Optimizing trading decisions of wind power plants with hybrid energy storage systems using backwards approximate dynamic programming

Benedikt Finnah,[*] Jochen Gönsch [†]

Mercator School of Management, University of Duisburg-Essen, Lotharstraße 65, 47057 Duisburg, Germany

Abstract

On most modern energy markets, electricity is traded in advance and a power producer has to commit to deliver a certain amount of electricity some time before the actual delivery. This is especially difficult for power producers with renewable energy sources that are stochastic (like wind and solar). Thus, short-term electricity storages like batteries are used to increase flexibility. By contrast, long-term storages allow to exploit price fluctuations over time, but have a comparably bad efficiency over short periods of time.

In this paper, we consider the decision problem of a power producer who sells electricity from wind turbines on the continuous intraday market and possesses two storage devices: a battery and a hydrogen based storage system. The problem is solved with a backwards approximate dynamic programming algorithm with optimal computing budget allocation. Numerical results show the algorithm's high solution quality. Furthermore, tests on real-world data demonstrate the value of using both storage types and investigate the effect of the storage parameters on profit.

Keywords: *renewable energy, hydrogen storage, hybrid energy storage system, backwards approximate dynamic programming, optimal computing budget allocation*

---

[*]Mail: benedikt.finnah@uni-due.de
[†]Corresponding author. Mail: jochen.goensch@uni-due.de, Tel.: +49-2033794369; Fax: +49-2033791760.

1

1 Introduction

Fossil resources are finite and their usage to produce electric energy causes CO_2 emissions. These are just two reasons why more and more countries strive to substitute conventional power plants with renewable energy sources. However, these environmental advantages play a subordinate role for investors and renewable energy sources must be able to compete with conventional energy sources in terms of profitability. In Germany, for example, the pioneering Renewable Energies Law (Erneuerbare-Energien-Gesetz) allowed to feed renewable energy into the grid at a fixed price, independent of the current supply and demand situation. As the share of renewables rises, this system is no longer viable and power producers are increasingly encouraged to sell their energy at new electricity markets. To date, these electricity markets exist in most developed countries. Due to technical properties of the power grid, energy is traded in advance. That is, power producers have to sell their energy some time (e.g., 30 minutes) before delivery.

However, production of renewable energy is usually stochastic (e.g., wind and solar). Thus, renewable power producers currently start to use energy storages like batteries to exploit daily and weekly price fluctuations and meet their delivery commitments. All batteries suffer from a certain self-discharge. For example, the widely-used sodium sulfur batteries lose up to $20\,\%$ of the energy stored per day and can only be used as short-term storages. A particularly large project is located in Aomori, Japan (NGK Insulators, 2019) and aims at leveling the output of a $51\,MW$ wind farm.

Another possibility to store electricity is electrolysis: An electrolyzer uses electricity to split water into hydrogen and oxygen. Later, the hydrogen can be used to generate electricity via a fuel cell. In addition – and this is the main difference to other electricity storage systems –, the hydrogen can be sold directly. Its applications include fuel for hydrogen cars, the production of ammonia for fertilizer and feed into the natural gas grid. This technology is used, for example, in Mainz, Germany (Energiepark Mainz, 2019). Such a hydrogen based energy storage (HES) can be used as long-term storage because it has nearly zero self-discharge. Although the electrolyzer and the fuel cell are expensive, the hydrogen storage itself is a simple and cheap gas tank. Accordingly, an HES is suitable to store a large amount of energy over a long period of time and can help, for example, to exploit seasonal effects of the electricity prices.

The problem considered in this paper is the following: The decision maker is a power producer with stochastic production (e.g., wind or solar) who aims to maximize the net present value of his or her revenue. The power producer trades energy in the continuous intraday market 30 minutes before delivery. To reduce the risk of energy shortages (e.g., due weather changes) and to increase his or her flexibility, the power producer uses a battery as short-term storage and an HES as long-term storage. While the battery can only be charged or discharged, the hydrogen in the HES can be used to produce electricity again or be sold directly. At each point in time, the power producer has to decide on the amount of energy to sell or buy, the amount of hydrogen to sell, and where to store excess energy (if any). The planning horizon should consist of several months (i.e., we will use one year in the numerical study) because the HES as a long-term storage can exploit the ups and downs in the energy prices during the different seasons.

The structure of the considered virtual power plant (combination of production and storages) is visualized in Figure 1. A formal description of the decision problem can be found in Section 4. To show the different behavior of the power producer in different seasons and to show the benefits from using two storage technologies, we consider a planning horizon of one year. The common modeling approaches for such a stochastic decision problem with timely distributed decisions are multi-stage stochastic programming and dynamic programming. Because the computational burden of multi-stage stochastic programs increases exponentially in the number of decision stages, we model the problem as a dynamic program. For dynamic programs, the computational burden increases only linearly. We consider the constraints of all components (e.g., maximum power of the electrolyzer or maximum charge/discharge rates of the battery). Since the proposed dynamic program cannot be solved

exactly and the state, action, and outcome spaces are multidimensional and continuous, we use approximate dynamic programming (ADP) to solve this problem heuristically. More precisely, we use a backwards ADP (BADP) algorithm where the values of the post-decision states are approximated with a linear architecture of basis functions.
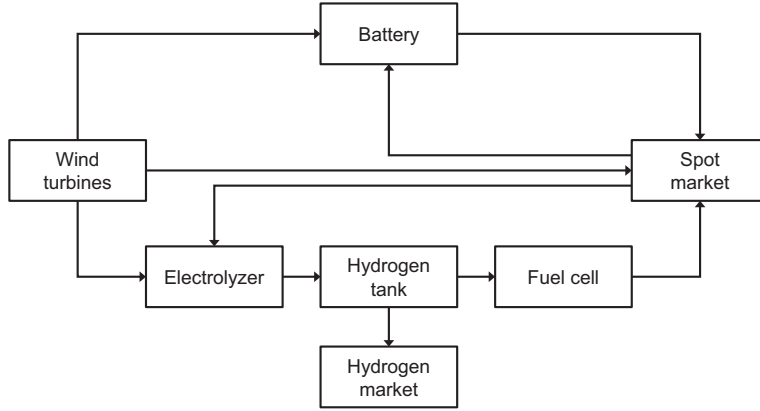

Figure 1: Structure of the power plant

We make the following contributions to the existing literature: (1) From an application point of view, we model the decision problem of a profit maximizing wind power producer with a hybrid energy storage system as a dynamic program and solve it with a BADP approach. (2) From a methodological perspective, we compare different norms and sample weights to fit the basis functions and we present an optimal computing budget allocation (OCBA) algorithm for these approximations and show that such an OCBA can enhance the expected profit. (3) From a managerial perspective, we investigate the synergy effects between the long-term and short-term storages. In doing so, we also show that a joint optimization and operation of the wind turbine and the storages outperforms a separated optimization.

The remainder of this paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the considered energy market and introduces the assumptions we use to model reality for our research. The mathematical model and the formal optimization problem are given in Section 4. In Section 5, we present our BADP algorithm and its variants. More precisely, we introduce the OCBA which is a rule how to sample the states during the BADP algorithm. Section 6 contains the numerical study while Section 7 concludes the paper. The appendix contains an overview of the notation and parameter values and the validation of the proposed algorithm based on a simplified problem.

2 Literature review

The optimal management of a power plant with stochastic output combined with an energy storage is a highly active research field (e.g., van der Klauw et al. (2017), Schwarz et al. (2017)). Multiple literature reviews are dedicated to this topic (e.g., Fathima and Palanisamy (2015), Rahman et al. (2015), Weitzel and Glock (2018), or the references in Lust and Waldmann (2018)). Therefore, we restrict this literature review to the two key topics relevant for our contributions to distinguish this paper from the literature. First (Subsection 2.1), the profit maximizing optimization of a stochastic power producer combined with multiple energy storages (hybrid energy storage systems). Second (Subsection 2.2), BADP algorithms for solving energy storage problems.

3

2.1 Power plant with hybrid energy storage systems

The most recent literature review is Weitzel and Glock (2018). They categorize the literature, among others, based on storage types, planning and forecast horizon, mathematical model type and solution method. Multiple storage types are rarely considered in the literature. Weitzel and Glock (2018) find only 16 contributions with two or more storage types. Only six of them use an HES. In this section, we review these six papers.

Among these six, four consider microgrids without a connection to a larger grid. Accordingly, the power producer tries to satisfy the demand of the microgrid at minimal cost. Electricity and hydrogen cannot be bought or sold, and, thus, the HES works similar to a battery on an abstract level. The electricity is generated at least from solar and wind energy, and costs may arise, for example, due to wear of the system's components or consumption of diesel fuel to generate electricity. In particular, Dufo-López et al. (2007) optimize the configuration as well as the control of the power plant with genetic algorithms. García et al. (2013) optimize the control of their power plant using a fuzzy logic controller. Safari et al. (2013) study a similar problem and use a fuzzy logic controller with a particle swarm optimization algorithm. Trifkovic et al. (2014) maximize a weighted sum of the renewable energy used and the stored hydrogen penalized with startup costs of the system's components. To do so, they employ a two-layer model predictive controller.

Two papers consider a grid connection. Naraharisetti et al. (2011) consider a grid-connected power plant with multiple storages including an HES. An electricity and a hydrogen demand exist. The deterministic problem is modeled as a mixed integer linear program. Only Garcia-Torres and Bordons (2015) explicitly consider stochastic supply. Their microgrid includes solar and wind power and three different storage types including an HES. Again, cost is minimized using a multi-layer model predictive controller and a rolling horizon approach is used.

To the best of our knowledge, we are the first to capture the decision problem of a profit maximizing power producer with multiple storage technologies in a dynamic program.

2.2 Backwards approximate dynamic programming literature

Traditional ADP algorithms, which approximate the value function, traverse the state space forwards in time. By contrast, recently four BADP algorithms were developed in the context of energy storage problems. They mimic backwards recursion in going through the state space backwards in time. We review these four algorithms we are aware of. All of these approximate the value function of the pre-decision states and a dedicated analytic derivation or numerical calculation of a decision's expected value is necessary. The tractability of this approach depends on the problem and the architecture of the value function approximation (VFA). Because we approximate the value functions of the post-decision states we do not need such an explicit calculation of the expectation. Instead, the expectation is implicitly numerically calculated by fitting the approximation to a multitude of observations.

Hannah and Dunson (2011) present a BADP algorithm for their electricity storage problem. They only evaluate a discrete set of pre-decision states, where the expected value is numerically estimated. For this a weighted sum of the value function of the next period is used, where the weights are approximations to the transition probabilities. If a decision leads to a non discretized state, the value function is linearly interpolated.

Gönsch and Hassler (2016) present an approximate policy iteration with forward and backward steps to tackle an electricity storage problem with one storage type. Their hybrid approximate policy iteration iteratively computes new VFAs. This solves the problem of the large state space, but the expected value still has to be computed for every action and sampled state. They analytically derive the expectation and the optimal decisions.

Cheng et al. (2017) use a BADP algorithm for their discretized storage problem. The algorithm works like classical backwards recursion, except that not every state is evaluated, only a small sample is. They use a

low-rank approximation to fit the VFAs. Analogous to Gönsch and Hassler (2016), the expected values have to be computed explicitly for every decision. Considering the computational burden, this is only realistic for bigger instances if the expected value can be derived analytically for every state and decision, but no analytic derivations are mentioned.

Durante et al. (2020) formulate their BADP algorithm for discrete problems only. It also works like classical backwards recursion. A small fraction of the pre-decision states is evaluated to approximate the value function of these. After that every post-decision state is evaluated by computing the expected value of the VFA. Again, an analytic representation of the expected value is needed for every post-decision state. They briefly mention different methods to fit the approximation architecture, but apparently use the standard *l2*-norm without weights.

Another approach related to BADP is the approximate dual dynamic programming algorithm proposed by Löhndorf et al. (2013). This algorithm contains a forward pass and a backward pass to approximate the value function of the post-decision state with hyperplanes. During the forward pass, the approximate dual dynamic programming algorithm explores new states. During the backward pass, the algorithm revisits the states from the forward pass. For each state, it iterates over all states in the state space that share the same endogenous parts, i.e. that only differ in the exogenous information part to derive new hyperplanes. Therefore, this approach is only suitable for discrete or discretized exogenous information.

