
BATTERIES NOT INCLUDED: MODELLING GRID-SCALE BATTERY STORAGE REQUIREMENTS FOR THE UK NATIONAL GRID

Alan Clark, Max Bronckers, Alistair McLeay, and Katie Collins – [Project GitHub Organisation](#)

ABSTRACT

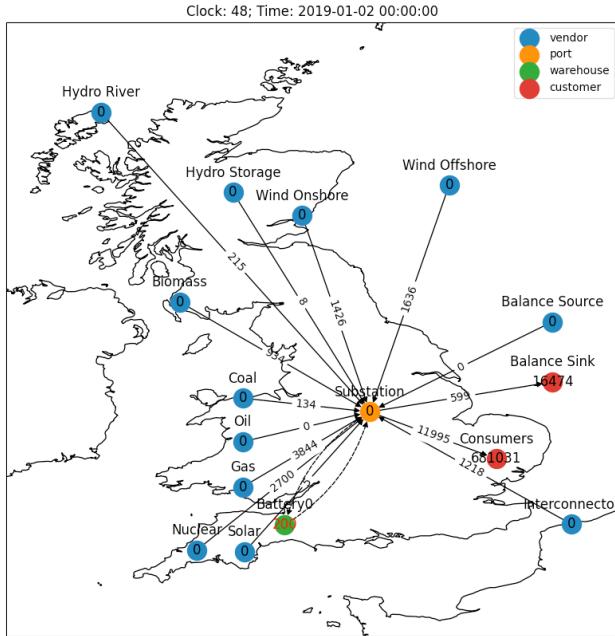
The UK’s transition towards renewable sources of electricity is introducing increasing levels of generation unpredictability, thereby posing new challenges for ensuring the stability of the UK National Grid. Further, the weather-dependence of renewable sources can result in a mismatch between periods of high supply and consumer demand. This work explores the potential to add grid-scale battery facilities to the electricity network in an attempt alleviate pressure on the Grid’s balancing mechanisms and maximise utilisation of renewable energy. Simplified supply and demand models were developed and incorporated into Amazon’s supply-chain modelling tool, `miniSCOT`, to produce a digital-twin of the National Grid, into which battery capacity could be introduced. Bayesian optimisation was employed to determine the optimal total capacity to incorporate under a range of scenarios. Under our baseline scenario, we found that 812 MWh of total capacity should be added to the Grid, increasing to 1864 MWh depending on the depth of discounting that battery operators may be able to offer compared to other generation sources. We did not find evidence to support increasing the total capacity to accommodate additional generation from surges in solar or wind activity. Although many simplifying assumptions inevitably had to be made given the short timeframe of the project, our work provides a strong and highly flexible foundation for future investigations to build upon.

1 Introduction

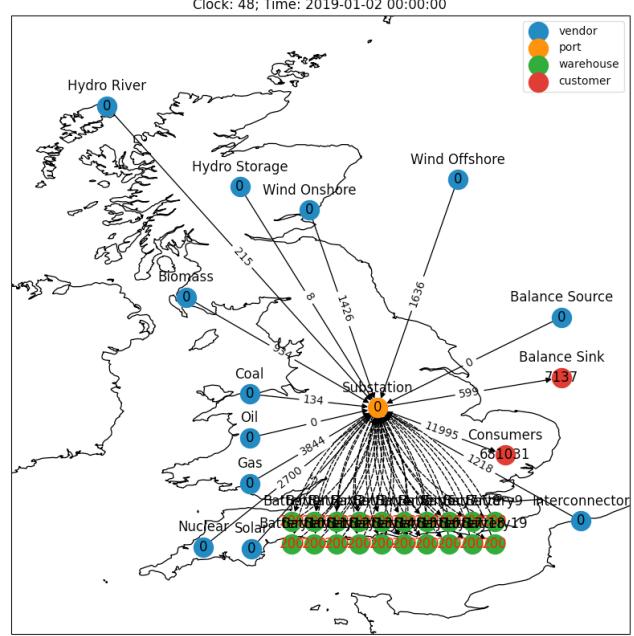
The UK government has legally committed to a target of net-zero greenhouse gas emissions by 2050 [1, 2]. Given its growing use for heating and transport, electricity will be increasingly important in supporting the delivery of net-zero emissions, potentially providing around half of final energy demand and requiring generation to as much as double [3, 4, 5, 6]. It is estimated that between 34 GW and 77 GW of new wind and solar generation could be required to meet demand in 2030, potentially quadrupling by 2050 [3, 7].

The existing electricity system was designed to operate with traditional generation sources: coal, gas, and nuclear [8]. However, since the mid-2000s, government policies which aim to reduce carbon dioxide emissions have instigated rapid growth in renewable power sources and small-scale, local energy generation – many of which are situated in different locations than traditional generation [4, 7, 9]. As such, network companies have needed to accommodate more than one million new installations [4, 8]. Managing the electricity system is subsequently becoming more complex; the increasing diversity of sources and growing demand creates new challenges for ensuring the stability of the Grid [8, 9]. The reliance of most renewables on current weather conditions naturally introduces uncertainty-laden dependencies into power generation, both epistemic (e.g., accurately modelling the impact of weather on supply) and aleatoric (e.g., the stochasticity inherent to weather dynamics). Significant increases in supply variability, most notably during periods of abnormally high winds, have led to instances of electricity wholesale prices going negative since 2019 [10, 11]. The events of February 2021 in Texas—which left 10 million people without electricity—highlight the impact that a failure to balance power grids can cause [12].

In this project, we asses how grid-scale battery facilities may be used to improve system robustness and maximise utilisation of renewable sources of electricity. The UK’s Electricity System Operator (ESO) has a number of mechanisms available to keep the grid balanced; we focused on the Bid-Offer mechanism, which is described in Appendix A.1. Simplified supply and demand models were developed using data from the UK Balancing Mechanism Reporting Service (BMRS) [13]. These were incorporated into Amazon’s supply-chain modelling tool, `miniSCOT` [14], to produce a digital-twin of the UK National Grid, into which we could introduce battery capacity and simulate surges of renewable generation (Section 2). A reward function was developed (Section 3) and `Emukit` [15] was used to employ Bayesian optimisation methods to determine optimal battery capacities to incorporate into the network (Section 4). Initial scenarios considered the battery storage needed to accommodate typical network imbalances that are not solely caused by the variability of renewables, but where having stored renewable capacity would reduce or remove the current requirements to, for example, vary gas turbine generation or pay consumers to modify their consumption (Section 5). Given the significant complexity of the real-life system being simulated, a broad range of simplifying assumptions inevitably had to be made. A sensitivity analysis was thus carried out to determine which inputs our results were most responsive to (Section 6). Finally, scenarios considering the impact of significant variations in renewable generation, driven by weather-based dynamics, on the optimal Grid battery capacity were considered (Section 7).



(a) Single battery in network



(b) 20 batteries in network

Figure 1: Visualisation of National Grid miniSCOT digital-twin with varying number of batteries. Values on edges indicate the MWh of electricity being transmitted during the current timestep. All edges are unidirectional; with the exception that those between the Substation and Batteries permit bidirectional flow and are solid when active. Values inside Consumers and Balance Sink nodes indicate total MWh of electricity received to date. Values inside batteries indicate current MWh of capacity; red text indicates the battery is at maximum capacity.

2 National Grid Simulator

miniSCOT [14], developed by members of Amazon’s Supply Chain Optimisation Technologies team, is a lightweight, Python-based supply chain simulation package which allows users to investigate supply chain architectures and algorithms at any level of fidelity. NetworkX [16] is used to create the supply network, with each node representing a Vendor, Port, Warehouse or Customer. Quantities of products can be held at Ports and Warehouses, as well as on edges to represent products that are in transit during a timestep. At each timestep, the Controller will send the State to configured Modules, and receive back a list of Actions. The Controller executes these Actions, and proceeds to the next timestep. Modules acquire any additional information needed (e.g. Demand Forecast, Prices/Costs) via a Service layer. Finally, the Controller sends back to each Module a Reward for that timestep [17].

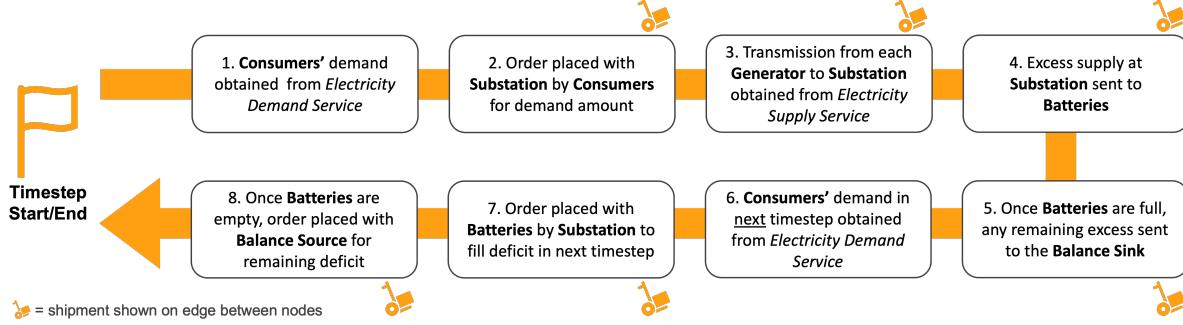
While miniSCOT was primarily designed for supply chains of physical products, here we adapted it to simulate the National Grid. We defined a single Electricity product with 1 unit representing 1 MWh, and applied the following node-type mappings:

- **Customer → Consumers:** A single Consumers node was used to represent total demand.
- **Vendor → Generator:** A Generator node was created for each fuel type reported on in the BMRS.
- **Port → Substation:** One Substation node was introduced to represent the bulk of the National Grid network. By disabling miniSCOT’s restriction that Ports must have non-negative inventory, the inventory at the Substation is used to indicate whether supply and demand in the network are balanced.
- **Warehouse → Battery:** Battery nodes have a maximum capacity, and capacities cannot be negative. At simulation initialisation, all batteries have 50 % capacity; poised to receive or discharge electricity. Note, we do not account for any loss of battery charge over time, nor a reduction in maximum battery capacity over time.

