presented previously. MDES (defined in the literature review as between 4 - 200 hours of an energy to power ratio) has been chosen to be studied here because of its interesting role in the FES which will be discussed at length later in this chapter. While the mismatch between supply and demand can be broken down into intra-day, inter-day, weekly and seasonal effects, MDES has increasingly found popularity and is projected (including in the FES) to proliferate

because it is needed for intra-day balancing of supply and demand. The intra-day effects are largely credited to the diurnal cycle of solar power. Understanding this mismatch is vital to understanding the potential value of installing an ESS. The structure of the chapter is as follows

- In order to understand an MDES bidding on the BM, historical settlement prices from 2022 are briefly described and analysed.
- An LPP is introduced that will make up the PFA for the MDES operating in 2022. The relative constraints and objective function are designed.
- The RLA for 2022 is presented, with a precursor analysing how exactly the observation and action spaces were chosen.
- The FES are described in detail and the role of MDES in the FES are explored. The main purpose of this section is to create another PFA; this one will act as training material for the RLA acting up to 2035.
- The RLA for 2035 is presented, including a description of the training regime.
- The results compare the performance of each PFA to each RLA and addresses the uncertainties in the previously outlined model.
- Finally there is a description of how imbalance pricing and rainflow counting can be applied to the results. Both measures are included as they add important insights into the ultimate operation decisions dictated by the RLAs.
- The LCOS is calculated for the results. The crux of this section reintroduces FFR operation and compares the LCOS of both.
- The results are discussed in the context of emerging energy market reform.

6.1 Historical settlement prices in 2022

The settlement prices in 2022 had a maximum of 652 £/MWh, a mean average of 57.1 £/MWh and a standard deviation of 22.3 £/MWh. Prices went over 100 £/MWh 7,728 times. The prices were consistently lowest in settlement period 30, which lasts from 15:00 - 15:30. Acceptance of bids came most from gas and wind.

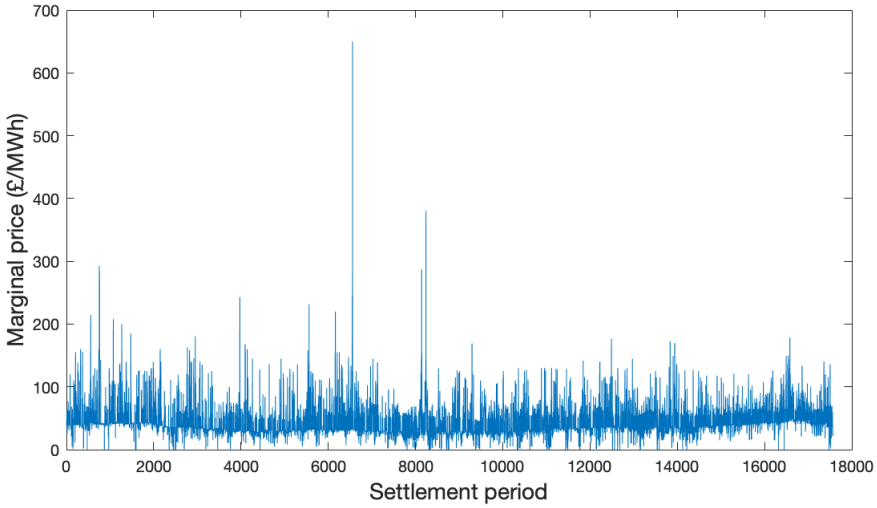

Figure 37: The settlement price over 2022.

The system of flagging and tagging etc. that was first introduced in chapter 4 has also been analysed for the year 2022. To reiterate, flagging is concerned with distinguishing between trades made for energy and system balancing, where system balancing concerns trades that meet physical constraints of the transmission grid (e.g. SO-flagging). The number of trades in 2022 flagged for being a system trade in 2022 came to 26.9%, with 0.4% being CADL flagged (76.2% from a bids by an ESS). 82% of first stage flags were also second stage flagged.

6.2 Establishing the linear programming problem

In order to examine the MDES arbitrage, a methodology for assessing the operation of the CBESS with perfect foresight is first presented. The objective of this approach is to minimise the cost to the CBESS operator, following the methodology outlined previously and fulfilling objective **O3**.

$$\mathbf{max}(b_t - c_t + d_t) \times price_t - b_t \times Opex_b - |\delta b_t| \times rc \tag{33}$$

subject to

$$0 \leq \sum_t (\eta_c c_t + \frac{d_t}{\eta_d})\tau \leq EScap \tag{34}$$

$$0 \leq c_t \leq p_c \tag{35}$$

$$0 \leq d_t \leq p_d \tag{36}$$

$$mpl \leq b_t \leq p_b \tag{37}$$

$$-rr \leq \delta b_t \leq rr \tag{38}$$

where b_t is the biomass dispatch at time t, c_t is the ES charging at time t, d_t is the ES discharging at time t, $EScap$ is the energy storage capacity, mpl is the minimum part load, $Opex_x$ is the operational cost of operating a asset x, $price_t$ is the remuneration for power delivered to National Grid ESO, p_x is the power capacity of asset x, rc is the cost of ramping a biomass plant and rr is the maximum ramping capability of the biomass generator.

6.3 Medium-duration storage through 2022

Having established the operation of the LPP it is now requisite to describe the RLAs that will do the learning. The RLAs are split into two groups; those designed to study MDES through 2022 and through 2035. These will then be compared and contrasted to demonstrate the evolving role of MDES up to 2035. The analysis ends in 2035 originally due to the general guidance given by Oudalov et al. (2007) that the lifetime of a 20 MWh/ 5 MW lithium ion battery is 13-15 years. Subsequently, Rahman et al. (2020) and Hossain et al. (2020) confirmed this. The RLAs in this section show similarities to those developed in the last chapter; particularly in the sense that the two RLAs are $Q_{Bio}(S, A)$ and $Q_{ES}(S, A)$ for both the biomass generator and ESS. First of all the observation space of $Q_{Bio}(S, A)$ was made up of 1) Biomass dispatch in the previous timestep and 2) BM settlement price in the previous timestep and 3) imbalance. Imbalance refers to the state of the system i.e. whether supply outweighs demand or vice versa. It allows for the inclusion of negative prices (getting paid to charge the MDES). For $Q_{ES}(S, A)$ the observation space was made up of 1) SoC in the previous timestep 2) aggregated wind and solar factor in

the previous timestep.

Both wind and solar dispatch will inform the best operation of the ES. However, to reduce the dimensionality of $Q_{ES}(S, A)$ the wind and solar factors were aggregated in the following way

$$ws = \sqrt{w_a^2 + s_a^2} \tag{39}$$

where ws is the aggregated factor, w_a is the normalised wind output and s_a is the normalised solar output. To avoid having either the wind or solar data having an outsized impact on the aggregated factor the solar and wind outputs are divided by the mean solar and wind output for the entire dataset, resulting in a new dataset with an average of 1.

Next is presented a justification of using an aggregated wind and solar factor in any case. To do this there is a discussion of the peak NDR dispatch and peak load (by which is meant all the demand across the grid). NDR dispatch peaks at some point during the day as does load. The daily weather cycle will influence the NDR dispatch peak and may also influence the load peak due to daily work cycles influenced by light and weather patterns leading to increased heating and/or cooling of homes. Figure 38, wherein the FES were examined, shows how the pattern generated by the interaction between NDR dispatch and load profiles can create a new profile of their own. Therefore and despite not necessarily coinciding, it would be reasonable to presume that NDR dispatch would be a indicator for our RLA to change its own dispatch decisions. On examination of the 2022 data, the peak load was most likely to coincide with peak NDR dispatch on days that 1) had a combination of generally high wind and generally high solar and 2) had a combination of generally low wind and generally low solar.

To understand why euclidean distances were used rather than the mean or other metric consider that:

- **When the numbers are very different**: The mean will be somewhere midway between the two numbers, showing a central tendency. The euclidean distance will be large, directly proportional to the difference between the numbers. This metric emphasises the disparity rather than a central point.
- **When the numbers are very similar or equal**: The mean will be very close to or equal to the numbers themselves while the euclidean distance will be small

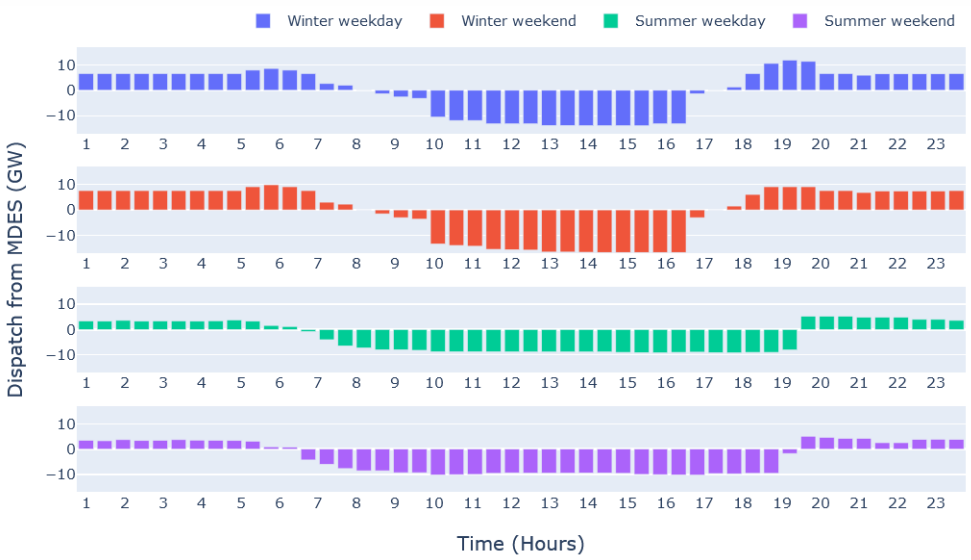

Figure 38: The daily dispatch of the medium-duration energy storage fleet in Leading the way for a winter weekday, weekend and summer weekday and weekend.

(approaching zero as the numbers become identical), indicating little to no variation between the numbers.

As was noted days with similar distances from the average as represented by w_a and s_a are more likely to have coinciding load peaks and stable prices and so it makes sense for the state action value function to be shaped in this way.

The aggregate load exceeds 37 GW approximately 2% of the time, although the aggregate load exceeds 37 GW in both summer and winter. It can be concluded from this first analysis that the BM price may be correlated to increased NDR dispatch, specifically from solar and wind. Pearson's product moment correlation was used in this instance to explore the data. Examining the correlations it is first notable that wind dispatch is more highly correlated with BM settlement price. The inter-relationships between all three variables were complicated. As an example the correlation coefficient relating BM settlement price to wind dispatch, varies with the level of solar dispatch. There is also some variation when including the breakdown between onshore and offshore wind.

Solar generally has a lower correlation coefficient to BM settlement price (other correlation methods have made similar findings for previous years (see Zakeri and Syri (2015) and Denholm et al. (2010))). One reason may be that during the night solar is consistently low while BM settlement price continues to fluctuate to the same degree as during the

day. According to National Grid ESO solar is more likely to be curtailed than wind (as it is higher in the official ranking order (*Operability Strategy Report* (2020))). Solar dispatch being correlated with load may explain why BM settlement price is relatively unaffected by solar, as the BM is designed to account for unexpected fluctuations. Jiang and Powell (2015) have reported similar effects on the US PJM

Having justified the use of an aggregated solar and wind factor, a retrospective look at other variables that could have been used are presented in table 16. All of which can be found with the BMRS data push API. These are consistent with the work presented in chapter 4 in that they are used by National Grid ESO to calculate the settlement price.

Indicator	Description	Correlation coefficient
Net imbalance volume	The volume difference between supply and demand in the settlement period (MW)	0.153
De-rated margin	The volume of generation over forecast demand (MW)	0.298
Forcast outages	Volume predicted to be offline in the settlement period (MW)	0.401

Table 16: Indicators ultimately discarded due to poor correlation to settlement price.

