

ORIGINAL RESEARCH

Open access

Coordinating Wind Power and Energy Storage in Wholesale and Ancillary Service Markets Using Deep Reinforcement Learning

Mohammed Al-Farsi^{1*}, Salim Al-Harthy¹, Nasser Al-Rawahi²

Abstract

The increasing deployment of wind power plays a key role in addressing climate change. Nevertheless, the intermittent nature of wind generation often leads to curtailment, resulting in substantial financial losses for wind farm operators. Battery energy storage systems (BESS) can help alleviate curtailment by providing local backup power. However, dedicating storage primarily to this supporting function can greatly diminish its profitability in energy markets. Optimal BESS operation requires carefully balancing local curtailment mitigation with participation in market bidding, yet achieving this in practice remains difficult owing to the intricate coordination required and the uncertainty of both electricity prices and wind output. This study explores a combined bidding strategy for a co-located wind-plus-battery setup operating in the spot market and the Regulation Frequency Control Ancillary Service markets. We introduce an innovative deep reinforcement learning method that separates the overall market engagement into two interconnected Markov decision processes—one for each asset. This structure allows the BESS to capture surplus wind energy that would otherwise be curtailed while simultaneously participating in joint bidding across markets to optimize total operational income. Validation was performed using real-world wind farm data. The proposed coordinated bidding approach outperformed a conventional optimization-based reference case, delivering roughly 25% higher revenue and achieving 2.3 times greater reduction in wind curtailment. Findings demonstrate that participating jointly in multiple markets substantially enhances the economic viability of wind-battery installations relative to engaging in individual markets in isolation. Additional simulations indicate that charging the BESS with otherwise curtailed wind power yields further economic benefits. Effective deployment of the developed algorithm is expected to promote greater co-location of renewable generation and storage facilities, thereby realizing broader advantages for the power system.

Keywords Wind-battery system, Wind curtailment, Electricity market, Deep reinforcement learning

*Correspondence:

Mohammed Al-Farsi
mohammed.alfarsi@gmail.com

¹ Department of Smart Grid Engineering, Faculty of Engineering, Sultan Qaboos University, Muscat, Oman

² Department of Energy Analytics and Digital Systems, Faculty of Engineering, German University of Technology in Oman, Muscat, Oman

Introduction

Over the last ten years, wind power has emerged as a critical contributor to decarbonizing electricity grids and supporting the global shift toward net-zero emissions [1]. In Australia, wind generation accounts for 9.9% of total electricity output [2] and is the primary utility-scale renewable source. At present, 9.7 GW of wind capacity is operational within the Australian National Electricity Market (NEM), while an additional 77 GW is slated for development in the coming decade [3]. Despite this expansion, the intermittent character of wind resources and limitations in forecasting accuracy complicate real-time

integration and dispatch, often forcing the market operator to curtail output to maintain grid stability and security [4]. Such curtailments impose direct financial burdens on wind generators. As wind penetration has risen, so too has the volume of curtailed energy, translating into notable economic setbacks for producers. In parallel, battery energy storage systems (BESS) have gained traction, with around 650 MW currently registered in the NEM and a further 34.3 GW projected over the next ten years [3].

Co-locating renewable generation with BESS is gaining momentum because integrated development of these assets can curb curtailment, broaden income sources, reduce exposure to market risks, and postpone costly grid infrastructure upgrades. This approach aligns with recommendations from the Australian Energy Market Operator (AEMO), which advocates pairing renewables and storage within designated renewable energy zones during long-term system planning [5]. Given the rapid growth of both wind and BESS deployments, it is essential to identify effective coordination mechanisms that enhance grid support and improve the economic performance of these assets as the world moves toward a low-carbon energy landscape.

The viability of utilizing curtailed wind power for storage hinges on the capital expenses of BESS and the quality of operational coordination between the generation and storage components. When sited together, BESS can absorb excess wind output during high-generation intervals and discharge it when wind production is lower, thereby cutting curtailment. For example, Ref. [6] introduced a scenario-based stochastic optimization technique to reduce wind curtailment, while Ref. [7] proposed a two-stage optimization framework, and Ref. [6] formulated a linear programming model to determine optimal BESS power and energy capacity for managing curtailed wind. Earlier investigations have generally relied on stochastic or robust optimization to size and operate BESS in such setups, necessitating detailed information on the probability distributions of wind uncertainty. However, assigning BESS primarily to this supporting function can constrain its capacity to earn revenue from electricity market activities, such as energy arbitrage in the wholesale spot market (purchasing at low prices and selling at high prices) and delivering frequency control ancillary services (FCAS) in dedicated FCAS markets—both of which represent vital profit channels for large-scale storage in today's power systems.

Optimization techniques have similarly been employed to examine coordinated bidding between wind farms and BESS in electricity markets. For example, Ref. [8] developed a market-focused optimal dispatch approach for a wind installation paired with a hybrid BESS. Works [9, 10] developed stochastic and robust optimization frameworks, respectively, for wind-battery hybrids operating in power markets. These efforts usually treat the wind farm and BESS as distinct participants that submit combined bids, yet they rarely incorporate explicit management of wind curtailment within the BESS bidding logic. Moreover, the effectiveness of such strategies depends heavily on accurate energy price predictions, which remain challenging owing to pronounced price volatility and the multifaceted factors that influence market prices [11].

Limited attention has been given to real-time bidding approaches for co-located wind-BESS configurations that move beyond traditional optimization. The present work fills this void by proposing a deep reinforcement learning (DRL) strategy tailored for co-located wind and battery systems. This method enables simultaneous curtailment reduction and revenue maximization via integrated bidding in the spot market and regulation FCAS markets. Being model-free, DRL learns the inherent uncertainties in wind output and price dynamics directly from historical data, eliminating the need for explicit forecasts or prior distributional assumptions. Furthermore, it supports real-time decision-making that dynamically weighs the competing priorities of market engagement and local curtailment mitigation. For brevity, we denote this aggregated joint-market bidding strategy for the co-located wind-BESS system via deep reinforcement learning as “AggJointDRL.” The primary contributions of this research are outlined as follows:

- Integrating curtailment reduction with market participation: We analyze the interplay between managing wind curtailment and conducting real-time market bids in a co-located wind-storage plant, emphasizing the value of adaptable, real-time coordination between generation and storage to enhance profitability.

- DRL-driven joint bidding framework: We investigate the optimal bidding policy for the wind-battery combination across spot and regulation FCAS markets, applying a model-free DRL technique—specifically the twin delayed deep deterministic policy gradient (TD3)—that structures participation through two linked Markov decision processes (MDPs) to maximize collective returns.
- Simulation results and practical insights: The AggJointDRL approach was tested on actual wind farm data sourced from the Australian National Electricity Market. Outcomes indicate superior performance in two respects: (1) delivering roughly 25% higher revenue relative to an optimization-based reference; and (2) capitalizing on flexibility across both spot and FCAS markets to achieve more lucrative joint bidding results compared with isolated market involvement. Coordinated operation further improves economics by enabling BESS charging from otherwise wasted wind energy.

Related Works

Numerous investigations have examined the use of battery energy storage systems (BESS) to mitigate wind energy waste. For instance, Ref. [6] suggested a scenario-driven stochastic optimization procedure to decrease wind curtailment and lower system operating expenses. In contrast, Ref. [7] proposed a two-stage optimization procedure, and Ref. [12] developed a linear programming framework to select suitable BESS power ratings and storage capacities to mitigate wind curtailment. In addition, Ref. [13] proposed a multi-objective scenario-based expansion planning model for BESS to reduce anticipated wind curtailment and societal costs, and Ref. [14] formulated a continuous-time, risk-aware scheduling model for sub-hourly BESS dispatch within day-ahead unit commitment. A key drawback of these techniques is that they position BESS mainly as a supporting asset in a coupled wind-BESS arrangement, redirecting surplus wind power to offset low-output periods and thereby reducing curtailment—at the expense of the storage unit's potential earnings in electricity markets. These approaches also presuppose detailed knowledge of wind power uncertainty distributions derived from historical records.

Research targeting revenue maximization for wind-battery hybrids has likewise been conducted. Notably, Ref. [15] presented a linear programming formulation to optimize BESS earnings via energy arbitrage, although it presumed complete foresight of spot market prices, limiting its applicability to real-time operations. Reference [8] introduced a market-centric optimal dispatch policy for a wind plant supported by a multi-stage hybrid BESS, prioritizing wind farm profit enhancement while incorporating an expensive hydrogen-based storage backup. Studies [9, 16] have advanced stochastic optimization models for wind-battery units operating in day-ahead and spinning reserve markets; however, simultaneous real-time bidding across markets is more demanding due to extreme price fluctuations. Similarly, Refs. [10, 14] constructed robust optimization schemes for concurrent participation in spot, regulation, and reserve markets, yet these remain sensitive to the precision of price forecasts amid high volatility and intricate price-forming mechanisms [11]. Additional efforts [17–19] explored online BESS scheduling algorithms grounded in Lyapunov optimization [17, 18] and competitive optimization [19], offering theoretical performance bounds. Such methods could, in principle, be extended to co-located wind-battery setups if comprehensive system model details are provided. By comparison, the current study pursues a purely data-driven approach that bypasses explicit system modeling and capitalizes on large volumes of historical data.

DRL-oriented bidding approaches, by contrast, have garnered comparatively modest interest. Works [20] and [21] applied variants of the deep Q-network (DQN) and state-action-reward-state-action (SARSA) algorithms, respectively, enabling a wind farm to engage in the spot market. In both cases, BESS functioned chiefly as a corrective tool to compensate for wind forecasting inaccuracies rather than as an active market participant. Nevertheless, the inherent constraints of DQN and SARSA necessitate discretizing the action space, resulting in overly simplistic bidding policies.

To overcome the shortcomings observed in prior literature and close the gap concerning robust coordination for co-located wind-BESS plants, we introduce the AggJointDRL framework. This solution empowers the combined system to navigate the tension between revenue optimization through joint participation in spot and regulation FCAS markets and the reduction of wind curtailment in a flexible, adaptive manner.

The distinctive elements of this contribution relative to existing studies can be summarized as: (1) a unified methodology that concurrently tackles wind curtailment abatement and market bidding for storage assets, moving beyond the single-focus emphasis common in earlier work; (2) a novel model-free DRL bidding strategy for multi-market participation that directly extracts uncertainties in wind generation and prices from historical data, without depending on forecasts or distributional priors typical of optimization techniques; (3) new evidence regarding the critical role of agile, real-time coordination between renewable generation and storage in securing economic gains, particularly when bidding across several markets simultaneously. This insight advances the case for tightly integrated renewable-storage operations. Through a concrete demonstration of a co-located renewable-storage configuration, the research offers practical implications that may inform policies promoting such pairings to deliver system-wide advantages. The following section details the underlying model.

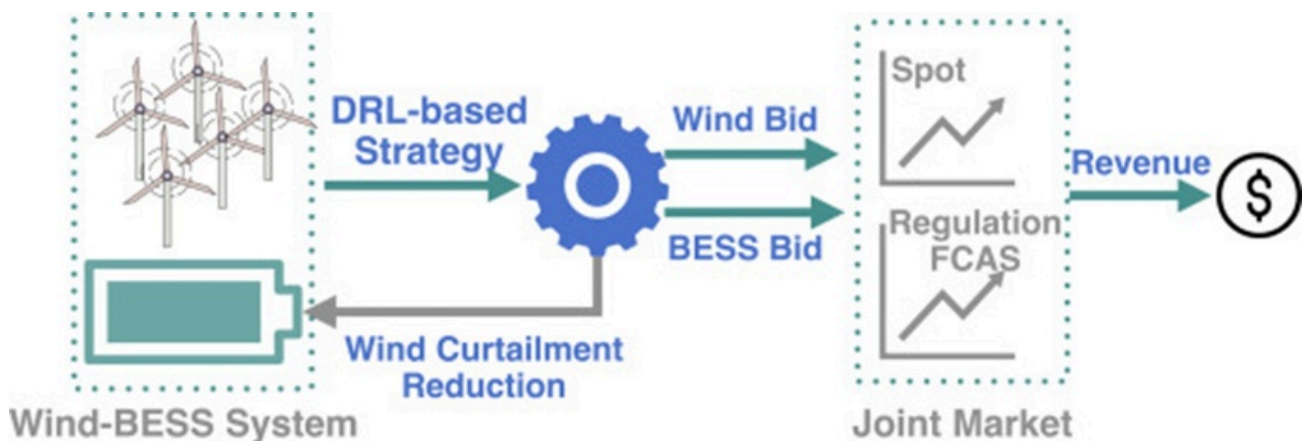


Figure 1. The framework of the AggJointDRL.

System Model

When constructing the AggJointDRL, we treat the wind-battery installation as a price taker whose bids exert no influence on other participants' decisions or on market-clearing outcomes. We further assume that one operator has unified ownership and control of both the co-located wind generators and the BESS. Consequently, the BESS can obtain any onsite wind energy that would otherwise be curtailed at zero cost. We also presume adequate substation and transmission capacity to support simultaneous power export from both assets. The BESS actively handles local wind curtailment while pursuing energy arbitrage opportunities in the Spot market and delivering frequency regulation services in the Regulation FCAS market. An overview of the system model is shown in [Figure 1](#).

The Australian national electricity market

The Australian National Electricity Market (NEM) spans approximately 40,000 kilometers of transmission infrastructure and serves approximately 9 million end users nationwide. Operating as a wholesale platform, it facilitates electricity transactions between generators and retailers, interconnects the eastern and southern regions, and meets about 80% of national electricity demand. The NEM functions as an energy-only market. The Australian Energy Market Operator (AEMO) gathers offers from all participants, assembles a bid stack, and schedules generation according to a least-cost principle [22]. Unlike electricity markets in the United States or Europe that include a day-ahead component, the NEM relies exclusively on its spot market for wholesale energy transactions.

The spot market

The Spot market in NEM serves as the real-time venue for trading bulk electricity between suppliers and consumers. Supply and demand are balanced instantaneously via a centrally managed dispatch process overseen by AEMO [23].

Generators submit price-quantity bids every five minutes. AEMO then ranks these bids from lowest to highest cost to construct the bid stack and dispatches resources efficiently. This mechanism resolves supply-demand mismatches at minimal cost. The resulting market-clearing price, commonly referred to as the Spot price, is set by the marginal generator required to satisfy final demand in the stack [24]. Generators whose bids are at or below the clearing price are dispatched for their offered volume and compensated at that price. Supply constraints frequently trigger sharp price spikes. Such volatility creates opportunities for participants to profit from buying low and selling high. It also encourages the use of financial hedging tools, such as contracts for difference, thereby increasing activity in the wholesale energy segment.

The FCAS market regulation

To sustain system frequency stability and counteract deviations driven by growing renewable penetration, the FCAS market procures essential system services—including inertia, system restart, and FCAS—from willing providers. Analogous to the spot market, AEMO ranks MW FCAS offers by cost in merit order, with the highest accepted offer establishing the clearing price for each FCAS category [25]. This market splits into Regulation and Contingency FCAS segments [25]. The present study concentrates on joint bidding within the Spot and Regulation FCAS markets.

Registered participants in the Regulation FCAS market can supply Regulation Raise (RR) and Regulation Lower (RL) services to keep system frequency inside the standard band of 49.85 Hz to 50.15 Hz. Rather than direct energy exchange as in the Spot market, regulation service delivery is governed by automatic generator control (AGC), which transmits successive normalized control signals to providers. These signals specify the required regulation direction: positive values call for frequency “Raise” and negative values indicate frequency “Lower.” For instance, when a wind farm or BESS offers RR-FCAS, it supplies power solely upon receipt of frequency-rise signals. We represent a sequence of AGC signals within one NEM dispatch interval as

$$s_t = [s_{t,1}, \dots, s_{t,l}, \dots, s_{t,L}] \in [-1, 1], \quad (1)$$

Where L denotes the number of signals in the sequence. The absolute magnitude $|s_{t,l}|$ of each signal corresponds to the power AEMO requests from the provider to correct instantaneous frequency imbalances. Each AGC signal lasts $\Delta s = 4$ seconds in the NEM.

Revenue of the wind-battery system

Wind farm

Wind farms in the NEM are generally classified as semi-dispatchable and must regularly submit updated availability forecasts (based on local wind measurements) to AEMO. Using these forecasts, the operator assigns a dispatch target (in MWh) for the wind farm to meet in the subsequent dispatch interval [26]. When the wind farm participates in both the Spot market and Regulation Raise FCAS (RR-FCAS), we employ a decision variable $v_{Wt} \in [0,1]$ to allocate the dispatch target between markets: $v_{Wt} p_{Wt}$ is offered to the Spot market and $(1 - v_{Wt}) p_{Wt}$ is allocated to the RR-FCAS sub-market.

Nevertheless, wind variability can cause discrepancies between the dispatch target and actual generation, p_{Wt} and Act_t , influencing exported power. When wind is abundant ($p_{Wt}^{W,Act} > p_{Wt}$), the dispatch target is assumed fully achievable. Conversely, during shortages arising from forecast inaccuracies or rule non-compliance [26], only partial fulfillment occurs ($p_{Wt}^{W,Act} < p_{Wt}$). We define actual dispatched wind power as $p_{Wt}^{W,Dis} = \min\{p_{Wt}^{W,Act}, p_{Wt}\}$. Penalties may apply if the dispatch target is not met. The combined revenue from joint-market participation for the wind farm is given by:

$$R^W = \Delta t \sum_{t=1}^T \left[v_t^W \rho_t^S + (1 - v_t^W) p_t^{RR} \right] \left(p_t^{W,Dis} - \lambda |p_t^{W,Dis} - p_t^W| \right), \quad (2)$$

Where Δt is the NEM dispatch interval (5 min), T is the total time horizon, ρ_t^S is the Spot price, p_t^{RR} is the RR-FCAS clearing price, and λ represents the penalty factor for deviations from the AEMO dispatch target [10].

BESS

Price fluctuations in the Spot market encourage the BESS to perform energy arbitrage. The BESS also earns compensation for supplying RR-FCAS and RL-FCAS in the Regulation FCAS market. Since simultaneous charging and discharging are not possible, we introduce binary variables v_{Ch_t} and v_{Dch_t} to enforce mutual exclusivity, expressed as

$$v_t^{Dch} + v_t^{Ch} \leq 1, \quad v_t^{Dch}, v_t^{Ch} \in \{0, 1\}, \quad (3)$$

The BESS remains idle when both variables equal zero.

The BESS revenue streams from the spot and regulation FCAS markets take the form.

$$R^{BESS} = \Delta t \sum_{t=1}^T \left[\rho_t^S \left(v_t^{Dch} \eta^{Dch} p_t^{BESS,S} - v_t^{Ch} \frac{1}{\eta^{Ch}} p_t^{BESS,S} \right) + \rho_t^{RR} v_t^{Dch} \eta^{Dch} p_t^{BESS,Reg} + \rho_t^{RL} v_t^{Ch} \frac{1}{\eta^{Ch}} p_t^{BESS,Reg} \right], \quad (4)$$

Where η^{Ch} / η^{Dch} denote charge and discharge efficiencies, p_t^{RR} and p_t^{RL} are the respective FCAS clearing prices, and $p_t^{BESS,S}$, $p_t^{BESS,Reg}$ represent offered power in the spot and regulation markets.

Beyond buying power in the spot market or absorbing grid energy via Lower Regulation FCAS, the BESS can charge using otherwise curtailed onsite wind generation, denoted $p_t^{BESS,WC}$. Subject to the charge/discharge constraint in Eq. 3, the BESS cannot draw from curtailed wind while simultaneously bidding to discharge in spot or regulation markets. This relationship is captured logically as

$$v_t^{Dch} p_t^{BESS,WC} = 0. \quad (5)$$

The volume of onsite wind curtailment is expressed as

$$p_t^{W,WC} = (p_t^{W,Act} - p_t^W) I(p_t^{W,Act} > p_t^W), \quad (6)$$

Where $I(p_t^{W,Act} > p_t^W)$ serves as the curtailment indicator. Consequently, the actual power drawn by the BESS from the wind farm becomes

$$\hat{p}_t^{BESS,WC} = \min \left\{ p_t^{BESS,WC}, p_t^{W,WC} \right\}. \quad (7)$$

Furthermore, because the BESS can utilize curtailed wind, the wind farm's forecasted availability p_t^W is adjusted according to

$$p_t^W \leftarrow p_t^W + \hat{p}_t^{\text{BESS,WC}}. \quad (8)$$

Repeated charge/discharge cycling accelerates BESS aging. We quantify degradation cost as

$$C^{\text{BESS}} = c \Delta t \sum_{t=1}^T v_t^{\text{Dch}} \left(p_t^{\text{BESS,S}} + p_t^{\text{BESS,Reg}} \right), \quad (9)$$

Assuming degradation arises solely from discharge operations [27, 28], with c as the technology-specific cost coefficient in AU\$/MWh [28].

Joint-market bidding of the wind-battery system

Accounting for the separate revenue flows of the wind farm and BESS from spot and regulation FCAS markets, together with BESS degradation expenses, we cast the optimal joint-market bidding of the wind-battery system as an optimization task whose objective is

$$R^{\text{Joint}} = R^W + R^{\text{BESS}} - C^{\text{BESS}}. \quad (10)$$

Real-time scheduling of the wind farm and BESS must respect installed capacities, formulated as

$$0 \leq p_t^W \leq P_{\max}^W, \quad (11)$$

$$0 \leq p_t^{\text{BESS,S}} \leq P_{\max}^{\text{BESS}}, \quad (12)$$

$$0 \leq p_t^{\text{BESS,Reg}} \leq P_{\max}^{\text{BESS}}, \quad (13)$$

$$0 \leq p_t^{\text{BESS,WC}} \leq P_{\max}^{\text{BESS}}, \quad (14)$$

$$0 \leq p_t^{\text{BESS,S}} + p_t^{\text{BESS,Reg}} + p_t^{\text{BESS,WC}} \leq P_{\max}^{\text{BESS}}, \quad (15)$$

Where $P_{\max}^W$ and $P_{\max}^{\text{BESS}}$ are the wind farm's installed capacity and the BESS's rated power, respectively. Eq. 11 limits the wind farm's forecasted availability. Eqs. 12 and 13 bound the BESS's offered power in spot and regulation FCAS markets. Eq. 14 ensures that the power drawn from the wind farm remains within the BESS's rated power. Eq. 15 guarantees that combined bids plus power from curtailed wind do not surpass BESS's rated power.