To aid reviewing and debugging the simulator, methods for visualising the network and product flows were also added to miniSCOT; Figure 1 illustrates the topology of our National Grid digital-twin. Our simulator did not take the physical location of nodes into account, however it would be possible to introduce distance-based transmission losses/penalties.

Our network includes two additional nodes: Balance Source (a pseudo-Generator) and Balance Sink (a pseudo-Consumer), which serve as proxies for the ESO’s Bid-Offer acceptances. If the the network (i.e. Substation) is over-supplied, then excess electricity can be "sold" to the Balance Sink, representing the acceptance of Bids for generators to reduce supply or consumers to increase demand. Conversely, if the network is under-supplied then the necessary electricity can be "purchased" from the Balance Source, representing the acceptance of Offers for generators to increase supply or consumers to decrease demand. These nodes are assumed to have infinite capacity to consume/generate electricity, and form a key part of our reward function.

For settlement purposes, the BMRS considers electricity to be generated, transported, delivered and used in half-hourly Settlement Periods, and therefore provides data with this granularity [18]. To align with this, our simulator steps forward in 30 minute time increments. The "delivery time" between every node was set to 1 increment. Thus, for example, Consumers' demand for timestep X is dispatched from the Substation and appears on the edge connecting the two nodes within timestep X. Then, at timestep X+1 it arrives at the Consumers. The sequence of actions carried out at each timestep is summarised in Figure 2. Supply models were developed for each fuel type reported by the BMRS for use by the Electricity Supply Service, along with a demand model for the Electricity Demand Service – these are discussed in Appendix A.2.



3 Reward Function

A reward function was required to enable estimation of the total battery capacity that should be introduced to the National Grid to minimise the requirement to use the Bid-Offer acceptance balancing mechanism. The miniSCOT cash accounting functionality was used to calculate timestep and episode rewards. Whilst we could not consider all factors of this complex scenario, it was vital to design a function that had a realistic and sensible relationship with the physical world and captured salient elements of the problem. After several iterations, Equation 1 highlights the final function we used to calculate timestep reward, R . Figure 3 provides an aid to visually interpret this function. Penalty and reward values used in the core scenarios for metered electricity are shown in Table 1. We let N_{bat} and C_{bat} represent the number and capacity (in MWh) of the batteries introduced. The reward approximated the cash accounting of the Bid-Offer acceptance balancing mechanism; a negative value indicates the cost of the mechanism, and the reward associated with 0 batteries is an approximation to the current cost. As will be seen in Section 5, the introduction of battery capacity can minimise this cost.

$$R = P_{bal_source} Q_{bal_source} + R_{bal_sink} Q_{bal_sink} + P_{bat_discharge} Q_{bat_discharge} + R_{bat_charge} Q_{bat_charge} + P_{bat_AM-CAPEX} C_{bat} N_{bat} \quad (1)$$

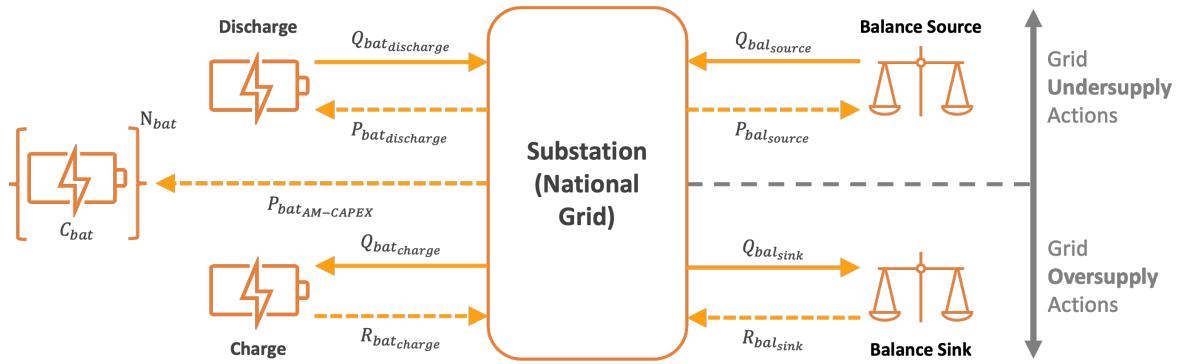


Figure 3: Reward function visualisation. Solid arrows indicate flow of electricity in MWh, dashed arrows indicate payments in £. Amortised battery CAPEX, assumed to be paid by National Grid, is also shown.

Data for the price of grid-scale batteries was sparse and highly variable; prices in the region of 110-330 £/kWh were found, with lower costs being cited in the most recent sources [19, 20, 21]. As such, a price of 200 £/kWh was chosen for our core scenarios; sensitivity to this value is discussed in Section 6. In reality, the relationship between battery capacity and £/kWh is unlikely to be flat; however, this was not factored into our design: 80 10-MWh batteries had the same reward as 10

Scenario	P_{bal_source}	R_{bal_sink}	Discharge Discount	$P_{bat_discharge}$	Charging Premium	R_{bat_charge}
Baseline	-36.65	27.63	0.20	-29.32	0.05	29.01
25 % Discount	-36.65	27.63	0.25	-27.49	0.05	29.01
33 % Discount	-36.65	27.63	0.33	-24.56	0.05	29.01
33 % Discount +	-36.65	27.63	0.33	-24.56	0.00	27.63
50 % Discount	-36.65	27.63	0.50	-18.33	0.05	29.01

Table 1: Metered electricity usage reward and penalty values used in the core scenarios; penalty and reward units are £/MWh.

80-MWh batteries. Consequently, our reward function effectively reports the optimal total capacity to introduce to the network. We observed that introducing a one-off battery CAPEX penalty at timestep 0 led to the optimal value being dependant on the simulation runtime – larger capacities were preferred once they had paid themselves off. Therefore, battery costs were amortised over a 15 year timeframe, with an equal payment made in each 30 minute time period ($P_{bat_{AM-CAPEX}}$). No discounting factor was applied, and no additional OPEX costs were considered. Our reward function assumed that National Grid is paying this cost; they may own the batteries themselves, or this may reflect licensing agreements and/or government incentives.

The average Offer accepted to increase generation/decrease demand in 2020 was 36.65 £/MWh, and the average Bid accepted to decrease generation/increase demand was 27.63 £/MWh [22]. These prices were used for P_{bal_source} and R_{bal_sink} in all scenarios. The “Baseline” scenario assumed that batteries are used to smooth network imbalances, and so reduced the number of Bid-Offer acceptances. We considered that batteries would be able to offer a 20 % discount on the average Bid price ($P_{bat_discharge}$) and would add a premium of 5 % to the average Offer price (R_{bat_charge}), outperforming other Generators/Suppliers. This would amount to batteries making only 0.31 £/MWh per charge/discharge cycle; however we reiterate that payments for their CAPEX are additionally being made in each time period. The remaining core scenarios consider battery operators being able to more deeply discount Offers to provide electricity. This assumed that they have been able to obtain sources of electricity for lower prices, either by having their own sources of renewables generation, and/or by exploiting time arbitrages.

4 Bayesian Optimisation

In our experiments, we sought to minimise the *negated* episode reward by introducing battery capacity into the Grid. Episode reward was obtained via running our simulator over a specified set of parameters, on a pre-defined domain. Our simulation function f is dependent on the number of discrete time steps (defined here as the “time horizon” of the simulation) and can therefore be expensive to evaluate. Moreover, the form of f is unknown and can be non-convex or multimodal. As such, our experiment was well-suited for the use of Bayesian optimisation (BO). Our BO loop implementation was built using the Emukit framework [15].

Surrogate Model BO employs a surrogate function to model beliefs about f . We opted for a Gaussian process (GP), which enabled us to flexibly condition and fit our model on a random set of simulation data points \mathcal{D} to get a predictive posterior of f . To fit our GP to a set of initial points \mathcal{D} , we maximised the log-posterior of the GP by varying the hyperparameters. We drew on GPy [23] as framework for our GP implementation.