The specifics of the value functions are presented. Again, a discretised approximation of the value function was used with

- **Biomass dispatch**: 10 boxes, 440 - 450 MW
- **BM settlement price**: 7 boxes 0-30, 30-40, 40-50, 50-60, 60-100, 100-150, 150+ £/MWh
- **SoC**: 10 boxes, 0 - 1
- **Aggregated wind and solar factor**: 10 boxes, 0 - 1.5

The BM data was gathered for 1st January 2018 - 31st December 2022, amounting to 87,840 data points. For the 2022 dataset, the 150+ £/MWh box only contains 1.8% of the dataset.

6.4 The future energy scenarios

The objectives for emission reduction as have been discussed previously have been studied extensively by National Grid ESO in *Future Energy Scenarios* (2021). They have published the FES complete with 4 scenarios for expansion of the electricity system to the year 2050. New capacity investments required throughout GB in the long run and their dispatch have been studied across the literature (subject to considerable uncertainty) to model the future grid. National Grid ESO have attempted to address this exact problem with their publication of the FES updated each year. Additionally National Grid ESO published an accounting of their sensitivity analyses carried out as applied to the FES.

In this thesis, the modeling done by National Grid ESO is combined with dispatch data from 2022 to show how MDES will be useful in Leading the way up to 2035. The result is a set of predictions for 2035 that takes the intermittent characteristics of wind, solar and demand into account and will be used in the next sections for developing the PFAs and RLAs.

The main advantage of using the FES in this capacity is the data available on the website by National Grid ESO (*Future Energy Scenarios (FES)* (2024)) providing capacities of renewables, hydropower, oil, gas, biomass and nuclear. Full details of their optimisation to reach these capacities can be found on their website but are based on a series of assumptions concerning the pathway to net zero and consumer behaviour in each scenario.

In Leading the way (the scenario focused on here) there is a strong security of supply and economic growth. Both qualities explain the high levels of NDRs (providing domestic supply and diversity of supply). Leading the way successfully reaches net zero by 2050 through its large NDR fleet and low societal opposition to nuclear. Other key assumptions in Leading the way relative to this thesis are:

- A 20% energy efficiency improvements by 2030 in grid-connected energy generation assets
- Vehicle to grid expansion adding to volume exchanges on the grid
- Little expansion of biomass capacity
- A high capacity of MDES

- Domestic demand side response limiting intra-day peaking relative to other scenarios. The response coming from a mixture of vehicle-2-grid and domestic household appliances

There are concerns that the FES contain too many uncertainties to be useful in this context but their use will be justified in the following section. Here, the uncertainties can be stratified into 2.

- **Stochastic uncertainties**: These uncertainties describe randomness in a dataset, they can be quantified in mathematical models given an understanding of the dataset dynamics. In the context of this thesis they cover weather effects, demand patterns (such as those introduced by electric vehicle charging), and unforeseen power outages that may occur. Given reliance on NDRs in Leading the way it would be reasonable to expect that these types of uncertainties would increase as the decade develops. Climate breakdown and unpredictable developments in weather patterns will exacerbate this. As previously mentioned correctly modelling these uncertainties can be achieved but rely on a correct estimation of the statistical parameters. The methodology of using an RLA handles this form of uncertainty naturally. Despite being built on a deterministic dataset, the RLA is designed to generalise and seek underlying patterns. If the RLA is successful in generalising to incorporate stochastic uncertainties this will be shown in the test results.

- **Social, economic and political uncertainties**: Uncertainties in government or other public policy. These uncertainties can be dismissed in the context of this thesis as 1) assumptions in Leading the way meet the legal obligation across the UK to reach net zero by 2050 and 2) Ofgem and BEIS have concluded that all the assumptions are reasonable (*Transitioning to a net zero energy system: Smart Systems and Flexibility Plan 2021* (2022)).

Having established that the framework of Leading the way is appropriate for use in this thesis, it is requisite to establish the methodology for exploring the specific insights of the FES. The methodology will show how the fleet of ES in Leading the way typically operates. The data from this will then be analysed as it will provide insights into the PFA operation of the CBESS' MDES. The methodology for demonstrating the operation

of MDES is as follows for each 30 minute period (the length of one BM settlement period) starting on 1st January

1. Demand for that half hour was taken from Leading the way
2. Historical wind data for 2022 was extracted and scaled according to the capacity dictated by Leading the way
3. Historical solar data for 2022 was extracted and scaled according to the capacity dictated by Leading the way
4. Capacities were dispatched according to the flowchart shown in figure 39

Finally, figure 39 shows the development of the dispatch decisions in the model of Leading the way. Starting with the demand at time t (d_t) the final supply (s_t) is calculated as a combination of dispatches from the generation types dictated from Leading the way. The demand data at time t is extracted and compared to the current modelled supply (which will be 0 at the start of the time period). A mismatch with supply less than demand will lead to a dispatch schedule of nuclear, wind, solar, biomass, MDES, gas and oil in that order. The dispatch is determined by using a scaling methodology. Scaling factors represent the ratio of capacities at two different points in time. Starting with dispatch data from 2022 and taken from the BMRS data push API. The method is used to determine the dispatch of National Grid ESO, given the FES capacities predicted. The scaling methodology is simply based on the proportionate share in the capacity of each generation type. The scaling factor was found through the division of each generation type's capacity by the total capacity installed. Ultimately, the scaling methodology was chosen because of the reliability of the FES data and the open-source nature, which allows future researchers to to understand their methodology. Additionally, National Grid did not use a linear projection of how technologies and capacities will develop, which avoids one of the main downfalls of relying in this type of scaling methodology and did so with knowledge of plans for future policy landscapes and market dynamics. The FES show interesting predictions concerning scaling factors and the subsequent mismatches.

Table 17 shows the scaling factors used in this thesis for scaling the dispatch from 2022. MDES shows the highest scaling factor and oil the lowest. The value for wind shows a combination calculated by National Grid ESO for onshore and offshore. The implications

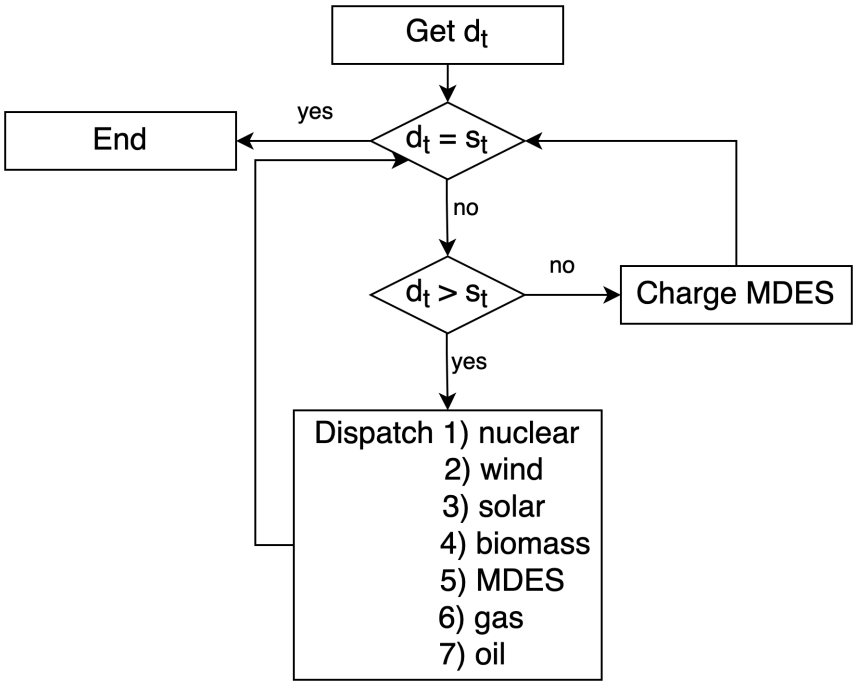

Figure 39: Flow chart for deciding the capacities dispatched.

Genertation type	Nuclear	Wind	Solar	Biomass	MDES	Gas	Oil
Scaling factor	0.6	3.9	3.17	0.72	7.17	0.55	0.47

Table 17: Scaling factors dictated by Leading the way in 2035 compared to 2022

of halving the capacities for gas and oil are significant as they contribute a large percentage to 2022's dispatch pattern.

An initial example is now presented to demonstrate how Leading the way can be explored. It comprises a whole year in order to capture seasonal effects on dispatch. The year is divided into two seasons representing winter and summer with a typical working day and a typical non-working day represented. Altogether there are four typical days. In figure 38, the four representative days were displayed so that the daily profile can be easily discerned with regard to the four seasons and differences between weekdays and weekends. Figure 38 shows the results of the above methodology for the four representative days. It displays the dispatch from the full fleet of MDES in GW in Leading the way as each representative day progresses. The bars delineated throughout the day signify the 48 BM settlement periods.

The general pattern is clear for all four days; dispatching into the grid in the morning,

gradually switching to charging in the day and charging again overnight. The summer days clearly exhibit a broader spectrum of settlement periods when the MDES fleet is charging. In both seasons, the weekend shows increased volume of charging. An interesting feature is the evening spike is discharge; most notable on the winter weekday. This comes from the fact that solar power in the evening hours is dispatching nothing, while demand remains relatively high. The changes in solar are also apparent across the seasons, where the switch to discharge happens later in the day.

In conclusion, this methodology can effectively model the operation of the ES fleet out to 2050.

6.5 Medium-duration storage through 2035

This RLA was based upon the foresight established in flowchart 39. The observation space for $Q_{ES}(S, A)$ was made up of 1) SoC and 2) aggregated wind and solar forecast. The observation space for $Q_{Bio}(S, A)$ was made up of 1) SoC and 2) total biomass dispatch in the previous timestep, with a 2-D action space of biomass dispatch to ES and biomass dispatch to the grid. The discretisations were as follows

- **Biomass dispatch**: 10 boxes, 440 - 450 MW
- **SoC**: 10 boxes, 0 - 1
- **Aggregated wind and solar forecast**: 10 boxes, 0 - 1.5

In both cases the reward functions are carried forward from chapter 5, namely equation 31 and equation 32. Additionally the ϵ-greedy policy is also employed, with the discount factor and learning rate set at 0.99 and 0.1 respectively. All weights in $Q(S, A)$ were initialised to 0. The learning was carried out on ARC3, part of the High Performance Computing facilities at the University of Leeds, UK.

Following the training the results were used to complete the rainflow counting analysis with the Matlab function `rainflow` for counting SoC cycles. The aim of this analysis was to capture in this chapter how operational decisions may have degraded the BESS; effectively a loss of the BESS capacity over its calendar life. High temperatures, high currents and high levels of cycling all affect degradation and rainflow counting address losses due to cycling. Rainflow counting examines the maxima and minima in the SoC

profile. The method employed 10 bins for the histogram set at 10% of the depth of discharge. The bin size was selected as it corresponds to the discretisation used in the observation space of the RLAs.

6.6 Results and analysis

In this section, the results of the preceding sections are presented. First of all, the metadata produced during learning was analysed showing that the most time consuming element of learning came during action selection and the least time consuming step the temporal difference update itself (30.4% and 8.2% respectively).

During learning it can be of some concern that there may be variation between different runs, leading to different learned policies. To establish some level of consistency the training process was repeated 5 times for each RLA. For all RLAs the learning curves were plotted during 5 attempts at learning with the standard parameters, which to reiterate were $\alpha = 0.1$ and $\gamma = 0.99$. The curves had a commensurate shape suggesting that the state-action value functions were developing in a like manner on each run. Importantly the curves began to plateau around the same learning step in the learning process, suggesting stability. The five runs were all tested against the same 500 settlement periods. Subsequently the range in their resultant profits was measured. The calculated average of the these ranges showed the similarity in performance of the five runs. For 2022 this came to 3.4 £/MWh and 5.9 £/MWh in 2035. The results showed a consistency of performance, indicating that the learning would not produce a wholly different state-action value function upon multiple runs and demonstrating reliability.