Summed up, to the best of our knowledge, no paper considers multiple norms or a rule for the distribution of samples (like OCBA) in the context of BADP.

3 Market structure and assumptions

In this section, we first (Subsection 3.1) explain the considered market framework. More precisely, we introduce the continuous intraday market and the balancing by the transmission system operator. Then (Subsection 3.2), we state all assumptions used to capture reality in the mathematical model formulation.

3.1 Market structure and continuous intraday market

Like in most countries, in Germany multiple energy markets coexist. All markets have in common that the energy is traded some time before the actual delivery, which is especially challenging for power producers with stochastic supply (e.g., wind or solar). For an overview and analysis of different electricity markets see, e.g., Koschker and Möst (2015), Bublitz et al. (2019), and the references therein. On the continuous intraday market, the market closure is closest to the physical fulfillment of the trades (30 minutes). This makes this market interesting for power producers with stochastic supply, as, for example, short-term weather forecasts are more reliable than long-term forecasts.

At the continuous intraday market, trading for delivery throughout the following day starts at 16:00. The products are commitments for electricity delivery in a specific time slot and can be traded traded up to 30 minutes before physical delivery. The continuous intraday market works similar to well-known stock markets, so traders can observe the current price while trading but large trading volumes can influence the market prices. Even if it is possible to trade products multiple hours ahead, the products are most liquid right before market closure. In the continuous intraday market, hourly, half-hourly and quarter-hourly products coexist. As the market is efficiency, these products are largely arbitrage free.

If the physical delivery/consumption of a power producer or consumer deviates from the commitment on the continuous intraday market (or, strictly speaking, the net commitment from all markets), these imbalances are balanced by the transmission system operator. Afterwards, the market participants are paid/billed for their over- or underproduction (or -demand) based on the current demand and supply situation. Since the transmission

5

system operator aims to reduce the risk of high energy excess and lack in the grid, the transmission system operator forbids power producers and consumers to intentional deviate from the commitments. Therefore, power producers have to fulfill their commitments as close as possible, even if it is more costly than the balancing by the transmission system operator.

3.2   Assumptions

Modelling inherently requires abstraction, generalization, and simplification. Also this is often done only more or less implicitly in the literature on energy storages, we explicitly state our assumptions that are the base for the mathematical model giiven in Section 4. Please note that we do not describe the entire setting, but concentrate on issues where we simplify reality to arrive at a tractable model. For each issue, we discuss why we think the respective assumption is reasonable, e.g. because it is only a slight deviation from reality (market rules etc.) and standard in the literature. Some simplify the model, others reduce the computational burden of the solution approach. Physical constraints of the components (e.g., maximum discharging rate) that are captured in the model will be introduced directly in Section 4 together with the formulae.

Assumption 1. *The power producer trades only in the continuous intraday market.*

Most authors consider single markets settings (e.g., Jiang et al. (2014), Hassler (2017), Cai et al. (2019), Cruise et al. (2019), Vivero-Serrano et al. (2019), Durante et al. (2020)). Since the continuous intraday market is the last opportunity to trade, it is the most interesting for power producers with stochastic supply. Therefore, we focus on this short-term market.

Assumption 2. *The power producer trades the hourly products right before market closure to the volume weighted average price.*

In most countries, only hourly products can be traded in a short-term market like the continuous intraday market. To keep our research comparable with the majority of the literature regarding modeling and solution techniques, we restrict ourselves to the hourly products. Since the hourly prices are close the the average prices of the substituting products (quarter-hourly or half-hourly) of the particular hour, this assumption is reasonable. Additionally we use the common assumption that the product is only traded right before market closure (e.g., Jiang et al. (2014), Cheng et al. (2017), Durante et al. (2020)).

Like in stock markets and in contrast to other energy markets (e.g., day-ahead market and intraday auction market), the traders in the continuous intraday market trade to individual market prices. Since we cannot reflect these individual trades (e.g., because of missing data and model complexity), we assume that the trades are priced at the published volume weighted energy prices. Since most volume is traded right before market closure, we think that (on average) the market price of the considered power producer is close to the volume weighted energy price.

Assumption 3. *The power producer is a price-taker.*

Like in stock markets, large trades in the continuous intraday market can influence the market prices. We follow the majority of the literature and consider a power producer which is small enough to be a price-taker (e.g., Löhndorf and Minner (2010), Kim and Powell (2011), Beraldi et al. (2018), Heredia et al. (2018)). This means that the small power producer cannot influence the market prices.

Assumption 4. *The prices at the continuous intraday market follow a Markov process.*

According to this assumption, the price process is memoryless and future energy prices depend only on the most recent market price. This assumption is in line with the literature (e.g., Löhndorf and Minner (2010), Cheng et al. (2017), Gönsch and Hassler (2016), Wozabal and Rameseder (2020)).

6

Assumption 5. *At any time, the power production during the next hour is known.*

For example, short-term weather forecasts are much more accurate than long-term forecasts. Therefore, the short-term energy production is reliable, in contrast to future productions. This is again in line with the literature (e.g., Jiang et al. (2014), Durante et al. (2020), Wozabal and Rameseder (2020)).

Assumption 6. *Hydrogen can be sold instantaneous to a known price.*

For hydrogen, there is no market like for electricity, but hydrogen can be fed into the natural gas grid. This is done by several German natural gas providers. Since short-term natural gas prices fluctuate significantly less than electricity prices, it can be assumed that these are known or at least very reliable to forecast.

Assumption 7. *If energy must be delivered, the demand will be satisfied from the three sources available according to the following action preference rule: (1) currently produced energy, (2) energy from the battery, and (3) energy from the hydrogen storage.*

To reduce the complexity of the decision problem, some decisions are governed by a straightforward heuristic action preference rule. If energy must be delivered, the demand will be satisfied preferably with currently produced energy. If this is not sufficient, the battery will be discharged. In case even more energy is needed, the fuel cell is used to reconvert hydrogen. If this still does not provide enough energy, balancing will take place. Similar rules are common for settings with one storage device (e.g., Kim and Powell (2011), Gönsch and Hassler (2016)). In addition, this decision rule adheres to the requirements of the transmission system operator. More precisely, our power producer tries to fulfill the commitments even if this is financially unattractive.

Assumption 8. *The power producer is risk-neutral.*

Since we assume that the considered power producer is risk-neutral, we maximize the expected revenue. The power producer trades hourly (8760 times per year) and the hourly returns are only a small fraction of the investment cost. Therefore, the risk (e.g., standard deviation) of the single hourly returns is less interesting. Especially in the dynamic programming literature, this assumption is common (e.g., Jiang et al. (2014), Cheng et al. (2017), Gönsch and Hassler (2016), Durante et al. (2020)).

4 Model formulation

In this section we formally model the power producer's decision problem as a dynamic program. To do so, in Section 4.1, we define the decision stages including the temporal resolution and the sequence on events. In Sections 4.2 to 4.6, we explain how we choose the five standard ingredients of the dynamic program: decision variables, exogenous information, state variables, transition function, and contribution function. Based on this, we formally state the power producer's optimization problem in Section 4.7. The notation is summarized in Table A1 in the appendix.

4.1 Sequence of events

In this section, we describe the temporal resolution of the considered decision problem as well as the sequence of events during the decision stages. This sequence of events is illustrated in Figure 2.

Figure 2: Sequence of events

The optimization horizon is discretized into T decision stages. These decision stages are indexed with t. Since the power producer trades the hourly products on the continuous intraday market right before market closure (Assumptions 1 and 2), there is a decision each hour and a period corresponds to one hour.

At the beginning of time period t, the power producer observes the state S_t and bases the decision x_t on this. After this, the power producer has to deliver (or store) the traded energy. The delivery to the energy grid is not related to the current decision, but to the commitment x_{t-1}^E, which was decided in the previous period. In addition, the power producer delivers the hydrogen and stores the excess energy according to the recent decisions x_t^H and $x_t^{B,H}$. The post-decision state S_t^x is located right after the energy and hydrogen is delivered or stored. Immediately afterwards, the power producer observes the new exogenous information. More precisely, he or she observes the new electricity price on the continuous intraday market and the transmission system operator's penalties for over- and underproduction. The hydrogen price and the next period's energy production also become known. The exogenous information transits the post-decision state into the new state S_{t+1}.

As already briefly mentioned in the introduction, dynamic programming is particularly suitable for the problem at hand. Contrary to multi-stage stochastic programming, computational complexity increases only linearly in the length of the time horizon. This is important as energy production and energy prices show seasonal effects (see Section 6) and because the HES is a long-term storage. Thus, we need to consider a long time horizon consisting of several seasons as the long-term storage can store hydrogen e.g., from spring till autumn. Accordingly, we consider a planning horizon of one year in the numerical study. Shorter horizons, e.g., only a few weeks, can not capture the long-term advantages of the HES. However, we would like to stress that this does not imply assuming a deterministic future or forgoing advantages of reoptimization: the longer the planning horizon in a dynamic program, the better, as end-of-horizon effects matter less. The strength of dynamic programming is that when solving the optimization problem for a whole year we anticipate future outcomes, new information and adjust the hourly decisions accordingly. And because the computational burden increases only linearly, solving one large optimization problem is as hard as solving multiple smaller problems.

4.2 Decision variables

In every period t, three decisions have to be made. Thus, the decision vector $x_t = \left(x_t^E, x_t^H, x_t^{B,H} \right)$ consists of three components:

8

- x_t^E: amount of electricity (MW) to trade with delivery during period $t+1$ ($x_t^E > 0$ sale, $x_t^E < 0$ purchase)

- x_t^H: amount of hydrogen (MWh) to sell with instant delivery at the end of period t

- $x_t^{B,H}$: share of excess energy to store in the battery during period t

According to Assumptions 1 and 2, the power producer trades hourly products on the continuous intraday market. Therefore only one decision for the electricity trading per period is needed. The traded energy x_t^E has to be fulfilled (delivered or received) during period $t+1$. Please note that the decision x_t^H is only necessary because the HES allows to sell "stored energy" directly without reconversion to electricity. In line with Assumption 6, the hydrogen is delivered instantaneous in period t. The decision $x_t^{B,H}$ is the share of excess energy to store in the battery during period t. The excess energy may be energy produced in period t or purchased electricity decided in period $t-1$. Following Assumption 7 other possible decisions are governed by a decision rule.

4.3 Exogenous information

The exogenous information

$$W_t = \left(P_t^E, Q_t^{E,u}, Q_t^{E,o}, P_t^H, Y_t^E \right) \tag{1}$$

becomes known during period t right before the decision x_t is made. P_t^E is the price (€/MWh) for electricity during period $t+1$, which is already observable during period t. P_t^H is the price (€/MWh) for hydrogen in period t. This becomes known based on Assumption 6. $Q_t^{E,u}$ is the penalty (€/MWh) for electricity shortage and $Q_t^{E,o}$ the penalty price (€/MWh) for electricity overproduction in period t. Based on Assumption 5, the electricity production Y_t^E becomes known. The model and the solution approaches described in Section 5 are independent of the price processes. Thus, the implemented processes for the exogenous information W_t are described in Section 6.1.3 with the data for the numerical study.

4.4 State variable

The state S_t is an exhaustive description of the system in period t right after the exogenous information W_t becomes known. On an intuitive level, a variable needs to be included in the state if it is necessary to make future decisions. More formally, each variable that influences the system's evolution (like commitments regarding storage levels) or contributes to the objective function (like past penalty prices) must be included. Thus, the state contains all information necessary to make the decision x_t and model the future evolution of the system. Here, the state of the system is given by

$$S_t = \left(R_t^B, R_t^H, x_{t-1}^E, W_t \right) = \left(R_t^B, R_t^H, x_{t-1}^E, P_t^E, Q_t^{E,u}, Q_t^{E,o}, P_t^H, Y_t^E \right) \tag{2}$$

To model the future evolution of the system, the state obviously contains the storage levels R_t^B and R_t^H which denote the amount of energy (MWh) stored in the battery and the hydrogen tank at the beginning of period t, respectively. The electricity commitment for period t, x_{t-1}^E, and the electricity production during period t, Y_t^E, have a direct influence on the storage levels at the beginning of the next period and are therefore necessary. The state contains the next electricity price P_t^E because the power producer can already observe it and obviously bases the decision on the next trading volume on it. Moreover, P_t^E must be included as it is needed for the electricity price process (Assumption 4). The decision is also influenced by the currently known hydrogen price P_t^H. Irrelevant for the future evolution, but necessary to evaluate the contribution function defined in Section 4.6, are the current penalty for shortage $Q_t^{E,u}$ and the penalized electricity price for overproduction $Q_t^{E,o}$. Please note that in general, dynamic programs are modeled such that the state variable has as few dimensions

9

as possible, which is not the case here. However, the above pre-decision state definition leads to a minimal post-decision state variable, which is the only one that matters for our solution approach (see Section 5.1).