Kernel For our GP surrogate model, we had a design choice concerning its kernel. We evaluated the following kernels: RBF, Matern32, Rational Quadratic (RQ), RBF + Matern32, and RBF + Bias. Since our domain was initially two-dimensional (see Section 5), we were able to visually inspect the consequences of kernel choice on the GP fit to $N = 10$ random initial data points. From our initial simulations, we uncovered that our f was smooth and had no unexplained discontinuities. This meant that RBF and RQ were reasonable options for interpolation between observed data points, in contrast to Matern32 which defaulted to approximately uni-dimensional interpolation (see Figure 9 in Appendix A.3). For the RBF, we also considered adding a Bias kernel as we observed that the posterior mean defaulted to 0. We found such an addition was beneficial to the fit, as this combined kernel yielded the best reward, all other things being equal.

Hyperparameters Moreover, in order to maximize our GP fit and generalisation ability, we conducted a hyperparameter search to find a consistently well-performing setting. In GPy, the fitting of the GP to \mathcal{D} is done by minimising the negative log posterior, i.e. setting the hyperparameters to the maximum a posteriori (MAP) estimate using Equation 2.

$$\theta^* = \arg \min -\log p(\theta|\mathcal{D}) \propto -\log p(\mathcal{D}|\theta) - \log p(\theta) \quad (2)$$

The minimisation was done via a L-BFGS gradient-based optimizer. We randomly restarted the optimizer 10 times and selected the best hyperparameter configuration over the initial dataset \mathcal{D} . We also noticed that splitting up the hyperparameters per dimension helped the GP fit (by setting ARD to true). As our evaluations of f did not contain noise, we set the noise variance effectively to 0 in the GP regression.

We observed that, without a prior specified, the consequent maximum likelihood estimate (MLE) led the GP to degenerate fits with extreme lengthscales and variances - overfitting to the initial \mathcal{D} . As such, we sought to specify a prior to avoid the

GP falling into these cases. In our prior specification, we wanted to retain a high level of entropy, while maximizing the informativeness of said prior (e.g., not uniform). Since the Gamma distribution is the maximum entropy distribution for a random variable with a fixed, positive mean, and the variance and lengthscale domains are $(0, \infty)$, we opted for a Gamma prior. We varied the shape α and rate β parameters (over $\{(0.01, 0.01), (1, 0.1), (1e2, 1), (1e4, 10)\}$) as to fix the variance but vary the mean of our prior. For the RBF variance, we observed that a greater α led to a lower RBF variance, fitting more closely to the observed points (lower marginal log-likelihood) but resulting in a higher posterior (which the GP seeks to minimize).

We note the challenge of balancing fit to observations and posterior uncertainty, as the best reward found by the BO loop under a fixed number of iterations also depends on the set of initial points \mathcal{D} . If posterior variance is too large it can detract from exploitation in the acquisition function, which is especially relevant when the simulation time horizon increases as evaluated points are more expensive to obtain. Conversely, minimal uncertainty entails unexplored areas not being evaluated. In our early experimental runs, we observed sufficient exploration in the BO loop and opted for a default prior RBF variance setting of $\text{Gamma}(\alpha = 1e4, \beta = 10)$. We kept the RBF lengthscale prior fixed (with the exception of Section 7) to a $\text{Gamma}(\alpha = 10, \beta = 0.1)$ distribution, as this tended to work well across various simulations.

Acquisition Functions Another key BO design aspect was the acquisition function. Acquisition functions in BO determine the exploration/exploitation trade-off as they decide where to obtain the next data point from f by maximizing some utility-value criterion. We evaluated the performance of five acquisition functions: expected improvement (EI), probability of improvement (PI), negative lower confidence bound (NLCB), integrated variance reduction (IVR), and uncertainty sampling (US). Plots of the fitted GPs after BO, over these different acquisition functions, can be found in Appendix A.4. After evaluating the first three functions, we observed that EI consistently found the lowest minimum in a fixed number of iterations—which is sensible given that it improves the current minimum value the most in expectation (see Equation 3 in Appendix A.4). Consequently, we opted for EI as our default acquisition function. We noticed in our experiments later on, however, that EI performed too little exploration and converged to exploring points with rewards within a $\epsilon = 0.01$ improvement. This led the BO loop to terminate under ϵ -stopping criterion. We avoided this problem by providing the GP with more (random) initial points. When we evaluated US and IVR at a later stage, we observed that the BO loop performed more exploration using the same set of initial points and stopping conditions. As such, with the benefit of hindsight, we could also have opted for US or IVR; thus avoiding having to provide large numbers of initial points. The downside of using these functions, however, is that they can overemphasize exploration and use the maximum number of iterations specified, which can be detrimental to the BO run time.

5 Core Scenarios

5.1 Optimisation Across a Single Dimension

We first leveraged BO to identify the optimal total capacity to introduce to the National Grid network for the core scenarios outlined in Table 1. We held the individual battery maximum capacity, C_{bat} , constant at 2 MWh and varied the number of batteries, N_{bat} . GP fits for selected scenarios are shown in Figure 4; further discussion can be found in Appendix A.5. As shown in Table 2, we observed that increasing the Discharge Discount led to a corresponding increase in the optimal total capacity. This was anticipated: as the price that batteries can supply electricity, $P_{bat, discharge}$, decreases it makes sense for the total available capacity to increase until counterbalanced by the amortised battery CAPEX costs, $P_{bat, AM-CAPEX}$, for under-utilised capacity. Most scenarios considered battery operators paying a small premium of 5 % on the Offer price, however this was removed in the "33 % Discount +" scenario, which resulted in a small reduction in optimal total capacity, as would be expected.

Scenario	1D Optimisation		2D Optimisation		
	Cumulative Reward (£)	Optimal Total Capacity (MWh)	Opt. Total. Cap. Median (MWh)	Opt. Total. Cap. Mean (MWh)	Opt. Total. Cap. Std. Dev. (MWh)
Baseline	-547,000	812	870	952	285
25 % Discount	-465,000	1128	1184	1535	652
33 % Discount	-318,000	1440	1953	1750	521
33 % Discount +	-388,000	1292	1326	1344	130
50 % Discount	-51,000	1864	1860	1720	433

Table 2: Optimal total battery capacity to introduce into the National Grid, for each core scenario. Results are shown for running BO with both 1D (N_{bat}) and 2D (N_{bat}, C_{bat}) inputs. In the 2D case, the median, mean and standard deviation across 1, 10, 30, 60, and 120 simulated days are shown.

5.2 Optimisation Across Two Dimensions

After confirming the simulator and emulator were performing as expected when varying only N_{bat} , we expanded our experiments to additionally vary the value of C_{bat} . In Figure 5b, we see that there is a "valley" of minimum negative cumulative reward function values. This provided further evidence that the simulator and emulator were behaving correctly – as discussed in Section 3, we were aware that the structure of our reward function meant that BO could only optimise for total network

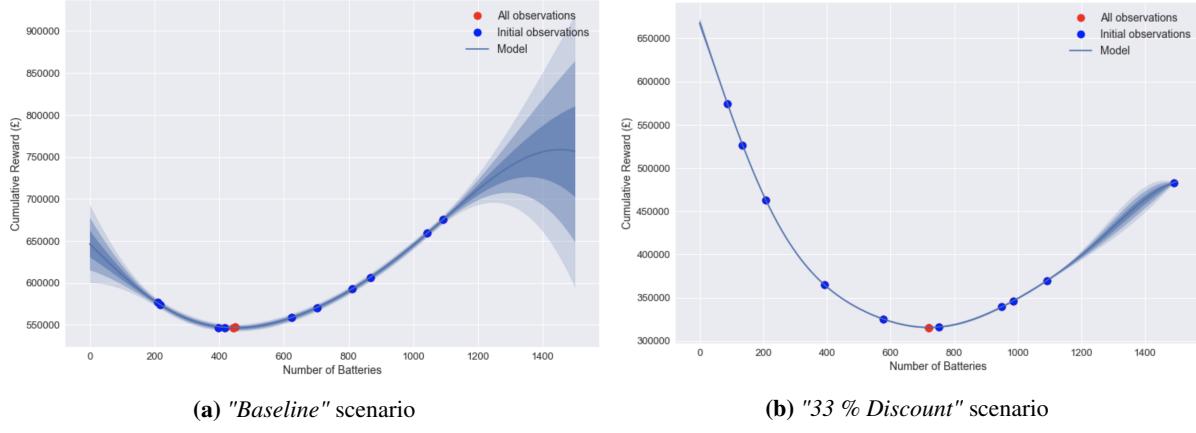


Figure 4: Results of running BO to find the optimal number of batteries for different scenarios, each with a fixed maximum battery capacity of 2MWh. Negative cumulative reward is shown.