Here, the operation according to the PFA is analysed. And secondly, it is compared to the RLA, to demonstrate that the RLA can effectively recreate the actions of the PFA.

It was found that the PFA can schedule the CBESS to make profit for the operator. Aggregating days from summer and aggregating days from winter (representing the sea-

Year	Increase in average profit since previous year (%)
2019	0.3
2020	0.7
2021	1.2
2022	1.1

Table 18: Percentage increase in average profits from the previous year.

sons) the battery starts discharging at 6:00 am in summer and 9:00 am in winter and reverts to charging again at 5:00 pm in the winter and 8:00 pm in summer. The pick-up in profit year on year was minimal (see table 18)

6.6.1 Behaviour of the 2022 reinforcement learning agent

The 2022 RLA well approximated the actions of the 2022 PFA under the assumptions discussed in chapter 4 of price taker status.

An interesting aspect of the RLAs to examine is the how the training and testing is divided among the year of data. For example, in the training presented above the training data was run from January 2018 and made up 80% of the dataset. This left the testing data comprising days taken only from a single year. Therefore, training was repeated 5 times, with the 20% testing period taken from a different year each time. Otherwise, another training was performed with days taken at random from throughout the dataset.

The various training regimes had an impact on profit across 500 test settlement periods and on the success rate. Most notable was the relatively low profit garnered when 2020 was chosen to be the testing year, perhaps precipitated by the low demand experienced during that year. The training regime taking random days from the dataset showed the most uniformity and insensitivity to price spikes. Despite this table 19 shows that using 2022 as the training set produced the highest success rate. Therefore, from this point in the thesis the RLA referred to and analysed will be the value function generated from the 2022 testing data.

The results show that the settlement periods most closely matching the PFA are clustered near the lower end of the overall profit spectrum, however this lower end is also

Testing data	Training data	Success rate
2022	2018-2021	0.82
2021	2018-2020, 2022	0.79
2020	2018-2019, 2021-2022	0.69
2019	2018 2020-2022	0.81
2018	2019-2022	0.78
20% of 2018-2022 chosen at random	2018-2022 excluding test data	0.75

Table 19: Success rate of the reinforcement learning agents for alternative combinations of training and testing data

where the biggest spread of results are.

Further insights into the operation of the CBESS can be garnered from the 500 test settlement periods. For example, considering the bottom of the spectrum, increasing the profit per settlement period of the PFA by 50% to the middle of the graph increases the RLA's profit per settlement period by 76%, demonstrating the dramatic rise in the RLA's accuracy when imitating a higher profit settlement period. A similar phenomenon occurs though to a lesser degree when going from 50% to 100%.

It would be reasonable to assume that the days containing the most volatility would be more profitable but in fact the most profitable are those with a consistently high settlement price. The shorter the periods of high settlement prices the lower the RLA's profit.

A closer examination of the details can reveal further insights. Overall, the biomass generator's capacity factor was 92.3%. In cases where it is profitable, ramping up is permissible. Daily averages of settlement price are correlated with the behavior of the biomass generator although it is rare for high average days to occur. A big advantage of the model is that it places a ceiling on the settlement price at which it will charge from the grid, relying instead on charging from the biomass generator.

Now presented is an analysis of how this RLA behaves in the months since the original work, which was focused on the year 2022. Regardless of season (as there were only 3 months to consider), a clear rule emerged which was that when the difference between daily average settlement price and peak settlement price exceeded 20 £/MWh average profit in each settlement period was in the top 50% most profitable settlement periods. The place in the settlement period rankings then increased monotonically with this differential between daily average settlement price and peak settlement price. When this differential between daily average settlement price and peak settlement price exceeds 92 £/MWh the improvements to profit are no longer apparent. This trend was first noted after studying the correlation of the differential to average settlement period profit per day. A clear advantage of this type of analysis is that many predict that the differential described here will increase with further NDR dispatch (Akram et al. (2020), Mostafa et al. (2020)). In 2022, nuclear, coal, oil and gas prevent excessive peakiness. The 2035 RLA should shed more light on this phenomenon. The RLA operating in January - March 2023 showed no significant deviation from the results across the 500 test settlement periods from 2022.

The SoC was unsurprisingly correlated with the time of day. Periods of low SoC

corresponded to late evening 19:30 - 22:00 when the MDES has been largely drained supporting peak late afternoon demand.

6.6.2 Behavior of the 2035 reinforcement learning agent

Meanwhile, the results for the 2035 RLAs examining the MDES arbitrage in 2035 overall have much higher profits on the scale of an extra £1000 - 2000 per settlement period. The Mean Absolute Error (MAE) between the PFA and RLA over the tested settlement periods came to 156.50 £/MWh, 2.45% of the mean profit garnered by the PFA across the test settlement periods. Additionally the spread of settlement periods where the RLA can meet the PFA are less clumped around the bottom of the range relative to the 2022 RLA.

In conclusion, MDES arbitrage has more value as the century develops by some quite significant degree. Despite the cost of ES technologies being subject to considerable uncertainty (Jin and Greaves (2021)) this thesis has shown that investment in an MDES will give immediate returns on the grounds of FFR and returns will in fact grow, as the operator switches to arbitrage, as the ES ages and the electricity system evolves to meet its legal obligations. In tests of the settlement periods presented above the gains made by the 2035 RLA by can be as much as an additional 41.2%.

In analysing the results of this RLA, a distinction should be drawn between two styles of arbitrage; high-value (taking advantage of demand spikes or opportunities to charge rapidly) and low-value (arbitraging over daily fluctuations). Low-value arbitrage has the advantage of being more frequent and also more predictable. Evidently, high-value trades deliver a higher instantaneous reward. Additionally it would be helpful to define days that minimise biomass ramping called low-ramping days. The following section addresses the uncertainties that may have influenced the results.

6.6.3 Addressing uncertainty in conclusions

It should be noted that this chapter has focused exclusively on the only FES capable of reaching net zero by 2050, an ambitious but legally binding goal. It is important to enumerate the few points that may reduce the profitability of the ES out to 2035 should the assumptions fail to hold.

- Having battery storage located by synchronous generators (such as bioenergy) to

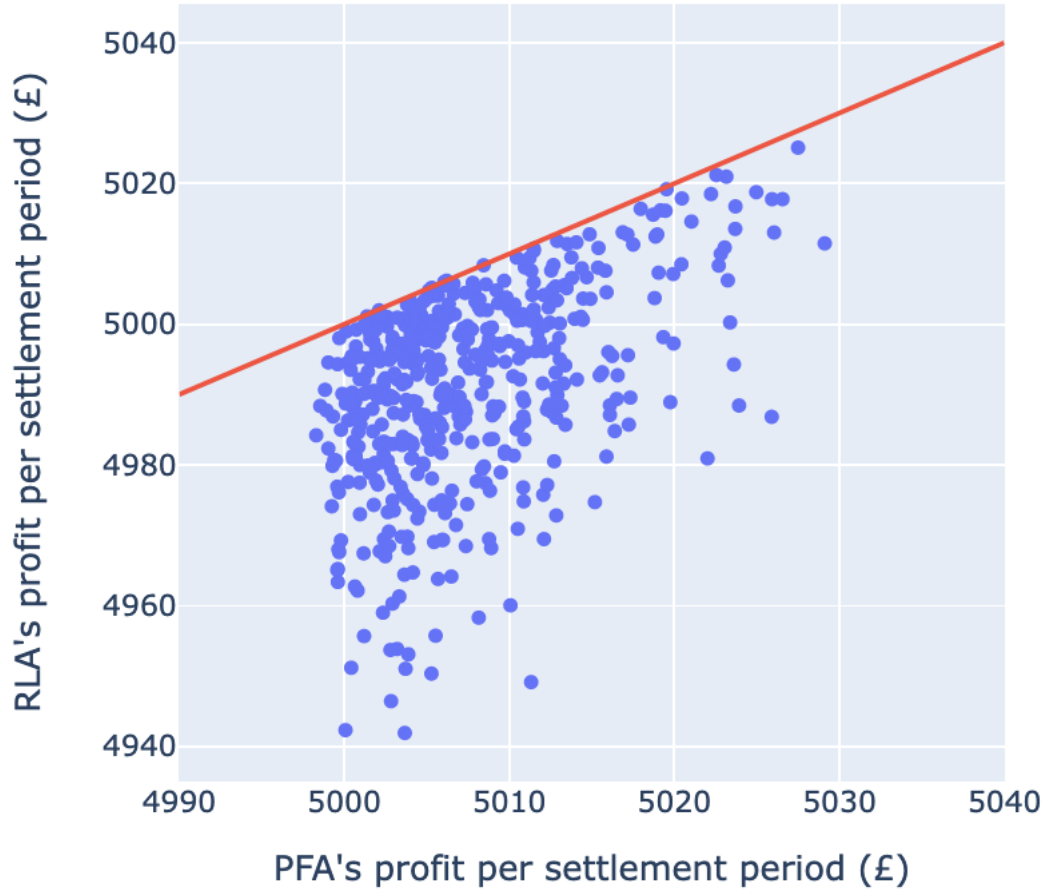

Figure 40: Profit accumulated per settlement period for 500 settlement periods. Red line shows theoretical maximum of the perfect foresight agent's profits.

reduce peaks across the grid may require capital expenditure in strengthening the distribution or transmission grid. Storage located in homes or use of electric vehicles that can manage demand directly at point of use by residents may be found financially preferable if they do not require such an expansion of the transmission grid.

- There is little data in the FES suggesting that there may be location dependant pricing though other sources have suggested this would be a pertinent measure (Schmidt et al. (2019), Zakeri and Syri (2015)). More recently there has been an ongoing Review of Electricity Market Arrangements (REMA) exploring the potential implementation of locational pricing. It raises concerns about potential price disparities between regions Drummond et al. (2022).

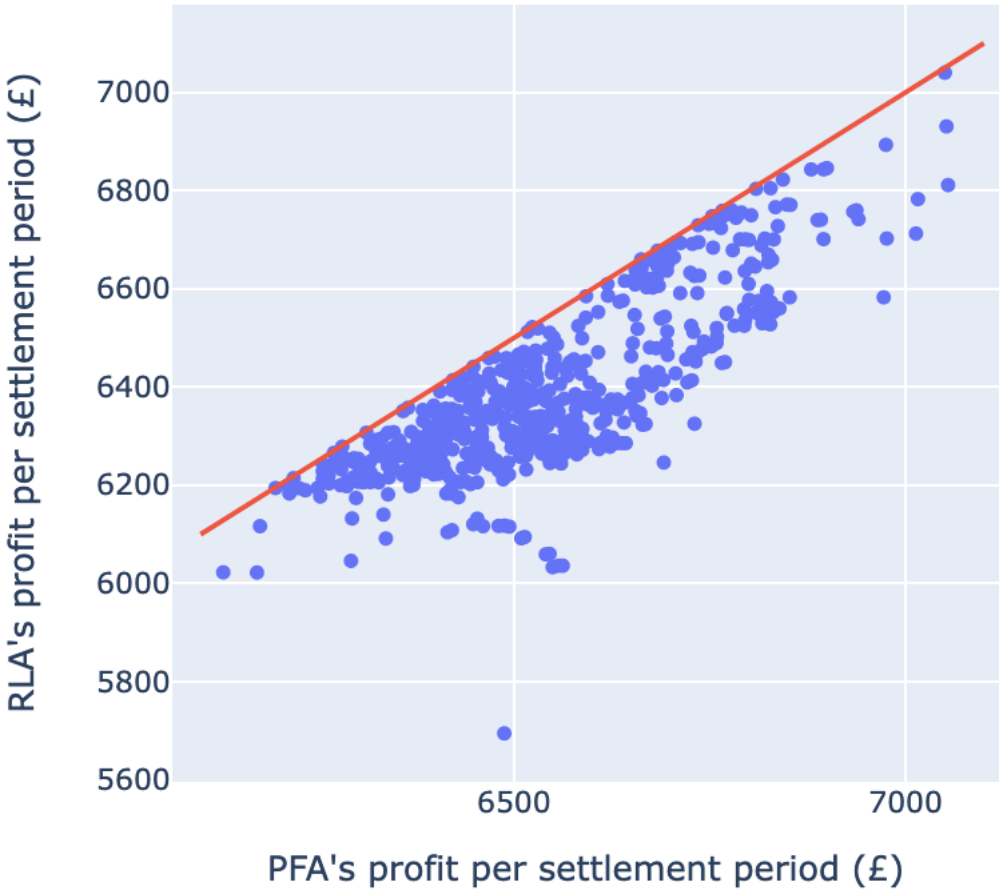

Figure 41: Profit accumulated per settlement period for 500 settlement periods. Red line shows theoretical maximum of the perfect foresight agent's profits.