4.5 Transition function

The transition function S_t^M defines the state of the next period S_{t+1} if decision x_t is made in state S_t and the exogenous information W_{t+1} arrives. In the transition function, we capture all characteristics of the power plant's components. More precisely, we take the capacity of the storages (maximum storage level) and the rated capacity (maximum power) of the battery, electrolyzer, and fuel cells into consideration.

$$S_{t+1} = S_t^M \left(S_t, x_t, W_{t+1}\right) = \left(R_{t+1}^B \left(\cdot\right), R_{t+1}^H \left(\cdot\right), x_t^E, W_{t+1}\right) \tag{3}$$

, where the new storage levels are defined by

$$R_{t+1}^B \left(S_t, x_t\right) = \left(1 - \gamma^B\right) \cdot \left(R_t^B + \Delta t \cdot \rho_{in}^B \cdot y_{t,in}^B \left(\cdot\right) - \frac{\Delta t}{\rho_{out}^B} \cdot y_{t,out}^B \left(\cdot\right)\right) \tag{4}$$

$$R_{t+1}^H \left(S_t, x_t\right) = \left(1 - \gamma^H\right) \cdot \left(R_t^H + \Delta t \cdot \rho_{in}^H \cdot y_{t,in}^H \left(\cdot\right) - \frac{\Delta t}{\rho_{out}^H} \cdot y_{t,out}^H \left(\cdot\right) - x_t^H\right) \tag{5}$$

with the electric power flows in and out of the storages described by the functions

$$y_{t,in}^B \left(S_t, x_t\right) = \begin{cases} \min \left\{x_t^{B,H} \cdot \left(Y_t^E - x_{t-1}^E\right), \frac{R_{max}^B - R_t^B}{\Delta t \cdot \rho_{in}^B}, y_{in,max}^B\right\} &, Y_t^E \geq x_{t-1}^E \\ 0 &, \text{else} \end{cases} \tag{6}$$

$$y_{t,out}^B \left(S_t\right) = \begin{cases} \min \left\{x_{t-1}^E - Y_t^E, R_t^B \cdot \frac{\rho_{out}^B}{\Delta t}, y_{out,max}^B\right\} &, Y_t^E < x_{t-1}^E \\ 0 &, \text{else} \end{cases} \tag{7}$$

$$y_{t,in}^H \left(S_t, x_t\right) = \begin{cases} \min \left\{\left(1 - x_t^{B,H}\right) \cdot \left(Y_t^E - x_{t-1}^E\right), \frac{R_{max}^H - R_t^H}{\Delta t \cdot \rho_{in}^H}, y_{in,max}^H\right\} &, Y_t^E \geq x_{t-1}^E \\ 0 &, \text{else} \end{cases} \tag{8}$$

$$y_{t,out}^H \left(S_t\right) = \begin{cases} \min \left\{x_{t-1}^E - Y_t^E - y_{t,out}^B \left(S_t\right), R_t^H \cdot \frac{\rho_{out}^H}{\Delta t}, y_{out,max}^H\right\} &, Y_t^E < x_{t-1}^E \\ 0 &, \text{else} \end{cases} . \tag{9}$$

Here, $\rho_j^i, i \in \{B, H\}, j \in \{in, out\}$ denotes the charge/discharge efficiency of the storages. The parameters $y_{j,max}^i$ represent the maximum power of the electrolyzer, the fuel cell and the battery and R_{max}^i the capacity of the storages. Note that the discharge of the storages does not depend on the current decision but on the previous decision x_{t-1}^E, which is included in the state. The length of a time period Δt is needed for the conversion of energy (MWh) to power (MW). The hydrogen is sold after the energy is converted. The self-discharge factors γ^i denote the fraction of stored energy lost by storage i over one period and are applied at the end of each period to the current storage levels.

4.6 Contribution function

The contribution function $C_t(S_t, x_t)$ describes the one-stage profit, that is, the net revenue the power producer obtains in period t if the power producer takes decision x_t in state S_t.

$$
\begin{aligned}
C_t(S_t, x_t) = {}& P_t^H \cdot x_t^H + \Delta t \cdot P_t^E \cdot x_t^E - \Delta t \cdot q^{gf} \cdot \max\left\{0, -x_t^E\right\} \\
& - \Delta t \cdot Q_t^{E,u} \cdot \max\left\{x_{t-1}^E - Y_t^E - y_{t,out}^B - y_{t,out}^H, 0\right\} \\
& + \Delta t \cdot Q_t^{E,o} \cdot \max\left\{Y_t^E - x_{t-1}^E - y_{t,in}^B - y_{t,in}^H, 0\right\}
\end{aligned}
\tag{10}
$$

C_t includes the hydrogen sales and the revenue of the electricity commitment for delivery in the next period which are decided in period t. If the power producer buys energy, the power producer has to pay the grid fees q^{gf}. In the case of electricity shortage in the current period, he or she has to pay the penalty $Q_t^{E,u}$. Likewise, if the electricity deliveries exceed the commitment x_{t-1}^E, the power producer obtains the penalized electricity price $Q_t^{E,o}$. Since the electricity commitment x_{t-1}^E was decided in period $t-1$, the over- and underproduction is independent of the decision in period t. According to the price-taker assumption (Assumption 3), the power producer does not influence the market price.

4.7 Problem statement and value function

Obviously, the decision x_t should depend on the current state S_t. Thus, we define a policy π as a function that maps every state to a decision: $X_t^\pi(S_t)$. Now, we have all necessary notation and can formally state the optimization problem faced by the power producer who maximizes the discounted (discount factor β) profit over an optimization horizon of length T.

$$
\max_\pi \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \cdot C_t(S_t, X_t^\pi(S_t)) \,|\, S_0\right]
\tag{11}
$$

$s.t.$

$$
\begin{aligned}
& S_{t+1} = S_t^M(S_t, x_t, W_{t+1}) && \forall t && (12) \\
& x_t^E \in \left[x_{min}^E, x_{max}^E\right] && \forall t && (13) \\
& x_t^H \in \left[0, R_t^H + \Delta t \cdot \rho_{in}^H \cdot y_{t,in}^H(\cdot) - \frac{\Delta t}{\rho_{out}^H} \cdot y_{t,out}^H(\cdot)\right] && \forall t && (14) \\
& x_t^{B,H} \in [0,1] && \forall t && (15)
\end{aligned}
$$

Following the risk-neutrality assumption Assumption (8), the objective function (11) maximizes expected profit discounted with a factor of β. The hourly discount factor has two effects. First, this reflects that returns in the near future are more important than returns in far future (e.g., these can be reinvested). Second, the standard deviation of the hourly returns increases in time (more precisely, the trend increases). Since many decision makers are slightly risk-averse, this is a simple method to include some risk- aversion. Therefore the discounting weakens Assumption 8.

Four groups of constraints have to be respected. First, the state transition is governed by the transition function. Second, the maximum electricity sale x_{max}^E reflects, for example, the maximum power of the power line connecting the producer to the grid. The lower bound x_{min}^E restricts the power producer while buying. For example, the power producer cannot buy more energy than the summed maximum charging powers of the storages. Third, the selling amount of hydrogen x_t^H has to be non-negative because buying hydrogen is not considered and cannot exceed what is available. Forth, the share of energy to store in the battery $x_t^{B,H}$ must obviously be in $[0,1]$. All additional storage constraints are already captured by the transition function. To

simplify the notation, we left out the straightforward definition of the initial state S_0.

An equivalent formulation of this problem is the Bellman equation

$$V_t\left(S_t\right) = \max_{x_t} C_t\left(S_t, x_t\right) + \beta \cdot \mathbb{E}\left[V_{t+1}\left(S_{t+1}\right) | S_t\right] \tag{16}$$

with the boundary conditions

$$V_T\left(S_T\right) = 0 \ \forall S_T. \tag{17}$$

Obviously, the same constraints have to be respected.

5 Backwards approximate dynamic programming

In this section, we present how we solve (12) to (17) using ADP. First, we reformulate the value function (16) using post-decision states (Section 5.1). Please note that the BADP algorithm presented in the remainder of this section is problem independent. It works on post-decision states, but there is a straightforward reformulation for every dynamic program. In Section 5.2, we explain the structure of the BADP algorithm developed. Then, we turn to the details of the regression for the VFA and discuss the norms, weighted regression, and the OCBA. The OCBA is a rule, how to evaluate the states during the BADP algorithm. In the numerical study, we show that these uneven allocation of the computing budget can enhance the BADP algorithm.

The notation of this section is also given in Table A2 in the appendix.

5.1 Reformulation using post-decision state

The standard approach to model a dynamic program is using the pre-decision value function (16). Accordingly, ADP usually approximates $V_t(S_t)$, the value of a pre-decision state, that is, the state of the system immediately before a decision is taken. By contrast, the post-decision states S_t^x describe the state of the system immediately after a decision is made, but before new exogenous information becomes known. In this paper, we approximate the values $V_t^x(S_t^x)$ of the post-decision states. This has an important computational advantage because the optimal decisions can be derived without an explicit calculation of the expected value for every decision (see, e.g., (Powell, 2011, Ch. 4.6)). With the post-decision state defined as

$$S_t^x\left(S_t, x_t\right) = \left(R_{t+1}^B, R_{t+1}^H, x_t^E, P_t^E\right) \tag{18}$$

the post-decision value function is given by

$$V_t^x\left(S_t^x\right) = \mathbb{E}\left[\max_{x_{t+1}}\left(C_{t+1}\left(S_{t+1}, x_{t+1}\right) + \beta \cdot V_{t+1}^x\left(S_{t+1}^x\right)\right) | S_t^x\right] \tag{19}$$

An additional advantage of this formulation is that the post-decision state has four dimensions less than the pre-decision state. Because the produced energy is already delivered or stored and the hydrogen is sold, Y_t^E, $Q_t^{E,u}$, $Q_t^{E,o}$ and P_t^H are not necessary anymore. The current commitment x_t^E is necessary, for example, to calculate future storage levels. Please note that it cannot be eliminated from the post-decision state as this calculation necessitates both x_t^E and the realization of production Y_{t+1}^E. We cannot eliminate P_t^E from the post-decision state as we need this for the Markovian price process (Assumption 4). For dynamic programs in general, the post-decision state space can be larger than the space of the pre-decision states.

For the approximation of the value function of the post-decision states V_t^x, we split x_t^E into a positive and negative part and choose a second degree polynomial. Because the positive times negative part of x_t^E is always

12

0 this results in $21 - 1 = 20$ coefficients/parameters for the regression. The resulting VFA is denoted by $\bar{V}_t^x$.

$$
\begin{aligned}
\bar{V}_t^x \left(S_t^x \right) = {} & \theta_1 \left(R_{t+1}^B \right)^2 + \theta_2 R_{t+1}^B R_{t+1}^H + \theta_3 R_{t+1}^B x_t^{E,pos} + \theta_4 R_{t+1}^B x_t^{E,neg} + \theta_5 R_{t+1}^B P_t^E + \theta_6 R_{t+1}^B \\
& + \theta_7 \left(R_{t+1}^H \right)^2 + \theta_8 R_{t+1}^H x_t^{E,pos} + \theta_9 R_{t+1}^H x_t^{E,neg} + \theta_{10} R_{t+1}^H P_t^E + \theta_{11} R_{t+1}^H \\
& + \theta_{12} \left(x_t^{E,pos} \right)^2 + \theta_{13} x_t^{E,pos} P_t^E + \theta_{14} x_t^{E,pos} + \theta_{15} \left(x_t^{E,neg} \right)^2 + \theta_{16} x_t^{E,neg} P_t^E + \theta_{17} x_t^{E,neg} \\
& + \theta_{18} \left(P_t^E \right)^2 + \theta_{19} P_t^E + \theta_{20}
\end{aligned}
\tag{20}
$$

with $x_t^{E,pos} = \max \left\{ x_t^E, 0 \right\}$ and $x_t^{E,neg} = \min \left\{ x_t^E, 0 \right\}$ and coefficients θ_i. This allows to capture the interactions between the four dimensions of the post-decision state space.