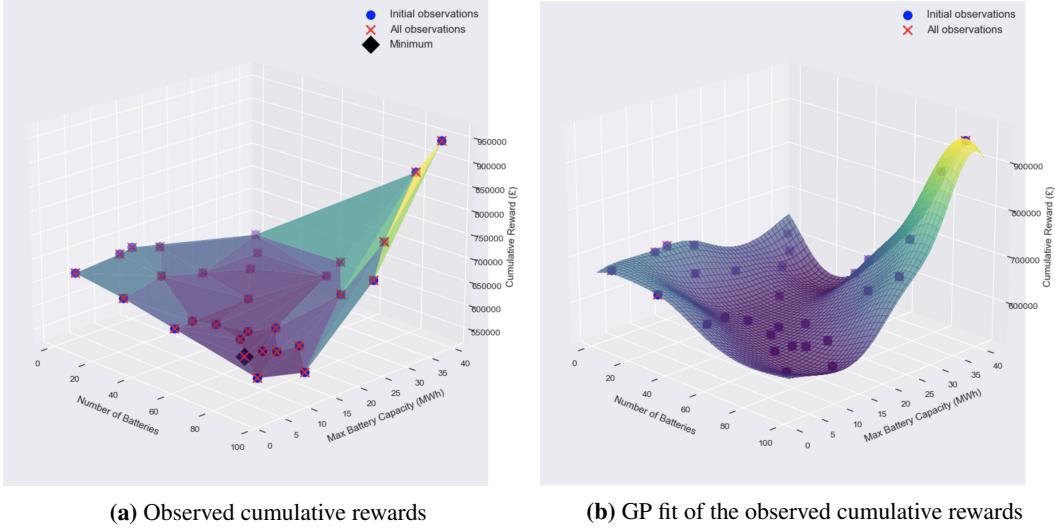


Figure 5: Results of running BO to find the optimal maximum battery capacity, C_{bat} , and number of batteries, N_{bat} , with a fixed battery penalty, $P_{bat_{AM-CAPEX}}$, of 200 £/kWh. Negative cumulative reward is shown.

battery capacity. We knew from the previous experiments that the optimal total capacity for the *Baseline* scenario was 812 MWh. Multiplying N_{bat} and C_{bat} values along the negative cumulative reward valley floor in Figure 5b indeed yields values in close proximity to 812 MWh.

6 Sensitivity Analyses

6.1 Simulation Time Horizon

When creating the simulator and reward function, design decisions were made to avoid the simulated time period length ("time horizon") impacting the optimal total capacity values obtained. Supply and demand models were designed to reflect a "typical day," and battery CAPEX was amortised. We assessed these design choices by evaluating the optimal capacities identified via BO across 1, 10, 30, 60, and 120 days. The results in Table 2 show that our design efforts had been largely successful. However, we found optimal capacities for the baseline scenario that were significantly higher than for the 1D optimisation. Investigation of the BO results revealed the expected area of the optimum capacity had not been examined, and a significantly higher value was returned instead (see Figure 12 in Appendix A.6). This highlights the aforementioned explore-exploit tradeoff of the BO acquisition function, and the possible inability to find the true minimum in a fixed number of iterations with an ϵ stopping condition of the EI acquisition function. Such behavior of EI is further evident upon inspecting the varied number of evaluated points across the different scenarios (see Table 5 in Appendix A.6). In light of computational constraints, we re-introduced the $\epsilon = 0.01$ stopping criterion for the non-baseline scenarios. However, this led to most BO loops only evaluating two iterations with batch size 5 and consequently under-exploring the domain (and explaining the larger standard deviation of the non-baseline scenarios).

6.2 Battery Penalty

As discussed in Section 3, a design choice was made in selecting a battery penalty, $P_{bat_{AM-CAPEX}}$, of 200 £/kWh. There was significant variability in the values found in literature, and as costs are likely to decrease over the coming years (driven by improvements in battery technologies and economies of scale [7, 20]) we desired to evaluate the sensitivity of optimal capacity to these costs. As would be expected, Figure 6 shows that, holding max battery capacity, C_{bat} , constant at 2 MWh, the optimal number of batteries, N_{bat} , increases as the battery penalty decreases. Thus, we obtained an emulator that would allow us to model the optimal total network capacity for various potential battery prices. Using the identified lower and upper bounds, 110 and 330 £/kWh, the optimal total battery capacities were found via BO to be 1672 MWh and 404 MWh. These results are in line with the optimal value of 812 MWh found at a price of 200 £/kWh in Section 5.1. Note, the lengthscale in Figure 6 appears too low, suggesting greater variability than we would envisage in reality, indicating the GP fit could be further improved.

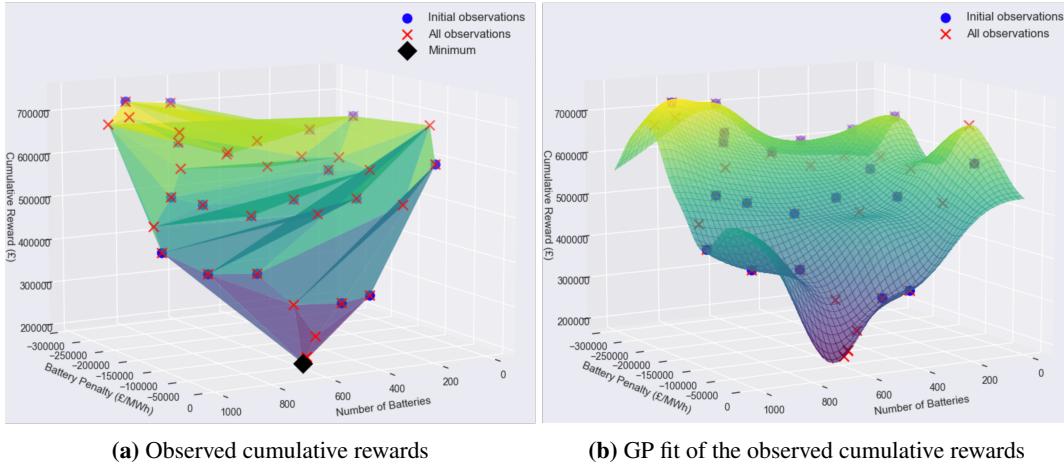


Figure 6: Results of running BO to find the optimal battery penalty, $P_{bat_{AM-CAPEX}}$, and number of batteries, N_{bat} , using a fixed max battery capacity, C_{bat} , of 2MWh. Negative cumulative reward is shown.

6.3 Global Sensitivity Analysis

A global sensitivity analysis was carried out to evaluate the overall sensitivity of the simulation function, f , to the inputs N_{bat} , C_{bat} , $P_{bat_{AM-CAPEX}}$ and the simulated time horizon. The input domain investigated and the results of the analysis are shown in Table 3. Given the time and compute expense of running the simulation itself, Monte Carlo estimates of the Sobol indices were obtained from a GP model fit to 100 points generated by the simulator using inputs randomly sampled across the entire domain. Whereas the previous two sections looked at the impact of $P_{bat_{AM-CAPEX}}$ and time horizon on the optimal total capacity, the results here relate specifically to the *reward value*. It can be seen that N_{bat} had the greatest main and total effect, followed by C_{bat} and time horizon, both of which displayed similar effects. From our previous experiments, it was expected that N_{bat} and C_{bat} would have a large impact on the reward value. Whilst the investigation in Section 6.1 confirmed that our design choices had made the optimal total capacity largely invariant to the time horizon, we did not attempt to normalise the reward function value across time. It was therefore not surprising that this should have such a large impact on the simulation function value. It was encouraging to note that $P_{bat_{AM-CAPEX}}$ had the lowest effects, given our high *a priori* uncertainty in its value.

Parameter	N_{bat}	C_{bat} (MWh)	Time Horizon (Days)	$P_{bat_{AM-CAPEX}}$ (£/MWh)
Lower Bound	1	1	1	110,000
Upper Bound	1000	100	120	330,000
Main Effect	0.229	0.210	0.210	0.083
Total Effect	0.401	0.364	0.358	0.178

Table 3: Global sensitivity analysis results.

7 Simulating Renewable Energy Surges

We next probed the impact of surges in renewables generation on the optimal total battery capacity. We chose solar and offshore wind supply sources as their generation dynamics are markedly different (see Figure 8 in Appendix A.7). For each generation type, we consider simulation time horizons of 10 and 30 days, over both “Baseline” and “33% Discount” scenarios. Looking at two time horizons allowed us to consider both short and longer-term impacts of such surges. We introduced surge by scaling the forecasted generation for each source by a scalar Surge Factor. As wind gusts in the physical world tend to be more sudden

and short-lived compared to bouts of excess sun or cloud coverage, we simulated surge in wind during the first 3 simulated days, compared to 7 days for solar. We seeded our BO simulation with 30 evenly-spaced initial observations. As in Section 5.1, C_{bat} was held constant at 2 MWh and N_{bat} varied to determine the optimal total capacity. Note, while we did not re-tune our base GP per scenario, we decreased the initial GP kernel lengthscale to 1 (from 10) to discourage overly smoothed fits.