- Interference by government or some public institution could change the trajectory of technology development and change the capacities predicted in Leading the way.

Other factors may strengthen the case for MDES out to 2035 such as the case that National Grid ESO has outlined plans to lower connection costs for generators willing to add ES to their on-site set-up.

6.6.4 Imbalance pricing

An intriguing aspect of the analysis yet unaddressed is the imbalance pricing offered by National Grid ESO, prices charged when a generator fails to meet its requirements. Imbalance pricing works thusly; after the settlement price has been calculated it will be charged as an imbalance price to any generator that has failed to provide its promised capacity. If a participant has a deficit (consumed more than contracted) or surplus (produced more

than contracted), they are considered imbalanced. Participants with a deficit will need to purchase additional electricity to balance their position, while participants with a surplus will receive compensation for the excess electricity they supplied. This system encourages generators who have not made their quotas to participate in the BM if the imbalance price is higher than the marginal cost of changing their output. In a post analysis imbalance pricing was applied to the RLA results when the limits on the SoC meant the CBESS could not meet its obligations. For example, in one instance the MDES was found at full capacity and had contracted a bid for buying from the grid. In this cases the CBESS was charged for its imbalance volume. In the 500 test settlement periods for the 2022 RLA this happened a total of 12 times, amounting to a penalty of £498.30.

6.6.5 Rainflow counting results

In this section, results for the rainflow counting algorithm are presented for operation in 2022 and 2035 for the testing data. The SoC profile from the RLAs were obtained in full during testing then rainflow counting began producing a count of cycles in the 10 discretised SoC boxes and a histogram of at which SoC the cycles most often take place. The 10 discretised boxes were taken from the SoC dimension of the ES observation space. The results showed an average of 622 cycles in each depth of discharge range in 2022 and 534 in 2035. The resultant histogram (see figure 42) showed that most cycles occurred in the 20 - 30% bin in both years. The results agree with the work of Camuñas García-Miguel et al. (2023) who also showed that a BESS operating for arbitrage, though in the Iberian electricity market, would cycle most with a 25 - 30% SoC (they used a smaller discretisation of SoC).

The amplitude of cycles tended to be greater in 2035 but, as was mentioned, less frequent, suggesting that the CBESS was time shifting energy on longer timescales.

Table 20 shows the capacity degradation incurred in cycling ranges researched by Gbadegesin et al. (2019). The table can be used as a guide to suggest how the capacity of the ESS in the CBESS has lost capacity. This information suggests that on average the battery is operating with 90.8% degradation by the end of the testing periods in both years.

Finally, this analysis suggests an opportunity for further analysis that incorporates the effect of battery degradation through rainflow counting.

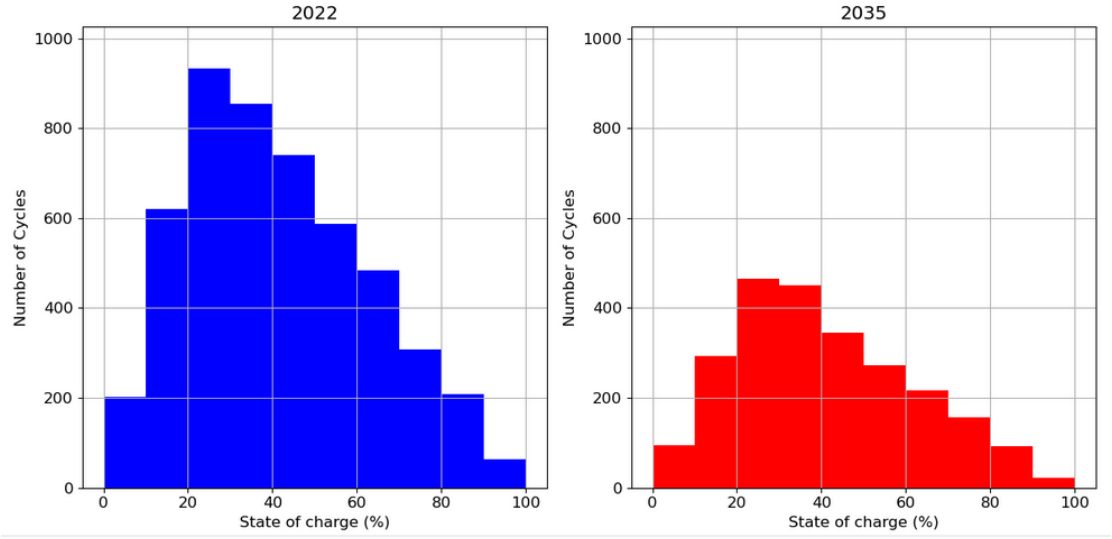

Figure 42: Histogram of rainflow counting in 2022 (right) and 2035 (left).

Number of cycles	Capacity degradation (%)
100	100
101-500	96.3
501-1000	90.8
1001-1500	85.4
1501-2000	80.1
2001-5000	75

Table 20: The degradation of a lithium ion battery taken from Gbadegesin et al. (2019) to demonstrate the effects of increased cycling on the technology

6.7 Levelised cost of storage

This section makes use of the LCOS as first introduced in section 2.2. For clarity the equation for calculating the LCOS is also presented here

$$LCOS = \frac{I + \sum_{n=1}^{N} \frac{\text{Operational cost}}{(1+r)^n} + \sum_{n=1}^{N} \frac{\text{Charging cost}}{(1+r)^n} + \frac{\text{End of life cost}}{(1+r)^{N+1}}}{\sum_{n=1}^{N} \frac{\text{Electricity delivered}}{(1+r)^n}} \tag{40}$$

showing the contributions of initial investment, operational cost, charging cost and end of life cost, r is the discount rate and N is the lifetime of the project. The LCOS in this section will be quoted in £/kW-yr.

Initial investment in equation 40 is quoted in £/kW, which is the unit commonly used throughout the literature (Evans et al. (2012)). While looking through the literature for an appropriate reference, papers were filtered to focus on grid-scale lithium ion batteries. A combination of methodologies was found across the literature including a focus on the

price of the physical technology and papers that abstracted to an analysis of market economics. A few papers included the possibilities of public policies influencing the initial investment costs as in Mauler et al. (2021). Despite the claims by some (Gür (2018)) that the literature underestimates true battery costs, this thesis will use that body of knowledge to set the initial investment costs in equation 40. In part this is motivated by the fact that very little industry expertise is publicly available. After examining the selected papers, ultimately it was decided to take the value from Cao et al. (2020) which was deemed most relevant to this thesis because of its focus on grid-scale lithium ion batteries being used for arbitrage between 2020 and 2025. Furthermore, National Grid ESO has outlined plans to lower connection costs for generators willing to add ES to their on-site set-up which will potentially lower the costs of initial investment.

The operational costs were discussed in several different papers but were ultimately taken from Cole et al. (2016) who projected lithium ion operational costs to 2050, with 6.36 £/kW-yr in 2022 and 3.18 £/kW-yr in 2035. This coincided with the NREL general advice that operational costs for a grid-scale lithium ion battery will come to 0.43% of the capital costs. The charging cost was simply that charged by National Grid ESO when buying. To work out the end of life cost He et al. (2020) was consulted which recommended a 6 £/MWh cost for recycling materials that made up the battery, disassembly and decommissioning (disconnection from the grid transmission system). Finally, the discount rate was set to 10% as it was set at this value for NPV calculations in chapter 5 and this provides a level of consistency. It should be noted that some of the above parameters have only been estimated in this analysis with the best data available.

The aim of this section is to minimise the LCOS of the CBESS, raising the possibility of frequency regulation services to be provided followed by energy arbitrage and maximise revenues. Across the literature publications such as Meng et al. (2020) have shown the possibility of an MDES being effectively used for a single purpose across its lifetime. Evidence of multiple uses such as stacking and sequencing are rare. This thesis considers that a better use of an MDES is to change its use as it ages and nature of the electricity market evolves.

For validation and assessment the LCOS has been calculated for the operating decisions developed for the testing data in the previous sections. The goal is to be able to model the frequency regulation and the energy arbitrage sequencing within this mathe-

matical formulation. The LCOS is a comprehensive metric that accounts for all relevant costs associated with energy storage, including capital expenditure, operational and maintenance costs, and the efficiency of the storage system over its lifetime. This makes it an ideal choice for evaluating the economic viability of battery systems in arbitrage operations. Compared to other metrics, such as simple payback period or net present value, LCOS offers a more nuanced understanding of the cost-effectiveness by incorporating the temporal aspects of energy storage and usage. While NPV provides a snapshot of profitability and payback period indicates the time required to recoup investment costs, neither fully captures the ongoing cost dynamics and operational efficiency of storage systems like LCOS does. Additionally, LCOS facilitates more straightforward comparisons across different technologies and storage durations, making it a superior metric for assessing the long-term economic performance of batteries in the context of energy arbitrage.

To get the LCOS from 2020 - 2035 the value function was applied to the FES in each year and processed through equation 40, as was the data from chapter 5.

In figure 43, the results for FFR and MDES arbitrage are shown for 2020-2035. It is most notable in the figure that the lowest LCOS crosses from functionalities between 2026 and 2027. Given this information, it can be concluded that the CBESS operator should switch their operational strategy after 2026. The reasons for this are many, but come, in part, from the increasing profits for delivering the MDES arbitrage. This result hinges on the fact that the ES purchased will have a lifetime longer than 5 years, which is a reasonable assumption for the lithium ion battery discussed at the start of this chapter. In conclusion, the LCOS came out at 3.09 £/kWh-yr assuming the lithium ion battery can live to 2035. A sensitivity analysis of the results explored the affects of charging costs and operational costs. Both parameters were varied uniformly by $\pm$ 5, 10 and 15%. The year of crossing from FFR to MDES arbitrage remained constant at 2027 for each combination, with the exception of increasing and decreasing operational costs by 15%. In those cases the year dropped to 2026 and rose to 2028 respectively. The changes had little impact on the 2035 LCOS.