5.2 BADP algorithm

The value functions are approximated using the BADP algorithm as follows (see Algorithm 1 for the pseudo-code). First, the VFA of the post-decision states in the terminal period is initialized according to the boundary conditions (17) (Step 1). Then, BADP goes backwards in time (Steps 2 – 14) and fits the VFA $\bar{V}_t^x$ for each time period t (Step 13). More precisely, in each time period, it samples M post-decision states (Step 4) and computes the average $\bar{v}_{m,t}^x$ over N evaluations for each state (Steps 5 – 11). The optimal decisions of the quadratic problem in Step 8 are computed via the Karush-Kuhn-Tucker conditions (not shown here). As we consider a multitude of constraints of the system's components (see Section 4), the optimal solution may be at a border and a large number of cases (45) has to be considered.

Algorithm 1 BADP

1: Set $\bar{V}_T^x = 0$
2: **for** $t = T - 1$ to 0 **do**
3:      **for** $m = 1$ to M **do**
4:          sample post-decision state $S_{m,t}^x$
5:          **for** $n = 1$ to N **do**
6:              sample exogenous information $\omega_{t+1} \in W_{t+1}$
7:              compute next pre-decision state $S_{m,t+1}$
8:              $x_{t+1}^* := \underset{x_{t+1}}{\arg\max} \left(C_{t+1} \left(S_{m,t+1}, x_{t+1} \right) + \beta \cdot \bar{V}_{t+1}^x \left(S_{t+1}^x \left(S_{m,t+1}, x_{t+1} \right) \right) \right)$
9:              $v_{m,t}^{x,n} := C_{t+1} \left(S_{m,t+1}, x_{t+1}^* \right) + \beta \cdot \bar{V}_{t+1}^x \left(S_{t+1}^x \left(S_{m,t+1}, x_{t+1}^* \right) \right)$
10:          **end for**
11:          $\bar{v}_{m,t}^x := \frac{1}{N} \cdot \sum_{n=1}^{N} v_{m,t}^{x,n}$
12:      **end for**
13:      compute regression $\bar{V}_t^x$ using $\bar{v}_{m,t}^x, m \in \{1, ..., M\}$
14: **end for**

Please note that the BADP algorithm primarily aims at learning the value functions to derive a policy (decision for each state), which is done analogous to Step 8 during operations. Although the VFAs can be close to the true values, they can obviously not be used for evaluation. In the context of ADP, the value of the policy learned is usually estimated by simulation.

5.3 Computing regression for the BADP algorithm

In this section we present multiple methods to compute the regression in Step 13 of the BADP algorithm. Section 5.3.1 surveys different norms for the regression, while Section 5.3.2 enhances these by a weighted regression.

In Section 5.3.3 we introduce our OCBA. This is a rule, how often the different states in the BADP algorithm have to be evaluated.

5.3.1 Norms for the regression

In this section, we survey the BADP algorithm outlined above with two norms for the computation of the regression of the value functions V_t^x (Step 13): the $l1$-norm and the $l2$-norm, but every other norm could be used as well. The $l2$-norm regression which also known as ordinary least squares (OLS) is the standard approach in ADP to fit VFAs.

A well-known shortcoming of OLS is the high sensitivity of the $l2$-norm to outliers. The values $\bar{v}_{m,t}^x$ obtained by the BADP algorithm and approximated in the regression are stochastic, so poor values $\bar{v}_{m,t}^x$ for a few post-decision states can considerably worsen the results. To reduce this problem, we also consider the $l1$-norm to compute the parameters of the regression.

$l1$-norm: Regressions with respect to the $l1$-norm, minimizes the sum of the absolute errors, that is, the differences between the values to be approximated and the resulting approximation. For M values to be approximated and p regression coefficients, the $l1$-approximation is a linear program with $2M + p + 1$ decision variables and $M + 1$ constrains. This program may be large but has a sparse coefficient matrix and can be solved by commercial solvers or with efficient algorithms dedicated to this problem structure (for the linear programming formulation and an efficient algorithm see, e.g., Barrodale and Young (1966)).

$l2$-norm: OLS minimizes the squared sum of the errors. The regression coefficients can be computed with the Gaussian normal equations. For M values to be approximated and p regression coefficients, the computational burden is a $p \times M$ to $M \times p$ matrix multiplication and a $p \times p$ equation system has to be solved.

5.3.2 Weighted approximations

In this section, we explain how we use weighted norms to fit the VFA. An important assumption for OLS to derive reliable regression parameters is the homoscedasticity. So if the variance of the regression errors is not nearly constant (heteroscedasticity) OLS can result in sub-optimal VFAs and therefore lead to sub-optimal policies. To tackle this problem, we use state-depending weights. Most forwards ADP algorithms try to visit the most promising area of the state space more often. The resulting uneven distribution of the evaluated states has a similar effect as a weighted approximation. By contrast, we evaluate the states multiple times to compute weights, which is only suitable for BADP algorithms. We use a general approach which is applicable to other problems as well. However, we evaluate it only in the context of the energy production problem considered in this paper.

For this approach, we also estimate the variance $\sigma_{m,t}^2$ (Step 11) of the estimator $\bar{v}_{m,t}^x$ for $V_t^x\left(S_{m,t}^x\right)$ and minimize the weighted errors with respect to the corresponding norm. For the weighted approximations, we choose the weights $\frac{1}{cv_{m,t}^2} = \frac{\bar{v}_{m,t}^2}{\sigma_{m,t}^2}$ (without normalization) with coefficient of variation $cv_{m,t}$ as they performed well in preliminary tests. This approach assumes that states with low coefficient of variation $cv_{m,t}$ are more reliably estimated than those with a higher coefficient of variation $cv_{m,t}$ and therefore should influence the regression more.

5.3.3 Optimal computing budget allocation

In the previous variants each post-decision state sampled in Step 4 is evaluated equally often (N times). However, to estimate the values $V_t^x\left(S_{m,t}^x\right)$ of all sampled post-decision states equally well, states with higher coefficient of variation need more evaluations. The OCBA algorithm is inspired by the OCBA algorithms for ranking and

selection. There, the aim is usually to maximize the probability to choose the best alternative (see, e.g., Chen et al. (2000), Jia (2012) and Al-Salem et al. (2017)), which leads to the allocation of the majority of simulations to a few alternatives. By contrast, our goal is an equally good estimation for all states, using a total of $N \cdot M$ simulations. Thus, the number of simulations $n_{m,t}$ allocated to the m-th state sampled in period t is chosen such that the summed squares of the ratios of the confidence levels' lengths $2z_{(1-\alpha)}\frac{\sigma_{m,t}}{\sqrt{n}}$ and the estimated mean $\bar{v}_{m,t}^x$ are minimized:

$$\operatorname*{arg\,min}_{n_{1,t},\cdots,n_{M,t}} \sum_{m=1}^{M} \left(2z_{(1-\alpha)} \frac{\sigma_{m,t}}{\sqrt{n_{m,t}} \cdot \bar{v}_{m,t}^x} \right)^2 = \operatorname*{arg\,min}_{n_{1,t},\cdots,n_{M,t}} \sum_{m=1}^{M} \frac{cv_{m,t}^2}{n_{m,t}} \tag{21}$$

This is performed as follows. First, all M states are evaluated n^{lb} times and the coefficients of variation $cv_{m,t}$ are estimated. Then, the following constrained problem can be formulated, which we solve heuristically:

$$\min_{n_{1,t},\cdots,n_{M,t}} \sum_{m=1}^{M} \frac{cv_{m,t}^2}{n_{m,t}} \tag{22}$$

$s.t.$

$$\sum_{m=1}^{M} n_{m,t} = N \cdot M \tag{23}$$

$$n_{m,t} \geq n^{lb} \qquad\qquad\qquad \forall m \tag{24}$$

$$n_{m,t} \in \mathbb{N} \qquad\qquad\qquad \forall m \tag{25}$$

If the integer constraints (25) are ignored, this is a continuous nonlinear resource allocation problem, which belongs to a well-researched problem class (see, e.g., Bretthauer and Shetty (2002), Patriksson and Strömberg (2015), van der Klauw et al. (2017)). Moreover, as the resource allocation problem (22–24) has a strictly convex objective function (22), is only bounded from below (24), and only has one linear constraint (23), the problem can be solved efficiently. We solve the continuous problem exactly with a pegging algorithm, which iteratively fixes infeasibly low decision variables to their lower bounds. For this the unbounded solution $n_{m,t} = \frac{cv_{m,t} \cdot N \cdot M}{\sum_{m'=1}^{M} cv_{m',t}}$ $\forall m$ is used. Finally, we round to the next integer solution.

In the literature, OCBA algorithms often ignore the lower bounds (24) to simplify the problem (e.g., Chen et al. (2000), Jia (2012), and Al-Salem et al. (2017)). Thus, too many simulations are allocated and the simulation budget is not respected properly. For our problem $3 - 5\,\%$ too many simulations would result from ignoring the lower bounds. While the effect on overall runtime is arguably very small, an identical number of simulations for all variants is the foundation for the numerical comparisons. After all samples are allocated, we use weighted approximations with weights $\frac{n_{m,t}}{cv_{m,t}^2}$ (without normalization). These weights reflect that states with more evaluations are more reliable.

Obviously it is possible to allocate not all $N \cdot M$ simulations at once but step by step to exploit that the estimate for $cv_{m,t}$ improves as more simulations are allocated and performed. For this the lower bounds n^{lb} would be state depending to account for the simulations already performed.

6 Numerical Study

This numerical study is structured as follows. First, we introduce the parameter values of the power plant and the stochastic processes in Section 6.1. Based on this we compare the BADP algorithm with the different regression techniques in Section 6.2. After this, we use the best approach to investigate the behavior of the power plant over the course of the year in Section 6.3.

6.1 Data for the numerical study

Subsection 6.1.1 states the time horizon. While Section 6.1.2 describes the parameters of the power plant's components, the parameters for the exogenous processes are defined in Section 6.1.3. The used notation is summarized in Table A3 and the parameter values are summarized in Table A4 in the appendix.

6.1.1 Time horizon

Energy production and energy prices are well known to show seasonal effects, which may be exploited by the long-term storage HES. To take advantage of HES' benefits, we thus need to consider a long time horizon consisting of several seasons as the long-term storage can store hydrogen e.g., from spring till autumn. Accordingly, we consider a planning horizon of one year ($T = 8760$). Shorter horizons, e.g., only a few weeks, can not capture the long-term advantages of the HES. As already discussed in Section 4, we would like to stress that this does not imply assuming a deterministic future or forgoing advantages of reoptimization. For the discount factor we choose $\beta = 1.05^{\frac{-1}{8760}}$ which corresponds to 5% per year. This is based on a realistic return on equity.

6.1.2 Parameters of the power plant's components

As there is to date no project which jointly uses an HES and another storage device, we base the configuration considered on representative projects that solely use an HES or a battery. For example, the German project RedoxWind combines a $2\,\text{MW}$ wind farm with a $20\,\text{MWh}$ redox-flow battery (Fraunhofer Institute for Chemical Technology, 2019). Larger projects are located in Japan, where a $51\,\text{MW}$ wind farm is combined with a $245\,\text{MWh}$ sodium sulfur (NaS) battery (NGK Insulators, 2019). These batteries have a roundtrip efficiency of 70 to 85% and a daily self-discharge up to 20% (e.g., Luo et al. (2015), Zhao et al. (2015)). The Energiepark Mainz in Germany uses three $2\,\text{MW}$ electrolyzers with a $26\,\text{MWh}$ storage for a $8\,\text{MW}$ wind park (Energiepark Mainz, 2019). An HES possesses a roundtrip efficiency of 35 to 65%, but it has almost no self-discharge (e.g., Zeng and Zhang (2010), Luo et al. (2015), Zhao et al. (2015)).