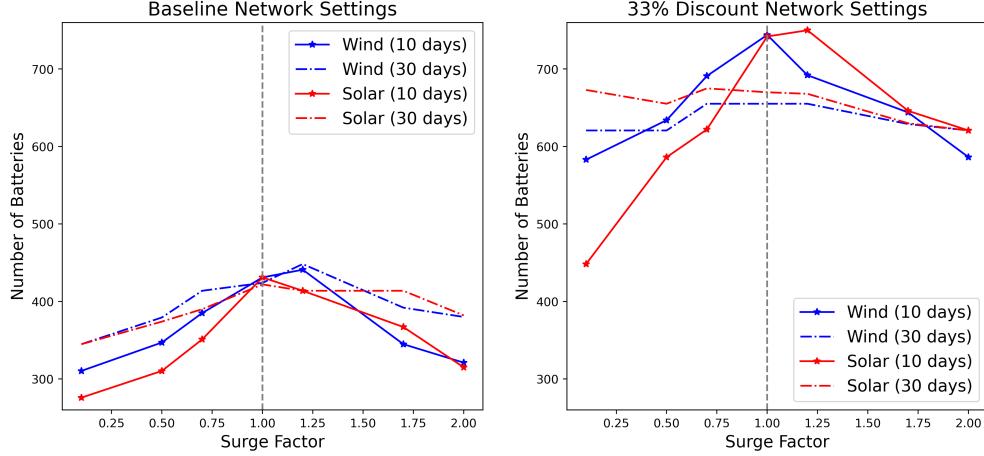


Figure 7: Optimal number of batteries to introduce into the Grid under wind and solar surges.

We focus first on the 10 day time horizon. As seen in Figure 7, in scenarios with a Surge Factor < 1 (i.e., supply *drops*), the optimal total capacity decreases, which is sensible to ensure battery utilisation remains high. Our reward function—which rewards batteries charging and penalises discharging—is able to correctly reflect what would be expected in reality. What is perhaps more surprising is that fewer batteries are desired for a Surge Factor > 1 in the majority of cases. Intuitively, having more batteries would allow more electricity from wind and solar surges to be stored, thereby enabling the economic benefit of the batteries to be enjoyed over a longer timeframe. Yet, it appears that the excess electricity is rapidly consumed by the Grid, at which point increased numbers of batteries are underutilised during normal weather conditions. We hypothesize that the CAPEX of the underutilised batteries may simply be too high to justify an increased total capacity.

We next compare the 10 and 30 day time horizons across both scenarios. As seen in Figure 7, we find that the impact of weather variations, even for fairly extreme weather events (such as a doubling in wind volume, or a 90% increase in cloud coverage for a week) diminishes over time. These results suggest that the optimal total capacity is robust to short-term weather events. However, it is important to note that we impose several assumptions in our simulations; for instance, we do not account for potential time arbitrages that could be exploited, which may increase the economic benefit of having a larger total capacity to store reserves during high winds/solar. Additionally, here we modelled wind and solar variability independently; yet in the real-world, such sources are often *correlated* (e.g. during a storm). For a preliminary investigation of joint surges see Appendix A.7.

8 Discussion and Conclusions

In conclusion, the results of our "*Baseline*" scenario indicate that the optimal total battery capacity to introduce into the Grid to minimise the use of the Bid-Offer balancing mechanism is 812 MWh. For comparison, in October 2021, Energy Superhub Oxford, a 50MWh lithium-ion system, became the first grid-scale facility to export to National Grid's transmission network [24, 25]. Our findings therefore suggest that a further 15 such sites would be required, based on batteries solely receiving charge when the network is oversupplied and assuming a very modest profit for battery operators. Should operators be able to provide a 50 % discount on average Bid prices (having sourced electricity cheaply from other sources, and/or through time arbitrage opportunities) optimal total capacity *increases* to 1864 MWh, or around 37 Energy Superhub Oxfords. Coincidentally, the same number of sites that the facility operator (EDF's Pivot Power) currently have in their development pipeline [26]. Moreover, we did not find sufficient evidence to support an increase in total capacity to accommodate temporary surges in renewables; the economics around any such increase are not pragmatic. This is in agreement with literature sources which envisage batteries being economical for providing short-term storage, whereas other energy storage mechanisms (such as the generation of hydrogen via electrolysis) are more suitable over long-term, inter-season time frames [7, 9].

We were highly encouraged that our experimental results appeared to make sensible inferences regarding the physical world, such as the way it responded to varying weather patterns. This supports the simulator and reward function design choices that had been made, as well as our efforts to perform robust Bayesian optimisation. Whilst there are many simplifying assumptions to be addressed, our work provides a strong and extensible foundation upon which future investigations can build, and over which policy decisions can begin to be formulated.

References

- [1] GOV.UK, “UK becomes first major economy to pass net zero emissions law,” jun 2019, Last Accessed: 2022-01-18. [Online]. Available: [https://www.gov.uk/government/news/uk-becomes-first-major-economy-to-pass-netzero-emissions-law](https://www.gov.uk/government/news/uk-becomes-first-major-economy-to-pass-net-zero-emissions-law)
- [2] HM Government, “Net Zero Strategy: Build Back Greener,” HM Government, Tech. Rep., oct 2021, Last Accessed: 2022-01-18. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1033990/net-zero-strategy-beis.pdf
- [3] Department for Business, Energy and Industrial Strategy, “Modelling 2050: Electricity System Analysis,” Department for Business, Energy and Industrial Strategy, Tech. Rep., dec 2020, Last Accessed: 2022-01-17. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943714/Modelling-2050-Electricity-System-Analysis.pdf
- [4] National Audit Office, “Electricity Networks,” National Audit Office, Tech. Rep., jan 2020, Last Accessed: 2022-01-17. [Online]. Available: <https://www.nao.org.uk/wp-content/uploads/2020/01/Electricity-networks.pdf>
- [5] Royal Academy of Engineering, “A critical time for UK energy policy: what must be done now to deliver the UK’s future energy system A report for the Council for Science and Technology,” Royal Academy of Engineering, Tech. Rep., oct 2015, Last Accessed: 2022-01-17. [Online]. Available: <https://www.raeng.org.uk/publications/reports/a-critical-time-for-uk-energy-policy>
- [6] National Infrastructure Commission, “Smart Power,” National Infrastructure Commission, Tech. Rep., oct 2015, Last Accessed: 2022-01-17. [Online]. Available: <https://nic.org.uk/app/uploads/Smart-Power-3.pdf>
- [7] National Grid ESO, “Future Energy Scenarios 2021,” National Grid ESO, Tech. Rep., jul 2021, Last Accessed: 2022-01-17. [Online]. Available: <https://www.nationalgrideso.com/document/202851/download>
- [8] National Infrastructure Commission, “Operability of Highly Renewable Electricity Sources,” National Infrastructure Commission, Tech. Rep., feb 2021, Last Accessed: 2022-01-17. [Online]. Available: <https://nic.org.uk/app/uploads/Operability-of-HRES-February-2021.pdf>
- [9] S. Bailey, A. Moffat, C. Smerdon, and P. Adams, “Strategy for Long-Term Energy Storage in the UK, Revision 14,” Jacobs, Tech. Rep., aug 2020, Last Accessed: 2022-01-17. [Online]. Available: <https://www.jacobs.com/sites/default/files/2020-10/Jacobs-Strategy-for-Long-Term-Energy-Storage-in-UK-August-2020.pdf>
- [10] Energy Live News, “UK’s power system emissions hit all-time low and wholesale power prices go negative,” may 2020, Last Accessed: 2022-01-18. [Online]. Available: <https://www.energylivenews.com/2020/05/26/uks-power-system-emissions-hit-all-time-low-and-wholesale-power-prices-go-negative/>
- [11] Smart Energy International, “The UK records negative wholesale prices for the first time,” dec 2019, Last Accessed: 2022-01-18. [Online]. Available: <https://www.smart-energy.com/renewable-energy/the-uk-records-negative-wholesale-prices-for-the-first-time/>
- [12] J. W. Busby, K. Baker, M. D. Bazilian, A. Q. Gilbert, E. Grubert, V. Rai, J. D. Rhodes, S. Shidore, C. A. Smith, and M. E. Webber, “Cascading risks: Understanding the 2021 winter blackout in Texas,” *Energy Research & Social Science*, vol. 77, p. 102106, jul 2021.
- [13] Elexon, “About the Balancing Mechanism Reporting Service,” Last Accessed: 2022-01-18. [Online]. Available: <https://www.elexon.co.uk/operations-settlement/bsc-central-services/balancing-mechanism-reporting-agent/>
- [14] P. A. Pérez-Fernández and J. Bell-Masterson, “miniSCOT,” <https://github.com/amzn/supply-chain-simulation-environment/>, 2021.
- [15] A. Paleyes, M. Pullin, M. Mahsereci, N. Lawrence, and J. González, “Emulation of physical processes with emikit,” in *Second Workshop on Machine Learning and the Physical Sciences, NeurIPS*, 2019.
- [16] A. A. Hagberg, D. A. Schult, and P. J. Swart, “Exploring Network Structure, Dynamics, and Function using NetworkX,” in *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, and J. Millman, Eds., Pasadena, CA USA, 2008, pp. 11–15. [Online]. Available: http://conference.scipy.org/proceedings/SciPy2008/paper_2/
- [17] J. Bell-Masterson, “miniSCOT Manual,” Draft kindly provided by author via email: 2021-12-06.
- [18] Elexon, “The Electricity Trading Arrangements - A Beginner’s Guide,” Elexon, Tech. Rep., 2019, Last Accessed: 2022-01-11. [Online]. Available: <https://www.elexon.co.uk/documents/training-guidance/bsc-guidance-notes/beginners-guide-2/>
- [19] E. Bellini, “A redox flow battery for MW-sized solar-plus-storage,” apr 2020, Last Accessed: 2022-01-21. [Online]. Available: <https://www.pv-magazine.com/2020/04/29/a-redox-flow-battery-for-grid-scale-solar-plus-storage/>
- [20] A. Macdonald-Smith, “AGL supplier Wärtsilä says batteries have bright future but cautions on costs,” sep 2021, Last Accessed: 2022-01-21. [Online]. Available: <https://www.afr.com/policy/energy-and-climate/battery-future-bright-but-caution-on-costs-20210917-p58sjj>
- [21] Tesla, “Order Megapack,” Last Accessed: 2022-01-21. [Online]. Available: <https://www.tesla.com/megapack/design>

