Applied Energy 270 (2020) 115130

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Effect of market design on strategic bidding behavior: Model-based analysis of European electricity balancing markets

Ksenia Poplavskaya[a,b,*], Jesus Lago[b,c,d], Laurens de Vries[b]

[a] AIT Austrian Institute of Technology, Center for Energy, Integrated Energy Systems, 1210 Vienna, Austria
[b] Delft University of Technology, Faculty of Technology, Policy, and Management, GA Delft, the Netherlands
[c] Algorithms, Modeling, and Optimization, Energyville, 3600 Genk, Belgium
[d] Energy technology, VITO Flemish Institute for Technological Research, 2400 Mol, Belgium

HIGHLIGHTS

- European regulatory change triggered substantial change in balancing market design.
- Opportunities for strategic behavior are identified using ABM with learning agents.
- Procurement of balancing energy in a standalone market increases market efficiency.
- Marginal pricing further enhances the economic efficiency of the balancing market.

ARTICLE INFO

Keywords:
Balancing market
Agent-based modeling
Bidding strategy
Reinforcement learning
Market efficiency

ABSTRACT

Market-based procurement of balancing services in Europe is prone to strategic bidding due to the relatively small market size and a limited number of providers. In the European Union, balancing markets are undergoing substantial regulatory changes driven the efforts to harmonize the market design and better align it with the goals of the energy transition. It is proposed to decouple the balancing *energy* (real-time) market from the (forward) balancing *capacity* market and the price of balancing energy will be based on the marginal bid. In this paper, the potential effects of these changes on market participants' strategies are analyzed using an agent-based model. This model compares the effects of a standalone balancing energy market with different pricing rules on economic efficiency with agents that apply naïve, rule-based and reinforcement-learning strategies. The results indicate that the introduction of a standalone balancing energy market reduces the cost of balancing, even in a concentrated market with strategic bidders. Marginal pricing consistently leads to lower weighted average prices than pay-as-bid pricing, regardless of the level of competition. Nevertheless, in an oligopoly with actors bidding strategically, prices can deviate from the competitive benchmark by a factor of 4–5. This implies that the introduction of a standalone balancing energy market does not entirely solve the issue of strategic bidding, but helps dampen the prices, as compared to the balancing market prior to the design change.

1. Introduction

To balance supply and demand, most European transmission system operators procure balancing services in a market-based way through a two-stage process, first reserving the necessary balancing capacity and then activating balancing energy when system deviations occur. However, market-based procurement is not necessarily efficient as the strict technical requirements limit the number of eligible balancing service providers (BSPs). Many European electricity balancing markets have design features that, along with market concentration, make them susceptible to gaming. With the help of an agent-based model (ABM) with artificial intelligence, we study opportunities for strategic behavior and assess whether expected balancing market design changes can improve its efficiency. As the EU intends to integrate growing shares of renewables into the European grids and markets, to harmonize balancing markets and facilitate cross-border procurement of balancing resources (cf. [1]), it is important to identify balancing market design features that facilitate market entry and increase robustness to strategic bidding. The first aspect has been addressed in detail in [2], while the second aspect requires quantification of the effects of bidding strategies

* Corresponding author at: AIT Austrian Institute of Technology, Center for Energy, Integrated Energy Systems, Giefinggasse 4, 1210 Vienna, Austria.
E-mail address: ksenia.poplavskaya@ait.ac.at (K. Poplavskaya).

https://doi.org/10.1016/j.apenergy.2020.115130
Received 30 January 2020; Received in revised form 28 April 2020; Accepted 3 May 2020
Available online 22 May 2020

under different market designs and is addressed in this paper using ABM.

To stabilize the system frequency, most European transmission system operators (TSOs) procure balancing services in a competitive, two-stage process. First the necessary balancing capacity is reserved; balancing energy is activated in real time, when actual system deviations occur. However, market-based procurement is not necessarily synonymous with efficient procurement [3,4]. Due to strict technical requirements, the current number of eligible balancing service providers (BSPs, parties who sell balancing services to the TSO) is limited. As a result, balancing markets are highly concentrated, which opens up room for opportunistic behavior and market inefficiencies.

The need for greater market integration [5] and the wish to remediate market inefficiencies led to the recent adoption of several European regulations and network codes [6,7]. Among them, the EU guideline on electricity balancing (EBGL, adopted in November 2017) defined the main features of harmonized European balancing markets [8]. Specifically, the balancing energy (BE) market is required to be decoupled from the balancing capacity (BC) market so that balancing energy bids are submitted in a separate auction close to real time. A review of balancing market design variables and their combinations is presented in [2]. The authors structured the design variables according to priority and showed that, in order to improve market access and performance, the splitting of the balancing capacity and energy markets is the necessary first step before addressing other design aspects as most other variables depend on it [2].

In order to analyze and study the expected behavior of market players under this new design, in this work we simulate a standalone BE market (hereafter *"split BC-BE market"*) with the help of an ABM. To this end, we implement naive and learning agents and compare their performance. The naïve agents bid their true short-term variable costs. The learning agents that are designed to represent different levels of market power take decisions either according to a pre-determined rule or by using a fitted Q-iteration algorithm (a class of reinforcement learning algorithms) to identify their bidding strategies. We investigate the potential efficiency gains from introducing a separate balancing energy market, as compared to a market where balancing capacity and energy are procured jointly (hereafter *"joint BC-BE market"*) used today. For this, we analyze the bidding behavior, profits of BSPs, and the cost of balancing in the face of this regulatory transformation.

This work provides an analysis of regulatory changes spurred by the EBGL with a new approach to modelling the balancing market, namely ABM with agents that apply learning strategies. Unlike other ABM-based studies of the balancing market, we focus on the market for balancing energy that is mandated by the EBGL. Our approach allows to represent individual elements of market design and their combinations in great detail, including different types of actors and technologies. Reinforcement learning allows agents to adapt their market strategies, which we compare with predefined strategies and with empirical observations. The combination of a detailed agent-based market model with artificial intelligence in the agents provides a powerful tool for analyzing the impact of market design on strategic behavior.

To the authors' knowledge, this is the first model-based study of the upcoming introduction of a standalone balancing energy market and marginal pricing and their effects on the bidding strategies of market actors. The model provides a deeper insight into the implications of these changes, helps to make market design more robust against gaming and to estimate the extent to which the actions of a single or few bidders can affect market outcome. This analysis is particularly relevant for the EU's harmonization efforts and energy policy goals. This paper provides useful conclusions for regulators, TSOs and policymakers and provides them with specific recommendations for improving balancing market design and efficiency.

We structure the paper as follows: Section 2 reviews the state of the art of the balancing market analysis and the use of ABM for electricity market modelling. Section 3 describes the functioning of the balancing market and the bidding process along with the main building blocks of its design. Section 4 presents the agent-based model, Elba ABM, its main features, key assumptions, design choices and agent strategies. Section 5 describes the simulation setup and scenarios. Section 6 provides and analyses model results and Section 7 concludes the paper.

2. Literature review

2.1. Balancing market analysis

Balancing markets in Europe have generally been a rather lucrative commercialization option for flexible generation. As a result, most of the current body of research has been focused on issues related to the portfolio optimization for participation in balancing markets (e.g. [9]). As the European countries have been gradually easing market access rules to new flexibility sources, recent research has extensively addressed the potential of distributed energy technologies, such as battery storage [10], heat pumps [11] household photovoltaic and storage systems[12], as well as demand response [13] for frequency support.

The relevance of the balancing market as performing a key function in the European electricity market design has been widely acknowledged in the literature. Research has addressed market design improvements [2,14], harmonization of market rules [15,16] and strategic bidding behavior [17,18], among others. The authors in [15] analyze possible future market design and argue for the use of asymmetric bidding in the balancing market and shortening the product length to enable the procurement of balancing reserves from renewables and other distributed energy resources. Positive effects of market integration and the possible cost savings that can be achieved with its help were addressed in [19]. Balancing market harmonization is however complicated by large national differences [20,21], which makes it important to identify the elements of an efficient market design. Currently, balancing markets are characterized by high entry requirements and therefore low competition levels [22,23]. Consequently, the conventional assumption that all participants behave competitively and bid their full available capacity at true costs seems rather unrealistic.

2.2. The use of agent-based modelling for the analysis of bidding strategies and electricity market design

Researchers widely use ABM to analyze the effects of policy and market design changes. As shown in [24] and [25], ABM is a suitable method for capturing balancing market complexity, including non-competitive behavior. For instance, authors in [24] used ABM to model the imbalance settlement and studied the effects of imbalance pricing on market actors. Researchers in [26] investigated the effect of different options for market clearing of interconnected day-ahead and balancing markets using ABM. In [27], ABM was applied primarily to analyze the effect of increasing shares of RES on electricity markets. Their model, MATREM, simulates the day-ahead and intraday markets as well as forward and bilateral markets and use complex agents able to interact with the user [27]. Researchers in [28] successfully combined agent-based modelling and optimization techniques to investigate the effect of demand response and storage systems in the electricity market as alternative to the capacity market. German electricity market design was analyzed in [25]; the authors found that the introduction of a capacity market can help solve the generation adequacy issue and is a viable alternative to the energy-only market in the long term. Bidding strategies were the main focus of [29] where the authors compared bid pricing rules in the DA market and the effects of price volatility. In [30], ABM was used to optimize bidding strategies of generating companies in the DA market and showed the suitability of this approach for modeling complex systems and interactions within them.

In an ABM, it is possible to equip the agents with learning capability [31,32]. For instance, Researchers in [33] and [34] developed PowerACE, an ABM that includes a spot and German balancing market. The authors in [34] provided a thorough assessment of several learning algorithms that can be integrated into ABM to represent agent behavior and showed that Q-learning produced better results than Erev-Roth-type reinforcement learning. Since market participants do not have access to complete information, they are bound to behave strategically in the face of uncertainty (e.g. [35]), optimizing their decisions by factoring in the risk associated with imperfect information. In [36], a short-term electricity market is modeled to teste agents' learning strategies and attitudes to risk. The authors showed that agent bidding strategies can be improved through more risk-averse strategies. Researchers in [23] developed an agent-based model of the German balancing market to study the bidding behavior of market actors and the effect of attitude towards risk on their bidding strategies and showed that ABM is an appropriate tool to analyze the balancing market [23].[1] The way the same market design can provoke different outcomes due to different agents participating in it incorporating agents' expectations and uncertainty was demonstrated with the help of ABM in [38].

2.3. Agent-based modelling and learning for the analysis of regulatory changes in the balancing market

ABM has proven to be a useful tool to capture market dynamics and complexity and account for the behavior of multiple actors and their reactions to market opportunities and incentives [39]. It further allows to analyze the effects of policy and market design changes considering adaptive behavior of participants [31,32]. The authors in [40] use empirical market data from Central Western Europe to emphasize that balancing market design has a direct effect on the strategies of flexibility providers. However, top-down optimization models cannot represent different bidding strategies and potential opportunistic behavior due to their intrinsic assumptions of perfect competition and foresight. Similarly, game theoretical approaches, while useful for identifying optimal strategies of market actors, lack flexibility in integrating multiple agents with different characteristics and strategies and do not scale up to include multiple players with a large number of decision variables (such as plants to dispatch). ABM allows for heterogeneity and a larger number of agents (e.g. [41]) and can help to understand and quantitatively assess the bidding behavior of the agents in repeated auctions. The relevance of the repeated nature of the balancing auction has been demonstrated e.g. in [42,43]. ABM makes it possible to evaluate the effects of actors' decisions (e.g. [27]), in particular types of bidding behavior, on the price levels and behavior of others by providing the agents with learning capabilities [44]. It is particularly suitable for exploration based on incomplete information (actual strategies of market participants are not disclosed) and multiple observations (market outcomes) [45].

3. European balancing markets

European balancing markets are rooted in the physical grid requirements and the TSOs' obligation to maintain the energy balance within their control system in order to maintain the network frequency in the interconnected system. System imbalances are caused by stochastic processes, uncertainty associated with generation and load forecasts, plant or line outages and the behavior of market participants. Balancing markets consists of several institutional arrangements, as is shown in Fig. 1.

Fig. 1 illustrates that the balancing market for electricity is a key link between the physical power system and the markets. The process in the balancing market starts with the procurement auction for the reservation of balancing capacity (BC), the goal of which is to ensure sufficient balancing capacity available for potential activation. It is followed by the activation of balancing energy (BE) in real time to resolve system imbalances, using the pool of balancing resources that were contracted during the previous stage. Finally, after real time, the costs of imbalances are settled between the TSO and the BRPs under the "polluter-pays" principle. Resulting imbalance prices are based on the cost of provision of balancing energy (although the methodologies differ among EU countries).

The bottom of this figure represents the Physical Layer of the system. The imbalances between electricity generation and consumption are controlled by the TSO in real time. The Actor Layer shows the players: the TSO is in the middle between the balancing services providers (BSPs), who obtain their resources from suppliers on the left, and the balancing responsible parties (BRPs), who are the cause the imbalances, on the right. In contrast to day-ahead and intraday markets, only market participants whose assets pass a stringent prequalification process may act as BSPs.[2] BRPs aggregate market actors (providers and consumers of electricity) into portfolios to achieve scale economies (on the right side in the Actor Layer, Fig. 1). BRPs submit planned load and generation schedules to the TSO day-ahead.

The Institutional Layer shows how the TSO handles imbalances through balancing services that it purchases from balancing service providers (BSPs) before real time in the *balancing capacity market*. System imbalances are corrected in real time by activating the regulating capacity that was purchased from the BSPs in the *balancing energy market*. Deviations from the required network frequency value can be both negative and positive. If the system imbalance is negative, i.e. the system is short, generation output must be increased (or demand reduced), activating positive balancing energy. Conversely, negative balancing energy is activated in case the system is long, i.e. oversupplied, and generation must be reduced (or demand increased). BRPs need to compensate the TSO for deviations from their schedules, e.g. caused by forecast errors of renewable generators.

Unlike spot markets, balancing markets are single-sided, with the TSO acting as the single buyer. TSOs use separate auctions for procuring the standardized balancing products. The EBGL defines four standard balancing products: Frequency Containment Reserve (FCR), automatic Frequency Restoration Reserve (aFRR), manual Frequency Restoration Reserve (mFRR) (Fig. 1, top), and Replacement Reserve (RR), which mainly differ according to their activation speed and duration of activation. The FCR is used to handle imbalances that are caused by so-called *"intra-dispatch interval variability"* [46], meaning that while demand changes continuously, schedules are submitted in discrete steps, most commonly of 15 min, and as a result there are continuous, small differences between supply and demand. We ignore this issue and focus on deviations between the actual and scheduled electricity generation or consumption per time interval, which are largely handled with aFRR (with mFRR and RR as backups). aFRR is used in all countries of the ENTSO-E area and has the highest trading volumes among the standard products (cf. [47]). In a series of interrelated electricity markets [48], the BC market is cleared before the day-ahead (DA) market (see also Fig. 2, top.) This may occur from one year to one day ahead of delivery time, depending on the country and the balancing product. The required BC for each product is determined by the TSO, whereas the demand for BE depends on actual imbalances.

The bid structure of aFRR (automatic frequency response reserve) includes the BC volume in MW and the respective BC price in €/MW. Commonly, the price for activation of BE in €/MWh must be provided at the time of the BC auction and only BSPs whose capacity bids have been accepted are considered for providing balancing energy (Fig. 1). A merit order based on the price of balancing capacity is created for clearing the BC auction, whereas another merit order is constructed

[1] In contrast, for highly competitive day-ahead markets fundamental optimization models have proven to yield better results [37].

[2] More information on the limits of access to the balancing market can be found in [2] and the detailed requirements can be found in the EBGL [8] as well as national prequalification documents.

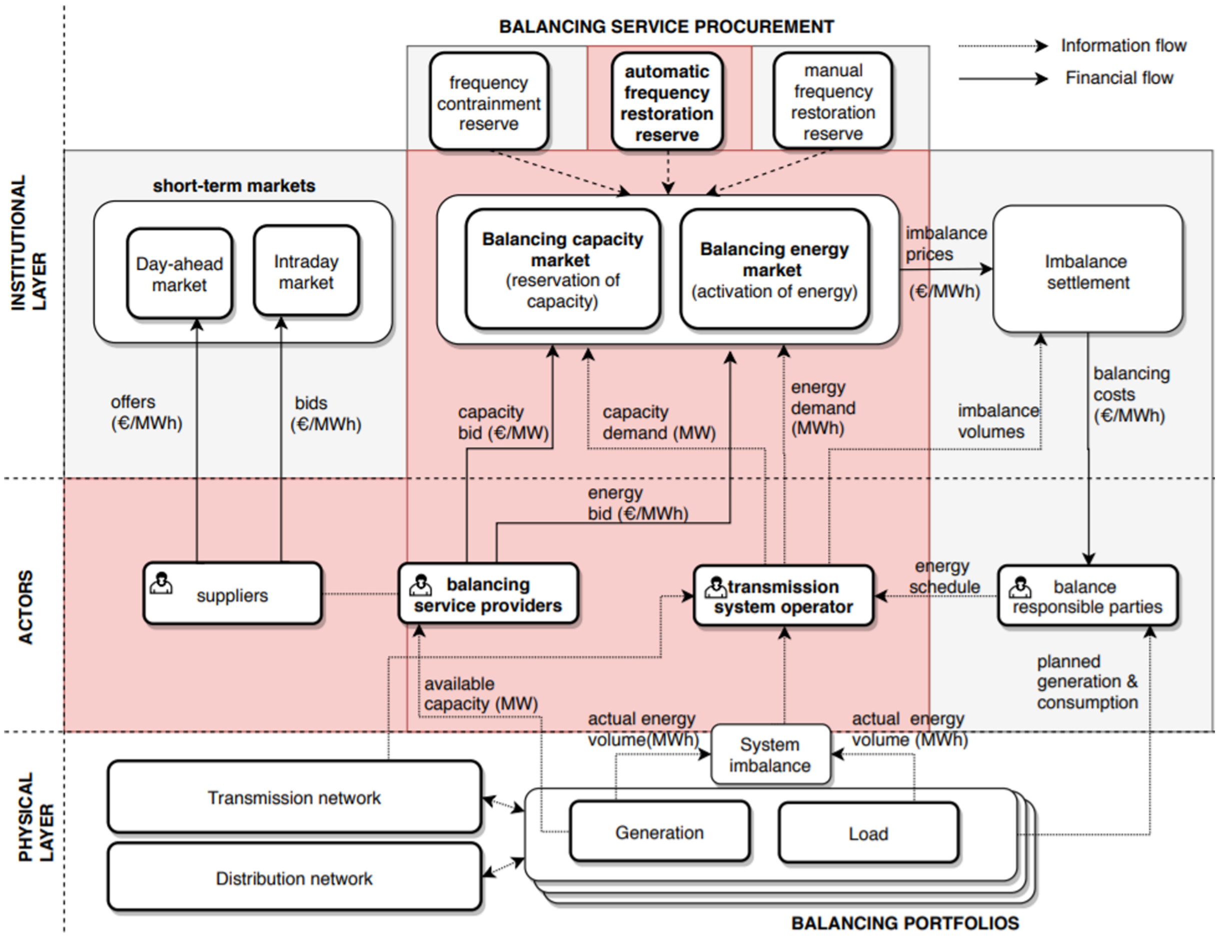

Fig. 1. Overview of the organization of European electricity balancing markets and their relation to short-term electricity markets. The focus of the Elba-ABM model in this paper is marked in red.

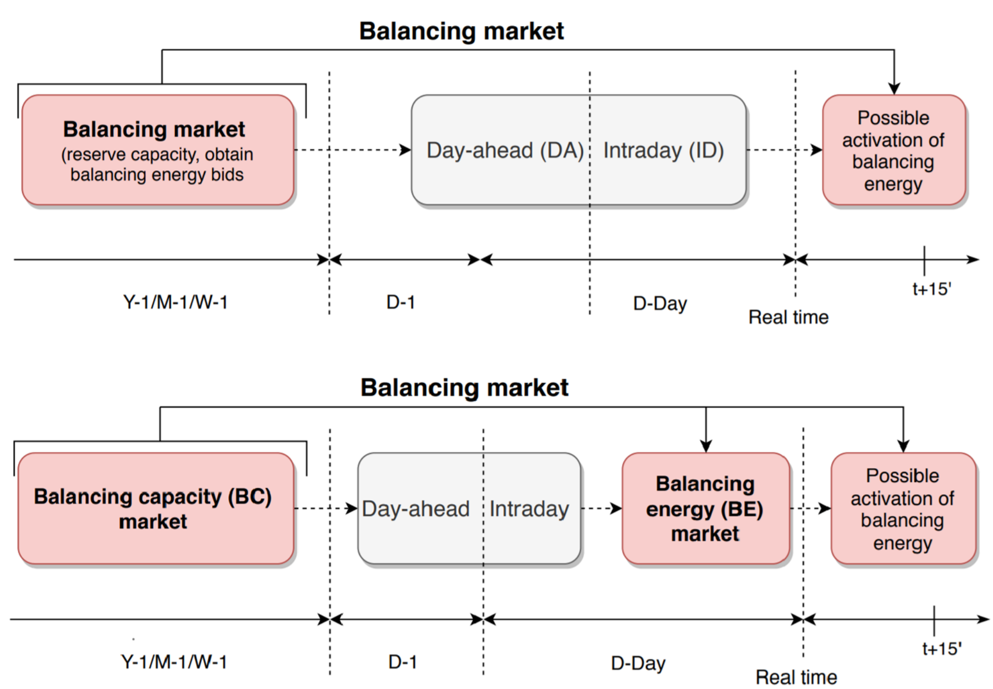

Fig. 2. Top: Current temporal sequence of the balancing and spot markets. Bottom: the change in the balancing market sequence proposed by the EBGL.

afterwards for the BE market by ranking the energy bids (from the accepted balancing capacity providers). In the market for positive regulation, the bids are ranked from the lowest to the highest, while in the market for negative regulation, a descending merit order is built: if a BSP submits a positive bid, he/she is willing to pay the TSO for reducing his/her output whereas the TSO must remunerate the BSP that submitted a negative bid and was awarded.

Under the European electricity market unbundling provisions, balancing services must be procured in a market-based way [1]. Yet, large differences in national balancing market designs exist among the EU countries [20]. In some balancing markets, BSPs are still required to submit symmetric bids, i.e. the same volumes of positive and negative regulation must be supplied, while in others asymmetric bids are allowed. The service provision can be remunerated according to a pay-as-bid rule or to a marginal price rule. The former implies that each generator receives the price they bid while in the latter case each awarded bid receives the same market clearing price. Balancing products are distinguished by the period during which they should be available for activation ranging from a day to an hour [20].

4. Elba ABM: Model overview

This section is divided into three parts. The first subsection introduces Elba-ABM (Agent-Based Model of ELectricity BAlancing market) and its main functionalities. The second subsection describes the modelled market design and the third subsection describes the three types of bidding behavior that are modeled.

4.1. Model introduction

Elba-ABM is a bottom-up agent-based model that simulates balancing market mechanisms and bidding decisions of individual BSPs. The main intention of the model is to represent key design features of European balancing capacity and balancing energy markets. The model makes it possible to adjust these design features in order to evaluate their impact on the strategies of BSPs and, consequently, on the market outcome. We focus on the effects of different combinations of market design variables on market efficiency in the presence of competitive and strategic bidding strategies.

Two versions of the model were developed that represent joint and split BC-BE markets, as will be described in Section 4.2. The models represent the process of bid submission, the market clearing processes and the financial settlement process (using either marginal or pay-as-bid pricing). The model can simulate bids per generator as well as portfolio bidding with generators of different technologies. In the model, the BSPs determine their bids individually based on their marginal costs[3] and/or prior experience (modeled through rule-based or reinforcement-learning (RL) agents). These strategies will be described in detail in Section 4.3. We use representative balancing market data that is based on data from the Austrian aFRR market [49].

The authors are aware of the strong connection between the balancing market and other short-term markets. Although the day-ahead market is not modeled explicitly, it is taken into account through day-ahead prices that are given to the BSP agents as an opportunity value. Secondly, the capacity that BSPs can bid in the balancing market is limited because it typically needs to consist of spinning reserve or fast-start units. The model uses a scenario generation technique proposed in [50] for developing realistic and correlated data for simulating the market. This technique generates realistic system imbalance scenarios that correlate with day-ahead market prices. For every yearly simulation, the Elba-ABM framework generates a new scenario of imbalances and prices.

[3] For the purpose of this analysis, we do not distinguish between variable costs and marginal costs and use the two terms interchangeably.

4.2. Joint versus split BC-BE markets

The model consists of a two-stage simulation, with the BC market setting the stage for the BE market. The bidding frequency for BC can be varied from once per year to daily. In the model version with a joint BC-BE market, the BE prices are set as part of the BC auction. In the split BC-BE market model, the BE market has either the same or a higher frequency. We implemented a frequency of once per hour. The time step for market clearing the BE market is set to 15 min, i.e. equal to the imbalance settlement period, so every hour, the BSPs offer their BE prices for the four 15-minute blocks of the delivery hour. Upward regulation and downward regulation are procured in two separate auctions (positive and negative markets, respectively). Each of the auctions can be cleared using a pay-as-bid (PaB) or marginal pricing (MP) rule.

The market clearing mechanism for the balancing market is the central element of the simulation model. The model procedures are summarized below and illustrated in Figs. 3 and 6. In the BC market, the TSO first announces the demand for balancing capacity; then, bidders submit BC bid volumes and prices based on their strategies; finally, the TSO awards bidders according to merit order results. In real time, when the awarded bidders participate in the BE market, the TSO determines imbalance volumes and clears the market per 15 min based on separate merit orders for + aFRR and –aFRR and then calculates and stores the results. Bidders obtain the market results *ex-post* and calculate their profits.

4.2.1. Joint BC-BE market

In the joint BC-BE market, the agents' bids do not change throughout a model run: each time step with a positive imbalance, agents submit the same positive bid, the same goes for steps with a negative imbalance. For instance, if we assume a product resolution of one day, the same BE price ladder (supply function) is used for all 96 time intervals of 15 min. The marginal clearing price (MCP) for each 15-minute interval varies only because of differences in the demand for balancing energy. Thus, the BC market determines *the frequency of change* of BE prices. The model flow of the joint market is illustrated in Fig. 3.

In the joint BC-BE market, the bid information must contain the BE prices for a given hour of the day. So if, for instance, hourly products for BE are assumed, then a BSP may submit, once a day, up to 24 BE bids, one for each hour, with optionally different BE prices. In the example below, a bidder offers its balancing resources by 23:00 for each hour of the next day (Fig. 4).

4.2.2. Split BC-BE market

In the split BC-BE market, a new merit order for BE is built every 15 min. The BE bids submitted on an hourly basis, i.e. the gate closure time (GCT) is assumed to be one hour ahead of delivery. The MCP is again determined by the actual imbalance volume, but in this case, BSPs have more room to adjust their bid strategies to generate a higher reward, as information is updated with a high frequency. The model runs the BE market for the 96 intervals per day (15 min interval). The simulation flow is illustrated in Fig. 5.

By way of example, assume that the gate for BE bids opens at 22:00 (GOT) and closes an hour later at 23:00 (GCT). Within this period, bids are submitted for potential activation between 00:00 and 01:00 of the next day. This means that the bidding period is from 22:00 to 23:00 whereas the delivery period is from 00:00 to 02:00. This is illustrated in Fig. 6.

4.3. Agent definition

To simulate the bidding behavior of BSPs, we consider three types of agents: naïve ones, rule-based and reinforcement learning (RL) agents. Their strategies are briefly summarized in Table 1.

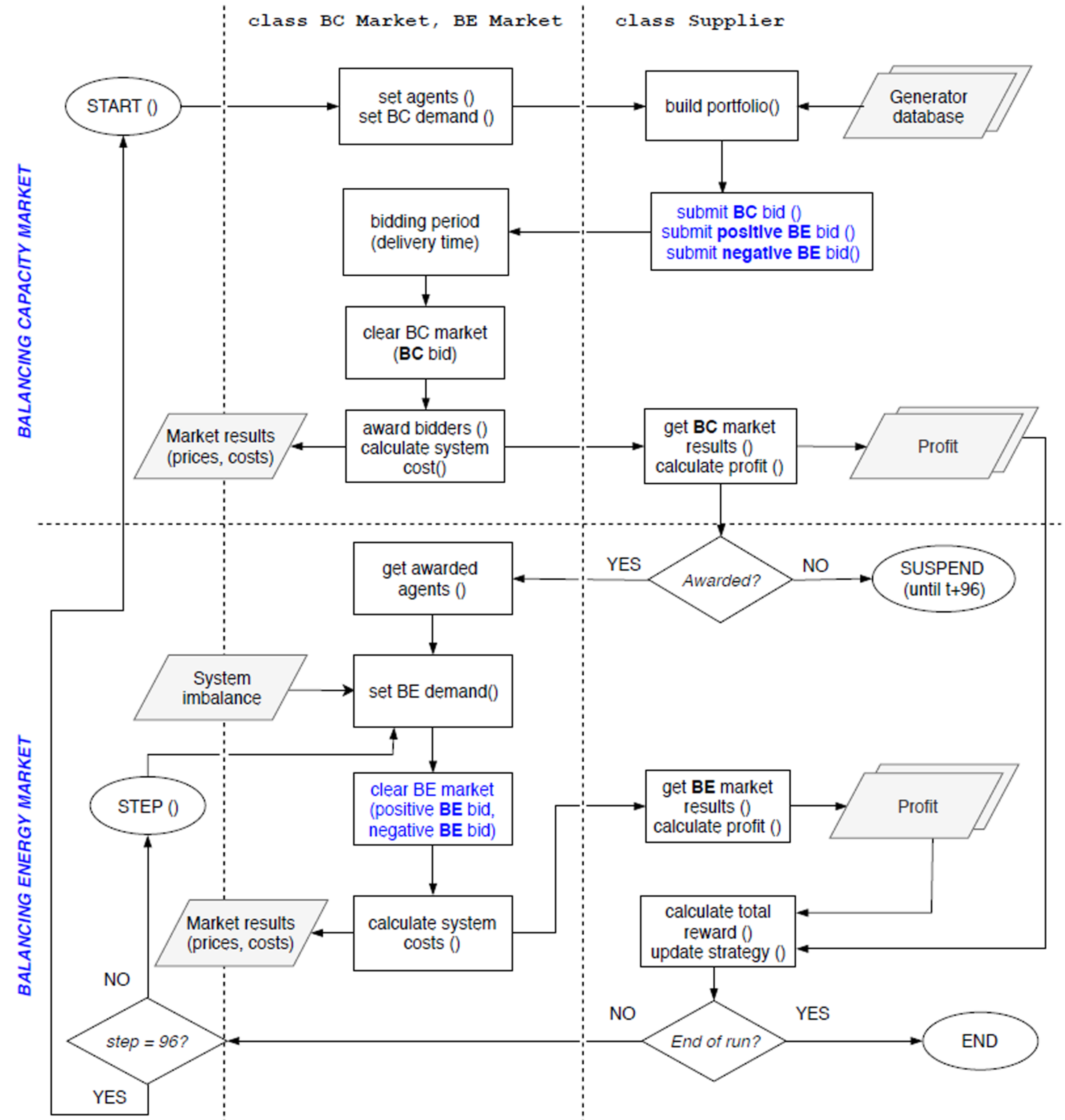

Fig. 3. General model structure diagram for a **joint** auction for balancing capacity and energy. The differences between the joint and split auction are marked in blue.

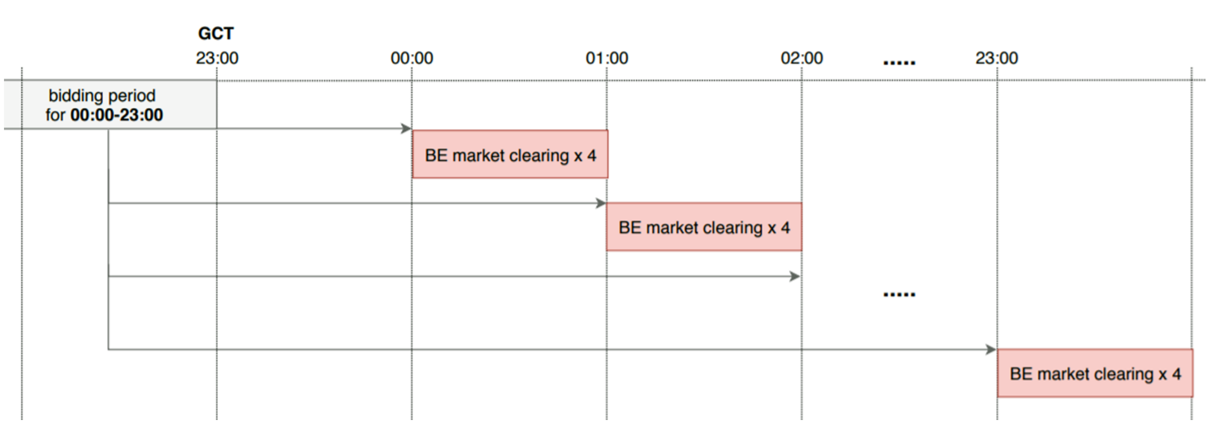

Fig. 4. Bidding procedure and market clearing in the joint BC-BE auction.

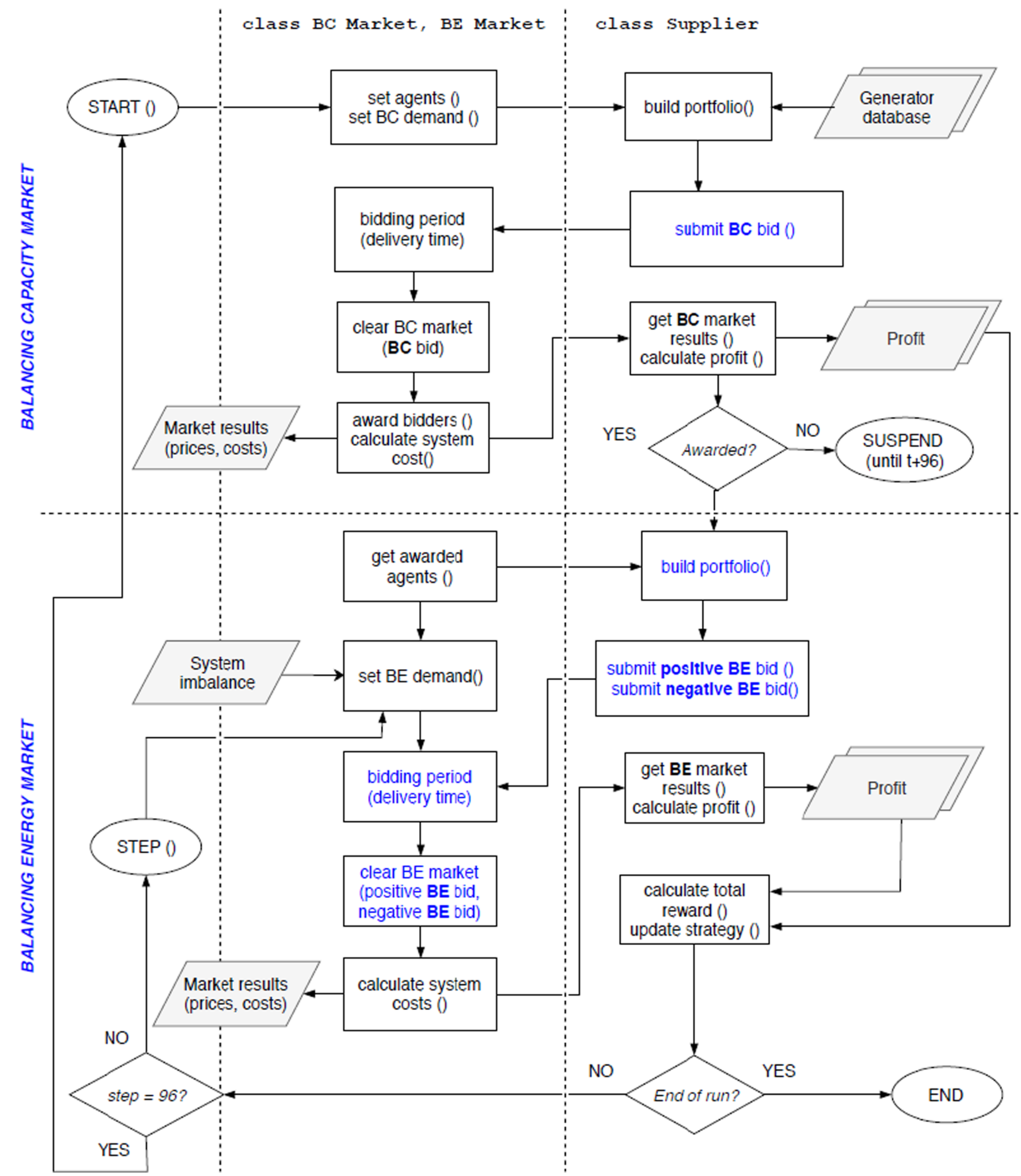

Fig. 5. General model structure diagram for a **split** balancing capacity and balancing energy auctions. The differences between the joint and split auction are marked in blue.

We use strategies 1 and 3 to compare the effects of market design changes under perfect competition and under strategic behavior whereas strategy 2 was introduced to calibrate RL agents' performance.

The analysis is based on the following hypotheses.

- f BSPs bid their true variable costs, as would be expected in a competitive market according to neo-classical economic theory, it would not matter if BC and BE markets are joint or not.
- In the market for downward regulation, if BSPs bid their true costs, they will offer to pay approximately their variable costs to the TSO in order to reduce generation output[4].
- As the number of market actors increases, the profits are expected to go down.

The performance of the agents is measured by their profits. Whereas in the BE market for + aFRR, the profit is calculated as revenue in a given delivery period minus the cost of producing additional energy, the calculation in the market for downward regulation (-aFRR) is less

[4] The cost structures of the bidders in the +aFRR and -aFRR market are different due to the fact that in the former agents increase output when the system is short, incurring generation costs, and in the latter decrease output, potentially saving costs. Thus, their bidding strategies in the two markets will also be different [48].

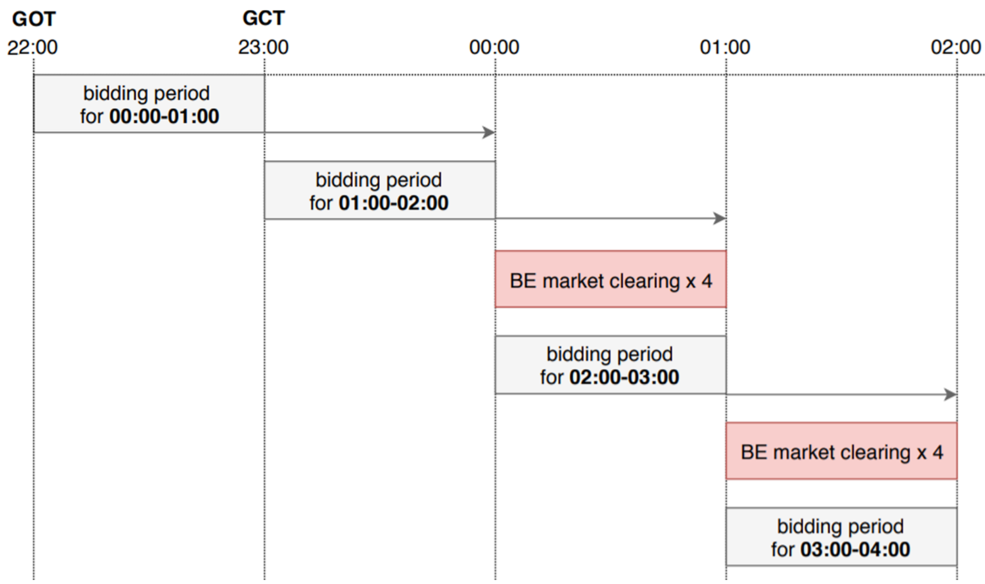

Fig. 6. Bidding procedure and market clearing in the joint BC-BE auction.

Table 1
Brief overview of the differences between the three agent types regarding their bidding strategies, use of learning and memory.

	1. **Naïve agents**	2. **Rule-based agents**	3. **RL agents**
Bid	True costs, i.e. bid is equal to generator marginal costs	A markup or markdown is added to marginal costs, depending on whether (a) the agent's bid was awarded at least once (a) in the last 2 h (split BC-BE market); (b) in the same hour on the previous day (joint BC-BE market). The size of markup is higher in peak periods.	Optimal policy for each generator in portfolio is determined based on the agent's own state, the system state, and memory dataset; actions are taken to maximize reward (see Annex B)
Learning	no	no	yes
Memory	no	short-term	long-term

straightforward. We assume that the BE bidder participates in the DA market and receives a uniform market price for the volume sold in the DA market. In the -aFRR market, BSPs are theoretically willing to pay the TSO a price up to their variable costs[5] since these costs are avoided by not having to generate the energy that they sold in the DA market [51]. Therefore, in a true-cost bidding strategy, the bid price for reducing output is equal to a generator's variable cost. As a consequence, a BSP still generates a net profit, because he/she saves his/her variable costs for the volume he/she was downward regulated, even if a he/she submits a positive bid in the BE market, i.e. pays to the TSO to reduce his/her output. Even if the profit in the BE market is zero, the BSP still generates an overall profit from the DA market. If a bidder places a bid below his/her marginal costs and the bid is accepted, he/she increases his/her profit in the BE market for -aFRR. Finally, if a BSP submits a negative price, i.e. demands to be remunerated for reducing his/her output, he/she receives an additional payment from the TSO for the balancing service. Due to minimum-load requirements, however, the volume that he/she can regulate downward is smaller than the total volume that he/she sold in the DA market.

4.3.1. True-cost bidding agents

True-cost (i.e. variable cost) bidding is expected according to neoclassical economic theory in case of perfect competition, when each actor is a price-taker. This provides a benchmark for the analysis but does not necessarily represent realistic behavior in a balancing market. Observed prices in Austria regularly reach several thousands of euro per MWh, which clearly points to strategies that significantly deviate from marginal-cost bidding [49]. To simulate strategic behavior in the balancing market, two other approaches are implemented, as described below.

4.3.2. Rule-based bidding agents

Rule-based agents bid according to a predefined rule: their variable costs are marked up or down by a coefficient that is adjusted as the model proceeds, separately for the positive and negative BE market. An agent considers whether the bidding period corresponds to a peak (from 8 am to 4 pm) or to an off-peak period (the remaining hours and weekends). By default, the value of the coefficient is equal to 1.0; for true-cost bidding agents, this is how it stays throughout the model run.

In the *split BC-BE market*, the results of two previous hours are stored. The coefficient is increased in the positive market and decreased in the negative market by 5% in an off-peak period and by 10% in a peak period if the generator was awarded at least 25% of those times, i.e. at least once in an hour (see Appendix B for details). Conversely, generators for which the condition is not fulfilled gradually revert to true-cost bidding. In the *joint BC-BE market,* the rule-based agents follow the same strategy but due to a lower bidding granularity, consider the results of the previous day for the same hour.

The strategy of true-cost and rule-based agents includes an additional consideration of situations when the marginal costs of a BSP participating in the –aFRR market happen to be higher than the DA market price. If awarded in the BC market, such a generator needs to be scheduled in the DA market to be available for downward regulation. He/She then places a negative bid for balancing energy equal to his/her marginal costs, which means that in case of activation, the TSO must pay an amount of the bid.

[5] According to game theory, optimal strategy for a BSP in the negative market would be to bid strictly negative. Yet in reality, the bidders' prices tend to be negative only in the first merit-order ranks and become positive and volatile very quickly [4].

4.3.3. Reinforcement learning agents

The learning agents use a reinforcement learning (RL) algorithm called fitted Q-iteration with which they adjust their bidding behavior to maximize their profit. The Q-iteration algorithm that has already been tested in many energy applications (e.g. [50,52]) was chosen for its relative simplicity and good performance. For instance, [53] uses fitted Q-iteration to control seasonal storage systems in the context of electricity markets. It is important to note that more advanced approaches were tested, e.g. approaches based on deep learning [54], such as double Q-learning [55], however, they were not as successful as fitted Q-iteration.

As in all RL algorithms, the method considers that the agent and the BE market can be modelled via a Markov decision process: the agent modelled by a state-action pair where each state is controlled with a discrete set of actions and transitions from one state to another are based on a probability distribution (see Appendix B). In addition, when transitioning states, the agent receives a reward representing how good the action taken was. The reward is not deterministic but generated from a probability distribution. During the training, the RL agent continuously updates and improves its policy that outputs, for each state, the optimal action that maximizes the expected value of the cumulative sum of rewards. After each round, the agent's information about its respective profits is updated. As the decision in the positive and negative balancing markets are independent from each other, separate policies are determined.

State space

To define the state space of the positive (negative) RL agent, we consider the following variables:

- The four most recently activated volumes for both the positive and negative BE market. The definition of *most recent* naturally depends on the specific gate closure times and on the market structure under study.
- The four most recent prices in the positive (negative) market
- The day-ahead market price and the corresponding hour.

It is important to note that selecting the number of recent values for the variables of interest is a design choice. We opted for four as a trade-off between computational complexity and method accuracy.

Action space

For the action space, we consider that, for each generator in its portfolio, each RL agent (BSP) bids its maximum available capacity at a variable price. Therefore, the action space is defined as a selection between a discrete set of prices for each of the agent generators.

- For the RL agent in the positive market, the action space for each generator is modeled as fifty prices log-uniformly distributed between 1 and 10 times the variable cost of the generator, i.e. the RL agent has 50 actions per generator. Then, for the total action space, the RL agent considers the set of all possible combinations (with replacement) of the fifty individual actions (see Appendix B).
- For the RL agent in the negative market, the action is space is similar. However, instead of the prices being discretized between 1 and 10 times the variable cost of each generator, they are discretize between 1 and -10 times the variable cost. The size of the action space scales similarly to the positive market.

The choices to select fifty values per generator and prices up to 10 times the variable cost are trade-offs between accuracy and computational cost.

Reward

The reward is the accumulated economic profit in the bidding period, e.g. for a balancing market with a four-hour product and market clearing of 15 min, the reward of a given state-action pair is the accumulated profit during the 16 market clearing steps.

Agent evaluation

After the initial training year (the exploration phase), the market performance is evaluated using a second simulated year (the exploitation phase). The agents' profit-maximizing bidding strategy is observed (see Appendix B for details).

4.4. Validation and sensitivity analysis

The market and the RL agent algorithms have been validated with multiple simplified scenarios to demonstrate that the agent's behavior is in line with what is expected from game theory (Bertrand competition). Bertrand competition implies competition on price and not on volume: as in this analysis, only those agents participate in the balancing energy market whose capacity was reserved in the previous market stage. Their capacity is therefore committed and cannot be changed in the balancing energy market. Validation tests replicated the main assumptions of Bertrand competition [56], two actors offering an identical product, in our case electrical energy, at the same location, balancing energy market, and a constant demand, in our case system imbalances.

Other factors in Bertrand competition that influence bidder strategies and whether they can reach a Nash equilibrium are whether the two actors have the same marginal costs and whether the demand can be covered by either actor entirely. Results of validation tests with a constant imbalance, i.e. demand for BE, show that when both agents have the same marginal costs and the demand can be covered by either of them, both agents bid their true costs, as expected from theory [56]. If their marginal costs are different, the agent with lower costs is incentivized to bid just below the (estimated) costs of the more expensive agent. The simulation results correspond in this case as well: the RL agent with marginal costs of 40€/MWh converges on a bid of 48,3€/MWh, slightly lower than the 50€/MWh bid of his/her true-cost bidding competitor, regardless of the pricing rule that is applied.

The situation is different if both agents are needed to cover the demand, ergo both of them have market power. In this case, the simulation results again correspond with theory and both agents bid high. Aside from total demand, other factors, such dynamic bidding, i.e. bidding in multiple consecutive runs, may cause agents to bid above their marginal costs due to learning effects from multiple rounds [57]. For instance, both reinforcement learning agents with the same marginal costs of 50€/MWh exploit multiple bidding rounds to develop very high bids and yet be awarded. As a result, they end up placing an average bid of 240 €/MWh despite limited demand. Our results are conservative with respect to price spikes because we consider a uniform imbalance within a 15-minute period. This excludes high but brief imbalances that may occur within the 15-minute periods.

In order to determine the best-performing RL strategy with respect to profit maximization, several configurations of the RL algorithm were tested, with regard to the number of choices when setting the bid price and the training time. Rule-based agents were used for the calibration of the RL agent. The results of sensitivity analyses showed that if RL agents could set the same maximum price in the positive/negative market of 500€/MWh/$-$500€/MWh, this produced poor results for the agents with cheap generation units due to the fact that the number of all available decisions is too broad for an agent to sufficiently test the performance of options closer to marginal costs. As a result, the RL agent is rarely awarded and has too little data about successful bids to take optimal decisions after training. Instead, the RL agent was set to be able to bid up to 10 times his/her marginal costs. Concerning training time, the RL agent is set up in such a way that it trains in the first year, whereas the following year it behaves optimally. Runs with two to five years were conducted and, since the performance of the RL agent didn't improve considerably with a greater number of training years, we used two-year simulations with one training year and one year when the RL agent behaves optimally.

5. Experiment design

In the model, reference data from the Austrian balancing market for aFRR was used [49]. Yet, the main goal of the study is not to imitate or make conclusions for this specific market. Rather, Elba-ABM is meant as a tool for testing different market results. The model is run for the split and joint BC-BE markets and market prices based on marginal bids (MP) or pay-as-bid (PaB). In each of these market designs, the following scenarios with regard to the agents were compared:

	baseline 3TC scenario				3RL scenario				1RL_5TC scenario				6RL scenario			
Description	Baseline scenario with 3 true-cost bidding agents				An oligopolistic scenario with 3 rein-forcement learning (RL) agents				Higher level of compete-tion with six agents*: 1 RL agent and 5 true-cost bidding agents				A higher level of competition with six agents: 6 RL agents			
BC-BE market	Split		Joint		Split		Joint		Split		Joint		Split		Joint	
Pricing rule	PaB	MP	PaB	MP	PaB	MP	PaB	MP	PaB	MP	PaB	MP	PaB	MP	PaB	MP

*This is a fair assumption for the number of participants as, according to the data of the Austrian TSO, the number of participants in a bidding round for aFRR varies between 5 and 10 BSPs [49].

As a baseline, the *3TC scenario* generates the prices and balancing costs that would be expected under the assumption of perfect competition. To estimate the impact of strategic bidding in an oligopoly on the market outcome, *3RL scenario* is used. These results are compared with the scenarios with a higher number of market actors to observe whether the presence of a single strategic bidder can significantly affect market efficiency (*1RL_5TC scenario*) and whether a higher number of competitors in a market with learning actors alone *(6RL scenario)* can improve market efficiency.

In order to compare market designs, similar generation portfolios were used in all scenarios in order to exclude the influence of portfolio differences on simulation results. Each agent has a portfolio of four generators with variable costs between 10 and 15, 30 and 35, 50 and 55, 70 and 75€/MWh[6]. This ensures that the results are not affected by large cost differences among agents while at the same time a stepwise merit order function can be built. In reality, one of the main prequalification requirements is a high speed of activation, which can be fulfilled only by few technologies such as hydropower, hard-coal and lignite, biomass, gas-fired power plants and CCGTs [23,47]. The variable costs of generation are approximated and assumed not to change for the period of simulation, so the different bid prices can occur only if an agent deviates from the true-cost bidding strategy. It is assumed that agents cannot split bid volumes, but can bid differently for each generator in their portfolio. The exact configuration of agent portfolios is detailed in Appendix C.

For our study, a series of assumptions related to the balancing market are made:

- The frequency at which the BE market is cleared is once per 15 min.
- Within a quarter of an hour, normally both positive and negative imbalances occur. For simplicity, only the net imbalance over 15 min (i.e. either positive or negative) is used.
- International cooperation (e.g. imbalance netting) is not considered, i.e. all imbalances are assumed to be handled within the control area.
- The BC market is assumed not to influence agents' bidding strategies because the profit in the BC market is considered negligible[7]. This assumption is based on the fact that that BE bid is independent of the BC bid [4] as well as on empirical evidence that balancing capacity prices tend to be low. BSPs bid low to secure their participation in the balancing energy market; the high balancing energy prices that are observed in practice make up for that [4]. The focus is therefore on the BE market.
- As BSPs are able to bid only a share of their total capacity for upward or downward regulation, a BSP is assumed to bid 10% of its total capacity in the balancing market [23] whereas the remainder is assumed to be bid in the DA market. The volume in the BE market is equal to the entire volume that is accepted in the BC market. BSPs are obliged to bid the entire committed volume throughout the delivery period.
- Agents are assumed to submit the same bid volume for both positive and negative generation[8].
- In order to specifically address the price levels and balancing costs under different market designs in the presence of learning agents, we use a single decision variable for the agents, their balancing energy price[9].

Many European markets are still characterized by a fairly low bidding frequency for aFRR [20]. However, the EBGL requires balancing energy to be procured as close as possible to real time. Consequently, balancing capacity auctions are expected to take place on a daily basis [8]. To account for these expected adjustments and to ensure that the design of the joint BC-BE market is comparable to the split BC-BE market, we apply a daily bidding frequency for balancing capacity.

6. Simulation results and discussion: The effect of balancing market design on the bidding behavior

The results of the 16 simulations are presented in this section; the agents and their portfolios are shown in Appendix B. Since the rule-based agents were mainly used to calibrate the RL agent, the results with rule-based agents are not included in this section. A scenario with all true-cost bidding agents is used as a baseline. The resulting market efficiency is of each market design in different scenarios is assessed based on the total cost of balancing and the weighted average prices.

In *3TC scenario*, the weighted average of the price-setting bids for + aFRR is 39 €/MWh and 48€/MWh for –FRR in both the split and joint markets and under both pricing rules[10]. The total cost of balancing for upward and downward regulation are lower under the pay-as-bid rule because there are no infra-marginal rents (see Fig. 7).

6.1. Oligopolistic scenario

In *3RL scenario* with strategic bidders (with all RL agents), the

[6] Assumptions about the costs of the generation technologies are based on the information provided in [58], [59].

[7] Interdependencies between BC and BE bids are disregarded in the current discussion and can be incorporates as a future step.

[8] Bids for + aFRR and –aFRR are submitted separately, so asymmetric bidding can be implemented easily in the model. For now, symmetric bidding is considered for simplification purposes. In practice, requirements for symmetric bidding are now considered unnecessarily restrictive with regard to the participation of new technologies, especially renewables and is expected to be substituted with asymmetric bidding, pursuant to the EBGL.

[9] It is important to note that the single decision variable and the exogenous day-ahead market prices, is not a limitation of the Elba-ABM framework. Instead, it is a design choice of the current study. The framework could in theory be used for more complex modeling, including multiple decision variables and interactions with other markets.

[10] A positive price for –FRR indicates the willingness of a BSP to pay to the TSO for reducing their output.

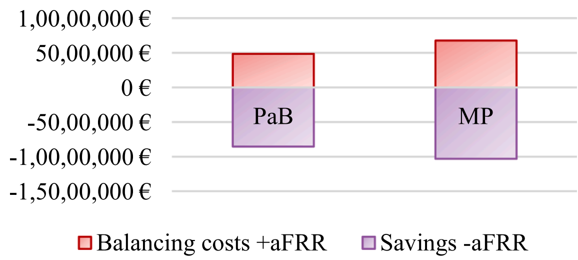

Fig. 7. Yearly balancing costs for positive balancing energy and savings from negative balancing energy under PaB and MP rules in the baseline scenarios with three true-cost bidding agents.

agents deviate considerably from the competitive strategy, notwithstanding the fact that none of them can cover the demand on their own. In the joint market with PaB pricing, the weighted average price of + aFRR is more than 7,5 times higher, at 294€/MWh, than the baseline, leading to a 3,5-increase in balancing costs. The weighted average price in the joint BC-BE market with marginal pricing also exceeds the weighted average price in the baseline, but less than the price in the scenario with the PaB rule, at 269€/MWh. For –aFRR, in turn, the weighted average marginal price falls to −73€/MW if PaB rule is applied and to −45€/MWh in case of MP, i.e. the agents make net profits from not producing and the TSO faces costs for downward regulation (Figs. 8 and 9).

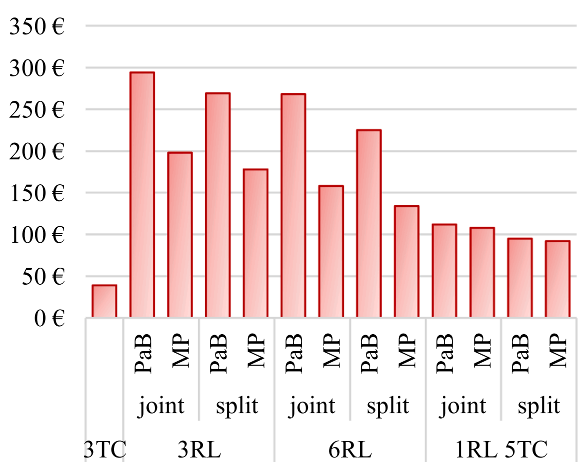

Fig. 8. Weighted average prices for + aFRR in 5 scenarios in joint and split BC-BE markets under PaB and MP rules.

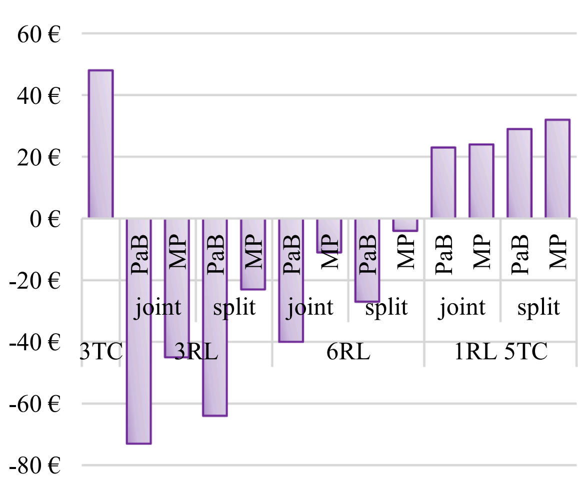

Fig. 9. Weighted average prices for −aFRR in 5 scenarios in joint and split BC-BE markets under PaB and MP rules.

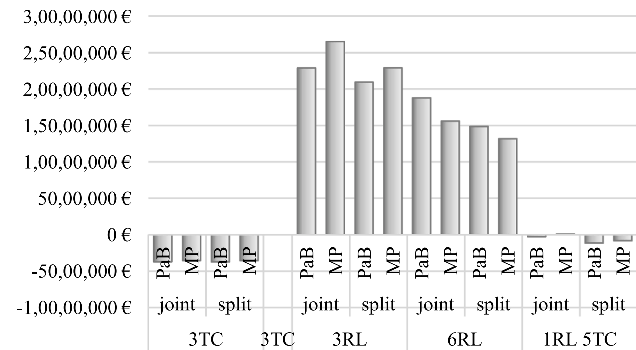

Fig. 10. Net balancing costs for upward and downward regulation for each scenario and market design option (negative values indicate net savings).

In the joint market, BSPs that bid opportunistically cannot affect the market outcome *within* the delivery period. However, this also means that if high BE bids are accepted, they apply for the entire product duration. The maximum marginal price for + aFRR regularly exceeded 700€/MWh, whereas the maximum –FRR price reached −700€/MWh 10 times in a year, largely corresponding to the times of high demand for –aFRR. In the split market, *3RL scenario* also produced average prices that were higher than the competitive benchmark, but less so than in the joint market. If the PaB rule is applied, the weighted average prices are 269€/MWh for + aFRR and −64€/MWh for –aFRR. If marginal pricing is applied, the prices decrease further: 178€/MWh for + aFRR and at –23€/MWh for –aFRR. This reduces overall balancing costs compared to the joint BC-BE market, but it still exceeds the cost of balancing in the baseline scenario by a factor of 2 to 3 for upward regulation. The total costs of balancing per scenario and market design option are shown in Fig. 10.

6.2. Scenarios with a higher degree of competitiveness

The results of the *6RL scenario* show that a more competitive market with six actors does not inoculate the market from fairly high prices if all six agents follow a RL strategy, i.e. learn from their experience and adjust their strategies in repeated auctions. The deviation from the baseline is particularly large if the PaB rule is applied: the weighted average price for + aFRR reaches 268€/MWh in the joint BC-BE market and 225€/MWh in the split BE market while –aFRR prices are −40€/MWh and −27€/MWh, respectively. Notably, the impact of a greater number of learning agents is greater for –aFRR, as is illustrated in Fig. 9. The observed cost of balancing, as compared to the oligopolistic *3RL scenario,* is more modest, yet it is still ca. 2–4 times higher than the baseline for + aFRR whereas savings in the –aFRR market go down by 76−92%, depending on the pricing rule applied (Fig. 10).

A scenario with all true-cost bidding agents and one RL agent, *1RL_5TC scenario*, was used to estimate the impact of a single learning agent on the market outcome. In this case, the RL agent is not able to deviate substantially from its marginal costs to increase its profit and does not affect the balancing costs significantly (Fig. 10). Yet, the weighted average price for + aFRR and –aFRR deviates from the competitive outcome, 92-108€/MWh for + aFRR and 23-32€/MWh for –aFRR (Figs. 8 and 9), in particular in the times of scarcity when all bidders are necessary to restore system balance. Balancing cost deviations from the competitive benchmark are the lowest in this scenario, as expected. The observed increase in total balancing costs is substantially lower, compared to the other scenarios, between 17% and 73%, where the split BC-BE market with marginal pricing produces the most cost-efficient result, as shown in Fig. 10.

The simulation results demonstrate that the balancing energy prices produced by Elba-ABM correspond to the prices observed in European balancing markets with the design modelled in *the joint BC-BE market* (e.g. in Germany and in Austria).[11] Previous research has demonstrated that the magnitude test is a useful approach to validate the results of agent-based models (cf. [61]). The real observed prices for balancing energy and the prices produced by the model both often deviate from marginal costs of the most expensive generation technologies, as is shown in Fig. 8. These simulation results confirm the argument that in concentrated balancing markets, players are able to coordinate their bids [57] and *"orientate their bids towards previous market results"* [62]. They also show how a single strategic bidder in a fairly competitive market can still at times affect the market result (*1RL_5TC scenario*). This implies that:

- Although a higher number of actors bidding competitively can dissuade their counterparts from bidding strategically by exposing them to a higher risk of not being awarded, the market is not immune to it, in particular in scarcity conditions However, a standalone BE market with marginal prices improves the incentive to place bids closer to marginal costs.
- Given these results as well as the fact that the need for larger balancing volumes is likely to grow to offset rapid integration of intermittent renewable generation, increasing the availability of balancing resources is essential. This can be achieved by easing prequalification conditions and facilitating cross-border procurement of balancing resources. The latter will in fact be enabled through EU platforms for cross-border exchange of balancing energy that are planned to be implemented by mid-2023 [63].

As the costs of balancing are at least partially recovered through network tariffs paid by consumers, the presence of strategic bidding will affect social welfare to a greater or lesser extent depending on the cost recovery scheme applicable in a given state. For instance, while the costs of reserving aFRR capacity are distributed among all grid users in most EU countries, the costs from activation of aFRR balancing energy are mostly recovered from the BRPs whose actions led to system imbalances [20].

Overall, the simulations of the split BE market consistently demonstrate more efficient market results; in the presence of true-cost bidding agents they approximate the competitive results in the baseline. At the same time, the differences in weighted average prices under the two pricing rules were observed in all scenarios and points to a tangible positive effect of marginal pricing (see Figs. 8 and 9).

In case of portfolio bidding, we find that RL agents apply a different strategy to generators with low to medium variable costs than to more expensive generators in their portfolio. Cheaper generators tend to be offered close to the variable costs while generators with higher variable costs are bid in at high prices. Consequently, they are rarely activated (2–10% of times in a year), but still allow RL agents obtain high profits during times of scarcity. Occasionally, they create price spikes of up to nine times the marginal cost of the most expensive generator.

A standalone BE market is likely to produce lower bid prices in the BE market for upward regulation and higher bid prices in the BE market for downward regulation. However, our experiments with learning agents show that also in the most efficient market design there is room for strategic behavior when the demand for balancing services is high. The effect of strategic bidding is significantly dampened if not all agents behave strategically, in particular if the uniform pricing rule is applied. The results consistently demonstrate a positive effect of the MP rule on the weighted average marginal prices in both positive and negative BE markets, especially if a standalone BE market is introduced pointing to the positive expected effect of the upcoming regulatory change. However, while the effects of these market design changes are significant, further measures to improve market access and competition are needed to make the balancing market robust against gaming.

[11] Specifically, the model results were compared with the prices for aFRR in Austria (time series of years 2017 and 2018 [60]).

7. Conclusions

We presented an agent-based model, Elba-ABM, to provide an insight into the effects of proposed changes to European balancing market design, in particular the introduction of a standalone balancing energy market and marginal-price settlement of energy bids, on strategic bidding in the balancing market. The agents are modelled with realistic generation portfolios and learning agents are equipped with reinforcement learning (using a neural network) to identify opportunities for strategic behavior. Using Elba-ABM, we assessed the results with respect to the profits of agents, the weighted average prices of positive and negative balancing energy and the total cost of balancing.

Testing the robustness of the new market design with a standalone balancing energy market, we came to the following conclusions:

(1) A split (standalone) balancing energy market reduces balancing costs and weighted average prices, compared to a joint BC-BE market. It is particularly helpful in case of an oligopoly, even though it does not solve the issue of market power in case of high market concentration entirely. Concerns that were raised about the negative effects of more frequent opportunities for learning leading to gaming [64] in case of highly granular markets were not supported by the simulation results.
(2) Marginal pricing performs better than pay-as-bid, regardless of whether the BE market is standalone or not.
(3) The fact that in more competitive scenarios the results of the joint and split balancing capacity and energy markets do not substantially differ from each other confirms the expectation that in a more competitive market, its exact design is less relevant and the results of different market designs are more likely to converge. But as long as balancing markets remain concentrated, a standalone balancing energy market is preferred since (a) in a closed setting of an oligopoly, a standalone BE market reduces agents' ability to affect market outcome; (b) it can be combined with voluntary bids, which can help dampen balancing energy prices.
(4) The new market design choices are likely to improve market performance but more new entrants are needed to obtain competitive prices. Therefore, particular attention should be given to market access conditions, such as reduction of minimum bid size, aggregated and asymmetric bidding (as pointed out in [2]), along with market design adaptations, in view of many new types of flexibility providers that are emerging.

Our methodological contribution consists of a novel combination of agent-based modelling with reinforcement learning techniques. Elba-ABM represents both a detailed model of the market and of the market actors. Their different characteristics, constraints and objectives, the absence of perfect foresight and other conditions of perfect competition are reflected in the model. Reinforcement learning techniques make it possible to emulate strategic behavior in a market in which actors explore opportunities for increasing their profits through different bidding strategies. We will build on this approach in future work to test other market design variables, integrate intertemporal constraints and to apply agent-based modelling to more complex cases with interrelated markets. A second tier of research should address approaches to the recovery of balancing costs and their effect on social welfare together with an investigation of links between balancing costs, distribution of imbalance costs and network tariffs.

CRediT authorship contribution statement

Ksenia Poplavskaya: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft. **Jesus Lago:** Methodology, Software, Writing - review & editing. **Laurens Vries:** Supervision, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank Werner Friedl for the valuable inputs and discussions that contributed to this paper as well as the reviewers for the thoughtful comments that helped us improve the quality of our work.

This research has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 675318 (INCITE).

Data Availability

Datasets related to this article can be found at https://www.apg.at/en/markt/netzregelung/statistik, hosted at the website of the Austrian transmission system operator, APG [60].

Appendix

Appendix A. List of abbreviations

ABM – agent based modeling
aFRR – automatic frequency restoration reserve
BC – balancing capacity
BE – balancing energy
BRP – balance responsible party
BSP – balancing service provider
DER – distributed energy resources
DSO – distribution system operator
EBGL – EU Regulation establishing a guideline on electricity balancing
FCR – frequency containment reserve
GOT – gate opening time
GCT – gate closure time
MCP – marginal clearing price
MP – marginal pricing
PaB – pay-as-bid
mFRR – manual frequency restoration reserve
TSO – transmission system operator
RES – renewable energy sources

Appendix B. Algorithms used to represent reinforcement learning strategies and rule-based

Primer on reinforcement learning (RL)

In general terms, the RL algorithm is formulated in line with the main principles of Markov decision processes, as per [65]. In particular, at each time step k the agent is defined by a state s_k, takes an action a_k, and transitions from $s_k$to s_{k+1} following some probabilistic dynamics $p(s_k, a_k)$. In the transition, it receives a reward r_k following a distribution $q(s_k, a_k)$ that represents the profit of taking action a_k at state s_k. The goal of the agent is to first learn the optimal policy $a^* = \pi^*(s_k)$ during an exploration phase, i.e. training, and then use that policy during an exploitation phase, i.e. regular operation.

During the exploration phase, the policy is improved based on the agent's memory M that contains tuples of state, transitioned state, action taken, and reward collected during each transition:

$$M = \{s_k, a_k, s_{k+1}, r_k\}_{k=1}^{T_e}$$

During this exploration phase, the actions are chosen both at random and by using the current best available policy; by doing so, the agent explores new combinations (s_k, a_k) of state and action pairs and ensures that the ones that seem optimal so far are indeed the best. After the training is completed, the agent's optimal policy, $\pi^*(s_k)$ attempts to maximize the expected cumulative sum of rewards, R, over the entire episode, T_e:

$$R = \sum_{k=1}^{T_e} \gamma^{T_e - k} E_{q(s_k, a_k)}\{r_k\}$$

where γ is the discount factor and E is expected value.

Reinforcement learning for the balancing market

The RL algorithm used in this study is based on [65,66] and adapted to the balancing market model, Elba-ABM. Agents are embedded in the market environment, as is shown in the flow diagrams in Figs. 3 and 5.

The actions represent bid prices that can be submitted by the RL agent, for each delivery period, k. Agent's step k, corresponds to the bidding period and is equal to one hour. Note that the system state, i.e. information the agent receives from the balancing market, is included in the agent state. As upward and downward regulation are procured in separate auctions, the agent's policies in these two markets are determined separately, i.e. we effectively consider a RL agent for the positive balancing market and another one for the negative balancing market.

For the sake of keeping a reasonable level of discretization and computation time, the maximum bid price is set to 10 times a generator's marginal costs (or 10 times less than a generator's costs in the negative market) whereas the action space is set to contain 50 actions per generator in the agent's portfolio: $A_g = \{a^1, \cdots a^{50}\} g \in G$. For an agent with n generators, the action space has a size of:

$$C^R(50, n) = \frac{(50+n-1)!}{n!(50-1)!}$$

For instance, for a portfolio consisting of 3 generators, 19,600 possible combinations are considered.

At every step k and for each market clearing *i,* the agent receives some reward $r_{k,i}$ where bids are submitted on an hourly basis and the market is cleared every 15 min, $i = 1, 2, 3, 4$ (see also Fig. 4). In particular, if the imbalance is negative, the agent receives a reward $r_{k,i}$ equal to the market price, $\lambda_{k,i}^{+\text{aFRR}}$ times the volume of awarded balancing energy $q_{k,i}^{+\text{aFRR}}$. If the imbalance is positive, the agent receives a reward $r_{k,i} = \lambda_{k,i}^{-\text{aFRR}} q_{k,i}^{-\text{aFRR}}$.

Then, to define the reward at time step k, the agent of the positive market considers the average income during the periods where the positive market was cleared:

$$r_k^{+\text{aFRR}} = \frac{1}{n^{+\text{aFRR}}} \sum_{i \in \mathcal{N}_p} \lambda_{k,i}^{+\text{aFRR}} q_{k,i}^{+\text{aFRR}}$$

where $\mathcal{N}_p = \{i | i = 1, \ldots, 4, I_{k,i} < 0\}$ and $I_{k,i}$ is the system imbalance for the market clearing i of time step k. This means that the reward in the positive market agent at times k is the average profit during the times when the system was short, so upward regulation was needed. The expression for the reward of the negative agents is the same. It is important to note that data in the memory $\boldsymbol{M}$ is only added if the specific (positive or negative) market is cleared. That is, if during transition from k to $k+1$ only the -aFRR market is cleared, the equation above would indicate that the profit of the positive market was zero and vice verse for the + aFRR market.

In order to train the agents, we consider a balancing market simulation period of a year, during which the memory is updated. The RL agents are trained in the presence of other, non-RL agents, in the market, if they are part of the scenario. The agents are trained with the fitted Q-iteration [50,52]. For the sake of simplicity and because the algorithm used is very standard, its mathematic details are not provided in this paper. However, the interested reader can consult [52] for further information.

During the exploitation phase in the second year, the agent uses the collected information to bid optimally. Using the optimal policy, the RL agent takes an optimal action a_k^* for each state, defining the agent's bidding strategy:

$$a_k^* = \pi^*(\lambda_{k-1,4}^{+\text{aFRR}}, \cdots \lambda_{k-1,1}^{+\text{aFRR}}, \lambda_{k-1,4}^{-\text{aFRR}}, \cdots \lambda_{k-1,1}^{-\text{aFRR}}, \lambda_k^{\text{DA}}, q_{k-1,4}^{+\text{aFRR}}, \cdots q_{k,i-1}^{+\text{aFRR}}, q_{k-1,4}^{-\text{aFRR}}, \cdots q_{k-1,1}^{-\text{aFRR}})$$

where λ_k^{DA} is the DA market price in the current hour.

Rule-based agent

The rule-based agent has a pre-defined strategy and is primarily used in the model for the calibration of the reinforcement learning agent.

The agent has a short-term memory of the previous success separately for each generator in his portfolio, expressed through binary variables, $d_{k,i}^{+\text{aFRR}} \in \{0, 1\}$, $d_{k,i}^{-\text{aFRR}} \in \{0, 1\}$, denoting whether the agent's generator was awarded in the positive or negative market per market clearing *i*in hour k. The strategy further differentiates between peak delivery hours, k^{peak}, and off-peak hours, k^{offpeak}: $K^{\text{peak}} = \{8, 9, \cdots 16\} \not\ni k^{\text{offpeak}}$.

Then, based on the information saved in memory (see below), each generator uses and updates four coefficients, $\beta_{k-1}^{+\text{aFRR,peak}}$, $\beta_{k-1}^{+\text{aFRR,offpeak}}$, $\beta_{k-1}^{-\text{aFRR,peak}}$, $\beta_{k-1}^{-\text{aFRR,offpeak}}$, to build new bids. The bid for each market and each period is built by multiplying the coefficient by the marginal cost of the generators, c_g. The former is equal to 1.0 by default. Each of the four coefficients can vary between [0.5, 1.5] and they are updated following the following rule: if during the last eight positive market periods, i, the generator was awarded at least once per hour on average, i.e. the success ratio, $\omega^{+\text{aFRR,peak}} \geq 0.25$, the coefficients increase the bid markup by 5% or 10% in the off-peak or peak bidding period, respectively . To sum up, the bid price in the positive market is determined as follows:

$$b_{g,k}^{+\text{aFRR}} = \begin{cases} c_g * \beta_{k-1}^{+\text{aFRR,peak}} + 0,1, & if\ k \in K^{\text{peak}} \\ c_g * \beta_{k-1}^{+\text{aFRR,offpeak}} + 0,05, & if\ k \in K^{\text{offpeak}} \\ c_g * \frac{1+\beta_{k-1}^{+\text{aFRR,peak}}}{2}, & if\ \omega^{+\text{aFRR,peak}} \leq 0.25 \end{cases}$$

Using the same 5% or 10% markups, the agent gradually reverts to true-cost bidding for those generator bids that were not awarded. Following a similar strategy in the –aFRR market, the coefficient is reduced by 5% or 10% if the condition is fulfilled.

For the rule-based agent in the *split BC-BE market*, the results of the previous two hours are memorized. For instance, for the peak periods in the positive market, the memory for each generator in the agent's portfolio contains:

$$M = \{\beta_{k-1}^{+\text{aFRR,peak}}, d_{i-8}^{+\text{aFRR,peak}}, \cdots d_{i-1}^{+\text{aFRR,peak}}\}$$

where $\beta_{h-1}^{+\text{aFRR,peak}}$is the last coefficient used in the positive market in the peak period.

In the *joint BC-BE market,* the strategy is slightly adapted to account for the lower bidding frequency. The algorithm remains the same but, due to daily bidding, the results of the same hour but of the previous day are considered for the rule-based agent. The results of previous Friday are considered if the bidding takes place on a Monday. All hours of Saturday and Sunday are considered off-peak.

Appendix C. Agents and portfolios used in the simulation scenarios

Total demand for aFRR capacity: ± 200 MW.

It is assumed that each generator can provide 10% of its capacity for balancing. The available balancing capacity of all generators is slightly exceeds the set total BC demand to ensure a specific number of participants in the subsequent BE market. Generator variable costs are approximated; the cost differences are kept low and each agent receives a generator in each cost category in order to eliminate portfolio effects on the market outcome.

Scenario with an oligopoly				Scenario with a higher competition level			
Agent ID	GeneratorID	Variable costs, €/MWh	Installed capacity MW	Agent ID	Generator ID	Variable costs, €/MWh	Installed capacity, MW
1	a	10	170	**1**	a	10	85
	e	32	170		h	31	85
	i	55	170		o	52	85
	j	70	170		v	73	85
2	b	12	170	**2**	b	11	85
	f	35	170		i	32	85
	g	50	170		p	53	85
	k	72	170		w	74	85
3	c	15	170	**3**	c	12	85
	d	30	170		j	33	85
	h	52	170		q	54	85
	l	75	170		x	75	85
				4	d	13	85
					k	34	85
					r	55	85
					s	70	85
				5	e	14	85
					l	35	85
					m	50	85
					t	71	85
				6	f	15	85
					g	30	85
					n	51	85
					u	72	85

References

[1] ENTSO-E. "All TSOs' proposal for the implementation framework for a European platform for the exchange of balancing energy from frequency restoration reserves with automatic activation in accordance with Article 21 of Commission Regulation (EU) 2017/2195 establishing a guideline on electricity balancing," ENTSO-E, Brussels, Belgium, Dec. 2018. Accessed: Dec. 21, 2018. [Online]. Available: < https://docstore.entsoe.eu/Documents/nc-tasks/EBGL/EBGL_A21_181218_ALL%20TSOs%20proposal_aFRRIF_for%20submission.pdf?Web=0 > .

[2] Poplavskaya K, De Vries LJ. Distributed energy resources and the organized balancing market: a symbiosis yet? Case of Three European Balancing Markets. Energy Policy 2019;126:264–76. https://doi.org/10.1016/j.enpol.2018.11.009.

[3] Clements AE, Hurn AS, Li Z. Strategic bidding and rebidding in electricity markets. Energy Econ 2016;59:24–36. https://doi.org/10.1016/j.eneco.2016.07.011.

[4] Ocker F, Ehrhart K-M, Ott M. Bidding strategies in Austrian and German balancing power auctions. WIREs Energy Environ 2018;7(6):e303. https://doi.org/10.1002/wene.303.

[5] Lago J, De Ridder F, Vrancx P, De Schutter B. Forecasting day-ahead electricity prices in Europe: the importance of considering market integration. Appl Energy 2018;211:890–903. https://doi.org/10.1016/j.apenergy.2017.11.098.

[6] Directive 2019/44/EU, "Directive (EU) 2019/44 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU (recast). OJ L 158, 14.06.2019, p. 125–195." 2019.

[7] Regulation 2019/943/EU, "Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity (recast). OJ L 158, 14.06.2019, p. 54-124.".

[8] European Commission. Commission Regulation (EU) 2017/2195 of 23 November 2017 establishing a guideline on electricity balancing; 2017.

[9] Möller C, Rachev ST, Fabozzi FJ. Balancing energy strategies in electricity portfolio management. Energy Econ 2011;33(1):2–11. https://doi.org/10.1016/j.eneco.2010.04.004.

[10] Merten M, Olk C, Schoeneberger I, Sauer DU. Bidding strategy for battery storage systems in the secondary control reserve market. Appl. Energy 2020;268:114951https://doi.org/10.1016/j.apenergy.2020.114951.

[11] Romero Rodríguez L, Brennenstuhl M, Yadack M, Boch P, Eicker U. Heuristic optimization of clusters of heat pumps: a simulation and case study of residential frequency reserve. Appl Energy 2019;233–234:943–58. https://doi.org/10.1016/j.apenergy.2018.09.103.

[12] Angenendt G, Merten M, Zurmühlen S, Sauer DU. Evaluation of the effects of frequency restoration reserves market participation with photovoltaic battery energy storage systems and power-to-heat coupling. Appl Energy 2020;260:114186https://doi.org/10.1016/j.apenergy.2019.114186.

[13] Muhssin MT, Cipcigan LM, Sami SS, Obaid ZA. Potential of demand side response aggregation for the stabilization of the grids frequency. Appl Energy 2018;220:643–56. https://doi.org/10.1016/j.apenergy.2018.03.115.

[14] Borne O, Korte K, Perez Y, Petit M, Purkus A. Barriers to entry in frequency-regulation services markets: review of the status quo and options for improvements. Renew Sustain Energy Rev 2018;81:605–14. https://doi.org/10.1016/j.rser.2017.08.052.

[15] Dallinger B, Auer H, Lettner G. Impact of harmonised common balancing capacity procurement in selected Central European electricity balancing markets. Appl Energy 2018;222:351–68. https://doi.org/10.1016/j.apenergy.2018.03.120.

[16] Doorman GL, Van der Veen R. An analysis of design options for markets for cross-border balancing of electricity. Util Policy 2013;27:39–48. https://doi.org/10.1016/j.jup.2013.09.004.

[17] Just S, Weber C. Strategic behavior in the German balancing energy mechanism: incentives, evidence, costs and solutions. J Regul Econ 2015;48(2):218–43. https://doi.org/10.1007/s11149-015-9270-6.

[18] Müsgens F, Ockenfels A, Peek M. Economics and design of balancing power markets in Germany. Int J Electr Power Energy Syst 2014;55:392–401. https://doi.org/10.1016/j.ijepes.2013.09.020.

[19] Farahmand H, Doorman GL. Balancing market integration in the Northern European continent. Smart Grids 2012;96:316–26. https://doi.org/10.1016/j.apenergy.2011.11.041.

[20] ENTSO-E, "Survey on Ancillary Services Procurement, Balancing Market Design 2016," ENTSO-E, Brussels, 2017. Accessed: Oct. 16, 2017. [Online]. Available: < https://www.entsoe.eu/Documents/Publications/Market%20Committee%20publications/WGAS_Survey_final_10.03.2017.pdf > .

[21] Ocker F, Braun S, Will C. Design of European balancing power markets. In: European Energy Market (EEM), 2016 13th International Conference on the, 2016, pp. 1–6, Accessed: Feb. 26, 2017. [Online]. Available: < http://ieeexplore.ieee.org/abstract/document/7521193/ > .

[22] Hartwig K, Kockar I. Impact of strategic behavior and ownership of energy storage on provision of flexibility. IEEE Trans Sustain Energy 2016;7(no. 2): p. 11, Apr. 2016, Accessed: Apr. 15, 2016. [Online].

[23] Maaz A. Auswirkungen von strategischem Bietverhalten auf die Marktpreise am deutschen Day-Ahead-Spotmarkt und an den Regelleistungsauktionen, PhD thesis, Institut für Elektrische Anlagen und Energiewirtschaft Forschungsgesellschaft Energie an der RWTH Aachen, Aachen, Germany; 2017.

[24] Van der Veen R. Designing Multinational Electricity Balancing Markets. Doctoral Thesis, Technische Universiteit Delft; 2012.

[25] Keles D, Bublitz A, Zimmermann F, Genoese M, Fichtner W. Analysis of design options for the electricity market: the German case. Appl Energy 2016;183:884–901. https://doi.org/10.1016/j.apenergy.2016.08.189.

[26] Abbasy A. National Design and Multinational Integration of Balancing Services Markets, PhD thesis, Delft University of Technology, Next Generation Infrastructures Foundation; 2012.

[27] Lopes F, Coelho H, editors. Electricity markets with increasing levels of renewable generation: structure, operation, agent-based simulation, and emerging designs, vol. 144. Cham: Springer International Publishing; 2018.

[28] Khan ASM, Verzijlbergh RA, Sakinci OC, De Vries LJ. How do demand response and electrical energy storage affect (the need for) a capacity market? Appl Energy 2018;214:39–62. https://doi.org/10.1016/j.apenergy.2018.01.057.

[29] Bagnall AJ, Smith GD. A multiagent model of the uk market in electricity generation. IEEE Trans Evol Comput 2005;9(5):522–36. https://doi.org/10.1109/TEVC.2005.850264.

[30] Li G, Shi J. Agent-based modeling for trading wind power with uncertainty in the day-ahead wholesale electricity markets of single-sided auctions. Appl Energy 2012;99:13–22. https://doi.org/10.1016/j.apenergy.2012.04.022.

[31] Gallo G. An integrated agent-based and production cost modeling framework for

renewable. Energy Stud 2016:2390–9. https://doi.org/10.1109/HICSS.2016.299.

[32] Helbing D. ed. Agent-based modelling. In: Social Self-Organization, Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 25–70.

[33] Sensfuß F. Assessment of the impact of renewable electricity generation on the German electricity sector. An agent-based simulation approach. Karlsruhe, Germany: University of Karlsruhe; 2007.

[34] Weidlich A, Veit D. A critical survey of agent-based wholesale electricity market models. Energy Econ 2008;30(4):1728–59. https://doi.org/10.1016/j.eneco.2008.01.003.

[35] Mielczarski W, Michalik G, Widjaja M. Bidding strategies in electricity markets. Presented at the International Conference on Power Industry Computer Applications, Santa Clara, California; May 1999.

[36] Esmaeili Aliabadi D, Kaya M, Sahin G. Competition, risk and learning in electricity markets: an agent-based simulation study. Appl Energy 2017;195:1000–11. https://doi.org/10.1016/j.apenergy.2017.03.121.

[37] Böttger D, Götz M, Theofilidi M, Bruckner T. Control power provision with power-to-heat plants in systems with high shares of renewable energy sources ? An illustrative analysis for Germany based on the use of electric boilers in district heating grids. Energy 2015;82:157–67. https://doi.org/10.1016/j.energy.2015.01.022.

[38] Bublitz A, Genoese M, Fichtner W. An agent-based model of the German electricity market with short-time uncertainty factors; May 2014. p. 1–5, doi: 10.1109/EEM.2014.6861215.

[39] Ventosa M, Baíllo Á, Ramos A, Rivier M. Electricity market modeling trends. Energy Policy 2005;vol. 33(7): p. 897–913. doi: 10.1016/j.enpol.2003.10.013.

[40] Brijs T, De Jonghe C, Hobbs BF, Belmans R. Interactions between the design of short-term electricity markets in the CWE region and power system flexibility. Appl Energy 2017;195:36–51. https://doi.org/10.1016/j.apenergy.2017.03.026.

[41] Sun J, Tesfatsion L. Dynamic testing of wholesale power market designs: an open-source agent-based framework. Comput Econ 2007;30(3):291–327. https://doi.org/10.1007/s10614-007-9095-1.

[42] Heim S, Götz G. Do Pay-as-bid Auctions Favor Collusion? Evidence from Germany's Market for Reserve Power. Zentrum für Europäische Wirtschaftsforschung, Berlin, Discussion Paper No. 13-035; 2013.

[43] Ocker F, Ehrhart K-M. The 'German Paradox' in the balancing power markets. Renew Sustain Energy Rev 2017;67:892–8. https://doi.org/10.1016/j.rser.2016.09.040.

[44] Aliabadi DE, Kaya M, Şahin G. An agent-based simulation of power generation company behavior in electricity markets under different market-clearing mechanisms. Energy Policy 2017;100:191–205. https://doi.org/10.1016/j.enpol.2016.09.063.

[45] Vermeulen B, Pyka A. Agent-based modeling for decision making in economics under uncertainty. Econ Open-Access Open-Assess E-J 2016. https://doi.org/10.5018/economics-ejournal.ja.2016-6.

[46] Hirth L, Ziegenhagen I. Balancing power and variable renewables: three links. Renew Sustain Energy Rev 2015;50:1035–51. https://doi.org/10.1016/j.rser.2015.04.180.

[47] regelleistung.net, "Official website of the German TSOs - Ausschreibungskalender." < https://www.regelleistung.net/ext/calendar/tender > .

[48] Poplavskaya K, Ocker F, Ehrhart KM. Impact of short-term market sequences on bidding behavior of market participants. In: Proceedings of the 3rd European Grid Service Markets Symposium (GSM), Lucerne, Switzerland; Jun. 2019. p. 73–86, doi: 10.5281/zenodo.3355399.

[49] APG Austrian Power Grid, "Official website of the Austrian TSO - Ausschreibungssystem - Regelenergiemarkt,"; 2019. < https://www.apg.at/emwebapgrem/startApp.do > (accessed Sep. 20, 2019).

[50] Lago J, Sogancioglu E, Suryanarayana G, De Ridder F, De Schutter B. Building day-ahead bidding functions for seasonal storage systems: a reinforcement learning approach. Presented at the Workshop on Control of Smart Grids and Renewable Energy Sources (CSGRES); 2019.

[51] Ocker F. Balancing Power Auctions. Theoretical and Empirical Analyses. Germany: Karlsruhe Institute of Technology (KIT), Karlsruhe; 2018.

[52] Ernst D, Geurts P, Wehenkel L. Tree-based batch mode reinforcement learning. J Mach Learn Res 2005;vol. 6(Apr): p. 503–56.

[53] Lago J, De Ridder F, Mazairac W, De Schutter B. A 1-dimensional continuous and smooth model for thermally stratified storage tanks including mixing and buoyancy. Appl Energy 2019;248:640–55. https://doi.org/10.1016/j.apenergy.2019.04.139.

[54] Lago J, De Ridder F, De Schutter B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Appl Energy 2018;221:386–405. https://doi.org/10.1016/j.apenergy.2018.02.069.

[55] Van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double Q-learning. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16), Phoenix, USA, Feb. 2016, [Online]. Available: < https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/download/12389/11847 > .

[56] Muendler M. "Oligopolisitic Competition 1. Industrial Organization Oligopolistic Competition. ECON 312, Lecture 4," University of California, Berkeley, 2018, Accessed: Nov. 20, 2019. [Online]. Available: < http://docplayer.net/4353818-Econ-312-oligopolisitic-competition-1-industrial-organization-oligopolistic-competition.html > .

[57] Ehrhart K, Ocker F. Allocation, prices, and costs in the electricity wholesale market and balancing power market – an integrated approach; 2018.

[58] Hirth L. "The European Electricity Market Model EMMA. Model documentation." 2017, [Online]. Available: < https://neon-energie.de/emma-documentation.pdf > .

[59] Pérez-Arriaga IJ. Regulation of the power sector. New York: Springer; 2013.

[60] APG Austrian Power Grid, "Balancing statistics in APG control area," APG; 2019. < https://www.apg.at/en/markt/netzregelung/statistik > (accessed Mar. 02, 2019).

[61] Weidlich A. Engineering interrelated electricity markets: an agent-based computational approach. Heidelberg: Physica-Verl; 2009.

[62] Ocker F. Design and performance of European balancing power auctions. In: presented at the European Energy Market (EEM), 2017 14th International Conference on the; Jun. 2017. p. 1–6, doi: 10.1109/EEM.2017.7981861.

[63] ACER, "Decision No 02/2020 of the European Union Agency for the Cooperation of Energy Regulators of 24 January 2020 on the Implementation framework for the European platform for the exchange of balancing energy from frequency restoration reserves with automatic activation." Jan. 24, 2020, Accessed: Feb. 02, 2020. [Online]. Available: < https://www.acer.europa.eu/Media/News/Pages/ACER-adopts-decisions-towards-a-single-EU-electricity-balancing-market.aspx > .

[64] Bundesnetzagentur, "Beschluss in dem Verwaltungsverfahren zur Festlegung von Ausschreibungsbedingungen und Veröffentlichungspflichten für Sekundärregelung." Az.: BK6-15-158, Jun. 13, 2017, Accessed: Jan. 10, 2018. [Online]. Available: < https://www.bundesnetzagentur.de/DE/Service-Funktionen/Beschlusskammern/1BK-Geschaeftszeichen-Datenbank/BK6-GZ/2015/2015_0001bis0999/BK6-15-158/BK6-15-158_download_Beschluss_vom_13_06_2017.pdf?__blob=publicationFile&v=3 > .

[65] Sutton RS, Barto AG. Reinforcement learning: an introduction, Nachdr. Cambridge, Mass: MIT Press; 2017.

[66] Lago J, Suryanarayana G, Sogancioglu E, De Schutter B. Optimal control strategies for seasonal thermal energy storage systems with market interaction. IEEE Trans Control Syst Technol 2020. in preparation.