In the case study, we consider a small wind farm with a rated capacity of $r = 20\,\text{MW}$. This corresponds to $4-8$ wind turbines. We also assume a connection to the grid with a maximum transmission capacity of $x_{max}^E = 25\,\text{MW}$. The sodium sulfur battery has a capacity of $R_{max}^B = 20\,\text{MWh}$ with symmetrical efficiencies of $\rho_{in}^B = \rho_{out}^B = \sqrt{0.75}$ and a self-discharge factor of $\gamma^B = 0.00925$, which corresponds to 20% per day. The maximum powers of the battery are $y_{in,max}^B = y_{out,max}^B = 5\,\text{MW}$. We model an HES with an electrolyzer and a fuel cell with $y_{in,max}^H = y_{out,max}^H = 5\,\text{MW}$, respectively. With $\rho_{in}^H = 0.8$, the electrolyzer has a higher efficiency than the fuel cell with $\rho_{out}^H = 0.6$. Compared to the other components, a hydrogen tank is cheap. Thus, we use a large hydrogen storage with a capacity of $R_{max}^H = 300\,\text{MWh}$ to exploit the higher electricity prices at the end of the year. As an HES has nearly zero self-discharge, we set $\gamma^H = 0$. While the power producer buys energy the power producer acts like a consumer and has to pay the obligatory grid fees (up to $30\,\text{€/MWh}$). However, because energy storages stabilize the grid, they only pay reduced fees. This reduction is a case by case decision (Sterner and Stadler, 2017, Ch. 15.4). We choose rather low fees of $q^{gf} = 5\,\text{€/MWh}$.

6.1.3 Exogenous price and wind processes

In line with Assumption 4, we model the electricity price as a Markov process. More precisely, we model the electricity price P_t^E as an Ornstein-Uhlenbeck process

$$P_{t+1}^E := \kappa \cdot P_t^E + (1 - \kappa) \cdot \mu_t + \tilde{P}_{t+1}^E, \quad \tilde{P}_{t+1}^E \sim N\left(0, \sigma^E\right), \quad (26)$$

where the price tends to revert to the mean price μ_t and the noise $\tilde{P}_{t+1}^E$ is normally distributed. Those autoregressive models are widely used to describe electricity prices in the context of trading problems (e.g.,

Löhndorf and Minner (2010), Kim and Powell (2011), Wozabal and Rameseder (2020)). For the time dependent mean price μ_t, we consider dependence on hour, weekday, month (triangular splines), and a trend. We use OLS to estimate the corresponding parameters with the prices of the German continuous intraday market of the years 2014 to 2018 (EPEX SPOT SE, 2019). Based on assumption Assumption 5 we used the volume weighted average price.

For the exogenous production of the wind turbines we use hourly wind data from a weather station in Hamburg-Billbrock in Northern Germany of the years 2012 to 2016 scaled to 100 m hub height. We choose Hamburg because many wind parks are located in its surroundings. In line with the literature (e.g., Shi et al. (2012), Gönsch and Hassler (2016), Biswas et al. (2017), Li et al. (2019)) we compute the production of the wind turbine based on Weibull distributed wind speeds.

We estimate time dependent Weibull distributions as follows:

- For each month and for each hour, we fit a Weibull distribution using all data of these month/hour combinations.

- For every hour, we linearly interpolate the mean and standard deviation of the monthly Weibull distributions to obtain daily distributions. For this we locate the monthly distributions in the middle of the month.

- Derive the time dependent scale parameter λ_t and shape parameter k_t based on the interpolated means and standard deviations.

To compute the hourly production Y_t^E from a wind speed ws_t, we use the production function of Gönsch and Hassler (2016):

$$Y_t^E = Y^E(ws_t) = \begin{cases} 0 & , ws_t < s_{ci} \ \ \text{or} \ \ ws_t \geq s_{co} \\ a + b \cdot ws_t^3 & , ws_t \in [s_{ci}, s_r) \\ r & , ws_t \in (s_r, s_{co}] \end{cases} \tag{27}$$

The properties of the wind turbine are based on four XEMC Darwind XE128-5MW turbines with a hub height of 100 m. The maximum production of $r = 20$ MW is generated at wind speeds above $s_r = 12$ m/s. The wind turbines do not produce energy if the wind speed is below the cut-in speed of $s_{ci} = 3$ m/s or above the cut-out speed of $s_{co} = 25$ m/s. The auxiliary parameters a and b are given by $a + b \cdot s_{ci}^3 = 0$ and $a + b \cdot s_{cr}^3 = r$.

Furthermore, we assume a constant hydrogen price of $P_t^H = 30$ €/MWh. This is approximately 50 % above the average natural gas price (2017 – 2018) at the EEX (2019). Such a higher hydrogen price is reasonable because several German natural gas providers charge their environmentally-conscious customers extra for this so-called wind gas.

In Germany, there is no electricity market after the continuous intraday market and the over-/underproduction is balanced by the transmission system operator. Afterwards, the power producer is paid/billed for the over-/underproduction where the penalty for shortage $Q_t^{E,u}$ is equal to the penalized price for overproduction $Q_t^{E,o}$. These penalties are so-called cross-control area balancing energy prices (reBAP, for German "regelzonenübergreifender einheitlicher Ausgleichsenergiepreis" (TransnetBW, 2019)). The reBAP is on average as high as the average electricity price. So power producers have nearly no financial incentive to fulfill their commitments. However, the transmission system operator forbids to intentional deviate from the commitments. Thus, power producers can not optimize their trading decisions based on the reBAP. To reflect this, we use the penalties for intentional deviation from the balancing group contract of the year 2006 (Bundesnetzagentur, 2006). Under this former regime, power producers were not paid for overproduction ($Q_t^{E,o} = 0$) and in case of underproduction had to pay the double price of the particular hour ($Q_t^{E,u} = 2P_{t-1}^E$). This assumption is

17

in line with the literature (e.g., Löhndorf and Minner (2010), Gönsch and Hassler (2016)). Other authors use penalties, which linearly depend on the electricity price (e.g., Kim and Powell (2011), Moghaddam et al. (2013), Kuznetsova et al. (2015), Hassler (2017)).

In Section 6.4, where we compare different settings, we evaluate these with the average reBAP of the years 2017 and 2018 ($Q_t^{E,o} = Q_t^{E,u} = 38.40 \text{€/MWh}$).

6.2 Numerical evaluation of the BADP algorithm

In this section we specify the parameters of the different regression techniques for the BADP algorithm. Additionally, we compare the different approaches with an intuitive heuristic. All approaches are implemented using Matlab version R2017a and run simulations on a standard PC with a 2.9 GHz Intel Core i5 processor and 6 GB of RAM.

- **$l1(M, N)$:** This is the BADP algorithm with the $l1$-norm as described in Section 5.3.1. For this and the following BADP approaches, the M post-decision states are sampled using a space-filling design (Sobol sequence) and each state is evaluated N times. Note that $l1(M, 1)$ is the standard approach that evaluates every state once. To guarantee comparability of the results, every BADP approach performs $M \cdot N = 5000$ evaluations of post-decision states per period.

- **$l2(M, N)$:** Analogously, this BADP approach uses the standard $l2$-norm as described in Section 5.3.1.

- **$l1$-$w(M, N)$:** In this BADP approach, the weighted $l1$-norm is used as described in Section 5.3.2. After preliminary tests we choose $M = 250$ and $N = 20$ for both weighted variants.

- **$l2$-$w(M, N)$:** Analogously, this BADP approach uses the weighted $l2$-norm (see again Section 5.3.2).

- **$l1$-$OCBA(M, N)$:** This is a BADP approach with OCBA and the $l1$-norm (Subsection 5.3.3). Here, N denotes the average number of evaluations per sampled post-decision state. After preliminary test we choose $n^{lb} = 0.9 \cdot N$ for both OCBA variants, so 90 % of the samples are allocated before we allocate the last 10 % using the OCBA algorithm.

- **$l2$-$OCBA(M, N)$:** Analogously, this is a BADP approach with OCBA and the $l2$-norm (see again Section 5.3.3).

- **ssl-$heuristic$:** If the electricity price is above an electricity price threshold, this benchmark approach offers the energy in the battery plus a constant offer parameter. If the price is below the threshold, this heuristic buys energy based on the free space in the battery minus the same offer parameter. All hydrogen except a given safety stock level (ssl) is sold. In the last period of the optimization horizon, all hydrogen and the maximum electricity is sold. To parametrize this heuristic, the optimal safety stock level, the electricity price threshold and the offer parameter are computed with the Matlab patternsearch algorithm. We are confident to obtain good parameter combinations, however, a more sophisticated procedure would probably be much faster.

We do not implement a forwards ADP approach as a benchmark. Since the revenues and penalties of over- and underproduction in the contribution function (10) depend only on the state but not on the current action, the state's transition is more important than the revenues of the current action. This leads to bad decision of forwards ADP algorithms and therefore to bad VFAs. Jiang et al. (2014) shows for an energy trading problem, that forwards ADP approaches have to exploit the monotony and concavity in the storage level to perform well.

Since the implemented approaches cannot solve the proposed dynamic program exactly, we validate our BADP algorithm in the Appendix B. Because there is no algorithm which can solve the dynamic program in an appropriate time, we validate the algorithm with a simplified and discretized version of our dynamic program.

In the following, we compare the approaches. Regarding days of the week for the price process, we use the calendar of the year 2018. Every policy is evaluated with 1,000 simulation runs. For each approach we use the same out of sample scenarios from our calibrated processes.

The length of the 99 % confidence intervals are below 0.5 % of the mean. The relative differences between the objectives, derived with the BADP variants, and the reported evaluations are below 15 %. As the policies do not depend on the VFA itself, but on the gradient, these differences are reasonable.
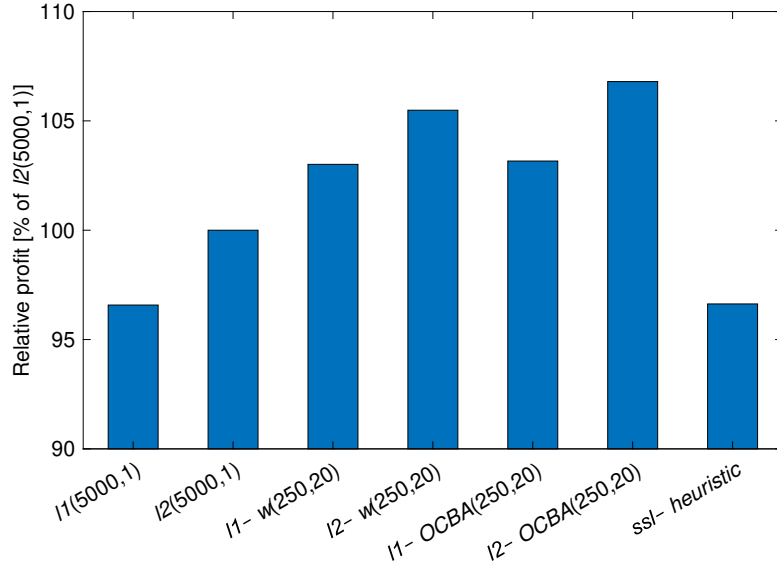
Figure 3: Comparison of variants on complete year

Figure 3 shows that the weighted approaches outperform the unweighted ones and that the OCBA variants further improve profit. All *l2* variants outperform their corresponding *l1* approach. *l1*(5000,1) is even not better than *ssl-heuristic*. Thus, the approximation with respect to the *l1*-norm is not suitable for the problem. As it obtains the highest profit, we use *l2-OCBA*(250,20) in the following.

The time to learn the policy is nearly the same for all BADP variants: approximately 5.5 h. Only *l2*(5000,1) and *l1*(5000,1) are a little bit slower, as the matrix multiplication and the linear program, that are used for the computation of the regression parameters, are larger. We tested the influence of M with these two approaches (not shown here) and choose $M = 5000$ as we observed only a small improvement in profit for higher values. For $M > 25,000$ there is no further improvement at all.