- [22] Elexon, “BMRS API and Data Push Version 6.0,” Elexon, Tech. Rep., nov 2021, Last Accessed: 2022–01–20. [Online]. Available: <https://www.elexon.co.uk/documents/training-guidance/bmc-guidance-notes/bmrs-api-and-data-push-user-guide-2/>
- [23] GPy, “GPy: A gaussian process framework in python,” <http://github.com/SheffieldML/GPy>, 2012.
- [24] Energy Superhub Oxford, “Battery energy storage,” Last Accessed: 2022–01–18. [Online]. Available: <https://energysuperhuboxford.org/technologies/battery-energy-storage/>
- [25] Pivot Power, “Pivot Power battery first tertiary connection to export onto the Grid,” Last Accessed: 2022–01–18. [Online]. Available: <https://www.pivot-power.co.uk/news-and-events/pivot-powers-cowley-battery-first-tertiary-connection-to-export-onto-the-grid/>
- [26] EDF Renewables, “Battery Storage,” Last Accessed: 2022–01–16. [Online]. Available: <https://www.edf-re.uk/battery-storage>
- [27] National Grid ESO, “Who we are,” Last Accessed: 2022–01–18. [Online]. Available: <https://www.nationalgrideso.com/who-we-are>
- [28] A. Bourn, “Elexon Data Portal,” <https://github.com/OSUKED/ElexonDataPortal>, 2021.
- [29] Elexon, “BMRS Reports Guide,” Last Accessed: 2022–01–20. [Online]. Available: <https://www.bmreports.com/bmrs/?q=help/overview>

A Appendices

A.1 Balancing Mechanism

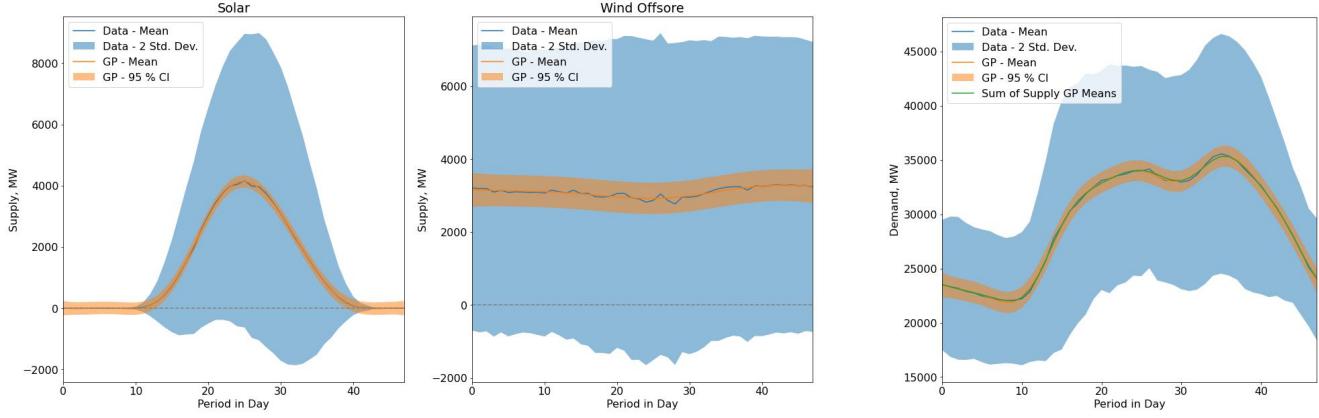
Balancing of the National Grid must happen constantly; supply and demand need to be exactly matched to keep system voltage and frequency stable [7][8][18]. The UK's Electricity System Operator (ESO), National Grid ESO, is responsible for taking actions to ensure the safe and efficient movement of power across the high-voltage transmission network [4][27][18]. For the purposes of trading and settlement (e.g., determining the amount of electricity used/generated and arranging payment for it), electricity is considered to be generated, transported, delivered and used in half hour chunks called Settlement Periods. Electricity Suppliers forecast in advance what the demand from their customers will be during a Settlement Period and contract with Generator(s) for that volume of electricity. Generators are expected to generate and deliver their contracted volume of electricity, and Suppliers are expected to consume their contracted volume of electricity. However, Grid imbalances can occur when: 1) Suppliers forecast their electricity requirements incorrectly, 2) a Generator is unable to generate the contracted amount, 3) and/or there are problems transporting electricity. To handle such scenarios, the ESO has a range mechanisms available to balance the system, a key one of which are Bid-Offer acceptances. An Offer Price indicates the price a Generator wants to be paid per MWh for an increase in generation or a Supplier wants to decrease demand, and a Bid Price indicates the price a Generator is willing to pay per MWh for a decrease in generation or an Supplier to increase demand. The ESO will, in real-time and as required, match supply and demand in each half hour by accepting Bids or Offers. Afterwards, metered volumes are collected for the half hour from Generators and Suppliers, and compared against their contracted volumes (which are adjusted for any Bids or Offers accepted). A not-for-profit entity called Elexon is responsible for working out a price for the difference and transferring funds appropriately – they oversee the strategic operation and day-to-day management of the Balancing and Settlement Code (BSC) and publish public information relating to the electricity market via the Balancing Mechanism Reporting Service (BMRS) [13][18].

A.2 Supply and Demand Models

Data from the BMRS was used to create models that could be employed by miniSCOT's supply and demand forecasting services during each timestep of the simulation [28][22]. A highly simplified approach was taken to derive models that provided sufficient functionality for this initial project. For supply, FUELHH and B1620 sources were used, which report average generation, in MW, for every 30 minute Settlement Period, broken down into the following fuel types: solar, onshore wind, offshore wind, hydro pumped storage, hydro run-of-river and poundage, biomass, fossil gas, fossil oil, fossil coal, nuclear and interconnector (i.e., flows to and from Europe) [29]. Data for the period 01/01/2020 - 31/12/2020 was used. For each fuel type, an average for each 30 minute period in the day across the full year was calculated and a GP model with RBF covariance function fit to the 48 resulting datapoints using GPy [23]. The initial lengthscale was set to 1 (i.e. a single Settlement Period), as was the variance. Additive Gaussian noise was assumed. Figure 8a shows the mean values and interval representing two standard deviations obtained for solar and offshore wind generation, along with the GP mean. Details of the GP fits are provided in table 4. The GP models had relatively small noise variance values, which is not surprising given the smoothness of the data, but is not reflective of the variance occurring intraday. However, the large data variance shown in Figure 8a is mainly reflective of seasonal variability. To keep things simple, the noise standard deviation was fixed to 10 % of the mean, yielding the orange 95 % confidence intervals in Figure 8a. An advantage of our approach was that models reflected a "typical day", and so reduced the impact of the time-of-year and length-of-time the simulator was run for.

Fuel Type/ Demand Model	RBF 1 Variance	RBF 1 Lengthscale	RBF 2 Variance	RBF 1 Lengthscale	Gaussian Noise Variance	Fixed Gaussian Noise Variance	Log Likelihood
Solar	1083565	9	-	-	737	8470	-269.4
Wind Offshore	2534490	35740	-	-	2940	42479	-277.9
Wind Onshore	2534471	35741	-	-	1797	32003	-267.4
Nuclear	18823509	1164811	-	-	0	7269	-192.7
Fossil Oil	0	0	-	-	0	0	59.9
Fossil Gas	11110575	3	-	-	4165	201622	-337.2
Fossil Coal	127975	667	-	-	95	6088	-216.8
Hydro Run-of-River and Poundage	41183	1270	-	-	51	516	-204.7
Hydro Pumped Storage	23660	3	-	-	213	404	-245.4
Biomass	48971	3	-	-	26	4948	-215.7
Interconnector	2538113	9661870	-	-	4142	13619	-301.1
Demand - 1 RBF	422009042	10	-	-	149835	190926	-386.1
Demand - 2 RBF	150162190	19653539	60656361	5	91348	190926	-370.6

Table 4: GPy model fit results for supply and demand models. Note that fossil oil was only used twice in 2020 and so having a model that will always return 0 is suitable.