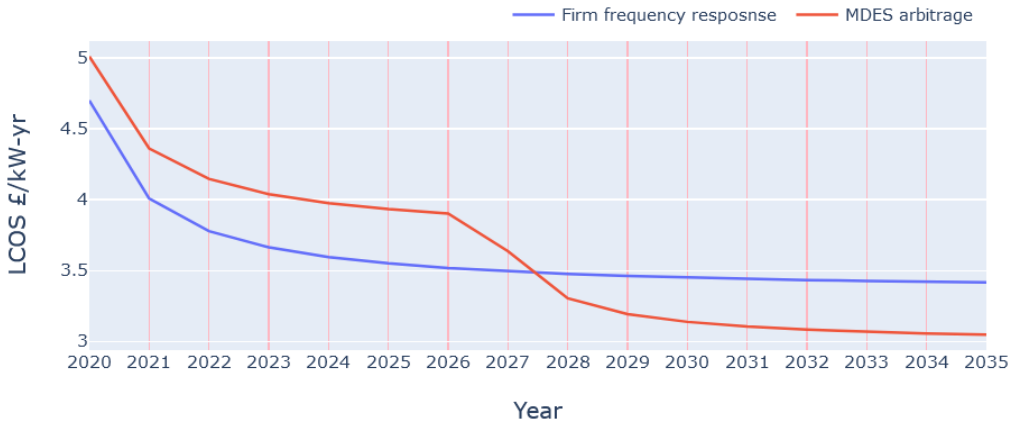

Figure 43: The levelised cost of energy storage of providing Firm Frequency Response compared to medium-duration energy storage arbitrage out to 2035

Although several optimisation models are available in the literature, this proposed extension provides a significant result, in particular by demonstrating that an exact year can be projected for a profitable trade-off between FFR and arbitrage. The results suggest that other plants such as coal and oil would also benefit from an ES installation.

6.8 Energy market reform

The RoC tax credit scheme discussed in the literature review incentivises the use of bioenergy generators. The future of the RoC scheme or any such scheme that may replace it is unclear. Presented here is an analysis of how the such a scheme would apply to the CBESS.

The previous work has delineated arguments both for and against the deployment of a government scheme expanding CfDs. On the one hand, the premise of this thesis hinged on the eventual decommissioning of coal, oil and gas plants. The politically motivated green schemes can potentially achieve this. Papers have been published to refute this argument (Grubb and Newbery (2018)). The aim of the CfD scheme brought on by energy market reform was to invest in load following CCGT including refurbishment and new builds (exact capacities to be decided by the relevant government minister under advice from the energy ombudsman).

There are avenues to making the CBESS profitable, although this depends on the time to installing the ESS. Grubb and Newbery (2018) have claimed that to increase investment, investors must be confident that prices are predictable in the long and short

term with constant government policies. The development of techniques such as these show that predictive models allow operators to change their operation in the middle of changing political and policy landscapes. Claims that it necessarily takes a very stable and predictable environment can be upended as is needed in a century likely to be a characterised by upheaval. Instead operators can use changing projections to constantly re-evaluate the best operational strategy.

The models developed could likely account for carbon pricing, such as the UK Emissions Trading System (ETS). It has already been shown that this methodology can be adjusted as projections for NDRs change.

7 Conclusion

This thesis has addressed the challenges bound to affect the electricity system in the coming decades by suggesting techniques for the operator of a CBESS. In particular, MDES for arbitrage and FFR have been studied. In all cases the benefits of investing an ESS have been verified, as they increase profits from FFR and proof the bioenergy generator against a changing electricity system.

Conclusions from the chapters 3, 4, 5 and 6 will be presented and then the links between them discussed subsequently.

For example chapter 3 established a novel methodology for an ESS making operational decisions without foresight. The chapter started with a description of the linear programming problem and the solver. The cplex solver (version 12.8 developed by IBM) was justified for its speed and efficiency compared to other solvers available open source.

Chapter 3 established a novel methodology for employing temporal difference learning in the context of battery management. The chapter started with a comprehensive introduction to temporal difference learning, detailing its mechanism for learning from the environment by taking statistics about the world around it and updating predictions based on the difference between expected and actual outcomes.

The rationale for using temporal difference learning was its suitability for problems where future states depend on current actions. This methodology was also chosen for its efficiency in processing large datasets and dynamically adjusting to changing conditions.

The chapter delved into the reward function, which it will be shown in subsequent

chapters will be tailored to align with the specific objectives and constraints. This function plays a crucial role in guiding the learning process. The mathematical basis for using the ϵ-greedy policy was also presented in this chapter.

Results from the PFA showed price point dependency, a correlation between dispatch and ES SoC and a tendency for the ESS to switch between rated charging and discharging capacity rapidly. The chief novelty in the chapter arose from using the resultant dispatch decisions as training data for an RLA. The success of mimicking the PFA was vindicated by test runs depicted in the scatter plots in that chapter. At times, such as when there are peaks of price on the BM, the RLA lost predictive power suggesting it would be better suited to a more stable energy market. According to that chapter, the RLA can recover 10-94.3% of the PFA's profit depending on time of day, price peaks in the settlement period, recent frequency events and hyperparameters used to train the RLA. In order to maintain the integrity of this thesis the full range is reported but it should be noted that the range does not reflect the spread of the results but, in fact, the average success rates reported in each chapter show that the 10% result is rare. Furthermore the full range is reported because it raises further research questions that could allow more insight into why this RLA occasionally behaves in such a way that produces a poor comparison to the PFA. This idea is explored further in section 7.1.

At the end of the chapter, it was decided to use the rainflow counting methodology so that conclusions about battery longevity could be drawn in the subsequent chapters.

Chapter 4 explored the flagging, classification and tagging method of arriving on a market price for the BM. National Grid ESO makes publicly available both bids on the BM and the resultant imbalance price, allowing for the modelling of the flagging, classification and tagging process and its subsequent use in examining future bidding patterns. The results of the methodology suggested that the market as it is currently configured can be modelled successfully. The conclusions here, were that adding an RLA bid to the modelled calculation of settlement price only impacted the final price in 7% of cases with a mean average of 2.05 £/MWh. The results confirmed the original hypothesis that the agent developed throughout this c=thesis can be treated as a price taker.

The method as set out in chapter 4 was effectively utilised in chapter 6 to show that good operation of an MDES can limit imbalance positions of the CBESS to 12 out of 500 test settlement periods.

The next chapter established the methodology applied to a CBESS providing FFR. The aim was to articulate in detail the best operational provision of FFR for a CBESS, receiving an availability price (£/hr) and a response energy payment (£/MWh).

The sizing of the ESS was considered carefully to avoid and under or over-sizing. In order to do this an NPV analysis was carried out that compared the profitability of several ESSs of varying energy to power ratios considering they could live up to 20 years. The NPV analysis suggested that the appropriately sized ESS for the 450 MW biomass generator used in that chapter was a 5 MW power and energy capacity 20 MWh battery ESS giving an energy to power ratio of 4 hours. The finding of the NPV analysis bestows an interesting place of an ESS deliberately designed for a biomass generated in the literature when compared to those designed for wind farms. For example those designed to reduce the curtailment of wind have generally been shown to require a longer energy to power ratio (Zhou et al. (2011)). Analysing the results of this thesis, 87% of the BM revenues were captured over 2-3 hours of continuous charge and discharge, perhaps suggesting a smaller capacity ESS would be equally effective, though more considering the downsides of going to extreme levels of SoC, too low or too high. Such a situation could incur penalties if it hindered the ESS from responding correctly to contractual obligations.

In the following chapter, the same CBESS was used in a Leading the way scenario to provide medium-duration storage. This setup was chosen to allow for a direct comparison to the FFR usage and to do so the same sizing and energy to power ratios were needed. Additionally, the 4 hour energy to power ratio fell into National Grid ESO's published categories for what can be considered medium-duration storage. With this setup, the CBESS' main source of revenue came from operating on the BM. The advantages of operating the CBESS in this context are

1. The biomass generator can participate in the BM without prohibitively expensive ramping. This is extremely important for fully decarbonising the grid and replacing the fast acting load-following oil and gas that are absent in Leading the way.

2. The BM price spikes when demand is larger than NDR capacity can be smoothed out, benefiting grid National Grid ESO and encouraging them to support new connections of ES. This can come in the way of additional subsidies and tax credits for a biomass generator willing to invest in ES and lowered permitting standards. These could

come at a time when government policy may be more hostile to biomass, depending on ultimate decisions about its green credentials.

3. The revenue of the whole combined system was increased by the edition of the ESS, represented by profits per settlement period, in the range of 8% in chapter 5, 5% for the MDES in 2022 and 10% for the MDES in 2035. Such variations represent only averages of the profits per settlement period. The results confirm the work of Xie et al. (2018) and Kintner-Meyer et al. (2012) which tracked a similar MDES to the 2050s but contrary to the results of Drury et al. (2011) which explored an MDES operating in the US who discovered a much greater improvement reportedly linked to higher volatility in that market. Drury et al. (2011) deliberately compared the market to the UK. The increased volatility predicted does explain the variation in results quoted in that paper rather than a significant disparity in the results of both analyses.

Both chapters 5 and 6 optimised SoC restoration, and revenue, having an average SoC in the 0.4-0.5 range. The main work to restore the nominal charge occurred at times correlated with frequency trips, utilising charge and discharge capacity to accelerate the SoC restoration. In chapters 5 and 6 the RLA achieved a success rate of 0.87 and 0.82 respectively. In chapter 5, the RLA struggled most in specific EFA blocks, achieving either a high or low success rate. Meanwhile in chapter 6, the RLA was consistent across settlement periods.

The CBESS can be used for FFR provision in an electricity grid with a low and developing NDR fleet and for medium-duration provision as NDRs proliferate. As was shown the CBESS could continue to provide FFR in the high NDR Leading the way scenario, but it would then be more profitable to switch to medium-duration storage. Following Leading the way's development, the selection of when to tender for FFR and when to supply medium-duration storage is largely dependant on how the frequency market may develop. For example, if the system operator cannot guarantee a sufficiently high availability price across every EFA block the CBESS should switch to working with medium-duration storage. The methodology has good adaptability and can be repurposed for other power systems and explains the discrepancies in recommendations for the appropriate operational strategies in chapters 5 and 6. The adoption of Leading the way in chapter 6 compli-

cates the results by making them more speculative as opposed to the historical data that formed the basis of chapter 5. Future work should focus on repurposing this methodology as National Grid ESO continues to adjust its projections.

In conclusion and reflecting on the original research objectives set at the beginning of this thesis. It can be delineated to what extent these objectives have been achieved. Objectives **O1** and **O2** were achieved in chapters 3 and 4 respectively. The results of objective **O3** have been discussed in this chapter previously with the metric of success rate being used to indicate how well the objective was met. The objective called for the ability to maximise profits so in this sense the objective was not met but was met to the extent that it could be when including the uncertainty of future events. The success rate fulfilled the objective set out in **O4**.

An examination of **O5** showed that at least two helpful case studies were identified detailed in chapters 5 and 6. Both case studies fulfilled the objective as they focus on a CBESS in two separate markets. The only failure to fulfill the objective came from the mandate to examine a variety of technologies whereas it was found ultimately more useful to follow the same technological setup through both chapters. Using the same 450 MW biomass generator and 5 MW / 20 MWh ESS allowed for the analysis in chapter 6 looking at when to change operational strategies.

It is worth commenting on how the results of this thesis met and did not meet initial expectations. The initial expectations were that the ES would help maintain the biomass generator particularly through times of electricity market reform. The motivation behind the thesis hung off the prediction that the biomass generator would benefit from this installation. The actual results confirmed the initial hypothesis, first of all with the 2019 FFR provision and secondly with the MDES revenue out to 2035. In the first example, the ESS was found to increase profits in 91% of the tested settlement periods and reduce ramping in 73% of the settlement periods (compared to the biomass generator providing FFR by itself). Chapter 6 also confirmed the hypothesis in the long run, with biomass ramping reduced. One point, unexpected at the start of chapter 6, was that there was still significant ramping in the winter months of 2023 as solar is subject to seasonal constraints. This point provides a need for further research with respect to ensuring a seasonal storage of biomass to burn as ramping requires more fuel per MWh. The consideration of fuel supplies is one limitation of the research presented throughout this thesis.

7.1 Future work

There are broad implications of this research, especially in the following areas which will then be expounded upon:

- **Relevance to the field**: The methodology fills a gap in the literature as the perfect foresight was established, addressing the specific points raised by Raineri et al. (2006), Tan et al. (2021) and Mitkovska-Trendova et al. (2014).
- **Practical applications**: The thesis has shown practical applications as the methodology can be employed by operators to manage their assets. Furthermore, it can inform political bodies designing policies to send clear price signals to investors.
- **Impact on future research**: The methodology has opened up the potential for research in a series of electricity markets.

Addressing item 2, an addition to aiding the practicalities of this work would be to include in the LPP complexities (such as the non-linearity in self-discharge, round-trip efficiency and optimal cycling depth for a BESS). This, for example, would greatly affect chapters 5 and 6. In this case the PFA teaching the RLA would have to take some form other an LPP. Further technical considerations would affect the practicalities; the choice of storage method, transmission connection costs, the possibility of several smaller ESSs with a much higher total capacity working in tandem.

In chapter 6 the power output from 2022 was scaled for the capacities in 2035. For one, this assumption implies uniform wind patterns across the country or that any new capacity experienced the same weather events as the capacity already installed to 2022. This reflects a distortion of the practical implications of additional capacity; the same applying to additional capacity of solar power. Exacerbating the problem in terms of wind, the capacities retain the onshore/offshore ratio of wind, when offshore wind is expected to expand more rapidly (*Future Energy Scenarios* (2022)). Future research could include projections of possible future builds and incorporate this into this model.

Finally, repeating a single year's training data can be detrimental to the performance of an RLA due to the inherent variability in weather patterns. Training an agent on data from just one year fails to capture this variability, leading to an overfitting problem where the agent performs well under the specific conditions of the training year but struggles

to adapt to different or extreme weather patterns in other years (Collins et al. (2018)). Therefore, incorporating multiple years of training data is crucial for developing a RLA which can be addressed in the future.

References

Aburub, H., Basnet, S. and Jewell, W. T. (2019), 'On the use of adjustable-speed pumped hydro storage operation in the U.S. electricity market', *Journal of Energy Storage* **23**, 495–503.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X19300234*

Ackermann, T., Prevost, T., Vittal, V., Roscoe, A. J., Matevosyan, J. and Miller, N. (2017), 'Paving the way: A future without inertia is closer than you think', *IEEE Power and Energy Magazine* **15**(6), 61–69.

Adetokun, B. B., Oghorada, O. and Abubakar, S. J. (2022), 'Superconducting magnetic energy storage systems: Prospects and challenges for renewable energy applications', *Journal of Energy Storage* **55**, 105663.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X22016516*

Akram, U., Nadarajah, M., Shah, R. and Milano, F. (2020), 'A review on rapid responsive energy storage technologies for frequency regulation in modern power systems', *Renewable and Sustainable Energy Reviews* **120**, 109626.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032119308330*

Amzallag, C., Gerey, J., Robert, J. and Bahuaud, J. (1994), 'Standardization of the rainflow counting method for fatigue analysis', *International Journal of Fatigue* **16**(4), 287–293.
URL: *https://www.sciencedirect.com/science/article/pii/0142112394903433*

Andersson, J. and Grönkvist, S. (2019), 'Large-scale storage of hydrogen', *International Journal of Hydrogen Energy* **44**(23), 11901–11919.
URL: *https://www.sciencedirect.com/science/article/pii/S0360319919310195*

Argyrou, M. C., Christodoulides, P. and Kalogirou, S. A. (2018), 'Energy storage for electricity generation and related processes: Technologies appraisal and grid scale ap-

plications', *Renewable and Sustainable Energy Reviews* **94**, 804–821.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032118304817*

Arnold, M. and Andersson, G. (2011), Model predictive control of energy storage including uncertain forecasts, *in* '17^{th} Power Systems Computation Conference (PSCC 2011)', Stockholm, Sweeden.

Augustine, C. and Blair, N. (2021), 'Storage futures study: Storage technology modeling input data report'.
URL: *https://www.osti.gov/biblio/1785959*

Azad, M. L., Khursheed, A. and Kumar, V. (2015), Mitigating power oscillations in wind power plants using ESS, *in* '2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)', pp. 67–72.

Baumann, M., Domnik, T., Haase, M., Wulf, C., Emmerich, P., Rösch, C., Zapp, P., Naegler, T. and Weil, M. (2021), 'Comparative patent analysis for the identification of global research trends for the case of battery storage, hydrogen and bioenergy', *Technological Forecasting and Social Change* **165**, 120505.
URL: *https://www.sciencedirect.com/science/article/pii/S0040162520313317*

Bazdar, E., Sameti, M., Nasiri, F. and Haghighat, F. (2022), 'Compressed air energy storage in integrated energy systems: A review', *Renewable and Sustainable Energy Reviews* **167**, 112701.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032122005901*

Biancardi, A., Di Castelnuovo, M. and Staffell, I. (2021), 'A framework to evaluate how European Transmission System Operators approach innovation', *Energy Policy* **158**, 112555.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421521004250*

Biggins, F., Homan, S., Ejeh, J. and Brown, S. (2022), 'To trade or not to trade: Simultaneously optimising battery storage for arbitrage and ancillary services', *Journal of Energy Storage* **50**, 104234.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X22002638*

Biggins, F., Homan, S., Roberts, D. and Brown, S. (2021), 'Exploring the economics of large scale lithium ion and lead acid batteries performing frequency response', *Energy Reports* **7**, 34–41. 5th Annual CDT Conference in Energy Storage and Its Applications.
URL: *https://www.sciencedirect.com/science/article/pii/S2352484721001566*

BMRS API and Data Push User Guide (2019), Technical report, National Grid.

Botha, C. and Kamper, M. (2019), 'Capability study of dry gravity energy storage', *Journal of Energy Storage* **23**, 159–174.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X18308211*

Bracale, A., Caramia, P., Carpinelli, G., Mancini, E. and Mottola, F. (2015), 'Optimal control strategy of a DC micro grid', *International Journal of Electrical Power & Energy Systems* **67**, 25–38.
URL: *http://www.sciencedirect.com/science/article/pii/S0142061514006851*

Brandon, N., Edge, J., Aunedi, M., Bruce, P., Chakrabarti, B., Esterle, T., Somerville, J., Ding, Y., Fu, C., Grant, P., Hall, P., Huang, C., Leng, G., Li, Y., Martins, V., Navarro, M., Gil Posada, J., Rennie, A., Rogers, D., Strbac, G., Perez Villar, S., Yufit, V., Wang, J. and Worsley, D. (2016), Uk research needs in grid scale energy storage technologies, Technical report.

Bruce, A. and Ruff, L. (2018), Carbon intensity forecast methodology, Technical report, National Grid.

Bunn, D. and Yusupov, T. (2015), 'The progressive inefficiency of replacing renewable obligation certificates with contracts-for-differences in the UK electricity market', *Energy Policy* **82**, 298–309.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421515000038*

Camuñas García-Miguel, P. L., Alonso-Martinez, J., Arnaltes Gómez, S. and Rodríguez-Amenedo, J. L. (2023), 'Impact of risk measures and degradation cost on the optimal arbitrage schedule for battery energy storage systems'.
URL: *http://dx.doi.org/10.2139/ssrn.4356270*

Cao, J., Harrold, D., Fan, Z., Morstyn, T., Healey, D. and Li, K. (2020), 'Deep rein-

forcement learning based energy storage arbitrage with accurate lithium-ion battery degradation model', *IEEE Transactions on Smart Grid* pp. 1–1.

Chatenet, M., Pollet, B. G., Dekel, D. R., Dionigi, F., Deseure, J., Millet, P., Braatz, R. D., Bazant, M. Z., Eikerling, M., Staffell, I., Balcombe, P., Shao-Horn, Y. and Schäfer, H. (2022), 'Water electrolysis: from textbook knowledge to the latest scientific strategies and industrial developments', *Chemical Society Reviews* **51**, 4583–4762.

Cho, J. and Kleit, A. N. (2015), 'Energy storage systems in energy and ancillary markets: A backwards induction approach', *Applied Energy* **147**, 176–183.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261915001506*

Cole, W. J., Marcy, C., Krishnan, V. K. and Margolis, R. (2016), Utility-scale lithium-ion storage cost projections for use in capacity expansion models, *in* '2016 North American Power Symposium (NAPS)', pp. 1–6.

Collins, S., Deane, P., Gallachóir, B. , Pfenninger, S. and Staffell, I. (2018), 'Impacts of inter-annual wind and solar variations on the european power system', *Joule* **2**(10), 2076–2090. Open Access.
URL: *https://doi.org/10.1016/j.joule.2018.06.020*

Cost Projections for Utility-Scale Battery Storage: 2020 Update (2020), Technical report, National Renewable Energy Laboratory.

Cross, S., Welfle, A. J., Thornley, P., Syri, S. and Mikaelsson, M. (2021), 'Bioenergy development in the UK & Nordic countries: A comparison of effectiveness of support policies for sustainable development of the bioenergy sector', *Biomass and Bioenergy* **144**, 105887.
URL: *https://www.sciencedirect.com/science/article/pii/S0961953420304219*

Crude oil and natural gas database from Trading Economics (2022).
URL: *https://tradingeconomics.com/commodity/crude-oil*

Denholm, P., Cole, W. and Blair, N. (2023), 'Moving beyond 4-hour li-ion batteries: Challenges and opportunities for long(er)-duration energy storage'.
URL: *https://www.osti.gov/biblio/2000002*

Denholm, P., Ela, E., Kirby, B. and Milligan, M. (2010), Role of energy storage with renewable electricity generation, Technical report, United States.
URL: *https://www.osti.gov/biblio/972169*

Do, H. X., Nepal, R. and Jamasb, T. (2020), 'Electricity market integration, decarbonisation and security of supply: Dynamic volatility connectedness in the irish and great britain markets', *Energy Economics* **92**, 104947.
URL: *https://www.sciencedirect.com/science/article/pii/S0140988320302875*

Drummond, P., Grubb, M., Price, J., Gajardo, S. M. and Page, K. (2022), 'Review of electricity market arrangements consultation', *The UCL Institute for Sustainable Resources* .

Drury, E., Denholm, P. and Sioshansi, R. (2011), 'The value of compressed air energy storage in energy and reserve markets', *Energy* **36**(8), 4959–4973. PRES 2010.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544211003665*

Duggal, I. and Venkatesh, B. (2015), 'Short-term scheduling of thermal generators and battery storage with depth of discharge-based cost model', *IEEE Transactions on Power Systems* **30**(4), 2110–2118.

Durante, J. L., Nascimento, J. and Powell, W. B. (2017), 'Backward approximate dynamic programming with hidden semi-markov stochastic models in energy storage optimization'.

End of coal power to be brought forward in drive towards net zero (2020), Technical report, Department for Business, Energy & Industrial Strategy.

Enhanced Frequency Response: Frequently asked questions (2016), Technical report, National Grid.

Enhanced Frequency Response Market Information Report (2018), Technical report, National Grid.

Eser, P., Singh, A., Chokani, N. and Abhari, R. S. (2016), 'Effect of increased renewables generation on operation of thermal power plants', *Applied Energy* **164**, 723–732.
URL: *https://www.sciencedirect.com/science/article/pii/S0306261915015901*

Evans, A., Strezov, V. and Evans, T. J. (2012), 'Assessment of utility energy storage options for increased renewable energy penetration', *Renewable and Sustainable Energy Reviews* **16**(6), 4141–4147.
URL: *http://www.sciencedirect.com/science/article/pii/S1364032112002316*

Facchini, A., Rubino, A., Caldarelli, G. and Liddo, G. D. (2019), 'Changes to Gate Closure and its impact on wholesale electricity prices: The case of the UK', *Energy Policy* **125**, 110–121.
URL: *http://www.sciencedirect.com/science/article/pii/S0301421518306992*

Fan, A., Huang, L., Lin, S., Chen, N., Zhu, L. and Wang, X. (2018), Performance comparison between Renewable Obligation and Feed-in Tariff with Contract for Difference in UK, *in* '2018 China International Conference on Electricity Distribution (CICED)', pp. 2761–2765.

Firm Frequency Response Interactive Guidance (2017), Technical report, National Grid.

Frequency Response Market Information Report: Monthly Report January 2022 (2022), Technical report, National Grid.

Fuller, W. A. (1976), *Introduction to Statistical Time Series*, 2 edn, John Wiley & Sons inc.

Future Energy Scenarios (2021), Technical report, National Grid.

Future Energy Scenarios (2022), Technical report, National Grid.

Future Energy Scenarios (FES) (2024).
URL: *https://www.nationalgrideso.com/future-energy/future-energy-scenarios*

Gbadegesin, A. O., Sun, Y. and Nwulu, N. I. (2019), 'Techno-economic analysis of storage degradation effect on levelised cost of hybrid energy storage systems', *Sustainable Energy Technologies and Assessments* **36**, 100536.
URL: *https://www.sciencedirect.com/science/article/pii/S2213138819301626*

Green, A. (2024), 'ADGEfficiency / energy-py-linear'. Accessed: 2024-05-21.
URL: *https://github.com/ADGEfficiency/energy-py-linear?tab=readme-ov-file*

Grubb, M. and Newbery, D. (2018), 'UK Electricity Market Reform and the energy transition: Emerging lessons', *The Energy Journal* **39**.
URL: *https://doi.org/10.5547/01956574.39.6.mgru*

GSR023: Clarification of the applicability of the N-1-1 criterion (2018), Technical report, National Grid.

Guerra, O. J., Zhang, J., Eichman, J., Denholm, P., Kurtz, J. and Hodge, B.-M. (2020), 'The value of seasonal energy storage technologies for the integration of wind and solar power', *Energy Environ. Sci.* **13**, 1909–1922.
URL: *http://dx.doi.org/10.1039/D0EE00771D*

Guidance for generators: Co-location of electricity storage facilities with renewable generation supported under the Renewables Obligation or Feed-in Tariff schemes (2020), Technical report, Ofgem.

Gür, T. M. (2018), 'Review of electrical energy storage technologies, materials and systems: challenges and prospects for large-scale grid storage', *Energy & Environmental Science* **11**(10), 2696–2767.
URL: *https://doi.org/10.1039/c8ee01419a*

Gutierrez-Alcaraz, G. and Sheble, G. B. (2004), Electricity market price dynamics: Markov process analysis, *in* '2004 International Conference on Probabilistic Methods Applied to Power Systems', pp. 14–19.

Haji Bashi, M., Yousefi, G., Gharagozloo, H., Khazraj, H., Leth Bak, C. and Fariada Silva, F. (2018), A comparative study on the bidding behaviour of pay as bid and uniform price electricity market players, *in* '2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I CPS Europe)', pp. 1–6.

He, G., Ciez, R., Moutis, P., Kar, S. and Whitacre, J. F. (2020), 'The economic end of life of electrochemical energy storage', *Applied Energy* **273**, 115151.
URL: *https://www.sciencedirect.com/science/article/pii/S0306261920306632*

Ho, W. S., Hashim, H. and Muis, Z. A. (2012), Integrated biomass power plant and storage for peak load management, *in* I. A. Karimi and R. Srinivasan, eds, '11th International

Symposium on Process Systems Engineering', Vol. 31 of *Computer Aided Chemical Engineering*, Elsevier, pp. 1000–1004.
URL: *http://www.sciencedirect.com/science/article/pii/B9780444595065500316*

Hogan, W. W. (2022), 'Electricity market design and zero-marginal cost generation', *Current Sustainable/Renewable Energy Reports* **9**(1), 15–26.
URL: *https://doi.org/10.1007/s40518-021-00200-9*

Hossain, E., Faruque, H. M. R., Sunny, M. S. H., Mohammad, N. and Nawar, N. (2020), 'A comprehensive review on energy storage systems: Types, comparison, current scenario, applications, barriers, and potential solutions, policies, and future prospects', *Energies* **13**(14).
URL: *https://www.mdpi.com/1996-1073/13/14/3651*

Huang, Q., Xu, Y., Wang, T. and Courcoubetis, C. A. (2018), 'Market mechanisms for cooperative operation of price-maker energy storage in a power network', *IEEE Transactions on Power Systems* **33**(3), 3013–3028.

Imbalance Pricing Guidance (2019), Technical report, National Grid.

Jiang, D. R. and Powell, W. B. (2015), 'Optimal hour-ahead bidding in the real-time electricity market with battery storage using approximate dynamic programming', *INFORMS Journal on Computing* **27**(3), 525–543.
URL: *https://doi.org/10.1287/ijoc.2015.0640*

Jin, S. and Greaves, D. (2021), 'Wave energy in the UK: Status review and future perspectives', *Renewable and Sustainable Energy Reviews* **143**, 110932.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032121002240*

Junyent-Ferr, A., Pipelzadeh, Y. and Green, T. C. (2015), 'Blending HVDC-link energy storage and offshore wind turbine inertia for fast frequency response', *IEEE Transactions on Sustainable Energy* **6**(3), 1059–1066.

Jülch, V. (2016), 'Comparison of electricity storage options using levelized cost of storage (LCOS) method', *Applied Energy* **183**, 1594–1606.
URL: *https://www.sciencedirect.com/science/article/pii/S0306261916312740*

Khosravi, M., Afsharnia, S. and Farhangi, S. (2021), 'Optimal sizing and technology selection of hybrid energy storage system with novel dispatching power for wind power integration', *International Journal of Electrical Power & Energy Systems* **127**, 106660.
URL: *https://www.sciencedirect.com/science/article/pii/S0142061520342058*

Kintner-Meyer, M. C., Balducci, P. J., Colella, W. G., Elizondo, M. A., Jin, C., Nguyen, T. B., Viswanathan, V. V. and Zhang, Y. (2012), National assessment of energy storage for grid balancing and arbitrage: phase 1, WECC, Technical report, United States.
URL: *https://www.osti.gov/biblio/1131386*

Kirli, D., Parzen, M. and Kiprakis, A. (2021), 'Impact of the COVID-19 Lockdown on the Electricity System of Great Britain: A Study on Energy Demand, Generation, Pricing and Grid Stability', *Energies* **14**(3).
URL: *https://www.mdpi.com/1996-1073/14/3/635*

Klæboe, G., Eriksrud, A. L. and Fleten, S.-E. (2015), 'Benchmarking time series based forecasting models for electricity balancing market prices', *Energy Systems* **6**(1), 43–61.
URL: *https://doi.org/10.1007/s12667-013-0103-3*

Klumpp, F. (2016), 'Comparison of pumped hydro, hydrogen storage and compressed air energy storage for integrating high shares of renewable energies—potential, cost-comparison and ranking', *Journal of Energy Storage* **8**, 119–128.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X16301712*

Lai, C. S. and Locatelli, G. (2021), 'Valuing the option to prototype: A case study with Generation Integrated Energy Storage', *Energy* **217**, 119290.
URL: *https://www.sciencedirect.com/science/article/pii/S0360544220323975*

Lian, B., Yu, D., Wang, C., Le Blond, S. and Dunn, R. W. (2014), Investigation of energy storage and open cycle gas turbine for load frequency regulation, *in* '2014 49th International Universities Power Engineering Conference (UPEC)', pp. 1–6.

Löhndorf, N. and Minner, S. (2010), 'Optimal day-ahead trading and storage of renewable energies—an approximate dynamic programming approach', *Energy Systems* **1**(1), 61–77.
URL: *https://doi.org/10.1007/s12667-009-0007-4*

Luerssen, C., Verbois, H., Gandhi, O., Reindl, T., Sekhar, C. and Cheong, D. (2021), 'Global sensitivity and uncertainty analysis of the levelised cost of storage (LCOS) for solar-PV-powered cooling', *Applied Energy* **286**, 116533.
URL: *https://www.sciencedirect.com/science/article/pii/S0306261921000854*

Luo, B., Yang, Y., Wu, H. and Huang, T. (2019), 'Balancing value iteration and policy iteration for discrete-time control', *IEEE Transactions on Systems, Man, and Cybernetics: Systems* pp. 1–11.

Madahi, S. S. K., Gokhale, G., Verwee, M.-S., Claessens, B. and Develder, C. (2024), 'Control policy correction framework for reinforcement learning-based energy arbitrage strategies'.

Maldet, M., Revheim, F. H., Schwabeneder, D., Lettner, G., del Granado, P. C., Saif, A., Löschenbrand, M. and Khadem, S. (2022), 'Trends in local electricity market design: Regulatory barriers and the role of grid tariffs', *Journal of Cleaner Production* **358**, 131805.
URL: *https://www.sciencedirect.com/science/article/pii/S0959652622014159*

Martins, J. and Miles, J. (2021), 'A techno-economic assessment of battery business models in the uk electricity market', *Energy Policy* **148**, 111938.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421520306492*

Mauler, L., Duffner, F., Zeier, W. G. and Leker, J. (2021), 'Battery cost forecasting: a review of methods and results with an outlook to 2050', *Energy & Environmental Science* **14**(9), 4712–4739.
URL: *https://doi.org/10.1039/d1ee01530c*

McGlade, C., Pye, S., Ekins, P., Bradshaw, M. and Watson, J. (2018), 'The future role of natural gas in the UK: A bridge to nowhere?', *Energy Policy* **113**, 454–465.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421517307693*

McKone, R. P. P. and Wolfs, P. J. (2019), Analysis of historical spot prices for the australian national energy market, growing concerns about negative prices postulations on energy storage, *in* '2019 29th Australasian Universities Power Engineering Conference (AUPEC)', pp. 1–6.

Meng, L., Zafar, J., Khadem, S. K., Collinson, A., Murchie, K. C., Coffele, F. and Burt, G. M. (2020), *IEEE Transactions on Smart Grid* **11**(2), 1566–1581.
URL: *https://doi.org/10.1109/TSG.2019.2940173*

Ming, Z., Kun, Z. and Daoxin, L. (2013), 'Overall review of pumped-hydro energy storage in China: Status quo, operation mechanism and policy barriers', *Renewable and Sustainable Energy Reviews* **17**, 35–43.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032112003589*

Mirzapour, O., Rui, X. and Sahraei-Ardakani, M. (2024), 'Grid-enhancing technologies: Progress, challenges, and future research directions', *Electric Power Systems Research* **230**, 110304.
URL: *https://www.sciencedirect.com/science/article/pii/S0378779624001925*

Mitkovska-Trendova, K., Minovski, R. and Boshkovski, D. (2014), 'Methodology for transition probabilities determination in a Markov decision processes model for quality-accuracy management', *Journal of Engineering Management and Competitiveness* **4**, 59–67.

Mostafa, M. H., Abdel Aleem, S. H., Ali, S. G., Ali, Z. M. and Abdelaziz, A. Y. (2020), 'Techno-economic assessment of energy storage systems using annualized life cycle cost of storage (LCCOS) and levelized cost of energy (LCOE) metrics', *Journal of Energy Storage* **29**, 101345.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X19316925*

National Grid ESO (2022), 'Balancing services and volumes reports', `https://www.nationalgrideso.com/industry-information/balancing-services/demand-side-response-dsr`.

National Renewable Energy Laboratory (2005), A method and case study for estimating the ramping capability of a control area or balancing authority and implications for moderate or high wind penetration, Technical report, National Renewable Energy Laboratory.

Oh, E. and Wang, H. (2020), 'Reinforcement-learning-based energy storage system oper-

ation strategies to manage wind power forecast uncertainty', *IEEE Access* **8**, 20965–20976.

Operability Strategy Report (2020), Technical report, National Grid.

Oudalov, A., Chartouni, D. and Ohler, C. (2007), 'Optimizing a battery energy storage system for primary frequency control', *IEEE Transactions on Power Systems* **22**(3), 1259–1266.

Pandžić, K., Pandžić, H. and Kuzle, I. (2018), 'Coordination of regulated and merchant energy storage investments', *IEEE Transactions on Sustainable Energy* **9**(3), 1244–1254.

Park, M., Ryu, J., Wang, W. and Cho, J. (2016), 'Material design and engineering of next-generation flow-battery technologies', *Nature Reviews Materials* **2**, 16080.
URL: *https://doi.org/10.1038/natrevmats.2016.80*

Pereira, S., Ferreira, P. and Vaz, A. (2014), 'Short-term electricity planning with increase wind capacity', *Energy* **69**, 12–22.
URL: *https://www.sciencedirect.com/science/article/pii/S0360544214000450*

Pfenninger, S. and Staffell, I. (2016), 'Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data', *Energy* **114**, 1251–1265.
URL: *https://www.sciencedirect.com/science/article/pii/S0360544216311744*

Pierrehumbert, R. (2022), 'Plant power: Burning biomass instead of coal can help fight climate change—but only if done right', *Bulletin of the Atomic Scientists* **78**(3), 125–127.

Pimm, A. J., Palczewski, J., Morris, R., Cockerill, T. T. and Taylor, P. G. (2020), 'Community energy storage: A case study in the UK using a linear programming method', *Energy Conversion and Management* **205**, 112388.
URL: *http://www.sciencedirect.com/science/article/pii/S0196890419313950*

Pupo-Roncallo, O., Campillo, J., Ingham, D., Ma, L. and Pourkashanian, M. (2021), 'The role of energy storage and cross-border interconnections for increasing the flexibility of future power systems: The case of Colombia', *Smart Energy* **2**, 100016.
URL: *https://www.sciencedirect.com/science/article/pii/S2666955221000162*

Pérez-Díaz, J. I., Chazarra, M., García-González, J., Cavazzini, G. and Stoppato, A. (2015), 'Trends and challenges in the operation of pumped-storage hydropower plants', *Renewable and Sustainable Energy Reviews* **44**, 767–784.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032115000398*

Rahman, M. M., Oni, A. O., Gemechu, E. and Kumar, A. (2020), 'Assessment of energy storage technologies: A review', *Energy Conversion and Management* **223**, 113295.
URL: *https://www.sciencedirect.com/science/article/pii/S0196890420308347*

Rahman, M. M., Oni, A. O., Gemechu, E. and Kumar, A. (2021), 'The development of techno-economic models for the assessment of utility-scale electro-chemical battery storage systems', *Applied Energy* **283**, 116343.
URL: *https://www.sciencedirect.com/science/article/pii/S0306261920317256*

Raineri, R., Ríos, S. and Schiele, D. (2006), 'Technical and economic aspects of ancillary services markets in the electric power industry: an international comparison', *Energy Policy* **34**(13), 1540–1555.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421504003696*

Rancilio, G., Rossi, A., Di Profio, C., Alborghetti, M., Galliani, A. and Merlo, M. (2020), 'Grid-Scale BESS for ancillary services provision: SoC restoration strategies', *Applied Sciences* **10**(12).
URL: *https://www.mdpi.com/2076-3417/10/12/4121*

Razmi, A. R., Heydari Afshar, H., Pourahmadiyan, A. and Torabi, M. (2021), 'Investigation of a combined heat and power (CHP) system based on biomass and compressed air energy storage (CAES)', *Sustainable Energy Technologies and Assessments* **46**, 101253.
URL: *https://www.sciencedirect.com/science/article/pii/S2213138821002630*

Research, development and innovation roadmap (2016), Technical report, European Network of Transmission System Operators for Electricity.

Salvini, C. and Giovannelli, A. (2022), 'Techno-economic comparison of diabatic CAES with artificial air reservoir and battery energy storage systems', *Energy Reports* **8**, 601–607. The 8th International Conference on Energy and Environment Research –"Devel-

oping the World in 2021 with Clean and Safe Energy".
URL: *https://www.sciencedirect.com/science/article/pii/S2352484722007259*

Sandhu, H. S., Fang, L. and Guan, L. (2016), 'Forecasting day-ahead price spikes for the Ontario electricity market', *Electric Power Systems Research* **141**, 450–459.
URL: *https://www.sciencedirect.com/science/article/pii/S0378779616302978*

Schmidt, O., Melchior, S., Hawkes, A. and Staffell, I. (2019), 'Projecting the future levelized cost of electricity storage technologies', *Joule* **3**(1), 81–100.
URL: *https://www.sciencedirect.com/science/article/pii/S254243511830583X*

Schmidt, O. and Staffell, I. (2023), *Monetizing Energy Storage: A Toolkit to Assess Future Cost and Value*, 1 edn, Oxford: Oxford University Press.

Schoenung, S. M. and Hassenzahl, W. V. (2003), 'Long- vs. short-term energy storage technologies analysis: a life-cycle cost study (a study for the DOE energy storage systems program)'.
URL: *https://www.osti.gov/biblio/918358*

Sioshansi, R., Denholm, P., Jenkin, T. and Weiss, J. (2009), 'Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects', *Energy Economics* **31**(2), 269–277.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988308001631*

Sneezer, S. (2021), Modeling & optimization of a renewable energy supply chain with a focus on energy storage, Master's thesis, Texas A&M University.
URL: *https://hdl.handle.net/1969.1/195151*

Song, H., Liu, C. ., Lawarree, J. and Dahlgren, R. W. (2000), 'Optimal electricity supply bidding by Markov decision process', *IEEE Transactions on Power Systems* **15**(2), 618–624.

Staffell, I. and Pfenninger, S. (2016), 'Using bias-corrected reanalysis to simulate current and future wind power output', *Energy* **114**, 1224–1239.
URL: *https://www.sciencedirect.com/science/article/pii/S0360544216311811*

Stefan, T. and Rodney, W. (2007), 'Outlier treatment and robust approaches for modeling electricity spot prices '.

Steward, D., Saur, G., Penev, M. and Ramsden, T. (2009), 'Lifecycle cost analysis of hydrogen versus other technologies for electrical energy storage'.
URL: *https://www.osti.gov/biblio/968186*

Strassheim, A., de Haan, J. E. S., Gibescu, M. and Kling, W. L. (2014), Provision of frequency restoration reserves by possible energy storage systems in Germany and the Netherlands, *in* '11th International Conference on the European Energy Market (EEM14)', pp. 1–5.

Strbac, G., Aunedi, M., Pudjianto, D., Sanders, D., Hart, A., Ravishankar, M. and Brunert, J. (2016), 'An analysis of electricity system flexibility for Great Britain'.

Sundararagavan, S. and Baker, E. (2012), 'Evaluating energy storage technologies for wind power integration', *Solar Energy* **86**(9), 2707–2717.
URL: *https://www.sciencedirect.com/science/article/pii/S0038092X12002253*

Sutton, R. S. and Barto, A. G. (2020), *Reinforcement learning: An introduction*, Vol. 2.

System Needs and Product Strategy (2017), Technical report, National Grid.

Tan, K. M., Babu, T. S., Ramachandaramurthy, V. K., Kasinathan, P., Solanki, S. G. and Raveendran, S. K. (2021), 'Empowering smart grid: A comprehensive review of energy storage technology and application with renewable energy integration', *Journal of Energy Storage* **39**, 102591.
URL: *https://www.sciencedirect.com/science/article/pii/S2352152X21003340*

Tang, Y. and Agrawal, S. (2020), 'Discretizing continuous action space for on-policy optimization', *Proceedings of the AAAI Conference on Artificial Intelligence* **34**(04), 5981–5988.
URL: *https://ojs.aaai.org/index.php/AAAI/article/view/6059*

The Electricity Trading Arrangements (2019), Technical report, National Grid.

Transitioning to a net zero energy system: Smart Systems and Flexibility Plan 2021 (2022), Technical report, Ofgem.

UK electricity production (2022), `https://electricityproduction.uk/plant/oil/`.

Van den Bergh, K. and Delarue, E. (2015), 'Cycling of conventional power plants: Technical limits and actual costs', *Energy Conversion and Management* **97**, 70–77.
URL: *https://www.sciencedirect.com/science/article/pii/S0196890415002368*

van Kooten, G. C. and Johnston, C. M. (2016), 'The economics of forest carbon offsets', *Annual Review of Resource Economics* **8**(1), 227–246.
URL: *https://doi.org/10.1146/annurev-resource-100815-095548*

Vivero-Serrano, G. D., Bruninx, K. and Delarue, E. (2019), 'Implications of bid structures on the offering strategies of merchant energy storage systems', *Applied Energy* **251**, 113375.

Walawalkar, R., Apt, J. and Mancini, R. (2007), 'Economics of electric energy storage for energy arbitrage and regulation in New York', *Energy Policy* **35**(4), 2558–2568.
URL: *http://www.sciencedirect.com/science/article/pii/S0301421506003545*

Way, R., Ives, M. C., Mealy, P. and Farmer, J. D. (2022), 'Empirically grounded technology forecasts and the energy transition', *Joule* **6**(9), 2057–2082. Open Access.
URL: *https://doi.org/10.1016/j.joule.2022.08.009*

Welfle, A. and Röder, M. (2022), 'Mapping the sustainability of bioenergy to maximise benefits, mitigate risks and drive progress toward the sustainable development goals', *Renewable Energy* **191**, 493–509.
URL: *https://www.sciencedirect.com/science/article/pii/S0960148122004463*

Wogrin, S. and Gayme, D. F. (2015), 'Optimizing storage siting, sizing, and technology portfolios in transmission-constrained networks', *IEEE Transactions on Power Systems* **30**(6), 3304–3313.

Wu, D., Kang, J., Yang, F. and Yang, L. (2024), *Main Challenges and Countermeasures for New Energy Development in China Under the Construction of New Power System*, Springer Nature Singapore, Singapore, pp. 373–392.
URL: *https://doi.org/10.1007/978-981-99-7289-0*$_2$1

Xie, C., Hong, Y., Ding, Y., Li, Y. and Radcliffe, J. (2018), 'An economic feasibility assessment of decoupled energy storage in the uk: With liquid air energy storage as a

case study', *Applied Energy* **225**, 244–257.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261918306263*

Ye, Y., Qiu, D., Li, J. and Strbac, G. (2019), 'Multi-period and multi-spatial equilibrium analysis in imperfect electricity markets: A novel multi-agent deep reinforcement learning approach', *IEEE Access* **7**, 130515–130529.

Zakaria, N. J., Shapiai, M. I. and Wahid, N. (2021), 'A study of multiple reward function performances for vehicle collision avoidance systems applying the dqn algorithm in reinforcement learning', *IOP Conference Series: Materials Science and Engineering* **1176**(1), 012033.
URL: *https://dx.doi.org/10.1088/1757-899X/1176/1/012033*

Zakeri, B. and Syri, S. (2015), 'Electrical energy storage systems: A comparative life cycle cost analysis', *Renewable and Sustainable Energy Reviews* **42**, 569–596.
URL: *https://www.sciencedirect.com/science/article/pii/S1364032114008284*

Zhang, F., Tokombayev, M., Song, Y. and Gross, G. (2014), Effective flywheel energy storage (FES) offer strategies for frequency regulation service provision, *in* '2014 Power Systems Computation Conference', pp. 1–7.

Zhang, Z., Lohr, L., Escalante, C. and Wetzstein, M. (2010), 'Food versus fuel: What do prices tell us?', *Energy Policy* **38**(1), 445–451.
URL: *https://www.sciencedirect.com/science/article/pii/S0301421509007174*

Zhou, Y., Scheller-Wolf, A., Secomandi, N. and Smith, S. F. (2011), Managing wind-based electricity generation with storage and transmission capacity.

Zhu, J., Zhang, H., Yuan, W., Zhang, M. and Lai, X. (2013), Design and cost estimation of superconducting magnetic energy storage (SMES) systems for power grids, *in* '2013 IEEE Power & Energy Society General Meeting', pp. 1–5.