6.3   System behavior

In this section, we analyze the behavior of the power plant operating with a policy derived from *l2-OCBA*(250,20) over a period of one year. Again, the year 2018 was used for days of the week. First, Section 6.3.1 focuses on system behavior over the course of the year and discusses seasonal and weekly patterns. Then, Section 6.4 focuses on the importance of the different storage devices.

19

6.3.1 Behavior over the course of the year
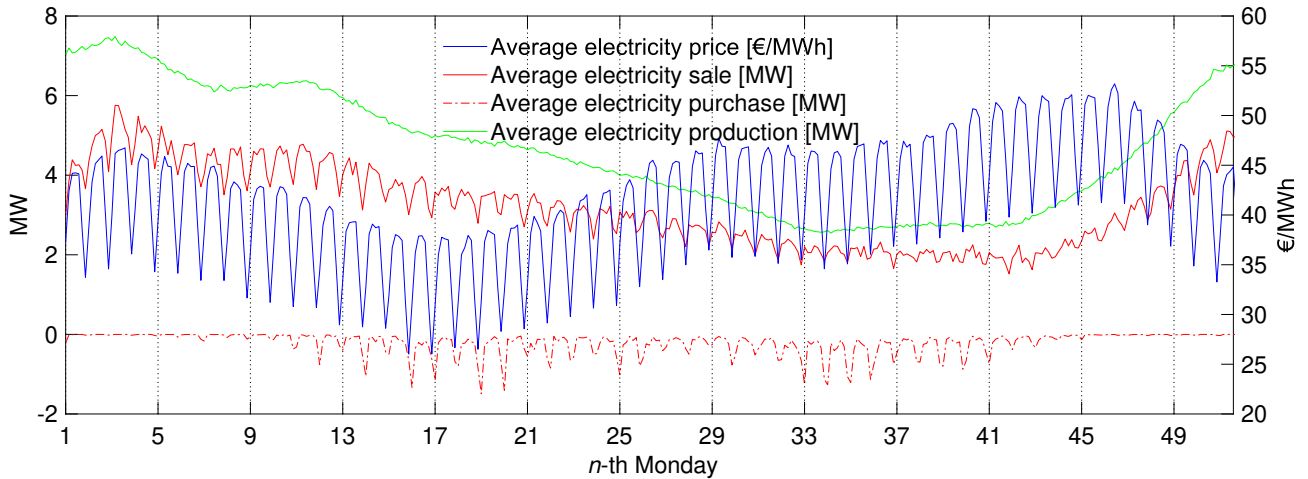
Figure 4: Electricity price, trades and production over the course of the year

Figure 4 shows the averages of the (simulated) electricity price, sales (on the positive vertical axis), purchases (on the negative vertical axis), and production over the course of the year (daily average). Since the values shown are averages over 1,000 simulations, the power producer appears to simultaneously sell and buy energy, which never happens in a single simulation run. Please note that actual electricity deliveries can fall short of sales in case of shortage, but this rarely occurs with more than $99.3\,\%$ of sales actually delivered.

As expected, the electricity price shows a strong weekly pattern, which is reflected in electricity sales. On weekends, prices are on average $16\,\%$ lower than on weekdays and $12\,\%$ less electricity is sold, but $44\,\%$ more electricity is bought.

There is also a clear yearly pattern. The price process successfully replicates the fact that electricity prices tend to be lower in the market region during spring. Likewise, average production in September and October is about only half compared to December and January. In months with lower production, sales are also lower. Sales are always constantly considerably lower than production. This is driven by the avoidance of shortages, losses in the system and hydrogen sales. The difference between production and sales is smaller in September and October when prices are high and production is low. Most electricity is purchased in months with low production (April till October). In the other months the wind turbine produces enough energy to reach the storages' maximum power inflow.

20
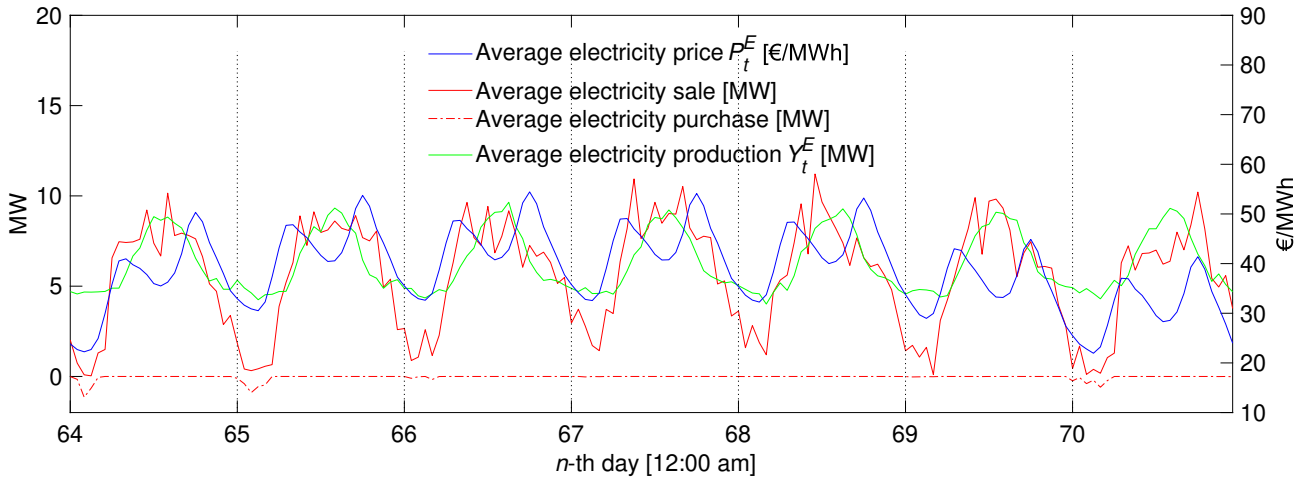
Figure 5: Electricity price, trades and production over the course of the 10th week

Figure 5 shows the (simulated) electricity price, trades, and production over the course of the 10th week. In the high price hours the average energy sales are as high as the average production. In the night hours much less energy is sold. The rare energy purchases only occur in the night where production and prices are low.
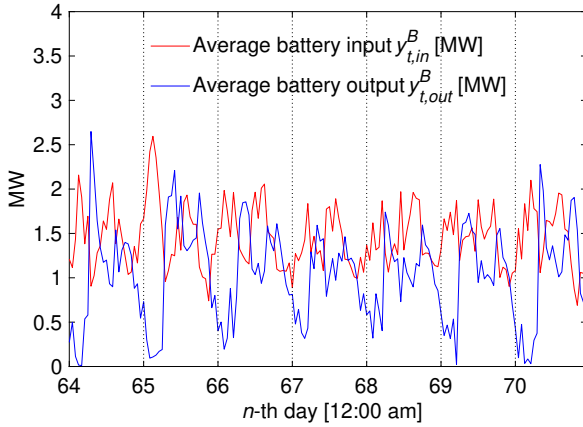
Figure 6: Battery usage
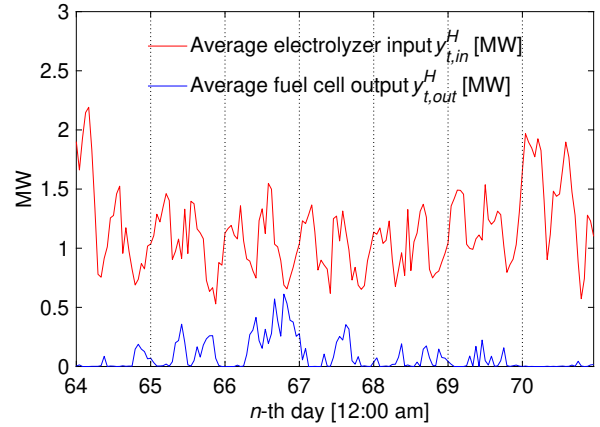
Figure 7: Hydrogen usage

Those weekly patterns are also observable in the storage levels and the power flows. Thus, we show the battery's power flows (Figure 6) and the HES' power flows (Figure 7) for the 10th week. As the hydrogen sales have no weekly or seasonal pattern we do not show these. Since the values shown are averages, a storage appears to be simultaneously charged and discharged, which never happens in a single simulation run.

Both storages are mostly charged in the night hours and between the two price peaks at day. Only on Monday morning and Sunday, where the prices are low, the HES stores as much energy as the battery. The stored energy in the battery is used for the daily electricity trading, while HES discharges are much lower. Further investigations show (not shown here) that the hydrogen production peaks during the first four month where electricity prices are low. The hydrogen reconversion peaks at the end of the year, where electricity prices are high. This indicates, that the hydrogen is stored for multiple weeks to reconverse the hydrogen when the electricity prices are higher or sold directly. Accordingly, the weekly average hydrogen storage level in week 30 is over thrice as high as in week ten. At the end on the year, the hydrogen storage is empty.
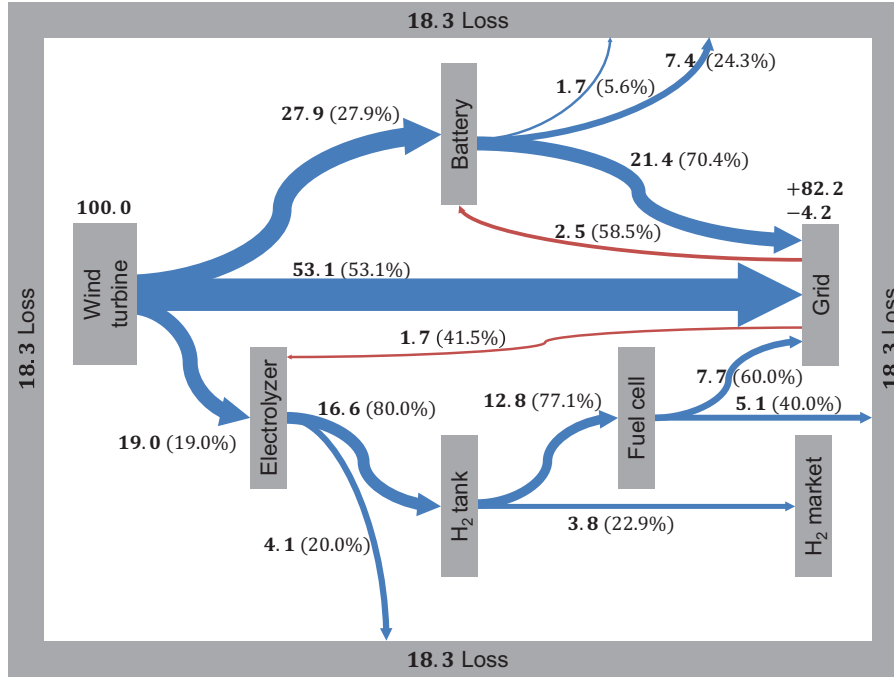
Figure 8: Average energy flows (MWh, bold), normalized to a total production of 100 MWh

Figure 8 displays all average energy flows. To ease the presentation, total energy production is normalized to 100 MWh. For each component of the power plant, the energy flows are given in absolute values in bold (in MWh) and as a percentage of the energy flowing into the component.

Overall, 53.1 % of the electricity produced is delivered to the grid directly. About a fifth of the production and 1.7 MWh of the purchased energy goes to the electrolyzer, of which slightly above half reaches a market (about two-thirds as electricity and a third as hydrogen). Likewise, 27.9 % of production and 2.5 MWh of the purchases go to the battery, of which about two-thirds reaches the market, the other third is lost. 1.7 MWh are lost to the self-discharge of the battery and 7.4 MWh are lost due to the roundtrip efficiency of the battery. In total, 18.3 % of the energy produced is lost, mostly because of the storages' efficiencies. Nonetheless, the storages are valuable. First, the policy would (hopefully) just not use them if their losses exceeded the benefits. Moreover, we will see in the next section that the system performs indeed much worse without storages.

6.4 Importance of the storage devices

In this section, the importance of the power plant's storage devices is investigated. To do so, we compare the basic setting with alternative settings.