(a) Supply data for solar and offshore wind generation

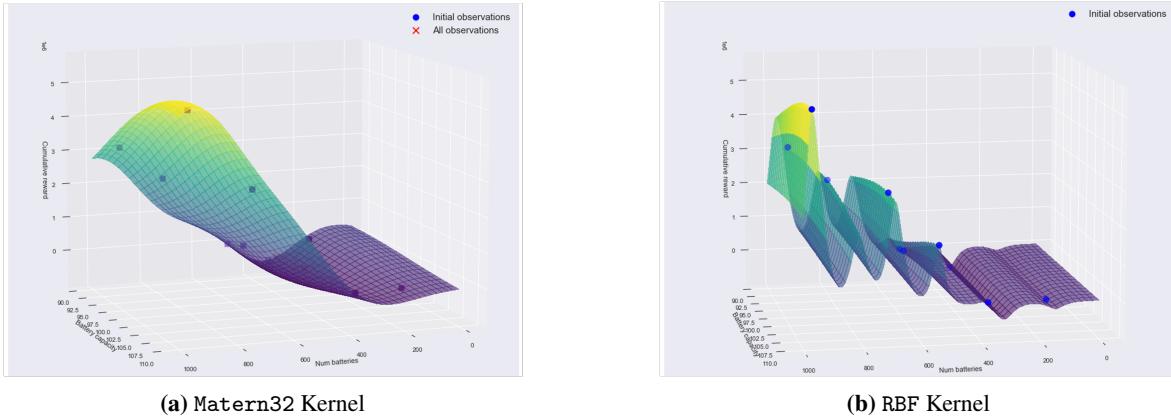
(b) Demand Data

Figure 8: Intraday supply and demand data and models. Demand data is the sum of all supply data.

A GP demand model was fit to the sum of the intraday mean values for all fuel types, with the noise standard deviation again fixed to 10 % of the mean. The data, GP fit, and sum of supply models is shown in Figure 8b. The details of the GP fit are again provided in 4.

When called, the supply and demand services load the appropriate GPy model and call the predict function for the given Settlement Period. The returned mean and variance values were used to sample from a univariate Gaussian. This provided an estimate for the average MW of supply/demand during the 30 minute period, which was multiplied by 0.5 to yield MWh values. Simply creating a lookup table of mean and variance values would have offered similar functionality to our final use of GP models, however our approach smoothed the mean values, allowed exploration of different settings, and enabled the models used by miniSCOT to be easily exchanged. Supply and demand imbalances were driven by the stochasticity of the models, and were highly reliant on the chosen Gaussian noise variance.

A.3 Kernels



(a) Matern32 Kernel

(b) RBF Kernel

Figure 9: GP fit to the same set of points with RBF and Matern32 kernels.

A.4 Acquisition Functions

Expected Improvement

The Expected Improvement acquisition function decides the next point x^* to evaluate that maximizes the improvement in reward the most in expectation, which is equivalent to selecting the point with the maximum expected utility, as per Equation 3.

$$\begin{aligned}
x^* &= \arg \max a_{\text{EI}}(x) \\
u(x) &= \max (0, f' - f(x)) \\
a_{\text{EI}}(x) &= \mathbb{E}[u(x) | x, \mathcal{D}] = \int_{-\infty}^{f'} (f' - f) \mathcal{N}(f; \mu(x), K(x, x)) df \\
&= (f' - \mu(x)) \Phi(f'; \mu(x), K(x, x)) + K(x, x) \mathcal{N}(f'; \mu(x), K(x, x))
\end{aligned} \tag{3}$$

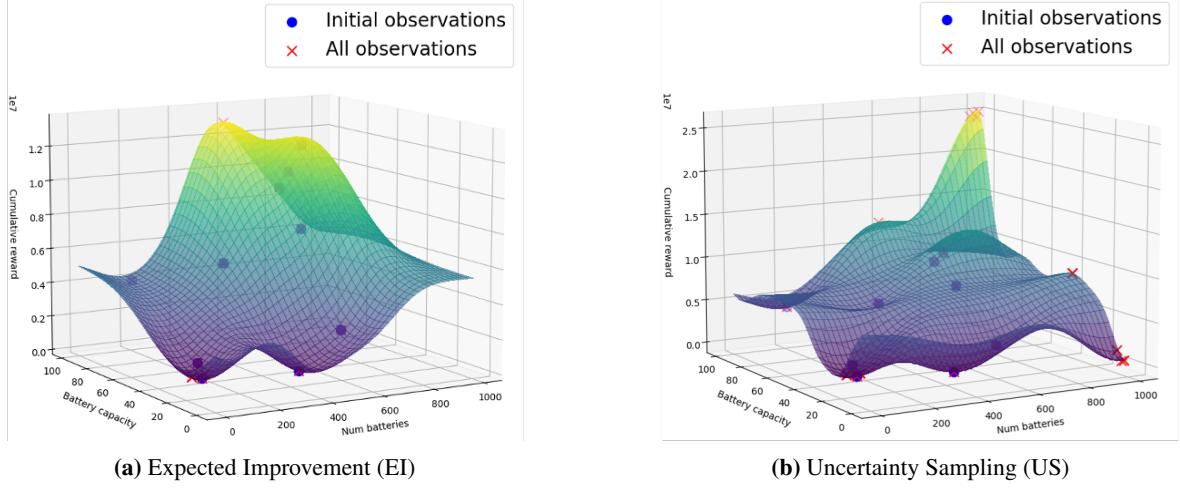


Figure 10: Exploration/exploitation trade-off of different acquisition functions.

A.5 GP Fits for 1D Bayesian Optimisation Experiments

Figure 4 in Section 5.1 shows two of the results of running Bayesian Optimisation on two of the five scenarios outlined in Table 1. The results from the other three scenarios; "25 % Discount", "33 % Discount +", and "50 % Discount +" are shown in Figure A.5. To achieve the fits for all five scenarios the setup described in section 4 was used. While the 'base' hyperparameters worked for most of the 1D experiments we ran (including many which are not included here), they did need to be tuned for some experiments. In particular, the length scale and ϵ (in the ϵ -stopping criterion) both had to be tuned (downward) to achieve GP fits that were not degenerate for certain scenarios.

A.6 Time Horizon Experiments

Table 5 shows the number of simulated points from running the BO loop under a fixed number of max iterations (and ϵ -stopping criterion for all but the baseline scenario) across various time horizons and scenarios. The EI acquisition function noticeably opts for exploiting rather than exploring under our surrogate GP model as it consistently converges to minimum reward improvements with a difference of less than $\epsilon = 0.01$ within 2 iterations. Figure 12 shows the 10 day time horizon GP fit (left) and its explored areas of the domain (right): the corner of $[0, 20]$ battery capacity and $[0, 200]$ batteries—the area where we expect the optimum capacity to be—remained unexamined. Figure 13 shows the GP model fit to the observed points using the "33% Discount" scenario across two different time horizons. The difference in order of magnitude of the cumulative reward (Y axes) of the plots (10^7 vs 10^8) highlights the variability in reward explained by the time horizon.

Number of Days	1	10	30	60	120
Baseline (20 % Discount)	25	25	25	25	25
25 % Discount	15	25	15	15	10
33 % Discount	10	10	15	15	10
33 % Discount +	15	10	20	10	10
50 % Discount	10	10	10	10	10

Table 5: Time Horizon - number of evaluated points (BO)

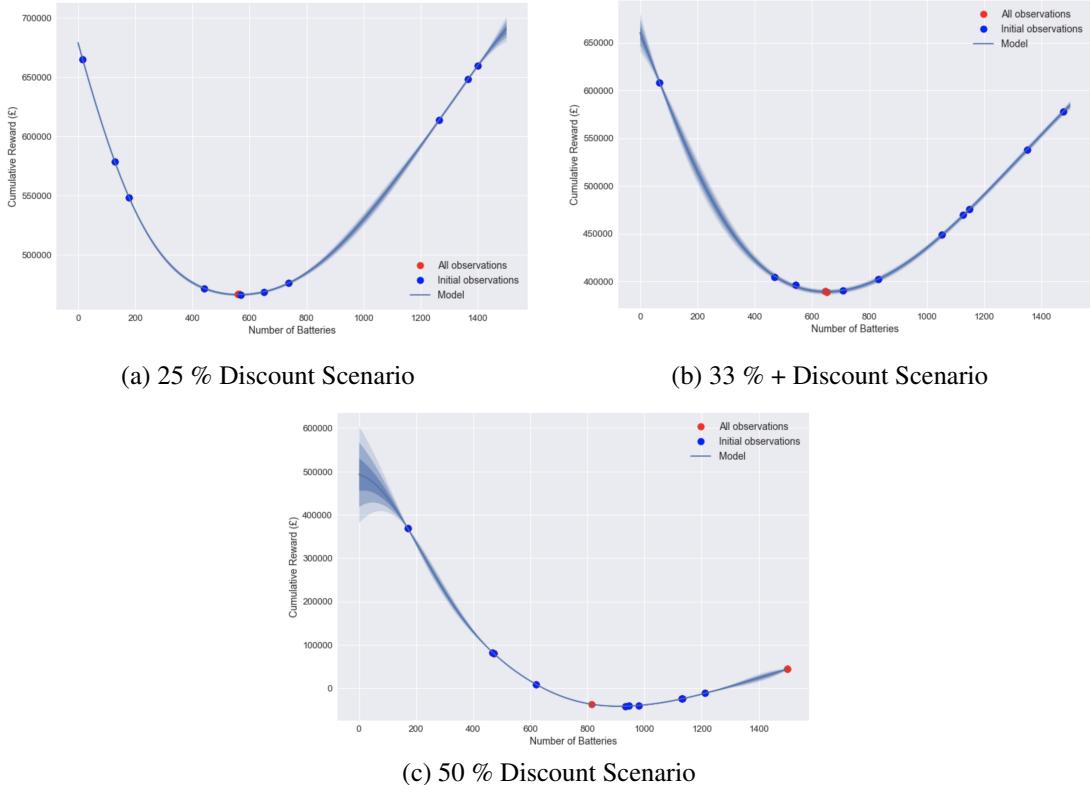


Figure 11: Results of running Bayesian Optimisation to find the optimal Number of Batteries for different scenarios, each using a fixed maximum battery capacity, C_{bat} , of 2MWh.

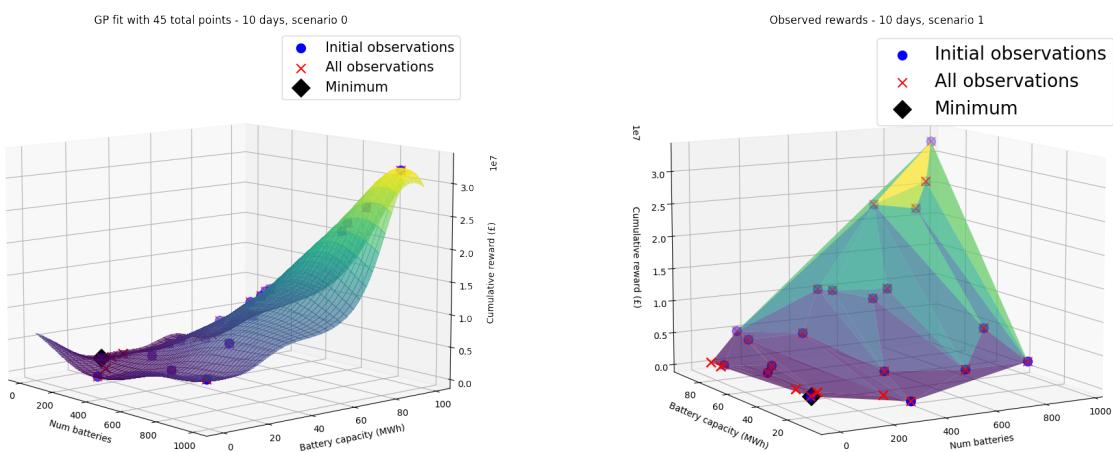


Figure 12: GP under explores domain - time horizon of 10 days, baseline scenario

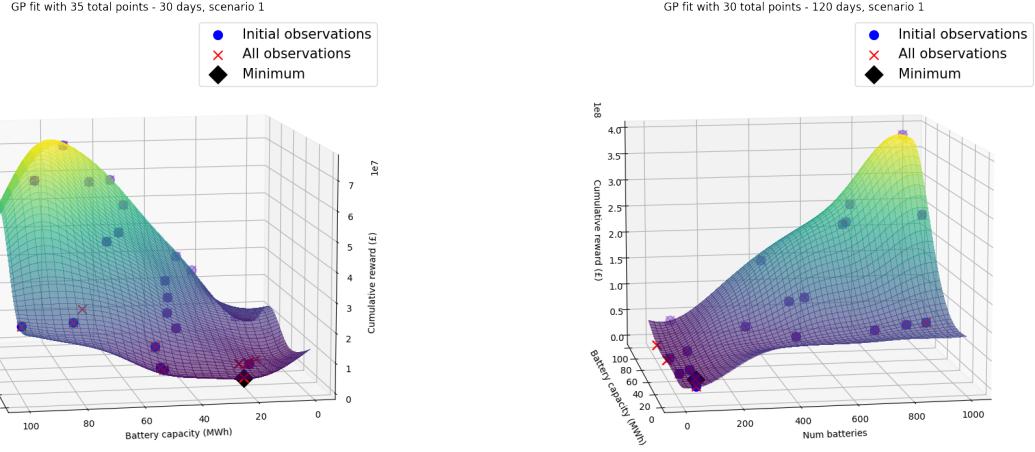


Figure 13: Reward magnitude difference between time horizons (30 days vs 120 days)

A.7 Joint Surge: Simulating Storms

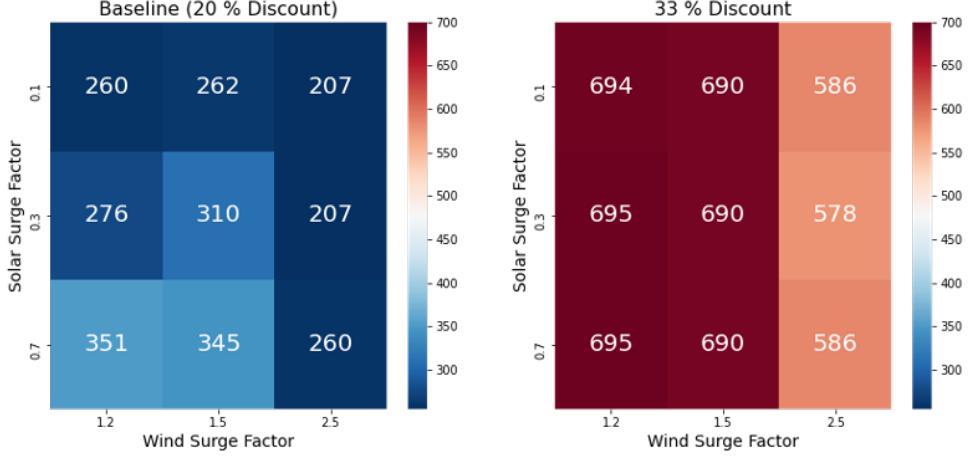


Figure 14: Optimal number of batteries to introduce into the Grid, under joint wind and solar surge (10 day time horizon).

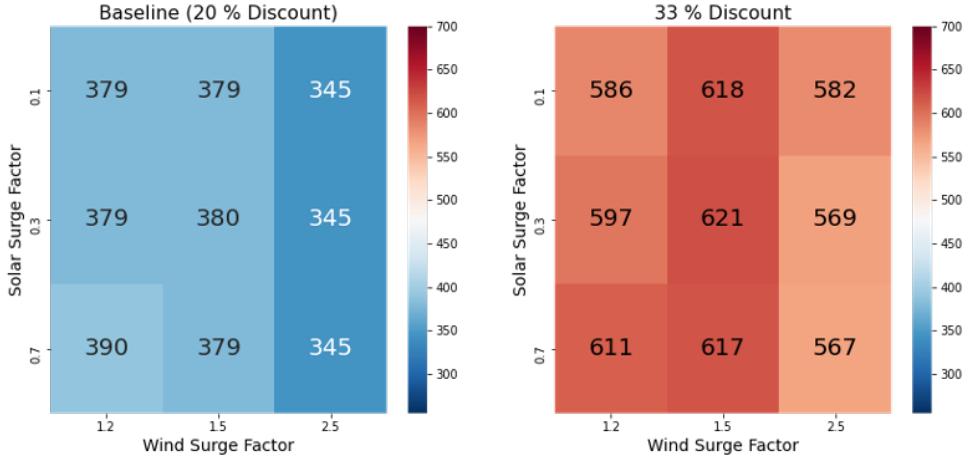


Figure 15: Optimal number of batteries to introduce into the Grid, under joint wind and solar surge (30 day time horizon).

To address the real-world interdependence between wind and solar supply, we simulated varying degrees of *joint* surges in wind and solar activity. Specifically, we modelled a 3 day increased off-shore wind supply (over Surge Factors {1.2, 1.5, 2.5})

during a simultaneous 7 day drop in solar supply (Surge Factors $\in \{0.1, 0.3, 0.7\}$). We chose such a setting given its parallels to possible storms that may occur; e.g., high winds coupled with thick cloud cover. To determine whether the optimal total capacity to introduce into the Grid is robust to such varying severity "storms," we conducted a similar 2x2 cross-product study (time horizon = [10 days; 30 days] x scenario = ["Baseline"; "33% Discount"]) similar to Section 7 – running BO with fixed C_{bat} at 2 MWh and varied N_{bat} .

For extended time horizons (30 days compared to 10 days - see Figure 15 vs. Figure 14), we find that, as in the univariate surge setting, with sufficient time the impact of the surge diminishes. However, we note that the baseline number of batteries in the no surge "*Baseline*" and "*33% Discount*" are approximately 420 and 720 batteries, respectively. We see that, while the impact of the surge has diminished after 30 days, even under the most mild joint surge modelled, the optimal number of batteries remain around 30 to 100 batteries lower, depending on the scenario. Further investigation is needed. Moreover, we emphasize that, as mentioned in Section 7, our simulation makes a number of assumptions that would need to be addressed before drawing concrete conclusions about such combined weather events.