BESS energy levels are further constrained by its prior state e_{t-1} plus the current change Δe_t , remaining between lower and upper bounds $E_{\min}$ and $E_{\max}$:

$$E_{\min} \leq e_{t-1} + \Delta e_t \leq E_{\max}. \quad (16)$$

BESS energy varies due to three factors: (1) spot market transactions, (2) regulation FCAS service delivery, and (3) absorption of curtailed onsite wind energy. The energy shift from spot market activity is:

$$\Delta e_t^S = \Delta t (v_t^{\text{Ch}} - v_t^{\text{Dch}}) p_t^{\text{BESS},S}. \quad (17)$$

Because regulation FCAS enablement follows AGC signals, the associated energy change aggregates contributions from each signal:

$$\Delta e_t^{\text{Reg}} = \Delta s \left[v_t^{\text{Ch}} \sum_{l=1}^L |s_{t,l}| \mathbb{I}(s_{t,l} < 0) - v_t^{\text{Dch}} \sum_{l=1}^L s_{t,l} \mathbb{I}(s_{t,l} \geq 0) \right] p_t^{\text{BESS},\text{Reg}}, \quad (18)$$

Where $\mathbb{I}(s_{t,l} < 0)$, $\mathbb{I}(s_{t,l} \geq 0)$ flag frequency-lower and frequency-raise directions, respectively.

The energy contribution from onsite wind curtailment is:

$$\Delta e_t^{\text{WC}} = \hat{p}_t^{\text{BESS},\text{WC}} \Delta t, \quad (19)$$

Therefore, the net energy variation for the BESS during the current dispatch interval is:

$$\Delta e_t = \Delta e_t^S + \Delta e_t^{\text{Reg}} + \Delta e_t^{\text{WC}}. \quad (20)$$

The complete mathematical formulation of the joint-market bidding optimization is presented as follows:

$$\begin{aligned} \max \quad & R^{\text{Joint}} \\ \text{s.t.} \quad & \text{Eq. (3), (5), (11)–(16)}. \end{aligned} \quad (21)$$

Materials and Methods

To maximize total revenue for the wind-battery installation, as expressed in [Eq. 21](#), we split the continuous bidding optimization task into two interconnected Markov decision processes (MDPs)—one for the wind farm and one for the BESS. Subsequently, we apply the TD3 algorithm [\[29\]](#) to maximize the expected cumulative returns for these MDPs. This separation supports comprehensive optimization of the entire revenue-driven bidding challenge.

MDP modeling

As outlined earlier, numerous factors, including variable wind output and fluctuating electricity prices, shape bidding decisions for the integrated wind-battery configuration in the spot and regulation FCAS markets. To more accurately model the joint-market bidding procedure, we divide it into two distinct MDPs (one dedicated to the wind farm and one to

the BESS). Each MDP comprises four fundamental constituents: the state space $\mathcal{S}^W / \mathcal{S}^{\text{BESS}}$, the action space $\mathcal{A}^W / \mathcal{A}^{\text{BESS}}$, the probability space $\mathcal{P}^W / \mathcal{P}^{\text{BESS}}$, and the reward space $\mathcal{R}^W / \mathcal{R}^{\text{BESS}}$.

State Space S: Every piece of internal data (for example, wind generation levels and BESS energy storage) together with external information (for example, prevailing energy prices) concerning the wind-battery system is encapsulated in a state vector $\mathbf{s}_t$. To assist the BESS in effectively addressing instances of wind curtailment, we incorporate the frequency of curtailment events recorded across the most recent M dispatch intervals, indicated as f^{WC}_t , directly into the BESS state representation. The state definitions for the wind farm and the BESS are given as:

$$\mathbf{s}_t^W = [p_{t-1}^{W,\text{Act}}, \rho_{t-1}^S, \rho_{t-1}^{\text{RR}}], \quad (22)$$

$$\mathbf{s}_t^{\text{BESS}} = [e_{t-1}, f_{t-1}^{\text{WC}}, p_{t-1}^{W,\text{Act}}, \rho_{t-1}^S, \rho_{t-1}^{\text{RR}}, \rho_{t-1}^{\text{RL}}]. \quad (23)$$

Action space A: Actions taken by the wind farm comprise its forecasted power availability a_t^W (normalized relative to P_{max}^W) and the splitting variable v_t^W that apportions the offered power across the separate markets. For the BESS, actions cover the quantities bid into the joint market ($a_t^{\text{BESS},S}$, $a_t^{\text{BESS},\text{Reg}}$) as well as the amount withdrawn from the local wind resource ($a_t^{\text{BESS},\text{WC}}$); all of these quantities are normalized by the rated power P . The BESS actions additionally encompass the binary indicators for charging and discharging, v_t^{Ch} / v_t^{Dch} . We specify the action sets for the wind farm and the BESS as:

$$\mathbf{a}_t^W = [a_t^W, v_t^W], \quad (24)$$

$$\mathbf{a}_t^{\text{BESS}} = [v_t^{\text{Dch}}, v_t^{\text{Ch}}, a_t^{\text{BESS},S}, a_t^{\text{BESS},\text{Reg}}, a_t^{\text{BESS},\text{WC}}]. \quad (25)$$

Probability space P: This element denotes the collection of probabilities governing transitions to the following state once a deterministic action has been executed, expressed as $\mathcal{P}(\mathbf{s}_{t+1} | \mathbf{s}_t, \mathbf{a}_t)$.

Reward Space R: Once action $\mathbf{a}_t$ is executed from state $\mathbf{s}_t$, both the wind farm and the BESS obtain immediate rewards that measure how well the chosen bid performs. To achieve maximum overall revenue, it is vital to craft appropriate reward signals for each component to encourage better bidding decisions.

To properly accommodate uncertainty in wind generation, maintain precise dispatch targets, and harness the advantages offered by joint-market flexibility, the reward function assigned to the wind farm is structured as:

$$r_t^W = [v_t^W \rho_t^S + (1 - v_t^W) \rho_t^{\text{RR}}] \left(\min \left\{ a_t^W, a_t^{W,\text{Act}} \right\} - \lambda \left| a_t^W - a_t^{W,\text{Act}} \right| \right), \quad (26)$$

Where the normalized value of actual wind generation $a_t^{W,\text{Act}}$, is computed as:

$$a_t^{W,\text{Act}} = \frac{p_t^{W,\text{Act}}}{P_{\text{max}}^W}. \quad (27)$$

To facilitate advantageous energy arbitrage operations by the BESS within the spot market, we define two charge/discharge indicator terms I_t^{Ch} , I_t^{Dch} , as follows:

$$I_t^{\text{Ch}} = \text{sgn}(\bar{\rho}_t^S - \rho_t^S), \quad (28)$$

$$I_t^{\text{Dch}} = \text{sgn}(\rho_t^S - \bar{\rho}_t^S), \quad (29)$$

Where $\text{sgn}(\cdot)$ stands for the sign function and $\bar{\rho}_t^S$ corresponds to the exponential moving average of the spot price, calculated via:

$$\bar{\rho}_t^S = \tau^S \bar{\rho}_{t-1}^S + (1 - \tau^S) \rho_t^S, \quad (30)$$

Where τ^S is a smoothing parameter. The charge/discharge indicators incentivize the BESS to buy energy at low prices ($\rho_t^S < \bar{\rho}_t^S$) and sell at high prices ($\rho_t^S > \bar{\rho}_t^S$). Any bids that violate this guideline will be penalized. Hence, the arbitrage rewards can be formulated as:

$$r_t^{\text{BESS,S}} = a_t^{\text{BESS,S}} \left| \rho_t^S - \bar{\rho}_t^S \right| \left(I_t^{\text{Ch}} v_t^{\text{Ch}} \frac{1}{\eta^{\text{Ch}}} + I_t^{\text{Dch}} v_t^{\text{Dch}} \eta^{\text{Dch}} \right). \quad (31)$$

Furthermore, the rewards associated with the Regulation FCAS market are defined as:

$$r_t^{\text{BESS,Reg}} = a_t^{\text{BESS,Reg}} \left(v_t^{\text{Ch}} \frac{1}{\eta^{\text{Ch}}} \rho_t^{\text{RL}} + v_t^{\text{Dch}} \eta^{\text{Dch}} \rho_t^{\text{RR}} \right). \quad (32)$$

In addition, the BESS gains positive rewards when it mitigates onsite wind curtailment, which can be expressed as:

$$r_t^{\text{BESS,WC}} = \lambda \left[v_t^{\text{W}} \rho_t^S + (1 - v_t^{\text{W}}) \rho_t^{\text{RR}} \right] \min \left\{ a_t^{\text{BESS,WC}}, a_t^{\text{W,WC}} \right\} \frac{1}{\eta^{\text{Ch}}}, \quad (33)$$

Where the onsite curtailed wind power is expressed in normalized form as:

$$a_t^{\text{W,WC}} = \frac{p_t^{\text{W,WC}}}{P_{\text{max}}^{\text{W}}}. \quad (34)$$

The BESS reward function integrates both bidding-related incentives and rewards for reducing wind curtailment, and is formulated as:

$$r_t^{\text{BESS}} = r_t^{\text{BESS,S}} + r_t^{\text{BESS,Reg}} + r_t^{\text{BESS,WC}}. \quad (35)$$

Learning optimal bidding strategy via TD3

We employ an advanced off-policy deep reinforcement learning method, TD3 [29], to solve the constructed MDPs (using the same TD3 architecture for both). For the sake of brevity, the description of TD3 is provided without attaching specific MDP-related notation. TD3 seeks to discover an optimal policy $\pi(a_t | s_t)$ capable of maximizing expected cumulative returns across a finite time window, which can be stated as:

$$J_{\pi} = \mathbb{E}_{s_t \sim \mathbb{P}, a_t \sim \pi(s_t)} [R_1], \quad (36)$$

Where the discounted cumulative return is expressed as:

$$R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'}, \quad (37)$$

With γ representing the discount factor.

TD3 relies on two primary components arranged in an actor-critic architecture: the actor function $\pi(a_t | s_t)$, which selects actions that produce state transitions, and the critic function (commonly called the Q-function) $Q(a_t, s_t)$, which assesses the quality of each state-action pair. The Q-function takes the form:

$$Q(a_t, s_t) = \mathbb{E}_{s_t \sim \mathbb{P}, a_t \sim \pi(s_t)} [R_t | s_t, a_t]. \quad (38)$$

Leveraging the Q-function, the objective function of the actor from [Eq. 36](#) can be restated as:

$$J_{\pi} = \mathbb{E}_{s \sim \mathbb{P}} [Q(s_1, \pi(s_1))]. \quad (39)$$

We utilize neural networks to approximate both the actor function π_{ϕ} and the Q-function Q_{θ} , with ϕ and θ denoting the respective network parameters. Updates to these networks are performed via gradient descent using the Adam optimizer [\[30\]](#).

Update π_{ϕ} : The actor network undergoes improvement by maximizing the objective specified in [Eq. 39](#). The associated gradient expression is:

$$\nabla_{\phi} J_{\pi}(\phi) = \nabla_{\phi} \mathbb{E}_{s \sim \mathbb{B}} [Q_{\theta}(s_1, \pi_{\phi}(s_1))], \quad (40)$$

Where a replay buffer $\mathbb{B}$ stores sequences of MDP transitions, each represented as the tuple $\{s_t, a_t, r_t, s_{t+1}\}$.

Specifically, to guarantee that the combined volume of all bids plus any power extracted from onsite curtailed wind remains below the BESS rated capacity as stipulated in [Eq. 15](#), we introduce a supplementary loss term into the TD3 implementation for the BESS, defined as:

$$L_{\pi}(\phi) = a_t^{\text{BESS, sum}} \mathbb{I}(a_t^{\text{BESS, sum}} > 1), \quad (41)$$

With the aggregate of BESS actions expressed as:

$$a_t^{\text{BESS, sum}} = a_t^{\text{BESS, S}} + a_t^{\text{BESS, Reg}} + a_t^{\text{BESS, WC}}. \quad (42)$$

Consequently, the gradient applied to the BESS actor network is modified to:

$$\nabla_{\phi} J_{\pi}(\phi) \leftarrow \nabla_{\phi} J_{\pi}(\phi) - \beta_L \nabla_{\phi} \mathbb{E}_{a \sim \mathbb{B}} [L_{\pi}(\phi)], \quad (43)$$

Where β_L serves as the weighting factor for the added loss term. The standard gradient descent step for updating the actor network is then given by

$$\phi \leftarrow \phi - \eta^{\phi} \nabla_{\phi} [-J_{\pi}(\phi)], \quad (44)$$

Where η^{ϕ} indicates the learning rate applied to the actor network. Furthermore, TD3 employs delayed actor updates to reduce variance in Q-function estimates [29], which are achieved through a delay coefficient d that governs the staggered update schedule between the actor and critic networks.

Update Q_{θ} : The Q-network is refined by reducing the temporal difference residual associated with the Bellman equation. The Bellman equation [31] reads as:

$$Q(s_t, a_t) = r_t + \gamma \mathbb{E}_{s_{t+1} \sim \mathbb{P}, a_{t+1} \sim \pi(s_{t+1})} [Q(s_{t+1}, a_{t+1})], \quad (45)$$

In which the left-hand term is computed by the current Q-network Q_{θ} while the right-hand term is supplied by a target Q-network $Q_{\hat{\theta}}$ (where $\hat{\theta}$ denotes the frozen parameters of the target network). As a result, the loss objective for the Q-network becomes:

$$J_Q(\theta) = \mathbb{E}_{s_t \sim \mathbb{B}} \left[\frac{1}{2} \left[Q_{\theta}(s_t, \pi_{\phi}(s_t)) - Q_{\hat{\theta}}(s_t, \pi_{\hat{\phi}}(s_t)) \right]^2 \right], \quad (46)$$

A corresponding target actor network $\pi_{\hat{\phi}}$ is introduced to promote training stability and to aid the target critic in computing value estimates.

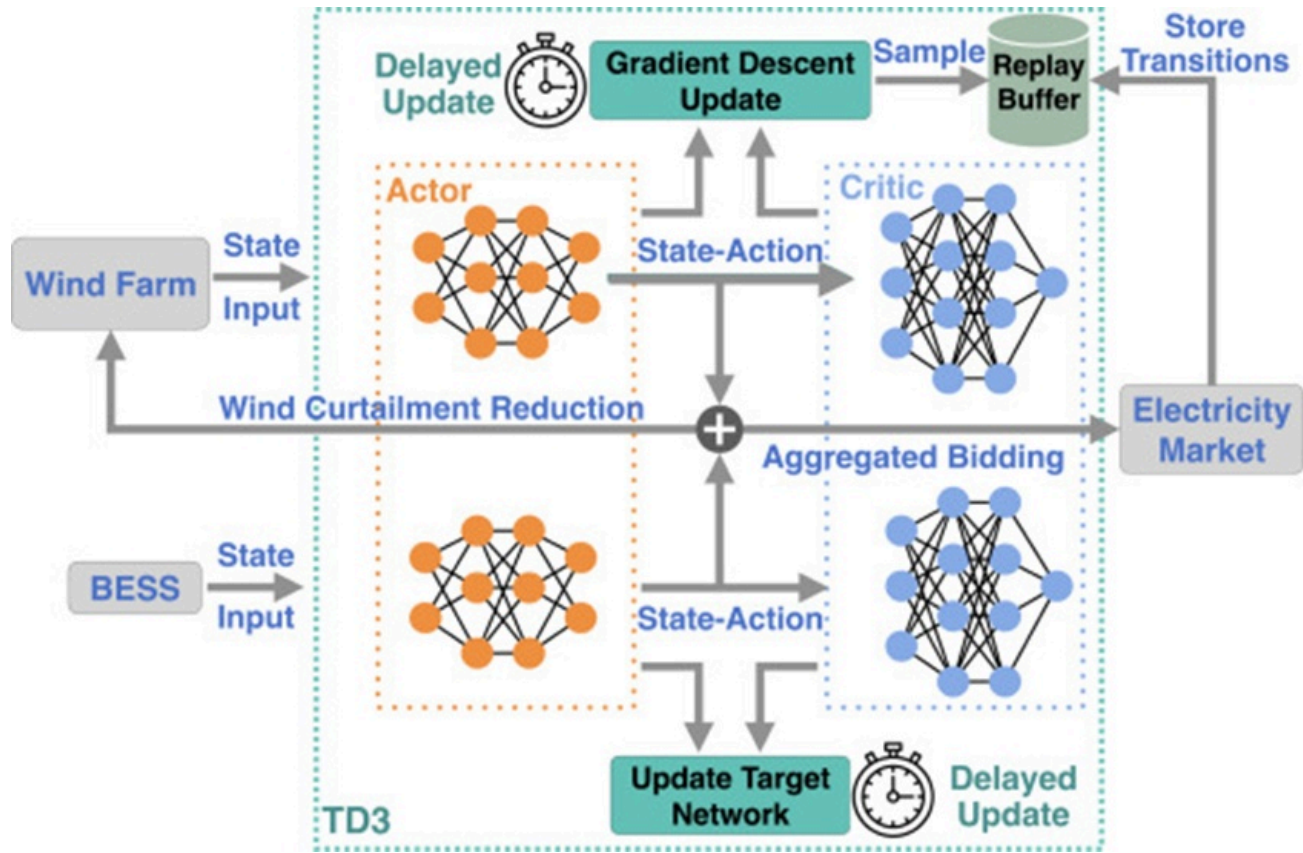


Figure 2. The architecture of the TD3-powered AggJointDRL.

Similar to the actor network, the gradient descent update for the Q network can be expressed as

$$\theta \leftarrow \theta - \eta^\theta \nabla_\theta J_Q(\theta), \quad (47)$$

Where η^θ also denotes the learning rate of the Q network.

To tackle the overestimation problem [32] commonly encountered in the Bellman equation, TD3 incorporates a clipped double Q-learning approach. In this method, an extra Q network, referred to as $Q\theta2$ (with the original Q network now denoted as $Q\theta1$), is employed to reduce estimation bias. Consequently, the Q-value estimation using the target networks (i.e., $Q\theta$ and $\pi\hat{\phi}$) is reformulated as

$$Q_{\hat{\theta}}(s_t, \pi_{\hat{\phi}}(s_t)) = r_t + \gamma \min_{i=\{1,2\}} \mathbb{E}_{s_{t+1} \sim \mathbf{P}} [Q_{\hat{\theta}_i}(s_{t+1}, \pi_{\hat{\phi}}(s_{t+1}))], \quad (48)$$

Where $\theta1$ and $\theta2$ represent the parameters of the target Q networks.

Update of target networks: In the TD3 algorithm, all target networks—including the target actor network $\pi\hat{\phi}$ and the two target Q networks $Q\theta1$ and $Q\theta2$ —are synchronously updated with the actor network using a moving-average (soft update) mechanism, which can be written as

$$\hat{\phi} \leftarrow \tau^\phi \phi + (1 - \tau^\phi) \hat{\phi}, \quad (49)$$

$$\hat{\theta}_i \leftarrow \tau^\theta \theta_i + (1 - \tau^\theta) \hat{\theta}_i, i = \{1, 2\}, \quad (50)$$

Where $\tau\phi$ and $\tau\theta$ are the smoothing coefficients.

To summarize, the overall structure of our TD3-empowered AggJointDRL is illustrated in [Figure 2](#), and the detailed algorithmic procedure is provided in [Algorithm 1](#).

Initialize actor and Q networks of both the wind farm and BESS $\phi^W, \theta_1^W, \theta_2^W, \phi^{\text{BESS}}, \theta_1^{\text{BESS}}, \theta_2^{\text{BESS}}$.
 Initialize target networks with their corresponding network parameters.
 Initialize the replay buffers $\mathbb{B}^W, \mathbb{B}^{\text{BESS}}$.
for $t = 1, \dots, T$ **do**
 Prepare the current states s_t^W, s_t^{BESS} .
 Get actions a_t^W, a_t^{BESS} through action networks and receive associated rewards r_t^W, r_t^{BESS} .
 if the BESS bids violate its energy limits **then**
 $a_t^{\text{BESS}} \leftarrow 0$.
 end if
 Transit into the next states $s_{t+1}^W, s_{t+1}^{\text{BESS}}$ via $\mathbb{P}^W$ and $\mathbb{P}^{\text{BESS}}$.
 Store transitions $\{s_t^W, a_t^W, r_t^W, s_{t+1}^W\}, \{s_t^{\text{BESS}}, a_t^{\text{BESS}}, r_t^{\text{BESS}}, s_{t+1}^{\text{BESS}}\}$ into replay buffers $\mathbb{B}^W$ and $\mathbb{B}^{\text{BESS}}$.
 Update Q networks via gradient descent.
 if $t \bmod d = 0$ **then**
 Update actor and target networks.
 end if
end for

Algorithm 1. The TD3-empowered AggJointDRL.

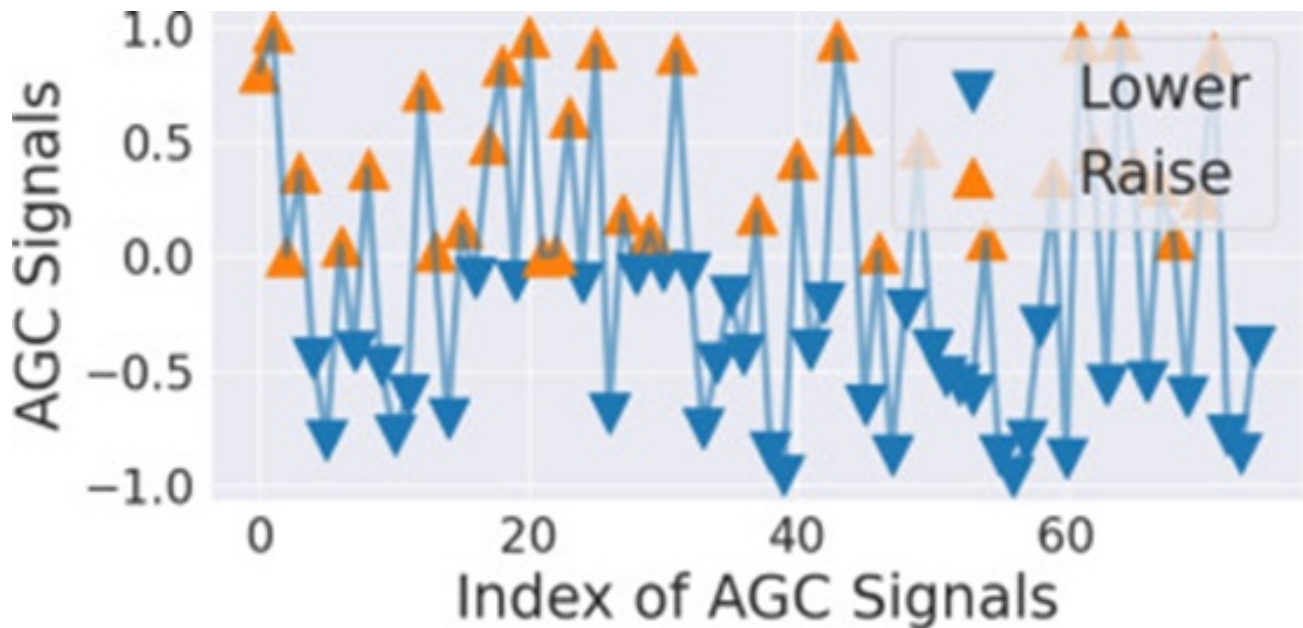


Figure 3. Illustration of AGC trajectories over a single dispatch period, where negative values correspond to down-regulation (frequency decrease) and positive values indicate up-regulation (frequency increase).

Experiments and Results

Experimental settings

Wind production data analyzed in this study originate from the Oaklands Hill Wind Farm situated in Victoria, a region within the National Electricity Market of Australia. For model development and validation, spot price records from 2018 [33] were employed, with data spanning January to November used for training and December data reserved for performance assessment.

The battery energy storage system (BESS) was modeled with a total energy capacity of 10 MWh and state-of-charge constraints of 5%-95%. Training of the proposed algorithm was executed on an NVIDIA TITAN RTX, with detailed initialization settings provided in Table 1.

To evaluate bidding behavior, three distinct participation configurations were considered: involvement limited to the spot market, engagement restricted to the Regulation FCAS market, and a combined participation strategy across both markets. Frequency control actions were emulated by generating AGC signals for both wind generation and the BESS using a uniform random distribution, $U \sim (-1, 1)$, to capture both upward and downward regulation requirements. A sample representation of these AGC signals within one dispatch interval is presented in Figure 3.

While the experimental analysis is grounded in the Australian market context, the proposed coordinated bidding approach for integrated wind–battery systems is not geographically constrained and can be adapted for application in other power markets, including those across the United States and Europe.

Table 1. The initialized parameters.

Δs	4 s	L	75
Δt	5 min	λ	1.5
η_{Dch}, η_{Ch}	0.95	c	1 AU\$/MWh
P_{maxW}	67 MW	$P_{maxBESS}$	10 MW
E_{min}	0.5 MWh	E_{max}	9.5 MWh
M	10	τ^S	0.9
γ	0.99	β^L	10
$\eta\phi, \eta\theta$	0.0003	τ^ϕ, τ^θ	0.01

Table 2. The evaluation revenue (in AU\$) and time costs (in minutes) of the P&O benchmark and our AggJointDRL in all three bidding scenarios.

		Strategy	Spot	Reg FCAS	Joint
Revenue	Wind	P&O	829,604 979,784	451,156	889,672
		Ours	18%	540,690	1,098,686
		Boost		20%	23%
	BESS	P&O	62,444	178,643	412,363
		Ours	113,292	247,759	524,917
		Boost	81%	39%	27%

Total		P&O	892,048	629,799	1,302,035
		Ours	1,093,076	788,449	1,623,603
		Boost	23%	25%	25%
Time Cost	Train	P&O	1.8	3.4	3.7
		Ours	92.4	89.7	91.9
	Evaluate	P&O	36.3	71.2	94.6
		Ours	0.1	0.1	0.1
	Total	P&O	38.1	74.6	98.3
		Ours	92.5	89.8	92.0

Effectiveness of the AggJointDRL

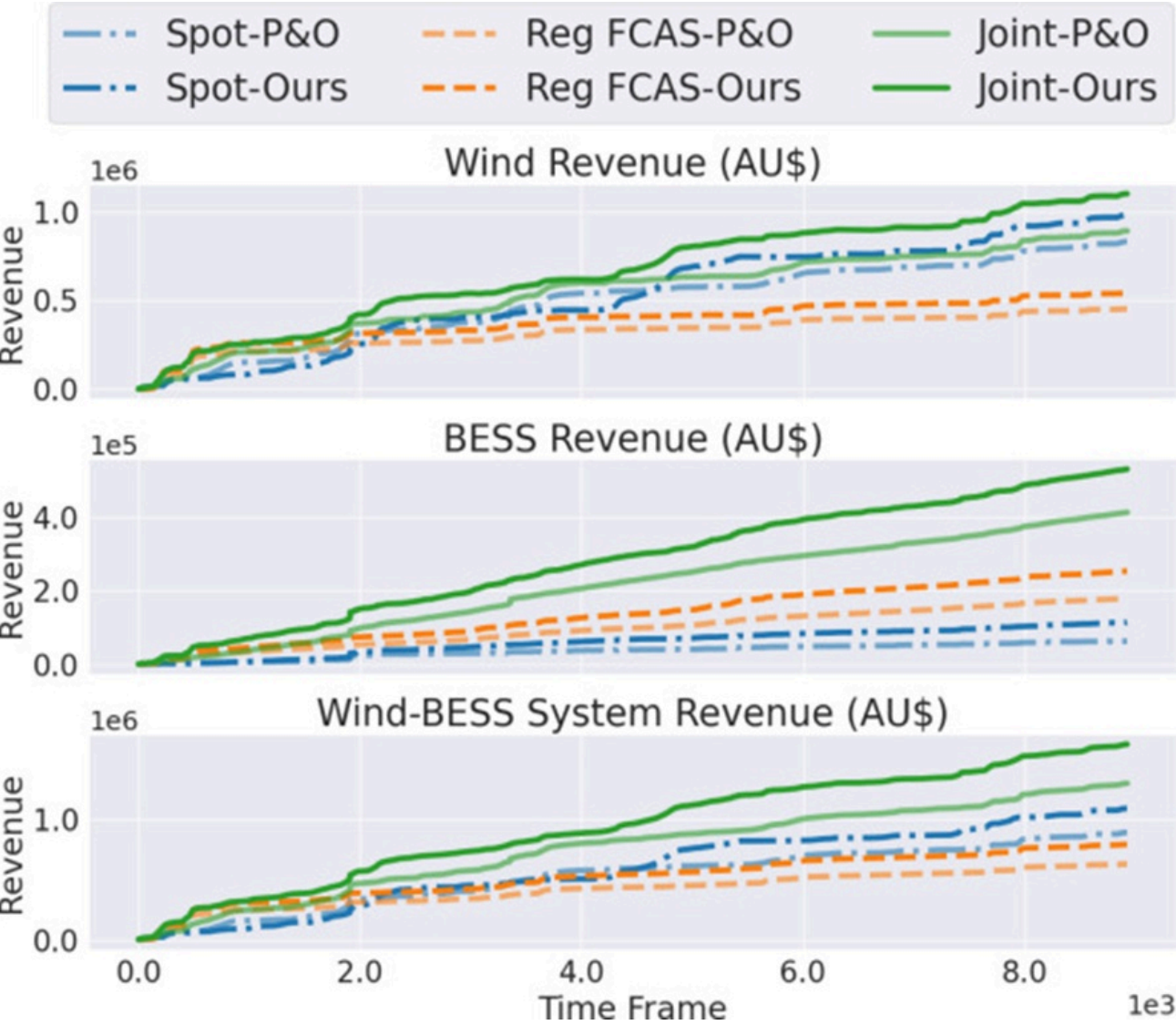


Figure 4. Comparative evaluation of total revenue achieved by the P&O baseline and the proposed AggJointDRL approach across the three bidding configurations.

Benchmark comparison

To assess the performance of the proposed AggJointDRL framework, it was benchmarked against a conventional predict-and-optimize (P&O) strategy. In this baseline approach, long short-term memory (LSTM) models are employed to estimate wind power availability and electricity prices, after which a revenue-maximizing bidding problem is solved using a mixed-integer linear programming formulation implemented via the PuLP library [34]. The comparative outcomes—covering revenues obtained under all three bidding settings—are presented in **Figure 4** and further detailed in **Table 2** for direct comparison.

The findings reveal a consistent advantage of the AggJointDRL method over the P&O approach, yielding revenue improvements of 23%, 25%, and 25% across the three respective scenarios. As highlighted in **Table 2**, the enhancement is particularly notable for BESS operations, with gains of 81% in the spot market, 39% in the Regulation FCAS market, and 27% under joint participation. Moreover, wind farm revenues increased markedly when the proposed strategy was used compared to the baseline.

Figure 5 compares predicted wind availability and actual spot-market generation by showing their absolute deviations. The reduced mean absolute error (MAE) achieved by AggJointDRL indicates a stronger capability to model the inherent variability of wind generation. This improved accuracy enables more precise adjustment of dispatch targets, thereby minimizing the risk of penalties from mismatches between scheduled and actual output.

Despite requiring more extensive training time, as indicated in **Table 2**, the AggJointDRL model demonstrates significantly faster execution during operation, requiring only 10 seconds to determine bidding strategies for a one-month horizon. In contrast, the computational burden of the P&O approach grows substantially with the number of markets involved. Consequently, once trained, AggJointDRL is better suited for real-time market participation, where rapid and reliable decision-making is critical.

Furthermore, the AggJointDRL framework was implemented using several deep reinforcement learning (DRL) algorithms, including proximal policy optimization (PPO) [35], deep deterministic policy gradient (DDPG) [36], and soft actor-critic (SAC) [37]. The comparative results, illustrated in **Figure 6**, show that the off-policy methods (DDPG, SAC, and TD3) deliver closely aligned performance and clearly outperform the on-policy PPO algorithm. This disparity can be attributed to the absence of a replay buffer in on-policy methods, which limits their ability to reuse historical experiences and fully explore the state spaces (i.e., SW and SBESS) associated with the wind farm and the BESS. Overall, these results indicate that the proposed joint-market bidding strategy is well-suited for integration with a range of off-policy DRL techniques.

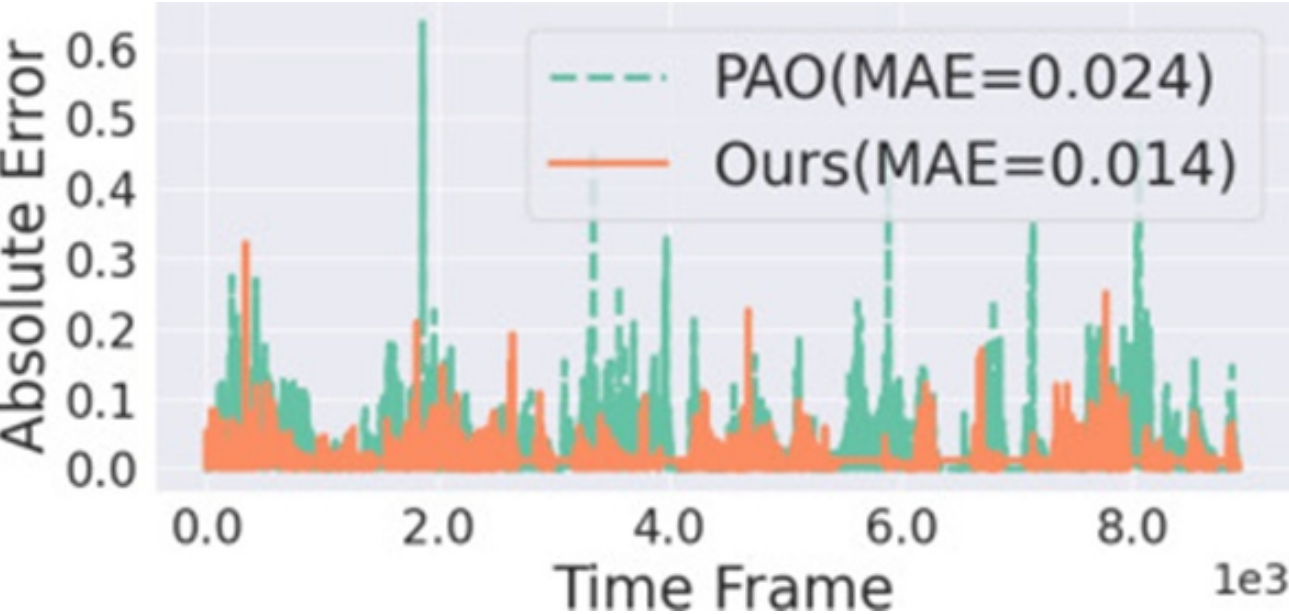


Figure 5. The absolute errors between the dispatch targets and the actual wind generation.

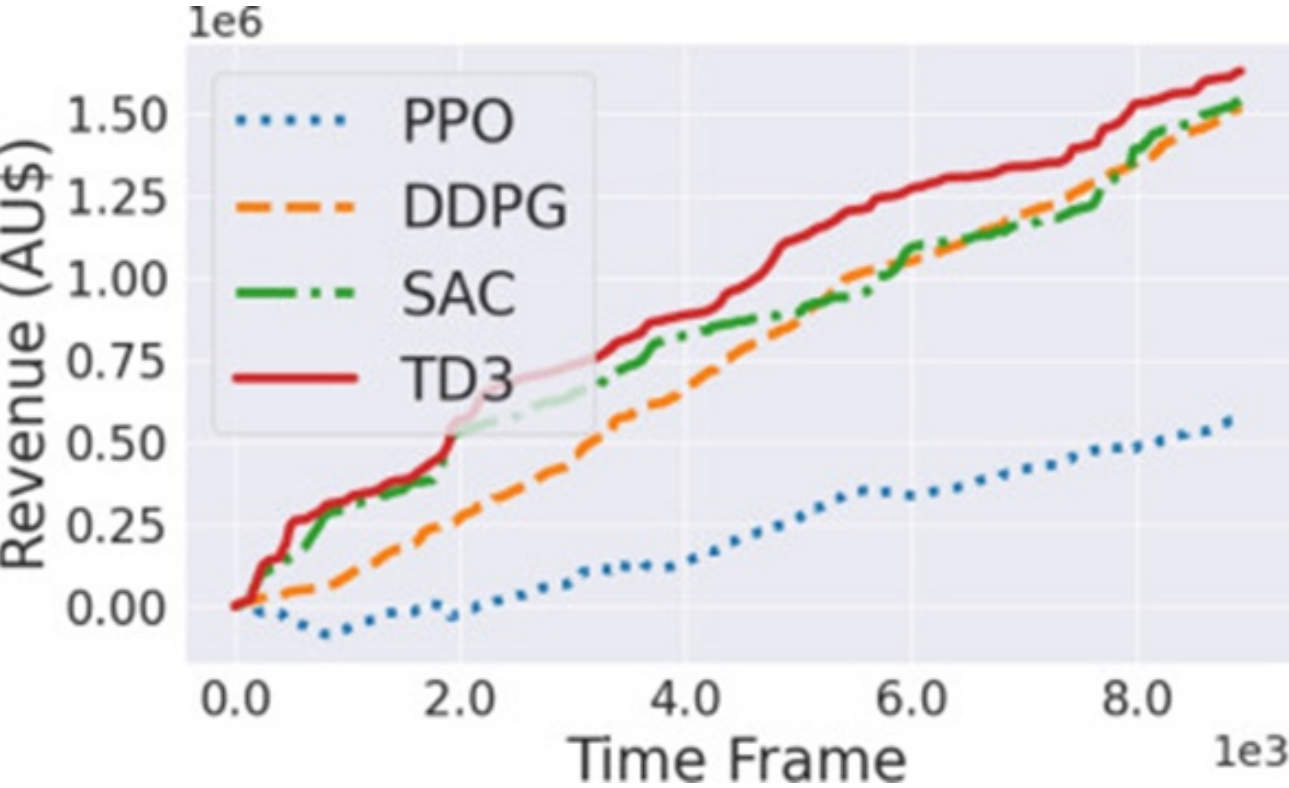


Figure 6. The revenue generated by the wind-BESS system through joint-market bidding using various algorithms.

Profitability analysis of the regulation FCAS market

Results presented in [Table 2](#) demonstrate that the battery energy storage system earns considerably more income by offering services in the Regulation FCAS market than by trading exclusively in the spot market. The key reason is that registration in the Regulation FCAS market allows the BESS to be compensated for reserving capacity to deliver both raise regulation (RR-FCAS) and lower regulation (RL-FCAS) services, regardless of whether the actual regulation energy

is dispatched. In the spot market, however, the BESS frequently needs to purchase electricity to recharge, which often generates net losses when spot prices remain at or above zero. This fundamental difference in payment mechanisms explains the large revenue disparity observed between the two market participation strategies.

Table 3 summarizes key statistics on past clearing prices in the Regulation FCAS market from 2016 to 2021. The widespread and rapid expansion of renewable generation has added significant uncertainty and power system disturbances, prompting the market operator to procure steadily larger quantities of ancillary services—particularly FCAS—from available resources. Consequently, the Regulation FCAS market has experienced notable growth, characterized by higher average clearing prices and markedly increased price fluctuations, as indicated by the rising annual standard deviations shown in **Table 3**. This heightened volatility reveals strong economic incentives in the Regulation FCAS segment and motivates participants, especially fast-acting lithium-ion battery systems that can respond within one second, to focus on delivering regulatory services rather than pursuing energy arbitrage opportunities in the spot market.

Table 3. Statistics on historical Regulation FCAS market-clearing prices.

Empty Cell		2016	2017	2018	2019	2020	2021
RR-FCAS	mean (AU\$/MWs)	13	30	24	39	22	19
	std (AU\$/MWs)	14	31	68	62	192	110
RL-FCAS	mean (AU\$/MWs)	6	24	11	19	10	13
	std (AU\$/MWs)	6	16	9	11	7	54

Effectiveness of joint bidding across both markets

As noted earlier, the Regulation FCAS market offers substantial profit potential. By enabling the wind-battery system to operate simultaneously in the spot market and the regulation FCAS market, the strategy can harness the complementary strengths of both marketplaces. This combined approach yields markedly superior total revenue compared with bidding into either market in isolation, as clearly shown in **Table 2**. Notably, the income earned by the BESS under joint-market participation exceeds the sum of the revenues obtained from the two markets when pursued separately.

Revenue from Regulation FCAS bidding tends to be higher whenever FCAS clearing prices exceed spot prices, while the reverse holds when spot prices are more attractive. **Figure 7** illustrates how the wind farm can capitalize on price volatility between the two markets under the joint bidding framework. Specifically, the wind farm increases its offered power volume in the raise regulation (RR-FCAS) sub-market—corresponding to lower values of the decision variable vW_t —precisely when the RR-FCAS clearing price rises above the prevailing spot price. This behavior highlights the advantage of coordinated participation: the system can dynamically allocate resources to whichever market provides the greater return at any given time, thereby maximizing overall profits by effectively exploiting price differences across both markets.

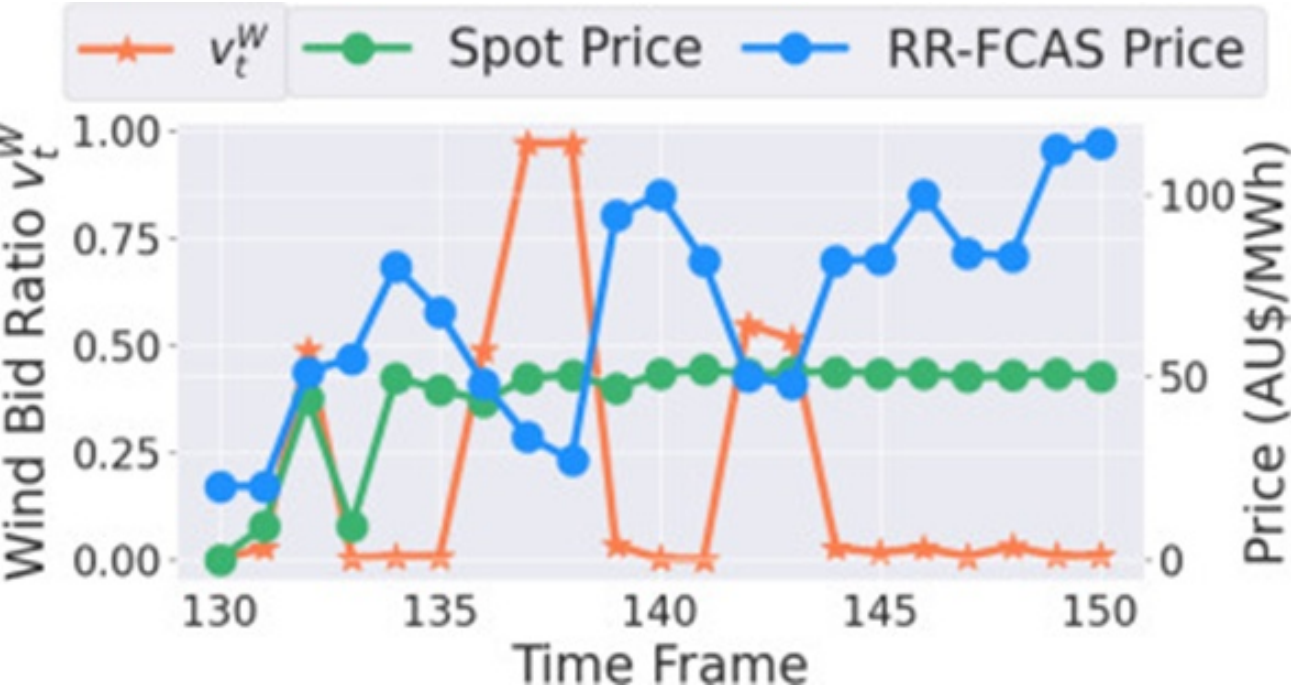


Figure 7. The wind farm bidding patterns in the joint market.

Table 4. The wind curtailment responses (in MWh) in the P&O and AggJointDRL.

	Spot	Reg. FCAS	Joint
P&O	130/623	53/623	78/623
AggJointDRL	313/437	235/371	257/415

Table 5. The aggregated revenue (in AU\$) with (i.e., coupled)/without (i.e., uncoupled) wind curtailment management in all three bidding scenarios.

	Spot	Reg. FCAS	Joint
Uncoupled	945,980	530,150	1,068,566
Wind Coupled	979,784	540,690	1,098,686
Boost	4%	2%	3%
Uncoupled	94,126	252,815	530,219
BESS Coupled	113,292	247,759	524,917
Boost	20%	−2%	−1%
Uncoupled	1,040,106	782,965	1,598,785
Total Coupled	1,093,076	788,449	1,623,603
Boost	5%	1%	2%

Managing wind curtailment

The capacity of the combined wind-BESS system to minimize onsite wind curtailment is directly connected to its overall profitability. By reducing curtailment, the wind farm can partially avoid financial penalties for missing dispatch instructions and can also use otherwise wasted energy to charge the battery, which may later be sold back to the grid through the BESS's market offers.

Table 4 compares the wind curtailment-handling performance of the conventional P&O benchmark with that of the proposed AggJointDRL method. The results indicate that AggJointDRL is markedly more effective at utilizing curtailed wind energy. Furthermore, curtailed wind energy serves as a meaningful charging source for the BESS, as shown in **Figure 8**. In detail, it contributes roughly 10% of the BESS's total charging energy when participating only in the spot market, around 16% in the Regulation FCAS market, and approximately 8% under joint-market bidding.

To better quantify the economic value of curtailment management, an additional set of experiments was conducted in which the co-located wind farm and BESS operated as two completely separate, uncoupled entities without any curtailment mitigation. Their performance was then compared with that of the AggJointDRL strategy, and the outcomes are summarized in **Table 5**. In the spot market, enabling wind curtailment reduction improved bidding outcomes for both the wind farm and the BESS. However, in the Regulation FCAS and joint-market scenarios, although the wind farm continued to gain extra revenue from curtailment reduction, the BESS experienced modest revenue declines of 2% and 1%, respectively. These findings reveal that the BESS may encounter a trade-off: prioritizing its individual earnings versus supporting the collective profitability of the integrated wind-battery system through effective wind curtailment management.

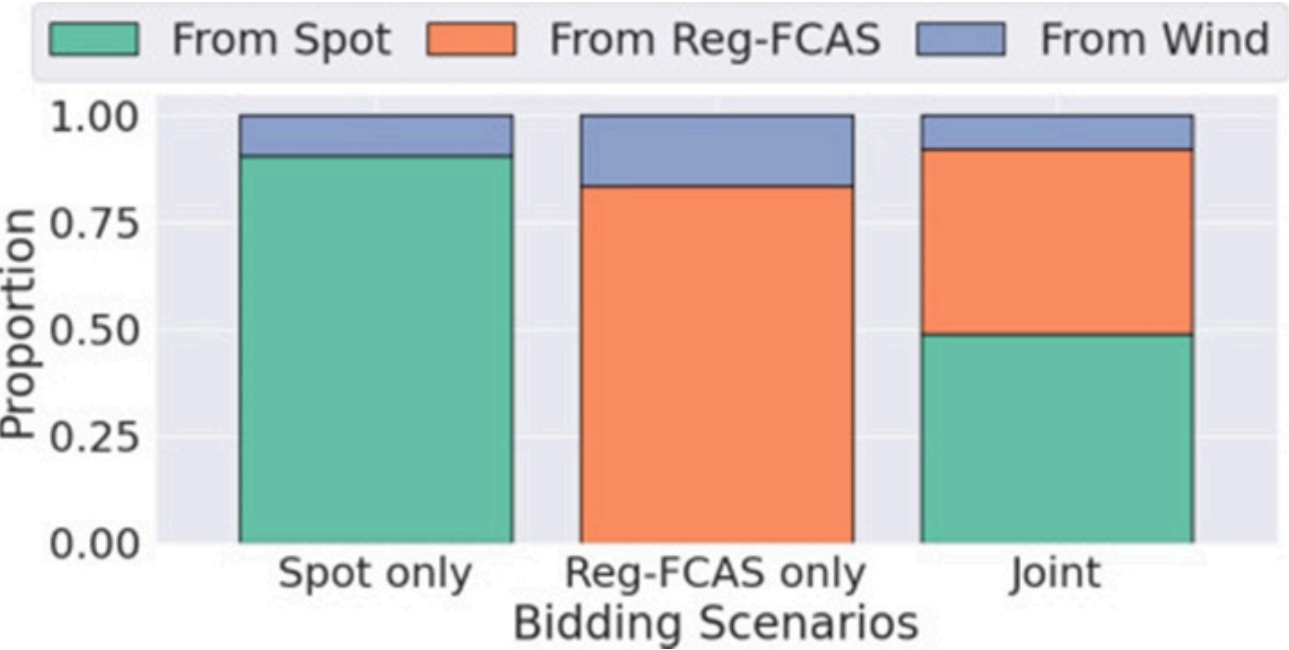


Figure 8. Composition of the BESS energy source.

The higher earnings achieved by the BESS in the spot market stem from the fact that charging with locally curtailed wind energy incurs no cost, whereas purchasing power at non-negative spot prices typically leads to financial losses. In contrast, when the Regulation FCAS market is available, offering lower regulation (RL-FCAS) services proves more attractive than relying on uncertain volumes of onsite curtailed wind power. Consequently, actively responding to wind curtailment results in a modest reduction in BESS profitability in both the Regulation FCAS and joint-market scenarios, as shown in **Table 5**.

Even though the BESS experiences these small revenue decreases, implementing wind curtailment management still boosts the total income of the integrated wind-battery system across all three bidding cases. This outcome underscores the importance of flexible, real-time coordination between wind generation and battery storage assets to maximize the economic performance of co-located renewable-plus-storage systems. A detailed case study further explores how the BESS navigates the balance between reducing wind curtailment and pursuing energy arbitrage opportunities in the spot market.

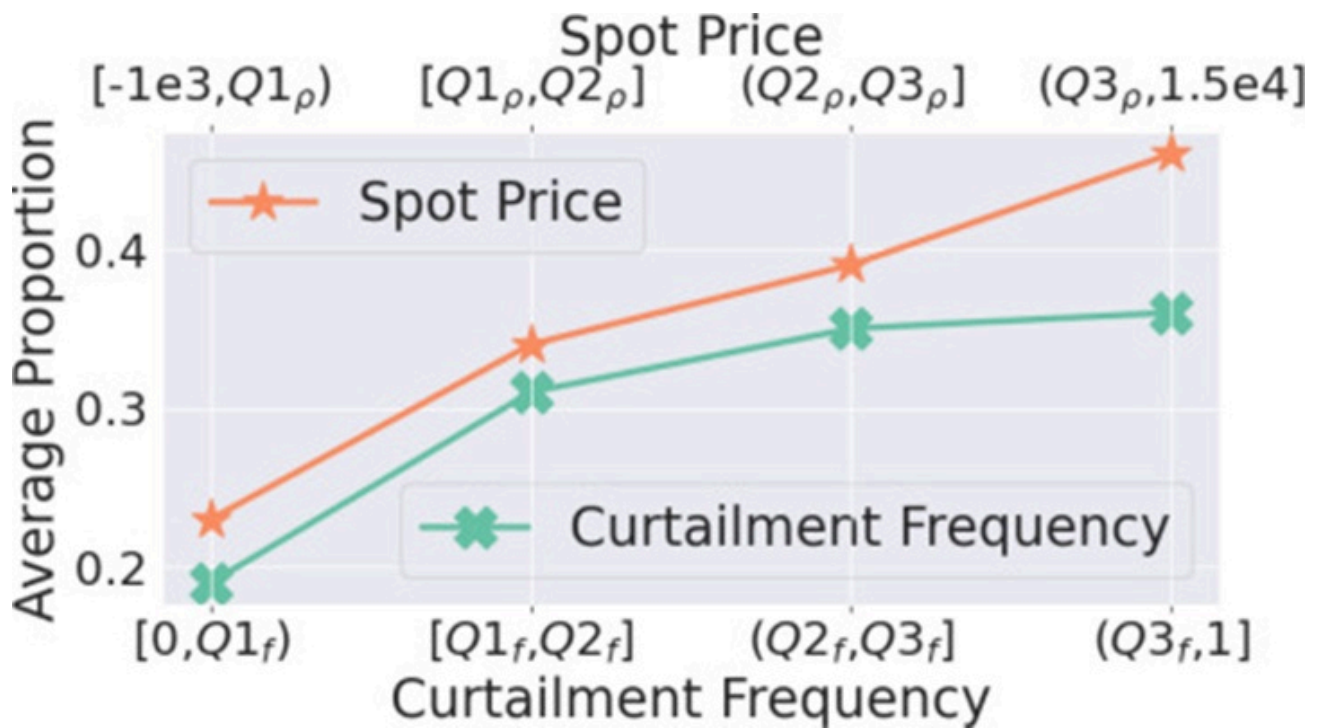


Figure 9. Influence of spot price variations and wind curtailment frequency on the bidding decisions of the BESS.

Case study in the spot market

Because BESS charging decisions respond dynamically to prevailing market conditions and the availability of curtailed wind energy, we examined the interplay between spot prices and the volume of curtailed wind energy drawn. Spot prices were divided into quartiles ($Q1_\rho$, $Q2_\rho$, and $Q3_\rho$) to reveal these patterns more clearly. As illustrated in **Figure 9**, the BESS adheres to arbitrage principles and predominantly charges from the spot market during periods of low prices. Conversely, when spot prices rise sharply, the system shifts its preference toward curtailed wind energy to avoid substantial financial losses from expensive grid purchases.

Although sourcing power from curtailed wind incurs no direct cost, its availability remains uncertain. It depends entirely on the volume of excess generation from the wind farm that is not dispatched to the grid. If insufficient curtailed energy is present, the BESS cannot rely on it and must seek alternative sources. Therefore, charging strategies must account for both the prevailing spot market cost and the probabilistic availability of free curtailed energy.

The frequency of wind curtailment events, denoted as f_{WC} , also exerts considerable influence on BESS charging behavior. To investigate this, curtailment frequency was segmented into quartiles ($Q1_f$, $Q2_f$, and $Q3_f$). Analysis shows that in the lowest-frequency quartile, the BESS utilizes approximately 19% of curtailed wind energy. As curtailment frequency rises, the proportion of charging from onsite curtailed wind increases, eventually stabilizing near 36%, as shown in **Figure 9**. This saturation effect contrasts with the behavior observed under high spot prices, where the BESS continues to

source over 50% of its charging energy from the spot market. The difference arises primarily from the inherent uncertainty of curtailment: even at high predicted frequencies, actual curtailment events may not materialize.

Conclusion

This research integrates renewable energy with artificial intelligence by introducing an innovative deep reinforcement learning approach for optimal bidding of co-located wind and battery storage systems. The proposed method effectively addresses wind curtailment challenges and market participation decisions, underscoring the transformative role of deep reinforcement learning techniques in advancing sustainable energy systems.

In conclusion, coordinated operation of a co-located wind farm and BESS substantially enhances joint bidding performance across both the spot and regulation FCAS markets. To capture this potential, we developed AggJointDRL—a model-free, real-time DRL-based bidding strategy that simultaneously optimizes revenue and mitigates wind curtailment. Results confirm that joint-market participation markedly improves the economic viability of the wind-battery system, delivering superior outcomes compared to bidding in either market separately. Moreover, AggJointDRL consistently outperforms the conventional P&O benchmark in both computational efficiency and financial returns for both the wind farm and the BESS. The study further demonstrates that otherwise wasted onsite wind energy serves as a valuable, low-cost charging resource for the BESS. By actively managing curtailment, the integrated system achieves additional revenue gains, emphasizing the critical need for agile coordination between generation and storage assets to maximize profitability.

A detailed examination of BESS operations under varying spot prices and wind curtailment frequencies reveals that the system increasingly relies on curtailed wind energy as both prices and curtailment likelihood increase. These insights offer practical value to wind farm operators seeking to refine their market strategies and inform investment decisions regarding battery storage deployment. Successful adoption of such strategies may also promote greater co-location of renewable generation and storage, supporting broader policy objectives for optimizing energy system planning and delivering system-wide benefits.

Future work will extend the joint bidding framework by incorporating transmission constraints. A primary difficulty lies in scaling to a sophisticated price-maker formulation, given the more competitive environment faced by larger co-located systems. In parallel, efforts will focus on improving the accuracy of wind farm availability forecasts to enable more effective, proactive wind curtailment management.

Nomenclature

ιW_t	Proportion of wind farm bid power in the spot market
ιDch_t	Binary variable indicating the BESS discharge operation
ιCh_t	Binary variable indicating the BESS charge operation
pW_t	Forecasted wind generation availability
pW,Act_t	Actual wind generation
pW,WC_t	Curtailed wind generation
$pBESS,S_t$	Battery bid power in the spot market
$BESS,Reg\ p_t$	Battery bid power in the Regulation FCAS market
$pBESS,WC_t$	Power that the BESS plans to draw from wind curtailment
$\bar{p}BESS,WC_t$	Actual power drawn by BESS from the onsite wind farm

s_t	A series of AGC signals in one dispatch interval
φS_t	Market-clearing price in the spot market
φRR_t	Market-clearing price (i.e., ancillary service price) in Regulation Raise FCAS market
φRL_t	Market-clearing price (i.e., ancillary service price) in the Regulation Lower FCAS market
e_t	Battery current capacity
L	Length of AGC signals in one dispatch interval
Δs	Duration of each AGC signal
Δt	Duration of one NEM dispatch interval.
T	The overall time frame
λ	Penalty coefficient for deviations between actual wind generation and prescribed dispatch target
ηDch	Battery discharging efficiency
ηCh	Battery charging efficiency
c	Battery degradation cost coefficient
	Wind installed capacity
	Battery rated power
E_{min}	The battery has a lower energy storage limit
E_{max}	Battery upper energy storage limit

Acknowledgements

None

Conflict of interest

The authors declare the following financial interests/personal relationships, which may be considered potential competing interests: Hao Wang reports that financial support and travel were provided by the Australian Research Council.

Financial support

This work has been supported in part by the FIT Academic Funding of Monash University, Australia, and the Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) under Grant [DE230100046](#).

Ethics statement

None

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

IPCC. Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK; New York, NY, USA: Cambridge University Press; 2022.
<https://doi.org/10.1017/9781009157926>.

Clean Energy Council. Clean Energy Australia Report. Clean Energy Council; 2022.

Australian Energy Market Operator. Australian National Electricity Market Generation Information. Australian Energy Market Operator; 2022.

Burke DJ, O'Malley MJ. Factors influencing wind energy curtailment. *IEEE Trans Sustain Energy*. 2011;2(2):185-93.
<https://doi.org/10.1109/TSTE.2011.2104981>.

Australian Energy Market Operator. Integrated System Plan for the National Electricity Market. Australian Energy Market Operator; 2022.

Sun Y, Zhong J, Li Z, Tian W, Shahidehpour M. Stochastic scheduling of battery-based energy storage transportation system with the penetration of wind power. In: 2017 IEEE Power & Energy Society General Meeting. 2017:1.
<https://doi.org/10.1109/PESGM.2017.8273785>.

Dui X, Zhu G, Yao L. Two-stage optimization of battery energy storage capacity to decrease wind power curtailment in grid-connected wind farms. *IEEE Trans Power Syst*. 2018;33(3):3296-305.
<https://doi.org/10.1109/TPWRS.2017.2779134>.

Zhang Z, Zhang Y, Huang Q, Lee W-J. Market-oriented optimal dispatching strategy for a wind farm with a multiple stage hybrid energy storage system. *CSEE J Power Energy Syst*. 2018;4(4):417-24.
<https://doi.org/10.17775/CSEEJPES.2018.00130>.

Akbari E, Hooshmand R-A, Gholipour M, Parastegari M. Stochastic programming-based optimal bidding of compressed air energy storage with wind and thermal generation units in energy and reserve markets. *Energy*. 2019;171:535-46.
<https://doi.org/10.1016/j.energy.2019.01.014>.

Xie Y, Guo W, Wu Q, Wang K. Robust MPC-based bidding strategy for wind storage systems in real-time energy and regulation markets. *Int J Electr Power Energy Syst*. 2021;124:106361.
<https://doi.org/10.1016/j.ijepes.2020.106361>.

Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int J Forecast*. 2014;30(4):1030-81.
<https://doi.org/10.1016/j.ijforecast.2014.08.008>.

Alanazi A, Khodaei A. Optimal battery energy storage sizing for reducing wind generation curtailment. In: 2017 IEEE Power & Energy Society General Meeting. 2017:1-5.
<https://doi.org/10.1109/PESGM.2017.8274599>.

Saber H, Moeini-Aghtaie M, Ehsan M, Fotuhi-Firuzabad M. A scenario-based planning framework for energy storage systems with the main goal of mitigating wind curtailment issue. *Int J Electr Power Energy Syst*. 2019;104:414-22.
<https://doi.org/10.1016/j.ijepes.2018.07.020>.

Nikoobakht A, Aghaei J, Shafie-Khah M, Catalão JPS. Minimizing wind power curtailment using a continuous-time risk-based model of generating units and bulk energy storage. *IEEE Trans Smart Grid*. 2020;11(6):4833-46.
<https://doi.org/10.1109/TSG.2020.3004488>.

Gill S, Ault GW, Kockar I. The optimal operation of energy storage in a wind power curtailment scheme. In: 2012 IEEE Power and Energy Society General Meeting. 2012:1-8.
<https://doi.org/10.1109/PESGM.2012.6344705>.

Khaloie H, Abdollahi A, Shafie-khah M, Anvari-Moghaddam A, Nojavan S, Siano P, et al. Coordinated wind-thermal-energy storage offering strategy in energy and spinning reserve markets using a multi-stage model. *Appl Energy*. 2020;259:114168.
<https://doi.org/10.1016/j.apenergy.2019.114168>.

Urgaonkar R, Urgaonkar B, Neely MJ, Sivasubramaniam A. Optimal power cost management using stored energy in data centers. In: Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems. New York, NY, USA: Association for Computing Machinery; 2011:221-32.
<https://doi.org/10.1145/1993744.1993766>.

Qin J, Chow Y, Yang J, Rajagopal R. Distributed online modified greedy algorithm for networked storage operation under uncertainty. *IEEE Trans Smart Grid*. 2016;7(2):1106-18.
<https://doi.org/10.1109/TSG.2015.2422780>.

Mo Y, Lin Q, Chen M, Qin S-ZJ. Optimal online algorithms for peak-demand reduction maximization with energy storage. In: Proceedings of the Twelfth ACM International Conference on Future Energy Systems. e-Energy '21. New York, NY, USA: Association for Computing Machinery; 2021:73-83.
<https://doi.org/10.1145/3447555.3464857>.

Yang J, Yang M, Wang M, Du P, Yu Y. A deep reinforcement learning method for managing wind farm uncertainties through energy storage system control and external reserve purchasing. *Int J Electr Power Energy Syst*. 2020;119:105928.
<https://doi.org/10.1016/j.ijepes.2020.105928>.

Oh E, Wang H. Reinforcement-learning-based energy storage system operation strategies to manage wind power forecast uncertainty. *IEEE Access*. 2020;8:20965-76.
<https://doi.org/10.1109/ACCESS.2020.2968841>.

Australian Energy Market Operator. Guide to Market Clearing. Australian Energy Market Operator; 2021.

Australian Energy Market Operator. How the National Electricity Market Works. Australian Energy Market Operator; 2020.

Australian Energy Market Operator. Five-Minute Settlement: High Level Design. Australian Energy Market Operator; 2017.

Australian Energy Market Operator. Guide to Ancillary Services in the National Electricity Market. Australian Energy Market Operator; 2015.

Australian Energy Market Operator. NEM Operational Forecasting and Dispatch Handbook for Wind and Solar Generators. Australian Energy Market Operator; 2022.

Bordin C, Anuta HO, Crossland A, Gutierrez IL, Dent CJ, Vigo D. A linear programming approach for battery degradation analysis and optimization in off-grid power systems with solar energy integration. *Renew Energy*. 2017;101:417-30.
<https://doi.org/10.1016/j.renene.2016.08.066>.

Anwar M, Wang C, de Nijs F, Wang H. Proximal policy optimization based reinforcement learning for joint bidding in energy and frequency regulation markets. In: IEEE Power & Energy Society General Meeting. 2022.

Fujimoto S, Hoof H, Meger D. Addressing function approximation error in actor-critic methods. In: International Conference on Machine Learning. 2018:1582-91.

Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.

Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. The MIT Press; 2018.

Hasselt H. Double Q-learning. In: Lafferty J, Williams C, Shawe-Taylor J, Zemel R, Culotta A, editors. Advances in Neural Information Processing Systems. Vol 23. Curran Associates, Inc.; 2010:1-9.

Australian Energy Market Operator. AEMO NEMWEB Data. Australian Energy Market Operator; 2022.

Mitchell S, O'Sullivan M, Dunning I. PuLP: A Linear Programming Toolkit for Python. Vol 65. The University of Auckland; 2011.

Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. 2017. arXiv:1707.06347.

Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, et al. Continuous control with deep reinforcement learning. 2015. arXiv:1509.02971.
<https://doi.org/10.48550/ARXIV.1509.02971>.

Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: Proceedings of the 35th International Conference on Machine Learning. Vol 80. PMLR; 2018:1861-70.