Table 1: Alternative settings

Setting	Evaluated with $Q_t^{E,o} = Q_t^{E,u} = 38.40$
basic setting	100.00 %
without hydrogen sales	98.09 %
without HES, doubled battery	95.76 %
without battery, doubled HES	96.40 %
independent optimization and operation	78.65 %

Table 1 shows the profit obtained with several alternative configurations. For a better comparison we

evaluated the profit based on the average reBAP $Q_t^{E,o} = Q_t^{E,u} = 38.40 \, \text{€}/\text{MWh}$. Without the opportunity to sell the hydrogen directly the power producer looses 2 pp of the profit. Obviously the hydrogen sales become more important with a higher hydrogen price (not shown here). If the power producer uses only one storage type (with doubled storage capacity and maximum power flows) profits shrink about 4-5 pp. This shows that a combination out of short and long-term storage can be beneficial. For the independent optimization and operation of production and the storages, the battery and the HES is not restricted by x_{max}^E. In this setting the wind turbine alone has to sell conservative to avoid the penalty payments, so the revenue decreases over 21 pp. If there would be no grid fees, and if the wind turbine would be allowed to deliver the production to the market price without previous trading, there should be no difference between the jointly and the separated optimization.

7   Conclusion

In this paper, we considered the decision problem of a power producer with stochastic production from a renewable source and a hybrid storage (battery and hydrogen-based) who trades in advance on an electricity market. The problem was modeled as a dynamic program and solved with a BADP algorithm where the values of the post-decision states are approximated with a linear architecture of basis functions. Numerical evaluations showed that the algorithm obtains good solutions.

Regarding methodology, we suggest to consider various norms and sample weights for BADP. While the numerical study shows that the traditional *l2*-norm outperforms the *l1*-norm, the new weighted approach leads to a big improvement in solution quality. OCBA further improves the results. Although we evaluated these algorithmic strategies only for the problem considered here, we strongly believe that these findings are more general and, thus, relevant also for the development of BADP algorithms for other problems.

Regarding managerial insights, we numerically analyzed the optimal operation of the power producer with real-world data from Germany. While the operation over time is as expected, it is interesting to see that almost one fifth of energy produced never reaches the grid but is lost due to the efficiencies of charging and discharging the storages. Furthermore, we analyzed the contribution of the storage systems and saw considerable synergy effects between the two storage types which act as short and long-term storage, respectively. Due to grid charges, an integrated optimization of production and storage management is essential.

There are several avenues for future refinements of the model. One extension is to allow to trade multiple hours ahead, but this extension would dramatically increase the dimensionality of the decision, state and outcome space. Another interesting avenue seems the inclusion of limit orders. With these the power producer can not only trade to the current price, but speculate for better prices.

23

A Notation and parameter values

Table A1: Notation introduced in Section 4

Parameters

R_{max}^B	Storage capacity of the battery
R_{max}^H	Storage capacity of the hydrogen tank
x_{max}^E	Maximum electricity commitment
x_{min}^E	Minimum electricity commitment
ρ_{in}^B	Efficiency when energy is stored in the battery
ρ_{out}^B	Efficiency when energy is taken from the battery
ρ_{in}^H	Efficiency when energy is stored as hydrogen
ρ_{out}^H	Efficiency when hydrogen is converted to electricity
$y_{in,max}^B$	Maximum electricity flow into the battery
$y_{out,max}^B$	Maximum electricity flow from the battery
$y_{in,max}^H$	Maximum electricity flow into the electrolyzer
$y_{out,max}^H$	Maximum electricity flow from the fuel cell
γ^B	Self-discharge per period of the battery
γ^H	Self-discharge per period of the hydrogen tank
Δt	Length of time intervals
T	Number of time intervals
β	Discount factor
q^{gf}	Grid fees

Decision variables $x_t = \left(x_t^E, x_t^H, x_t^{B,H} \right)$

x_t^E	Amount of electricity to trade (delivery during period $t+1$)
x_t^H	Amount of hydrogen to sell (instant delivery at end of period t)
$x_t^{B,H}$	Share of excess energy to store in the battery (during period t)

State variables $S_t = \left(R_t^B, R_t^H, x_{t-1}^E, P_t^E, Q_t^{E,u}, Q_t^{E,o}, P_t^H, Y_t^E \right)$

t	Discrete time index
R_t^B	Energy stored in the battery at start of period t
R_t^H	Energy stored in the hydrogen tank at the start of period t
x_{t-1}^E	Electricity which has to be delivered in period t
P_t^E	Electricity price observed in period t for delivery during period $t+1$
$Q_t^{E,u}$	Penalty for electricity shortage in period t
$Q_t^{E,o}$	Penalized electricity price for overproduction in period t
P_t^H	Hydrogen price in period t
Y_t^E	Energy production in period t

Exogenous information $W_{t+1} = \left(P_{t+1}^E, Q_{t+1}^{E,u}, Q_{t+1}^{E,o}, P_{t+1}^H, Y_{t+1}^E \right)$

P_{t+1}^E	Electricity price observed in period $t+1$ for delivery during period $t+2$
$Q_{t+1}^{E,u}$	Penalty for electricity shortage in period $t+1$
$Q_{t+1}^{E,o}$	Penalized electricity price for overproduction in period $t+1$
P_{t+1}^H	Hydrogen price in period $t+1$
Y_{t+1}^E	Energy production in period $t+1$

Functions

$y_{t,in}^{B}\left(S_{t}, x_{t}\right)$	Power flow into the battery in period t
$y_{t,out}^{B}\left(S_{t}\right)$	Power flow out of the battery in period t
$y_{t,in}^{H}\left(S_{t}, x_{t}\right)$	Power flow into the electrolyzer in period t
$y_{t,in}^{H}\left(S_{t}\right)$	Power flow out of the fuel cell in period t
$C_{t}\left(S_{t}, x_{t}\right)$	Contribution function in period t
$V_{t}\left(S_{t}\right)$	Value function in period t
$S_{t}^{M}\left(S_{t}, x_{t}, W_{t+1}\right)$	Transition function from period t to $t+1$

Table A2: Notation introduced in Section 5

$S_{t}^{x}\left(S_{t}, x_{t}\right)$	Post-decision state in period t
$V_{t}^{x}\left(S_{t}^{x}\right)$	Value function of the post-decision states in period t
$\bar{V}_{t}^{x}\left(S_{t}^{x}\right)$	Value function approximation of the post-decision states in period t
$x_{t}^{E,pos}$	Positive part of commitment x_{t}^{E}
$x_{t}^{E,neg}$	Negative part of commitment x_{t}^{E}
θ_{i}	Coefficients of the value function approximation
M	Number of post-decision states to be sampled
N	Number of evaluations per post-decision state (l2, l1, l2-w, l1-w)
$n_{m,t}$	Number of evaluations per post-decision state m in period t (l2-OCBA, l1-OCBA)
n^{lb}	Lower bound for the number of evaluations per post-decision state (l2-OCBA, l1-OCBA)
$v_{m,t}^{x,n}$	n-th evaluation of sampled post-decision state m in period t
$\bar{v}_{m,t}^{x}$	Estimated mean of sampled post-decision state m in period t
$\sigma_{m,t}$	Estimated standard deviation of sampled post-decision state m in period t
$cv_{m,t}$	Estimated coefficient of variation of sampled post-decision state m in period t
ω_{t}	Sample of exogenous information in period t
p	Number of regression coefficients

Table A3: Notation introduced in Section 6

κ	Mean-reversion parameter of the electricity price
μ_{t}	Mean parameter of the electricity price in period t
$\tilde{P}_{t+1}^{E}$	Noise of stochastic process of the electricity price from period t to $t+1$
σ^{E}	Standard deviation of the noise $\tilde{P}_{t+1}^{E}$
λ_{t}	Weibull scale parameter of the wind speed in period t
k_{t}	Weibull shape parameter of the wind speed in period t
ws_{t}	Wind speed sample in period t
r	Maximum power of the wind turbines
s_{ci}	Cut-in speed of the wind turbines
s_{co}	Cut-out speed of the wind turbines
s_{r}	Minimum wind speed with maximum power of the wind power plant

Table A4: Parameter values for the numerical study

Parameter	Value	Parameter	Value
R_{max}^B	$20\,\text{MWh}$	T	8760
R_{max}^H	$300\,\text{MWh}$	Δt	$1\,\text{h}$
x_{max}^E	$25\,\text{MW}$	x_{min}^E	$-10\,\text{MW}$
ρ_{in}^B	$\sqrt{0.75}$	κ	0.0518
ρ_{out}^B	$\sqrt{0.75}$	μ_t	$\in [-56.59, 163.72] \text{€}/\text{MWh}$
ρ_{in}^H	0.8	σ^E	$4.455\,\text{€}/\text{MWh}$
ρ_{out}^H	0.6	$Q_t^{E,u}$	$2P_{t-1}^E$
$Q_t^{E,u}$	$0.00\,\text{€}/\text{MWh}$	P_t^H	$30.00\,\text{€}/\text{MWh}$
γ^B	0.00925	q^{gf}	$5.00\,\text{€}/\text{MWh}$
γ^H	0	λ_t	$\in [0.1001, 0.2512]$
β	$1.05^{\frac{-1}{8760}}$	k_t	$\in [1.321, 3.195]$
$y_{in,max}^B$	$5\,\text{MW}$	r	$20\,\text{MW}$
$y_{out,max}^B$	$5\,\text{MW}$	s_{ci}	$3\,\text{m/s}$
$y_{in,max}^H$	$5\,\text{MW}$	s_{co}	$25\,\text{m/s}$
$y_{out,max}^H$	$5\,\text{MW}$	s_r	$12\,\text{m/s}$

B Validation of BADP using a discretized problem variant

The considered dynamic program is continuous and cannot be solved exactly. The standard approach to tackle this is to discretize the dynamic program and to solve the discretized problem with backwards recursion and lookup tables. In this Section, we show that our BADP algorithm can compete with this approach and is much faster. We specify the discretized backwards recursion approach as following.

- **Disc**: This is a standard benchmark approach and solves a discretized and slightly simplified version of the problem via backwards recursion. Every post-decision state is evaluated exactly, but as not all feasible decisions lead to a discrete state, the value functions are linearly interpolated.

  More precisely, the discretized prices and productions are each the 11 time dependent centroids of $10,000$ sample paths, computed with k-means clustering.

  The decisions x_t^H and x_t^E are discretized with 11 equidistant points ($[0, 5, 10, ..., 50]$ and $[-10, -6.5, ..., , 21.5, 25]$, respectively). By contrast, the decision regarding the allocation of excess energy to the storages $x_t^{B,H}$ is substituted by a simple and reasonable rule: If more energy is produced in a period than delivered, the battery is charged before hydrogen is produced.

  To obtain a discrete state space, it is also necessary to discretize the storage levels. We use again 11 equidistant points. Thus, in total 14,641 post-decision states have to be considered for each time period. Preliminary tests showed that these discretizations provide a reasonable trade-off between machine time and quality of the solution.

Strictly speaking, *Disc* tackles a slightly different problem than the BADP variants and, thus, a comparison on the power producer's problem as described in Section 6.1 would be too optimistic with regard to the BADP variants. To avoid this bias, we choose a conservative approach and use a discretized version of the problem to compare the approaches.

The discretized problem fully complies with the assumptions underlying *Disc* and outlined above (e.g., the decision $x^{B,H}$ is substituted with a decision rule). To save time, the length of the horizon was set to $T = 100$. Obviously, *Disc* is directly applicable to this problem, whereas the other approaches require straightforward modifications regarding the allocation of excess energy. They employ discrete stochastic processes outlined above and the discrete decision alternatives are enumerated. Apart from the linear interpolation, *Disc* solves the discretized problem exactly.

Every policy is evaluated with 1,000 simulation runs. For each approach we use the same out of sample scenarios from the calibrated processes.
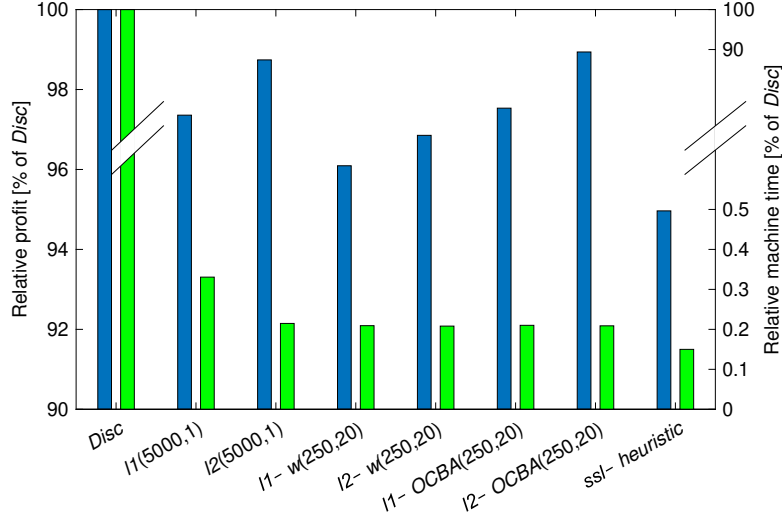
Figure 9: Profit and runtime of the approaches (discretized problem)

Figure 9 shows the approaches' performance relative to *Disc*. The left bars correspond to the relative revenue, while the right bars correspond to the relative machine time. As the length of the 99 % confidence intervals are below 0.5 % of the mean, we do not show these. As expected, the *Disc* approach has the highest revenue but also the highest machine time. While all BADP variants outperform the *ssl-heuristic* in terms of revenue, among these, the *l2-OCBA* is the most profitable one. Because of the extremely high machine time of *Disc*, cannot be used to solve the much more complex continuous problem in Section 6.2.

The profit gap between *l2-OCBA* and *Disc* is below 1.2 percentage points (pp) and this BADP variant is about 500 times faster than *Disc*, which needs over 60 hours. Even if we use a much rougher discretization for *Disc*, it is still slower, but with profits much lower than *l2-OCBA* (not shown here). This and the fact that all BADP variants are better than *ssl-heuristic* already shows that all approaches considered obtain reasonable results, even as we do not enforce the concavity of the VFAs. By contrast, Jiang et al. (2014) report that these properties have to be enforced in their energy storage problem to obtain good results with their forward ADP algorithm because the VFAs converge faster. By contrast, these properties are almost always satisfied with the BADP algorithm without enforcing them.

A comparison of the BADP variants shows that the weighted variants obtain lower profits than the standard variants with an unweighted norm and the OCBA variants. However, the unweighted variants are slightly slower than the weighted variants because the matrix multiplication and the linear program for the computation of the regression parameters are larger. The OCBA step itself is very fast and the OCBA variants have the same runtime as the weighted variants. The *l1* variants are always worse than the *l2* variants and there is no significant difference between *l2-OCBA* and *l2*(5000,1).

References

Al-Salem, M., Almomani, M., Alrefaei, M., and Diabat, A. (2017). On the optimal computing budget allocation problem for large scale simulation optimization. *Simulation Modelling Practice and Theory*, 71:149–159.

Barrodale, I. and Young, A. (1966). Algorithms for best L_1 and L_∞ linear approximations on a discrete set. *Numerische Mathematik*, 8(3):295–306.

Beraldi, P., Violi, A., Carrozzino, G., and Bruni, M. (2018). A stochastic programming approach for the optimal management of aggregated distributed energy resources. *Computers & Operations Research*, 96:200–212.

Biswas, P. P., Suganthan, P., and Amaratunga, G. A. (2017). Optimal power flow solutions incorporating stochastic wind and solar power. *Energy Conversion and Management*, 148:1194–1207.

Bretthauer, K. M. and Shetty, B. (2002). A pegging algorithm for the nonlinear resource allocation problem. *Computers & Operations Research*, 29(5):505–527.

Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., and Fichtner, W. (2019). A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Economics*, 80:1059–1078.

Bundesnetzagentur (2006). Bilanzkreisvertrag über die Führung eines Bilanzkreises. `https://www.bundesnetzagentur.de/DE/Service-Funktionen/Beschlusskammern/1_GZ/BK6-GZ/2006/2006_0001bis0999/2006_001bis099/BK6-06-013/BK6-06-013_GemeinsamerBKValle220NBId6386pdf.pdf?__blob=publicationFile&v=2`. Accessed 22 May 2019.

Cai, W., Mohammaditab, R., Fathi, G., Wakil, K., Ebadi, A. G., and Ghadimi, N. (2019). Optimal bidding and offering strategies of compressed air energy storage: A hybrid robust-stochastic approach. *Renewable Energy*, 143:1–8.

Chen, C.-H., Lin, J., Yücesan, E., and Chick, S. E. (2000). Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10(3):251–270.

Cheng, B., Asamov, T., and Powell, W. B. (2017). Low-rank value function approximation for co-optimization of battery storage. *IEEE Transactions on Smart Grid*, pages 1–1.

Cruise, J., Flatley, L., Gibbens, R., and Zachary, S. (2019). Control of energy storage with market impact: Lagrangian approach and horizons. *Operations Research*, 67(1):1–9.

Dufo-López, R., Bernal-Agustín, J. L., and Contreras, J. (2007). Optimization of control strategies for stand-alone renewable energy systems with hydrogen storage. *Renewable Energy*, 32(7):1102–1126.

Durante, J., Nascimento, J., and Powell, W. B. (2020). Backward approximate dynamic programming with hidden semi-markov stochastic models in energy storage optimization. Working paper, Princeton University, Princeton.

EEX (2019). GASPOOL. `https://www.eex.com/en/market-data/natural-gas/spot-market/gaspool`. Accessed 22 May 2019.

Energiepark Mainz (2019). Technical data about the Energiepark. `http://www.energiepark-mainz.de/en/technology/technical-data/`. Accessed 22 May 2019.

EPEX SPOT SE (2019). Market data. `https://www.epexspot.com/en/market-data`. Accessed 22 May 2019.

Fathima, A. H. and Palanisamy, K. (2015). Optimization in microgrids with hybrid energy systems – a review. *Renewable and Sustainable Energy Reviews*, 45:431–446.

Fraunhofer Institute for Chemical Technology (2019). Project RedoxWind. `https://www.ict.fraunhofer.de/en/comp/ae/rw.html`. Accessed 22 May 2019.

García, P., Torreglosa, J. P., Fernández, L. M., and Jurado, F. (2013). Optimal energy management system for stand-alone wind turbine/photovoltaic/hydrogen/battery hybrid system with supervisory control based on fuzzy logic. *International Journal of Hydrogen Energy*, 38(33):14146–14158.

Garcia-Torres, F. and Bordons, C. (2015). Optimal economical schedule of hydrogen-based microgrids with hybrid storage using model predictive control. *IEEE Transactions on Industrial Electronics*, 62(8):5195–5207.

Gönsch, J. and Hassler, M. (2016). Sell or store? an ADP approach to marketing renewable energy. *OR Spectrum*, 38(3):633–660.

Hannah, L. A. and Dunson, D. B. (2011). Approximate dynamic programming for storage problems. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, pages 337–344, USA. Omnipress.

Hassler, M. (2017). Heuristic decision rules for short-term trading of renewable energy with co-located energy storage. *Computers & Operations Research*, 83:199–213.

Heredia, F. J., Cuadrado, M. D., and Corchero, C. (2018). On optimal participation in the electricity markets of wind power plants with battery energy storage systems. *Computers & Operations Research*, 96:316–329.

Jia, Q.-S. (2012). Efficient computing budget allocation for simulation-based policy improvement. *IEEE Transactions on Automation Science and Engineering*, 9(2):342–352.

Jiang, D. R., Pham, T. V., Powell, W. B., Salas, D. F., and Scott, W. R. (2014). A comparison of approximate dynamic programming techniques on benchmark energy storage problems: Does anything work? In *2014 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*. IEEE.

Kim, J. H. and Powell, W. B. (2011). Optimal energy commitments with storage and intermittent supply. *Operations Research*, 59(6):1347–1360.

Koschker, S. and Möst, D. (2015). Perfect competition vs. strategic behaviour models to derive electricity prices and the influence of renewables on market power. *OR Spectrum*, 38(3):661–686.

Kuznetsova, E., Ruiz, C., Li, Y.-F., and Zio, E. (2015). Analysis of robust optimization for decentralized microgrid energy management under uncertainty. *International Journal of Electrical Power & Energy Systems*, 64:815–832.

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.

Löhndorf, N., Wozabal, D., and Minner, S. (2013). Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. *Operations Research*, 61(4):810–823.

Li, H., Liu, P., Guo, S., Ming, B., Cheng, L., and Yang, Z. (2019). Long-term complementary operation of a large-scale hydro-photovoltaic hybrid power plant using explicit stochastic optimization. *Applied Energy*, 238:863–875.

Luo, X., Wang, J., Dooner, M., and Clarke, J. (2015). Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied Energy*, 137:511–536.

Lust, A. and Waldmann, K.-H. (2018). A general storage model with applications to energy systems. *OR Spectrum*.

Moghaddam, I. G., Nick, M., Fallahi, F., Sanei, M., and Mortazavi, S. (2013). Risk-averse profit-based optimal operation strategy of a combined wind farm–cascade hydro system in an electricity market. *Renewable Energy*, 55:252–259.

Naraharisetti, P. K., Karimi, I., Anand, A., and Lee, D.-Y. (2011). A linear diversity constraint – application to scheduling in microgrids. *Energy*, 36(7):4235–4243.

NGK Insulators (2019). Rokkasho village, Aomori, Japan. https://www.ngk.co.jp/nas/case_studies/rokkasho/. Accessed 22 May 2019.

Patriksson, M. and Strömberg, C. (2015). Algorithms for the continuous nonlinear resource allocation problem—new implementations and numerical studies. *European Journal of Operational Research*, 243(3):703–722.

Powell, W. B. (2011). *Approximate Dynamic Programming*. John Wiley & Sons, Inc.

Rahman, H. A., Majid, M. S., Jordehi, A. R., Kim, G. C., Hassan, M. Y., and Fadhl, S. O. (2015). Operation and control strategies of integrated distributed energy resources: A review. *Renewable and Sustainable Energy Reviews*, 51:1412–1420.

Safari, S., Ardehali, M., and Sirizi, M. (2013). Particle swarm optimization based fuzzy logic controller for autonomous green power energy system with hydrogen storage. *Energy Conversion and Management*, 65:41–49.

Schwarz, H., Bertsch, V., and Fichtner, W. (2017). Two-stage stochastic, large-scale optimization of a decentralized energy system: a case study focusing on solar PV, heat pumps and storage in a residential quarter. *OR Spectrum*, 40(1):265–310.

Shi, L., Wang, C., Yao, L., Ni, Y., and Bazargan, M. (2012). Optimal power flow solution incorporating wind power. *IEEE Systems Journal*, 6(2):233–241.

Sterner, M. and Stadler, I. (2017). *Energiespeicher - Bedarf, Technologien, Integration*. Springer Berlin Heidelberg.

TransnetBW (2019). Balancing group billing. https://www.transnetbw.com/en/energy-market/balancing-group-management/balancing-group-billing. Accessed 22 May 2019.

Trifkovic, M., Marvin, W. A., Daoutidis, P., and Sheikhzadeh, M. (2014). Dynamic real-time optimization and control of a hybrid energy system. *AIChE Journal*, 60(7):2546–2556.

van der Klauw, T., Gerards, M. E. T., and Hurink, J. L. (2017). Resource allocation problems in decentralized energy management. *OR Spectrum*, 39(3):749–773.

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.

Weitzel, T. and Glock, C. H. (2018). Energy management for stationary electric energy storage systems: A systematic literature review. *European Journal of Operational Research*, 264(2):582–606.

Wozabal, D. and Rameseder, G. (2020). Optimal bidding of a virtual power plant on the spanish day-ahead and intraday market for electricity. *European Journal of Operational Research*, 280(2):639–655.

Zeng, K. and Zhang, D. (2010). Recent progress in alkaline water electrolysis for hydrogen production and applications. *Progress in Energy and Combustion Science*, 36(3):307–326.

Zhao, H., Wu, Q., Hu, S., Xu, H., and Rasmussen, C. N. (2015). Review of energy storage system for wind power integration support. *Applied Energy*, 137:545–553.

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek