Pumped Hydro Energy Storage: Structural Results and Solution Algorithms

Harun Avci[a,*], Ece Cigdem Karakoyun[b], Ayse Selin Kocaman[c], Emre Nadar[c], Parinaz Toufani[c]

[a]Department of Economics, Finance, and Quantitative Analysis, Kennesaw State University, Kennesaw, GA 30144, USA, havci@kennesaw.edu; [b]Econometric Institute, Erasmus University Rotterdam, 3062 PA, Rotterdam, The Netherlands, karakoyun@ese.eur.nl; [c]Department of Industrial Engineering, Bilkent University, 06800 Ankara, Turkey, selin.kocaman@bilkent.edu.tr, emre.nadar@bilkent.edu.tr, parinaz.toufani@bilkent.edu.tr

We study a pumped hydro energy storage facility that consists of two interconnected reservoirs – an upper and a lower – where one reservoir may be fed by natural inflow. The system operator makes real-time decisions regarding the amount of water to pump or release, thereby determining the amount of electricity to buy from or sell to the market. We model this problem as a Markov decision process, accounting for uncertainties in both streamflow rates and electricity prices. We prove the optimality of a state-dependent threshold policy under positive electricity prices: The state space can be partitioned into several disjoint domains, each associated with a different action type, such that it is optimal to bring the total amount of water in the facility and the amount of water in the upper reservoir to a different pair of state-dependent target levels in each domain. For more general data-calibrated instances in which prices can be negative, we use this threshold structure as the basis for policy-approximation algorithms. These algorithms yield near-optimal solutions up to 28 times faster than brute-force optimization, with no significant drain on total cash flows compared to the optimal solution.

Keywords: OR in energy; pumped hydro energy storage; streamflow; electricity price; Markov decision processes

1. Introduction

Pumped hydro energy storage (PHES) is the most mature large-scale energy storage technology, with a history dating back to the 1890s (DOE 2023a). The earliest PHES facilities were installed by state-owned utilities to support baseload power plants. The expansion of PHES installations slowed in the 1990s as the popularity of large baseload nuclear and fossil-fuel-fired power plants declined due to growing environmental and safety concerns. More recently, however, PHES has attracted renewed interest from private companies, driven by the rise of deregulated electricity markets and the growing share of intermittent renewable energy sources in power generation (IHA 2024b). As of 2024, PHES accounted for more than 90% of global utility-scale storage, with an installed capacity of 179 GW (IHA 2024a). Projects at various stages of development represent a total pipeline capacity of 214 GW (IHA 2025a).

A PHES facility, also referred to as a 'water battery,' typically consists of two reservoirs with an elevation difference; it stores energy in the form of hydraulic potential energy in the upper reservoir (Kunzig 2024). Water can be pumped from the lower reservoir to the upper reservoir when the operator wants to store energy during off-peak periods (i.e., when excess energy supply results in low electricity prices). During peak periods, the operator can generate energy by releasing water from the upper reservoir to the lower reservoir, and in cascade systems, even from the lower reservoir to the streambed. Natural streamflow may exist and feed either the upper or lower reservoir.

PHES facilities can be designed in various configurations depending on specific hydrological, geological,

* Corresponding author: havci@kennesaw.edu

The final version is available at https://doi.org/10.1016/j.eneco.2026.109568

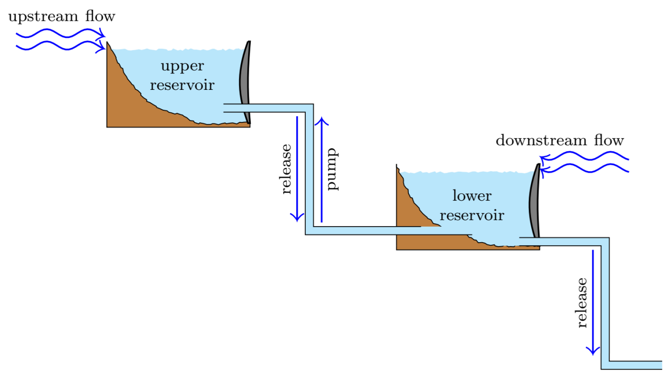
and topographical conditions. They are typically categorized into two broad classes: closed-loop and open-loop facilities. Closed-loop facilities are off-stream and have no natural inflow to either reservoir (DOE 2023b). Examples include the Turlough Hill facility in Ireland, the Ffestiniog facility in the UK, and the Marmora facility in Canada (IHA 2025b). In contrast, open-loop facilities are on-stream and have continuous natural inflow to either the upper or lower reservoir (DOE 2023b). Many PHES facilities in operation today fall into the open-loop category (Saulsbury 2020 and Blakers et al. 2021). The Tumut 3 facility on the Tumut River in New South Wales, Australia, and the Bleiloch facility along the Saale River in Germany are examples of PHES facilities having natural inflow to their upper reservoirs (Blakers et al. 2021 and Vattenfall 2025a). The Geesthacht and Goldisthal facilities in Germany and the Blenheim-Gilboa facility in the U.S. are examples of PHES facilities having natural inflow to their lower reservoirs (Vattenfall 2025b,c and The Center for Land Use Interpretation 2025). Open-loop facilities can be further categorized into two types: cascade facilities, which can release water from both reservoirs, and non-cascade facilities, which can release water only from the upper reservoir.

The distinction between closed-loop and open-loop PHES is not only hydrological but also economic. In a closed-loop facility, the total amount of water in the system is essentially fixed; therefore, the short-term operating problem mainly focuses on allocating a fixed storage resource across pumping and generation opportunities. In an open-loop facility, by contrast, natural inflows make the total amount of water in the system stochastic and can create spill risk when reservoir or releasing capacities become binding. Consequently, the marginal value of stored water depends not only on the current electricity price, but also on the reservoir in which the water is located, the unused capacity in each reservoir, and how future inflows may interact with current storage decisions. This economic difference motivates a model that tracks both the total amount of water in the PHES facility and its allocation between the upper and lower reservoirs.

We study the energy generation and storage problem for a generalized PHES configuration, as depicted in Figure 1. This configuration encompasses all previously discussed settings and examples, including many PHES facilities observed in practice and several examined in the literature (Toufani et al. 2022a,b and Yurter et al. 2024). In each time period, the PHES operator must decide how much energy to generate by releasing water from the reservoirs, how much to store by pumping water to the upper reservoir, and how much electricity to buy from or sell to the market. We formulate this problem as a Markov decision process (MDP) that incorporates the reservoir water-level dynamics through endogenous state variables and accounts for uncertainties in streamflow rate and electricity prices through exogenous state variables.

Despite the widespread use of PHES in practice and the extensive literature on its operational planning (see Parker et al. 2019 for comprehensive review for energy storage systems and Toufani et al. 2023 for a focused review on PHES under uncertainty), there is limited work characterizing the optimal policy structure for open-loop PHES systems. Toufani et al. (2023) classify the existing PHES optimization literature by problem type, PHES configuration, uncertainty paradigm, and solution method, and indicate that much

Figure 1 Illustration of a generalized PHES configuration.



of the literature relies on stochastic programming approaches. Structural results that exploit the specific properties of the underlying optimization problem remain limited, particularly for open-loop configurations with natural inflow. This study contributes to the literature by providing such a characterization.

The first part of our contribution is analytical. When the electricity price is always positive, we prove that the optimal value function is jointly concave in the endogenous state variables. This result leads to a state-dependent threshold policy for operating the open-loop PHES facility. The state space of our MDP can be partitioned into three disjoint domains, each associated with a different optimal action type that can be ‘pump,’ ‘release,’ or ‘keep unchanged.’ In each domain, it is optimal to bring the total amount of water in the PHES facility and the amount of water in the upper reservoir to a different pair of state-dependent threshold levels. These thresholds reflect the trade-off between the current market value of electricity and the future value of stored water. The threshold levels are higher for the ‘release’ decision than for the ‘pump’ decision. While higher water levels in any reservoir increase system cash flows, a higher water level in one reservoir is more desirable when the water level in the other reservoir is lower. Moreover, if both reservoirs have the same releasing capacity, a higher water level in the upper reservoir is more desirable than in the lower reservoir.

The second part of our contribution is algorithmic. When prices can be negative, the concavity argument used to prove the threshold characterization may no longer apply. We therefore develop a policy-approximation algorithm as a heuristic solution method that embeds the threshold structure established under positive prices into a backward induction framework. The algorithm follows the derived policy structure in positive-price states and uses the same structure as an approximation in negative-price states. In each period, it computes state-dependent threshold levels for the amounts of water in the PHES facility and the upper reservoir to guide energy generation and storage decisions, thereby reducing computational complexity. To further improve efficiency, we introduce a variant that employs partial-state-dependent threshold levels, which are independent of the total amount of water in the system. Our numerical results for baseline data-calibrated examples show that the policy-approximation algorithm yields solutions with a maximum deviation of 0.02% from optimality and is up to 11 times faster than the standard dynamic programming

(DP) algorithm. The variant yields solutions with an average deviation of only 1.32% from optimality and is up to 28 times faster than the standard DP algorithm.

The literature on the operational planning for PHES facilities is vast. Many studies derive optimal operational schedules for PHES facilities, either stand-alone or integrated with other renewable energy sources, under uncertainty (in electricity price, load demand, and/or renewable sources) via stochastic programming models (see Toufani et al. 2023). MDPs have garnered attention more recently. For example, Löhndorf et al. (2013) optimize the energy commitment and storage decisions of a PHES operator participating in a day-ahead electricity market. They formulate a multi-stage stochastic program for intraday decisions and an MDP for interday decisions, combining stochastic dual dynamic programming and approximate dynamic programming methods in their solution approach. Toufani et al. (2022a,b) employ MDP formulations to evaluate the short-term cash-flow benefits of transforming cascade hydropower stations into PHES systems and to compare the short-term cash-flow performance of different PHES configurations, respectively. These two studies do not characterize the structure of optimal policies, instead relying on brute-force optimization to obtain their results. This approach poses significant limitations for real-life implementation due to prohibitively long solution times. Liu et al. (2022b) explore the scheduling policy of a price-maker electricity merchant operating a closed-loop PHES facility together with a wind farm, while Liu et al. (2022a) examine the impact of production tax credits on dispatch decisions for a merchant operating the same hybrid system. In contrast to these studies, we characterize the optimal policy structure for open-loop PHES facilities. Steffen and Weber (2016) use continuous-time optimal control to derive efficient operation programs for a PHES facility operating under deterministic, time-varying electricity prices, distinguishing between cases with binding and non-binding reservoir constraints. Our setting differs in three key respects: prices are stochastic and can be negative in our data-calibrated instances; natural inflow makes the amount of water in the facility stochastic; and our PHES facility consists of two interconnected reservoirs with potentially independent inflows, necessitating two endogenous state variables rather than a single storage level.

A parallel strand of the PHES literature characterizes the facility from a technology and flexibility perspective rather than from an operations perspective. Zhao et al. (2022) develop a nonlinear model of a pump-turbine unit that couples hydraulic, mechanical, and electrical losses, showing that flexible operation, in which the unit's load varies widely to follow renewable output, can reduce conversion efficiency relative to operation near the design point. Zhao et al. (2024) extend this perspective by comparing several PHES technologies, including fixed-speed, variable-speed, ternary, and quaternary configurations, and find that operational flexibility, conversion efficiency, and cost performance depend on the unit configuration. While these studies establish that a facility's conversion efficiency fluctuates with its operating point and unit technology, we abstract from this dependence to maintain the tractability of our structural analysis, treating system efficiency as a constant.

Since closed-loop PHES facilities can be viewed as grid-scale storage systems similar to industrial batteries, the literature on the operational planning problem for storage facilities is also related to our work. In this literature, Zhou et al. (2019) consider a wind farm co-located with an industrial battery and model the energy generation and storage problem as an MDP. They prove the optimality of a partial-state-dependent threshold policy under positive electricity prices. However, our setting introduces additional challenges due to the open-loop nature of the PHES facility: (i) Unlike a battery with fixed total energy capacity (despite dynamically changing storage levels), the PHES facility – subject to natural inflows into either reservoir and water releases from the lower reservoir – acts as a storage unit with variable energy capacity. (ii) Water may spill from the PHES facility when natural inflows exceed reservoir capacities or when the operator pumps or releases excessive amounts of water. Operational decisions can thus affect the amount of water lost. (iii) While Zhou et al. (2019) track energy storage via a single endogenous state variable, our model requires tracking water levels in both reservoirs via two endogenous state variables. Consequently, unlike Zhou et al. (2019), we derive the optimal policy structure by identifying several multi-dimensional properties of the optimal value function, including *submodularity*. Recently, Karakoyun et al. (2026) extend the setting in Zhou et al. (2019) by adding hour-ahead energy commitment decisions and an extra state variable for the previous period's commitment level. However, their model still assumes a battery with fixed energy capacity, lacks spillover effects, and does not exhibit submodularity.

In what follows, §2 formulates the problem and §3 presents the optimal policy structure. §4 constructs the heuristic solution methods based on our structural results. §5 describes the time series models embedded into our MDP formulation. §6 presents the numerical results when the price can also be negative. §7 offers a summary and conclusion. Proofs of the analytical results are contained in the appendix.

2. Problem Formulation

The operating problem can be viewed as a repeated trade-off between immediate market cash flows and the future value of stored water. Releasing water generates current revenue, pumping incurs an electricity cost to increase future generation potential, and holding water steady preserves operational flexibility. In an open-loop facility, natural inflows make this trade-off state-dependent: high water levels elevate future spill risk, whereas low water levels restrict the operator's ability to execute price arbitrage during subsequent high-price periods. The MDP below formalizes this trade-off.

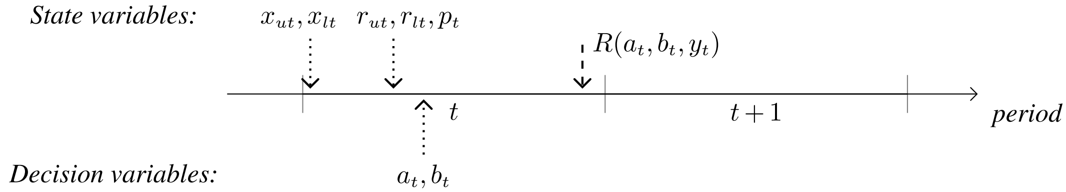
We consider the open-loop PHES facility shown in Figure 1. Such facilities can be constructed on a river by building two new reservoirs at different altitudes or replacing the turbines of a conventional hydropower station with reversible ones. Many investors in liberalized markets prefer to build PHES facilities by rehabilitating conventional hydropower stations (Ardizzon et al. 2014). The PHES facility we consider generates energy by releasing water from the upper reservoir to the lower reservoir and/or from the lower reservoir to the streambed. It can also store energy by pumping water from the lower reservoir to the upper reservoir. Both reservoirs receive natural inflows. Water flow between the reservoirs does not alter the total amount of water in the facility, unless the water level exceeds the capacity of either reservoir, in which case the excess water spills over. Similar PHES facilities have been examined by Toufani et al. (2022a,b) in the contexts described in Section 1. Our formulation closely follows the relevant material in these papers.

The PHES operator participates in a wholesale market where dispatch and purchase amounts determined by market participants in real time are always accepted, unlike forward markets where electricity is traded through bids or commitments made in advance. Focusing on energy generation and storage decisions that are free of interactions with commitment decisions allows us to more effectively capture the structural dynamics and stochastic nature of the problem. Similar approaches also appear in many related papers; see, for example, Castronuovo and Lopes (2004), Abbaspour et al. (2013), Shu and Jirutitijaroen (2013), Harsha and Dahleh (2014), Steffen and Weber (2016), and Zhou et al. (2016, 2019). We assume that the PHES operator is a price-taker in the wholesale market, which is reasonable when the facility contributes only a small share to the overall energy supply. This assumption is common in many PHES studies; see, for example, Garcia-Gonzalez et al. (2008), Vespucci et al. (2012), Löhndorf et al. (2013), Ding et al. (2014), Al-Swaiti et al. (2017), and Kusakana (2018). Furthermore, we treat the system efficiency as a constant, which is consistent with the modeling choice in much of the PHES literature, including Löhndorf et al. (2013), Steffen and Weber (2016), Toufani et al. (2022a,b), Liu et al. (2022b,a), and Karakoyun et al. (2026).

To simplify the illustration of our analysis, we construct the optimization model under the assumptions of equal head for both reservoirs and identical efficiencies in both generation and pumping modes. We denote the system efficiency by $\theta \in (0,1)$. To calculate the energy generated by releasing water from any reservoir, we multiply the volume of water released by the water density, the gravitational constant, and the potential head of the reservoir (under perfect efficiency). This conversion allows us to express the water levels in both reservoirs in energy terms. The PHES facility has finite energy and power capacities. The energy capacity of the upper (or lower) reservoir is the maximum amount of water (in energy units) that can be stored in the upper (or lower) reservoir. We denote the energy capacities of the upper and lower reservoirs by C_U and C_L, respectively. The power capacity is the maximum amount of water (in energy units) that can be released from the upper or lower reservoir or pumped from the lower reservoir in a single time period of length Δt. We denote the power capacities of the upper and lower reservoirs in releasing mode by K_{Ru} and K_{Rl}, respectively. We denote the power capacity in pumping mode by K_P. For notational convenience, we define $C_{Ru} = K_{Ru}\Delta t$, $C_{Rl} = K_{Rl}\Delta t$, and $C_P = K_P \Delta t$. The power capacities affect the optimization model only when $C_{Ru} \leq C_U$, $C_{Rl} \leq C_U + C_L$, and $C_P \leq C_L$.

We study the energy generation and storage decisions in this setting over a T-period planning horizon. Let $\mathcal{T} := \{1, 2, \ldots, T\}$ present set of periods. We define $x_{ut} \in [0, C_U]$ and $x_{lt} \in [0, C_L]$ as the amounts of water (in energy units) available in the upper and lower reservoirs, respectively, at the beginning of period t. Tracking the water levels in both reservoirs is important because the system's economic value exhibits a strict dependence on the specific allocation between the reservoirs, rather than merely the total amount of water available in the system. We also define r_{ut} and r_{lt} as the amounts of water runoff (external inflow)

Figure 2 Timeline of events in our model.



into the upper and lower reservoirs (in energy units) in period t, respectively. Lastly, we define p_t as the electricity price in period t. We include the variables x_{ut} and x_{lt} as well as the tuple $y_t := (r_{u\eta}, r_{l\eta}, p_\eta)_{\eta \leq t}$ in our state description. This tuple allows us to incorporate autoregressive time series models that use past values for forecasting. We assume that y_t follows an exogenous stochastic process obeying $\mathbb{E}_{p_\kappa|y_t}[|p_\kappa|] < \infty$, $\forall \kappa \geq t$.

At the beginning of period $t \in \mathcal{T}$, the operator observes first the amounts of water accumulated in the two reservoirs (x_{ut} and x_{lt}) and then the exogenous state variables (r_{ut}, r_{lt}, and p_t). With these observations, the operator decides how much water to release from or pump to the upper reservoir $a_t \in \mathbb{R}$ (in energy units) and how much water to release from the lower reservoir $b_t \in \mathbb{R}_+$ (in energy units). Water is released if $a_t > 0$ and water is pumped if $a_t < 0$. See Figure 2 for an illustration of the sequence of events. We define $\mathbb{U}(x_{ut}, x_{lt}, y_t)$ as the admissible set of action pairs (a_t, b_t) in state (x_{ut}, x_{lt}, y_t). The following conditions must hold for any action pair $(a_t, b_t) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$: $-C_P \leq a_t \leq \min\{x^r_{ut}, C_{Ru}\}$, $0 \leq b_t \leq C_{Rl}$, and $b_t - a_t \leq x^r_{lt}$, where for simplicity $x^r_{ut} := x_{ut} + r_{ut}$ and $x^r_{lt} := x_{lt} + r_{lt}$ denote the amounts of water available in the upper and lower reservoirs, respectively, upon observing the water runoff in period t. If the water runoff and the operator's decisions cause any water spillage from any reservoir, the excess amount spilling is lost. This implies that $x_{u(t+1)} = \min\{x^r_{ut} - a_t, C_U\}$ and $x_{l(t+1)} = \min\{x^r_{lt} + a_t - b_t, C_L\}$.

The PHES configuration we consider is flexible enough to cover various realistic scenarios: It reduces to a cascade system with only upstream (or downstream) natural inflow if $r_{lt} = 0$ (or $r_{ut} = 0$) for all t. It also reduces to a non-cascade system if $C_{Rl} = 0$ (i.e., releasing mode is not allowed in the lower reservoir). The non-cascade system further reduces to a closed-loop PHES facility if $r_{ut} = r_{lt} = 0$ for all t. Finally, our configuration reduces to a conventional hydropower station if $C_P = 0$ (i.e., pumping mode is not allowed).

The operator aims to maximize the expected total cash flow from selling or purchasing electricity during the finite planning horizon. We account for the operator's decision type to formulate the payoff in each period t: (i) water is pumped to the upper reservoir ($a_t \leq 0$) or (ii) water is released from the upper reservoir ($a_t > 0$). We call these decisions types P and R, respectively. The amount of energy generated by releasing water from the upper reservoir or required to pump water to the upper reservoir in period t can be written as a function of action a_t: $E(a_t) = \min\{\theta a_t, a_t/\theta\}$. Combined with the energy generated by releasing water from the lower reservoir, the amount of energy sold or purchased in period t is given by $E(a_t) + \theta b_t$. The payoff in period t is thus a function of action pair (a_t, b_t) and state tuple y_t: $R(a_t, b_t, y_t) = (E(a_t) + \theta b_t)p_t$. This payoff expression links the reservoir operating decisions to the operator's market cash flow. Releasing water generates current revenue, while pumping uses purchased electricity to increase future generation potential, subject to the conversion loss captured by θ. The term θb_t captures the additional generation value from the lower reservoir in cascade configurations.

A control policy π can be characterized by a sequence of decision rules $\{A_t^\pi(x_{ut}^\pi, x_{lt}^\pi, y_t)\}_{t\in\mathcal{T}}$, where $A_t^\pi(\cdot)$ maps the state $(x_{ut}^\pi, x_{lt}^\pi, y_t)$ to a feasible action pair $(a_t^\pi(\cdot), b_t^\pi(\cdot))$, and x_{ut}^π and x_{lt}^π denote the random state variables governed by policy π, $\forall t\in\mathcal{T}\backslash\{1\}$. We define Π as the set of all admissible control policies. For any initial state (x_{u1}, x_{l1}, y_1), the optimal expected total cash flow over the finite horizon is

$$\max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t\in\mathcal{T}} R(A_t^\pi(x_{ut}^\pi, x_{lt}^\pi, y_t), y_t)\middle| x_{u1}, x_{l1}, y_1\right].$$

The optimal value function $v_t^*(x_{ut}, x_{lt}, y_t)$ in each state (x_{ut}, x_{lt}, y_t) in each period $t\in\mathcal{T}$ can be calculated via a DP recursion given by

$$v_t^*(x_{ut}, x_{lt}, y_t) = \max_{(a_t,b_t)\in\mathbb{U}(x_{ut},x_{lt},y_t)} \left\{R(a_t,b_t,y_t) + \mathbb{E}_{y_{t+1}|y_t}\left[v_{t+1}^*\left(x_{u(t+1)}, x_{l(t+1)}, y_{t+1}\right)\right]\right\} \tag{1}$$

where $v_T^*(x_{uT}, x_{lT}, y_T) = 0$. Given the initial state (x_{u1}, x_{l1}, y_1), the optimal expected total cash flow equals $v_1^*(x_{u1}, x_{l1}, y_1)$. We define $(a_t^*(x_{ut}, x_{lt}, y_t), b_t^*(x_{ut}, x_{lt}, y_t))$ as the optimal action pair in period t. The DP recursion selects the operating decision that maximizes the immediate payoff plus the expected future value of the resulting water levels. Differences in the value function capture how the value of stored water depends on the water levels in the two reservoirs. The structural analysis below identifies the conditions under which this intertemporal trade-off leads to threshold levels that determine the optimal operating decision.

3. Structural Analysis

In this section, we establish several structural properties of the optimal value function and use these properties to characterize the structure of the optimal energy generation and storage policy. We first assume that the electricity price is strictly positive throughout the finite horizon. Consequently, the analytical results in this section apply to positive-price regimes. These foundational results are later integrated into a generalized DP algorithm, enabling our solution framework to fully accommodate negative electricity prices.

ASSUMPTION 1. $p_t > 0$, $\forall t\in\mathcal{T}$.

We require Assumption 1 to show that the payoff in period t is jointly concave in action pair (a_t, b_t): Since the minimum of affine functions is concave, $R(a_t, b_t, y_t) = \min\{(\theta a_t + \theta b_t)p_t, (a_t/\theta + \theta b_t)p_t\}$ is concave in a_t and b_t. We next prove the concavity of our optimal value function.

PROPOSITION 1. *Under Assumption 1,* $v_t^*(x_{ut}, x_{lt}, y_t)$ *is jointly concave in* (x_{ut}, x_{lt}), $\forall t\in\mathcal{T}$.

Proposition 1 shows that the incremental value of water declines as water levels in the reservoirs increase. While additional water expands the operator's generation capability and future operational flexibility, its marginal value declines as reservoirs fill due to impending capacity constraints and increasing spill risk.

We also assume that the amounts of water runoff into the system are bounded in each period.

ASSUMPTION 2. *$r_{ut} \leq C_{Ru}$ and $r_{lt} \leq C_{Rl} - C_{Ru}$, $\forall t \in \mathcal{T}$.*

Assumption 2 mitigates forced water spillage from the PHES facility. For instance, suppose both reservoirs are full at the beginning of a period (i.e., $x_{ut} = C_U$ and $x_{lt} = C_L$). If the inflow to the upper reservoir during this period exceeds its releasing capacity (i.e., Assumption 2 is violated), some water will spill from the upper reservoir regardless of the release decision. Otherwise, spillage may or may not occur depending on the actual release decision. Similarly, if the sum of the inflow to the lower reservoir and the releasing capacity of the upper reservoir exceeds that of the lower reservoir (i.e., Assumption 2 is violated), some water will spill from the lower reservoir when a large amount of water is released from the upper reservoir. Otherwise, spillage may or may not occur depending on the actual release decision for the lower reservoir. By avoiding scenarios where spillage is mechanically determined by inflows alone, this assumption keeps the structural analysis in a setting where inflows remain decision-relevant. The operational decisions still dictate whether water is kept in the system or lost, making it meaningful to compare the value of storing, pumping, or releasing an additional unit of water. We note that Assumption 2 implies that $r_{ut} + r_{lt} \leq C_{Rl}$: the total inflow to the system is bounded by the releasing capacity of the lower reservoir.

The assumptions allow us to bound the optimal amounts of water to be pumped or released in each period.

LEMMA 1. *Under Assumptions 1 and 2, $x^r_{ut} - C_U \leq a^*_t(x_{ut}, x_{lt}, y_t)$, $\forall t \in \mathcal{T}$.*

Lemma 1 implies that the amount of water pumped to (or released from) the upper reservoir can be restricted, without loss of optimality, to values that prevent water spillage. This is because water spillage from the upper reservoir causes a loss of water in the system and provides no benefit.

LEMMA 2. *Under Assumptions 1 and 2, $a^*_t(x_{ut}, x_{lt}, y_t) - b^*_t(x_{ut}, x_{lt}, y_t) \leq C_L - x^r_{lt}$, $\forall t \in \mathcal{T}$.*

Lemma 2 implies that the amounts of water pumped or released can be restricted, without loss of optimality, to values that prevent water spillage from the lower reservoir. This is again because water spillage from the lower reservoir causes a loss of water in the system and brings no benefit.

We also derive several additional structural properties of the optimal value function:

PROPOSITION 2. *Under Assumptions 1 and 2, the following properties hold for $\alpha > 0$ and $\beta > 0$:*

(a) $v^*_t(x_{ut}, x_{lt} + \alpha, y_t) \leq v^*_t(x_{ut} + \alpha, x_{lt}, y_t)$, $\forall t$, *if* $C_{Ru} = C_{Rl}$.

(b) $v^*_t(x_{ut} + \alpha, x_{lt}, y_t) - v^*_t(x_{ut}, x_{lt} + \alpha, y_t) \leq v^*_t(x_{ut} + \alpha, x_{lt} + \beta, y_t) - v^*_t(x_{ut}, x_{lt} + \alpha + \beta, y_t)$, $\forall t$.

(c) $v^*_t(x_{ut} + \alpha + \beta, x_{lt}, y_t) - v^*_t(x_{ut} + \beta, x_{lt} + \alpha, y_t) \leq v^*_t(x_{ut} + \alpha, x_{lt}, y_t) - v^*_t(x_{ut}, x_{lt} + \alpha, y_t)$, $\forall t$.

(d) $v^*_t(x_{ut} + \alpha, x_{lt} + \beta, y_t) - v^*_t(x_{ut}, x_{lt} + \beta, y_t) \leq v^*_t(x_{ut} + \alpha, x_{lt}, y_t) - v^*_t(x_{ut}, x_{lt}, y_t)$, $\forall t$.

We discuss below the implications of Proposition 2:

- Point (a) says that the system becomes more profitable as the amount of water in the upper reservoir grows while the total amount of water in the system remains constant. This result holds when both reservoirs have the same releasing capacity (i.e., $C_{Ru} = C_{Rl}$). When $C_{Ru} > C_{Rl}$, Assumption 2 is violated. When $C_{Ru} < C_{Rl}$, the same result may not hold. Consider, for instance, a case with zero natural inflow and high electricity price, where releasing the maximum possible amount of water is optimal in the last period. Suppose that the total water in the system equals the combined releasing capacity of the two reservoirs, i.e., $x_{u(T-1)} + x_{l(T-1)} = C_{Ru} + C_{Rl}$. If the water is distributed in accordance with the reservoirs' releasing capacities, i.e., $x_{u(T-1)} = C_{Ru}$ and $x_{l(T-1)} = C_{Rl}$, water can be released at full capacity from both reservoirs. By contrast, if all the water is stored in the upper reservoir, i.e., $x_{u(T-1)} = C_{Ru} + C_{Rl}$ and $x_{l(T-1)} = 0$, the amount of water used for generation from the lower reservoir is limited by the upper reservoir's smaller releasing capacity. Thus, having more water in the upper reservoir when the total amount of water is fixed need not be more profitable.

- Point (b) says that releasing a certain amount of water from the upper reservoir becomes less desirable as the water level in the lower reservoir increases. To explain this result, suppose the controller avoids releasing water from the lower reservoir to minimize water loss in the system (as previously noted, the optimal value function increases with total water volume). In this case, a high water level in the lower reservoir has undesirable outcomes: Holding a large amount of water in the lower reservoir limits the ability to capitalize on high electricity prices in future periods, when releasing water from the upper reservoir to sell energy is beneficial. On the other hand, pumping a large amount of water from the lower reservoir entails buying a significant amount of energy from the market.
- Point (c) says that releasing a certain amount of water from the upper reservoir becomes more desirable as the upper-reservoir water level increases. Although retaining water in the upper reservoir may be advantageous in anticipation of high electricity prices in the future, releasing some water when the upper-reservoir water level is high provides immediate returns without substantially compromising potential future revenues.
- Point (d) says that having extra water in the upper reservoir is more desirable when the lower-reservoir water level is low. Similarly, having extra water in the lower reservoir is more desirable when the upper-reservoir water level is low. Thus, the stored energy in the two reservoirs acts as an operational substitute. These properties reflect the submodularity structure in x_{ut} and x_{lt} as characterized by Topkis (1998). We note that points (b) and (d) imply the concavity in x_{lt}, whereas points (c) and (d) imply the concavity in x_{ut}.

Economically, these properties show that the marginal value of stored water in one reservoir is interdependent with the state of the other. Driven by this substitutability and the diminishing marginal returns of energy storage, the operator's strategy naturally follows a state-dependent threshold rule: if the marginal value of stored water deviates significantly from the immediate market payoff, adjusting the water level via pumping or releasing is optimal; otherwise, preserving the current allocation is best. In an open-loop PHES facility, this internal allocation is coupled with a system-wide decision regarding the total amount of water retained, which fluctuates due to natural inflows and releases from the lower reservoir. This dynamic gives rise to the two state-dependent target levels defined below.

We introduce the optimal state-dependent target levels $S_t^{(\nu)}$ and $S_{ut}^{(\nu)}$ that are associated with the total amount of water in the system and the amount of water in the upper reservoir, respectively, where $\nu \in \{\mathsf{P}, \mathsf{R}\}$ denotes the decision type:

$$\left(S_t^{(\nu)}, S_{ut}^{(\nu)}\right) := \underset{(z_t, z_{ut}) \in \Phi(x_t^r)}{\arg\max} \left\{ \left(E^{(\nu)}(-z_{ut}) - \theta z_t\right) p_t + \mathbb{E}_{y_{t+1}|y_t} \left[v_{t+1}^* \left(z_{ut}, z_t - z_{ut}, y_{t+1}\right)\right] \right\}, \tag{2}$$

where $E^{(\nu)}(-z_{ut}) = -z_{ut}/\theta$ for $\nu = \mathsf{P}$ and $E^{(\nu)}(-z_{ut}) = -\theta z_{ut}$ for $\nu = \mathsf{R}$, $x_t^r := x_{ut}^r + x_{lt}^r$ denotes the total amount of water in the system upon observing the water runoff in period t, $z_t := x_t^r - b_t$ and $z_{ut} := x_{ut}^r - a_t$ denote the amounts of water in the system and the upper reservoir, respectively, at the end of period t if the action pair (a_t, b_t) is taken in period t, and $\Phi(x_t^r) := \left\{(z_t, z_{ut}) \in \mathbb{R}_+^2 \mid x_t^r - C_{Rl} \le z_t \le x_t^r,\ 0 \le z_{ut} \le C_U,\ z_{ut} \le z_t \le z_{ut} + C_L\right\}$. (Recall that x_{ut}^r and x_{lt}^r are the reservoir water levels upon observing the water runoff in period t.) Notice that z_t and z_{ut} are the possible values of the optimal state-dependent target levels $S_t^{(\nu)}$ and $S_{ut}^{(\nu)}$, respectively. Although the target levels $S_t^{(\nu)}$ and $S_{ut}^{(\nu)}$ depend on x_t^r and y_t, we often suppress these dependencies for notational simplicity. We also introduce the optimal state-dependent target levels $S_t(z_{ut})$ for the total amount of water in the system conditional on the water level in the upper reservoir:

$$S_t(z_{ut}) := \underset{z_t \in \left[\max\left\{z_{ut}, x_t^r - C_{Rl}\right\}, \min\left\{z_{ut} + C_L, x_t^r\right\}\right]}{\arg\max} \left\{ -\theta z_t p_t + \mathbb{E}_{y_{t+1}|y_t} \left[v_{t+1}^* \left(z_{ut}, z_t - z_{ut}, y_{t+1}\right)\right] \right\}. \tag{3}$$

We again suppress the dependency of $S_t(z_{ut})$ on x_t^r and y_t for notational simplicity. We note that $S_t\left(S_{ut}^{(\nu)}\right) = S_t^{(\nu)}$ for $\nu \in \{\mathsf{P}, \mathsf{R}\}$.

Leveraging our analytical results, we are now ready to state the main result of this section:

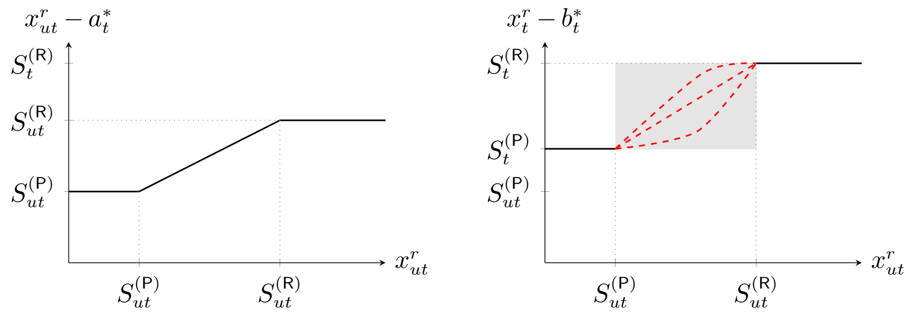
THEOREM 1. *Under Assumptions 1 and 2, the optimal energy generation and storage decisions in any period t can be specified as follows:*

$$a_t^*(x_{ut}, x_{lt}, y_t) = \begin{cases} -\min\left\{S_{ut}^{(\mathsf{P})} - x_{ut}^r, C_P\right\} & \text{if } x_{ut}^r \le S_{ut}^{(\mathsf{P})}, \\ 0 & \text{if } S_{ut}^{(\mathsf{P})} < x_{ut}^r \le S_{ut}^{(\mathsf{R})}, \\ \min\left\{x_{ut}^r - S_{ut}^{(\mathsf{R})}, C_{Ru}\right\} & \text{if } S_{ut}^{(\mathsf{R})} < x_{ut}^r, \end{cases}$$

and $b_t^*(x_{ut}, x_{lt}, y_t) = x_t^r - S_t(x_{ut}^r - a_t^*(x_{ut}, x_{lt}, y_t))$. The optimal state-dependent target levels obey: (i) $S_t(z_{ut})$ is nondecreasing in z_{ut} for $x_t^r - C_L \le z_{ut} \le x_t^r - C_{Rl}$ and (ii) $S_{ut}^{(\mathsf{P})} \le S_{ut}^{(\mathsf{R})}$ and $S_t^{(\mathsf{P})} \le S_t^{(\mathsf{R})}$.

To our knowledge, Theorem 1 provides the first characterization of the optimal policy structure for PHES facilities with natural inflow in the literature. The theorem has a direct interpretation in terms of the marginal value of stored water. If the water level in the upper reservoir is sufficiently high, i.e., $S_{ut}^{(\mathsf{R})} < x_{ut}^r$, the immediate generation revenue exceeds the future value of keeping this water in the upper reservoir, and thus

Figure 3 The optimal target levels vs. upper-reservoir water level when the system water level is fixed.



Note. The target levels $S^{(\nu)}_{ut}$ are assumed to be accessible and between $x^r_t - C_L$ and $x^r_t - C_{Rl}$. The function in the gray region is nondecreasing and depends on x^r_t and y_t.

it is optimal to release water until the level is as close as possible to $S^{(\mathsf{R})}_{ut}$. If the water level is sufficiently low, i.e., $x^r_{ut} \le S^{(\mathsf{P})}_{ut}$, the future value of additional water in the upper reservoir justifies pumping, and thus it is optimal to bring the water level as close as possible to $S^{(\mathsf{P})}_{ut}$. If the water level is between these two targets, i.e., $S^{(\mathsf{P})}_{ut} < x^r_{ut} \le S^{(\mathsf{R})}_{ut}$, it is optimal to keep it unchanged. Thus, the no-action region reflects states where neither immediate generation revenue nor the future value of increasing upper-reservoir water justifies adjusting the allocation.

The resulting optimal action for the upper reservoir is incorporated into the target level calculation for the total amount of water in the system. This target has a distinct role in the open-loop setting: because natural inflows and water release from the lower reservoir affect the total amount of water remaining in the facility, the operator must determine not only where water should be located but also how much water should remain in the system. Conditional on the total amount of water x^r_t and the exogenous state y_t, the target levels do not depend separately on the initial allocation of water across reservoirs. System inefficiencies lead to different marginal payoffs in the pumping and releasing modes. Therefore, the target levels for both the upper-reservoir water level and the system water level are higher in the releasing mode than in the pumping mode. See Figure 3 for an illustration.

Certain modeling assumptions used in our structural analysis are more readily relaxed than others. For example, different but constant heads across reservoirs, or different but constant efficiencies across the pumping and generation modes, can be incorporated by replacing the common efficiency parameter θ with reservoir- and mode-specific conversion coefficients. Under the natural condition that the energy required to pump one unit of water is at least as large as the energy recovered from releasing it, the concavity argument underlying the threshold-policy characterization remains valid. By contrast, variable-head systems would make the conversion rates state-dependent and could therefore disrupt the concavity and submodularity results. Similarly, relaxing the price-taking assumption would require modeling the operator's price impact through a residual demand or price-impact function (e.g., Liu et al. 2022b,a). In such settings, the threshold structure developed in this paper provides a natural starting point for approximation algorithms, though establishing formal optimality would require additional assumptions and new analytical work.

4. Heuristic Algorithm Development

Threshold policies are often easy to implement and, when informed by structural insights, can deliver performance comparable to that of optimal policies (e.g., Nadar et al. 2016, Zhou et al. 2019, and Karakoyun et al. 2026). Theorem 1 establishes the optimality of a state-dependent threshold policy when the electricity price is always positive. We now develop a heuristic solution algorithm by implementing the policy structure in Theorem 1 into a backward induction algorithm for the more general problem with possibly negative prices. Specifically, we exploit the policy structure in positive-price states and use an approximation in negative-price states. We present below two variants of our heuristic approach for the discrete-state and discrete-action version of our MDP. For each exogenous state, the first variant in Section 4.1 computes target levels conditioned on the total amount of water in the system, whereas the second variant in Section 4.2 derives these targets independent of the total water. We define $\mathcal{X}_u$, $\mathcal{X}_l$, $\mathcal{X}$, and $\mathcal{Y}$ as the discrete spaces of x_{ut}, x_{lt}, x_t^r, and y_t, respectively.

The time series models in Sections 5 formulate the stochastic component of each exogenous state variable as an AR(1) process. This allows us to reduce the computational burden of our MDP by redefining the exogenous state tuple in period t as $y_t = (f_t, \rho_t, j_t) \in \mathcal{Y}$, where f_t is the streamflow rate, ρ_t is the mean-reverting component of the price, and j_t is the spike component of the price. Although the spike component of the price is time-independent, we include it in the tuple y_t for calculation of the effective price in period t. For notational convenience, we also define $\bar{y}_t = (f_t, \rho_t) \in \overline{\mathcal{Y}}$ as the exogenous partial-state tuple in period t, where $\overline{\mathcal{Y}}$ is the discrete space of the exogenous state variables without the spike. To improve the algorithmic efficiency, we first calculate the value function $\bar{v}_{t+1}^{\mathsf{M}}(x_{u(t+1)}, x_{l(t+1)}, \bar{y}_{t+1}) := \mathbb{E}_{j_{t+1}}\left[v_{t+1}^{\mathsf{M}}(x_{u(t+1)}, x_{l(t+1)}, y_{t+1})\right]$ in state $(x_{u(t+1)}, x_{l(t+1)}, \bar{y}_{t+1})$ in period $t+1$, where M denotes the solution method implemented. We then calculate the action pair in state (x_{ut}, x_{lt}, y_t) in period t, using the function $\bar{v}_{t+1}^{\mathsf{M}}(x_{u(t+1)}, x_{l(t+1)}, \bar{y}_{t+1})$ and its expectation taken with respect to $\bar{y}_{t+1}|\bar{y}_t$.

4.1. Policy Approximation

We construct Algorithm 1 to implement this method that we call PA (the initials of policy approximation): To differentiate the quantities associated with the PA method from those of the optimal policy, we use the superscript PA. For example, $v_t^{\mathsf{PA}}(x_{ut}, x_{lt}, y_t)$ denotes the value function in state (x_{ut}, x_{lt}, y_t) in period $t \in \mathcal{T}$ calculated with this method. In positive-price states, the action a_t^{PA} is determined by state-dependent target levels $S_{ut}^{(\nu),\mathsf{PA}}$. See step 13 of Algorithm 1. We incorporate the property that the state-dependent target levels remain the same as long as the total amount of water in the PHES facility stays constant. See steps 6–8 of Algorithm 1. In positive-price states, the action b_t^{PA} is determined by state-dependent target levels $S_t^{(\nu),\mathsf{PA}}$ if

Algorithm 1 Policy approximation.

1: $\bar{v}_T^{\mathsf{PA}}(x_{uT}, x_{lT}, \bar{y}_T) \leftarrow 0, \forall (x_{uT}, x_{lT}, \bar{y}_T) \in \mathcal{X}_u \times \mathcal{X}_l \times \overline{\mathcal{Y}}$.
2: **for** $t = T-1, \ldots, 1$ **do**
3:   **for** $y_t \in \mathcal{Y}$ **do**
4:     **for** $x_t^r \in \mathcal{X}$ **do**
5:       $\left(S_t^{(\nu)}, S_{ut}^{(\nu)}\right) \leftarrow \arg\max_{(z_t, z_{ut}) \in \Phi(x_t^r)} \left\{ \left(E^{(\nu)}(-z_{ut}) - \theta z_t\right) p_t + \mathbb{E}_{\bar{y}_{t+1}|\bar{y}_t} \left[\bar{v}_{t+1}^{\mathsf{PA}}(z_{ut}, z_t - z_{ut}, \bar{y}_{t+1})\right] \right\}, \forall \nu$.
6:       **for** $(x_{ut}, x_{lt}) \in \mathcal{X}_u \times \mathcal{X}_l$ such that $x_t^r = x_{ut}^r + x_{lt}^r$ **do**
7:         $S_t^{(\nu),\mathsf{PA}}(x_{ut}, x_{lt}, y_t) \leftarrow S_t^{(\nu)}$ and $S_{ut}^{(\nu),\mathsf{PA}}(x_{ut}, x_{lt}, y_t) \leftarrow S_{ut}^{(\nu)}, \forall \nu$.
8:       **end for**
9:     **end for**
10:   **end for**
11:   **for** $(x_{ut}, x_{lt}, y_t) \in \mathcal{X}_u \times \mathcal{X}_l \times \mathcal{Y}$ **do**
12:     **if** $p_t > 0$ **then** ▷ The price is positive.
13:       Compute a_t^{PA} from Theorem 1 with $S_{ut}^{(\nu),\mathsf{PA}}(x_{ut}, x_{lt}, y_t), \forall \nu$.
14:       **if** $x_{ut}^r - a_t^{\mathsf{PA}} = S_{ut}^{(\nu),\mathsf{PA}}(x_{ut}, x_{lt}, y_t)$ for some ν **then**
15:         $b_t^{\mathsf{PA}} \leftarrow x_t^r - S_t^{(\nu),\mathsf{PA}}(x_{ut}, x_{lt}, y_t)$.
16:       **else**
17:         $S_t^{\mathsf{PA}}(x_{ut}^r - a_t^{\mathsf{PA}}) \leftarrow \arg\max_{z_t \in \mathcal{Z}} \left\{-\theta z_t p_t + \mathbb{E}_{\bar{y}_{t+1}|\bar{y}_t}[\bar{v}_{t+1}^{\mathsf{PA}}(x_{ut}^r - a_t^{\mathsf{PA}}, z_t - x_{ut}^r + a_t^{\mathsf{PA}}, \bar{y}_{t+1})]\right\}$, where $\mathcal{Z} = \left[\max\left\{x_{ut}^r - a_t^{\mathsf{PA}}, x_t^r - C_{Rl}\right\}, \min\left\{x_{ut}^r - a_t^{\mathsf{PA}} + C_L, x_t^r\right\}\right]$.
18:         $b_t^{\mathsf{PA}} \leftarrow x_t^r - S_t^{\mathsf{PA}}(x_{ut}^r - a_t^{\mathsf{PA}})$.
19:       **end if**
20:     **else** ▷ The price is negative.
21:       $a_t^{\mathsf{PA}} \leftarrow \begin{cases} -\min\left\{S_{ut}^{(\mathsf{P}),\mathsf{PA}} - x_{ut}^r, C_P\right\} & \text{if } x_{ut}^r \le Z_t, \\ \min\left\{x_{ut}^r - S_{ut}^{(\mathsf{R}),\mathsf{PA}}, C_{Ru}\right\} & \text{if } Z_t < x_{ut}^r, \end{cases}$ and $b_t^{\mathsf{PA}} \leftarrow 0$.
22:     **end if**
23:     $v_t^{\mathsf{PA}}(x_{ut}, x_{lt}, y_t) \leftarrow R(a_t^{\mathsf{PA}}, b_t^{\mathsf{PA}}, y_t) + \mathbb{E}_{\bar{y}_{t+1}|\bar{y}_t}\left[\bar{v}_{t+1}^{\mathsf{PA}}(x_{u(t+1)}, x_{l(t+1)}, \bar{y}_{t+1})\right]$.
24:   **end for**
25:   $\bar{v}_t^{\mathsf{PA}}(x_{ut}, x_{lt}, \bar{y}_t) \leftarrow \mathbb{E}_{j_t}\left[v_t^{\mathsf{PA}}(x_{ut}, x_{lt}, y_t)\right], \forall (x_{ut}, x_{lt}, \bar{y}_t) \in \mathcal{X}_u \times \mathcal{X}_l \times \overline{\mathcal{Y}}$.
26: **end for**

$S_{ut}^{(\nu),\mathsf{PA}}$ is physically accessible. See steps 14–15 of Algorithm 1. Otherwise, it is derived from the realized upper-reservoir water level. See steps 16–18 of Algorithm 1.

In negative-price states, the action a_t^{PA} is determined by step 21 of Algorithm 1, and $b_t^{\mathsf{PA}} = 0$ (i.e., release no water from the lower reservoir). The target level Z_t in step 21 is the maximal inventory level between $S_{ut}^{(\mathsf{R}),\mathsf{PA}}$ and $S_{ut}^{(\mathsf{P}),\mathsf{PA}}$ such that the producer is indifferent between action pairs $(-\min\{S_{ut}^{(\mathsf{P}),\mathsf{PA}} - x_{ut}^r, C_P\}, 0)$ and $(\min\{x_{ut}^r - S_{ut}^{(\mathsf{R}),\mathsf{PA}}, C_{Ru}\}, 0)$. This negative-price rule extends the threshold policy of Theorem 1 to states in which the concavity argument used to prove the theorem may no longer apply. It can be shown that our PA method would be optimal if $v_t^{\mathsf{PA}}(x_{ut}, x_{lt}, y_t)$ were jointly concave in $(x_{ut}, x_{lt}), \forall t \in \mathcal{T}$.

Our PA method accelerates the standard DP algorithm, without loss of optimality, if the price is always positive. For our PA method, in each period, the number of states in which the target levels are computed

Algorithm 2 Fixed-threshold policy approximation.

```
v̄_T^FPA(x_uT, x_lT, ȳ_T) ← 0, ∀(x_uT, x_lT, ȳ_T) ∈ 𝒳_u × 𝒳_l × 𝒴̄.
for t = T − 1, ..., 1 do
    for y_t ∈ 𝒴 do
        if p_t > 0 then                                                   ▷ The price is positive.
            (S_t^(ν),FPA, S_ut^(ν),FPA) ← arg max_{(z_t, z_ut)∈Φ} { (E^(ν)(−z_ut) − θz_t)p_t + 𝔼_{ȳ_{t+1}|ȳ_t} [v̄_{t+1}^FPA(z_ut, z_t − z_ut, ȳ_{t+1})] }, ∀ν.
            for (x_ut, x_lt) ∈ 𝒳_u × 𝒳_l do
                Compute a_t^FPA from Theorem 1 with S_ut^(ν) replaced by S_ut^(ν),FPA, ∀ν.
                If a_t^FPA ≥ 0, then S_t^FPA ← S_t^(R),FPA. Else, S_t^FPA ← S_t^(P),FPA.
                If x_t^r − S_t^FPA ≥ 0, then b_t^FPA ← min{x_t^r − S_t^FPA, x_lt^r + a_t^FPA, C_Rl}. Else, b_t^FPA ← 0.
                v_t^FPA(x_ut, x_lt, y_t) ← R(a_t^FPA, b_t^FPA, y_t) + 𝔼_{ȳ_{t+1}|ȳ_t} [v̄_{t+1}^FPA(x_u(t+1), x_l(t+1), ȳ_{t+1})].
            end for
        else                                                              ▷ The price is negative.
            for (x_ut, x_lt) ∈ 𝒳_u × 𝒳_l do
                a_t^FPA ← − min{x_lt^r, C_P, C_U − x_ut^r} and b_t^FPA ← 0.
                v_t^FPA(x_ut, x_lt, y_t) ← R(a_t^FPA, b_t^FPA, y_t) + 𝔼_{ȳ_{t+1}|ȳ_t} [v̄_{t+1}^FPA(x_u(t+1), x_l(t+1), ȳ_{t+1})].
            end for
        end if
    end for
    v̄_t^FPA(x_ut, x_lt, ȳ_t) ← 𝔼_{j_t} [v_t^FPA(x_ut, x_lt, y_t)], ∀(x_ut, x_lt, ȳ_t) ∈ 𝒳_u × 𝒳_l × 𝒴̄.
end for
```

is of order $O((|\mathcal{X}_u| + |\mathcal{X}_l|)|\mathcal{Y}|)$. In each state, the number of action pairs that should be considered for target-level computation is of order $O(|\mathcal{X}_u||\mathcal{X}_l|)$. Therefore, the total number of operations required for an exhaustive search over the action pairs is of order $O(T(|\mathcal{X}_u| + |\mathcal{X}_l|)|\mathcal{X}_u||\mathcal{X}_l||\mathcal{Y}|)$. For the standard DP algorithm, however, the number of states in each period is of order $O(|\mathcal{X}_u||\mathcal{X}_l||\mathcal{Y}|)$ and the number of action pairs in each state is of order $O(|\mathcal{X}_u||\mathcal{X}_l|)$. Hence, the total number of operations required for an exhaustive search is of order $O(T|\mathcal{X}_u|^2|\mathcal{X}_l|^2|\mathcal{Y}|)$. Consequently, our PA method has a distinct computational benefit over the optimal algorithm.

4.2. Fixed-Threshold Policy Approximation

Our PA method calculates the target levels $S_t^{(\nu),\mathsf{PA}}$ and $S_{ut}^{(\nu),\mathsf{PA}}$ in period t for each possible value of the total amount of water in the PHES facility. These extensive calculations are based on equation (2) that ensures the feasibility of the resulting decisions for the amount of water to be released from the lower reservoir. However, the feasibility of the action pairs can also be established after the target levels are calculated by ignoring the detailed state information. We now take this alternative route and consider a fixed-threshold version of our PA method that we call FPA. We construct Algorithm 2 to implement this method. Our FPA method calculates the target levels from equation (2) with $\Phi(x_t^r)$ replaced by $\Phi := \{z_t, z_{ut} \in \mathbb{R}^2 \mid 0 \le z_{ut} \le C_U, z_{ut} \le z_t \le z_{ut} + C_L\}$, denotes these target levels by $S_t^{(\nu),\mathsf{FPA}}(y_t)$ and $S_{ut}^{(\nu),\mathsf{FPA}}(y_t)$, and uses these target levels to determine the action pairs in positive-price states. The target levels $S_t^{(\nu),\mathsf{FPA}}(y_t)$ and $S_{ut}^{(\nu),\mathsf{FPA}}(y_t)$ no longer depend on the total amount of water in the PHES facility; only one pair of target levels is computed for each exogenous state tuple. The action a_t^{FPA} is computed from Theorem 1 with $S_{ut}^{(\nu),\mathsf{FPA}}(y_t)$ by taking into account the constraint that $a_t^{\mathsf{FPA}} \geq -x_{lt}^r$, which prevents pumping more water than the amount available in the lower reservoir. The target level S_t^{FPA} is set to $S_t^{(\mathsf{R}),\mathsf{FPA}}(y_t)$ if water is released from the upper reservoir (i.e., $a_t^{\mathsf{FPA}} \geq 0$), and it is set to $S_t^{(\mathsf{P}),\mathsf{FPA}}(y_t)$ otherwise. The action b_t^{FPA} is calculated as follows:

$$b_t^{\mathsf{FPA}} = \begin{cases} \min\{x_t^r - S_t^{\mathsf{FPA}}, x_{lt}^r + a_t^{\mathsf{FPA}}, C_{Rl}\} & \text{if } x_t^r - S_t^{\mathsf{FPA}} \geq 0, \\ 0 & \text{if } x_t^r - S_t^{\mathsf{FPA}} < 0. \end{cases}$$

If the total amount of water in the system is sufficiently high (i.e., $x_t^r \geq S_t^{\mathsf{FPA}}$), water is released from the lower reservoir until the total amount is as close as possible to S_t^{FPA}. Otherwise (i.e., $x_t^r < S_t^{\mathsf{FPA}}$), no water is released from the lower reservoir. In negative-price states, the action pairs $(a_t^{\mathsf{PA}}, b_t^{\mathsf{PA}})$ are obtained from the myopically optimal decisions. See step 14 of Algorithm 2. In these states, $a_t^{\mathsf{PA}} = -\min\{x_{lt}^r, C_P, C_U - x_{ut}^r\}$ (i.e., pump as much water as possible to the upper reservoir without causing any water spillage) and $b_t^{\mathsf{PA}} = 0$ (i.e., release no water from the lower reservoir). We choose to impose the myopic actions in negative-price states because the myopic actions are immediately obtained and are potentially optimal in many negative-price states. A similar approach for handling negative prices is presented in Karakoyun et al. (2026).

For our FPA method, in each period, the number of states in which the target levels are computed is of order $O(|\mathcal{Y}|)$. In each state, the number of action pairs considered for target-level computation is of order $O(|\mathcal{X}_u||\mathcal{X}_l|)$. To speed up the target-level computation, we take a coordinate descent-type approach: We first compute $S_{ut}^{(\nu),\mathsf{FPA}}(y_t)$ from equation (2) by setting the total amount of water in the PHES facility z_t to the upper reservoir capacity level C_U. We then compute $S_t^{(\nu),\mathsf{FPA}}(y_t)$ from equation (3) with z_{ut} replaced by $S_{ut}^{(\nu),\mathsf{FPA}}(y_t)$. We also ignore the dependence on x_t^r in this calculation step. With this approach, in each state, the number of action pairs considered for target-level computation is of order $O(|\mathcal{X}_u| + |\mathcal{X}_l|)$. Therefore, the total number of operations required for an exhaustive search over the action pairs is of order $O(T(|\mathcal{X}_u| + |\mathcal{X}_l|)|\mathcal{Y}|)$. Our FPA method thus provides a further substantial computational benefit.

In summary, both methods avoid an exhaustive search over feasible action pairs across all states by instead searching for target levels derived from the policy structure in Theorem 1. Our PA method conditions these target levels on the total amount of water in the PHES facility and thus stays closer to the optimal policy, whereas our FPA method omits this conditioning to prioritize computational speed.

5. Time Series Models and Discretization

To evaluate the proposed algorithms under realistic conditions, this section introduces the time series models and discretization methods used to generate our data-calibrated instances. The time series models that we consider to forecast the electricity price and streamflow rate utilize the historical data available from the State of New York (Sections 5.1 and 5.2). We embed these time series models into our MDP via exogenous state variables. We discretize the continuous space of these variables for computational purposes (Section 5.3). We set the period length to be one hour in our experiments; we use t as the index for one-hour length periods in the remainder of the paper. This assumption is common in the PHES literature; see, for example, Brown et al. (2008), Connolly et al. (2011), Duque et al. (2011), Kusakana (2016), Kocaman and Modi (2017), and Jurasz et al. (2018).

5.1. Time Series Model for the Electricity Price

We consider the electricity price data available for Albany in which the price has been recorded every five minutes between 2007 and 2019. For this data set from NYISO (2020), the average, median, minimum, and maximum values are \$45.14, \$34.09, $-$\$3678.02, and \$3393.33, per MWh, respectively. We define $\bar{t}$ as the index for five-minute length periods and $\overline{\mathcal{T}}$ as the set of these periods.

We construct our five-minute price model by adopting the iterative time series modeling approach of Zhou et al. (2019). This approach decomposes the time series into the components of seasonality ($s_{\bar{t}}$), mean reversion ($\rho_{\bar{t}}$), and spike ($j_{\bar{t}}$), by accounting for negative-price occurrences. Specifically, this approach sets the initial values of the estimated spikes $\{\hat{j}_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$ to zero and takes the following four steps in each iteration performed until the parameter estimates of the model converge: The first step calculates the estimated despiked prices $\{\hat{p}'_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$ by subtracting the estimated spikes from the observed prices (i.e., $\hat{p}'_{\bar{t}} = p_{\bar{t}} - \hat{j}_{\bar{t}}$). The second step accommodates *negative prices* by applying an inverse hyperbolic transformation to the despiked prices; the transformed estimated despiked price in period $\bar{t}$ is $\sinh^{-1}(\hat{p}'_{\bar{t}}/\ell)$, where $\sinh^{-1}$ is the inverse hyperbolic sine function and ℓ is the scale parameter that is chosen to be 30. It then eliminates the *seasonality* effect by deseasonalizing the transformed despiked prices. To obtain the estimated seasonality model $\{\hat{s}_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$, we fit the following linear regression to the transformed despiked prices:

$$s_{\bar{t}} = \gamma_1^{(p)} + \sum_{i=1}^{11} \gamma_{2i}^{(p)} D_t^{2i} + \sum_{j=1}^{6} \gamma_{3j}^{(p)} D_t^{3j} + \sum_{h=1}^{23} \gamma_{4h}^{(p)} D_t^{4h}, \ \forall \bar{t}, \tag{4}$$

where $\gamma_1^{(p)}$ is a constant and $\gamma_{2i}^{(p)}$, $\gamma_{3j}^{(p)}$, and $\gamma_{4h}^{(p)}$ are the respective coefficients of the dummy variables D_t^{2i}, D_t^{3j}, and D_t^{4h}, that are equal to one if the corresponding hourly period $t = \lceil \bar{t}/12 \rceil$ is in month i, week day j, and hour h, respectively, and zero otherwise. It calculates the estimated mean-reverting component $\{\hat{\rho}_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$ by removing the estimated seasonality effect from the transformed despiked prices (i.e., $\hat{\rho}_{\bar{t}} = \sinh^{-1}(\hat{p}'_{\bar{t}}/\ell) - \hat{s}_{\bar{t}}$). Table 1 exhibits the parameter estimates of the seasonality model. The third step captures the *mean-reverting* behavior via an AR(1) process given by

$$\rho_{\bar{t}} = \left(1 - \kappa^{(\bar{p})}\right) \rho_{\bar{t}-1} + \sigma^{(\bar{p})} \epsilon_{\bar{t}}, \ \forall \bar{t}, \tag{5}$$

where $\kappa^{(\bar{p})}$ is the speed of mean reversion, $\sigma^{(\bar{p})}$ is the volatility of white noise, and $\epsilon_{\bar{t}}$ is the white noise term. The error terms $\{\epsilon_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$ are independent standard normal random variables. The parameter estimates of this AR(1) process are $\hat{\kappa}^{(\bar{p})} = 0.1360$ and $\hat{\sigma}^{(\bar{p})} = 0.1687$. The last step identifies the

Table 1 Parameter estimates of the despiked price seasonality model.

$\hat{\gamma}_1^{(p)}$	0.9652											
i	1	2	3	4	5	6	7	8	9	10	11	
$\hat{\gamma}_{2i}^{(p)}(\cdot 10^{-3})$	222.0	143.1	-26.5	-98.2	-174.0	-164.6	-68.8	-83.4	-134.5	-153.1	-85.7	
j	1	2	3	4	5	6						
$\hat{\gamma}_{3j}^{(p)}(\cdot 10^{-3})$	-6.4	-10.9	-11.2	-11.6	-1.4	2.4						
h	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{\gamma}_{4h}^{(p)}(\cdot 10^{-3})$	-34.4	-73.4	-84.0	-78.3	-62.3	1.7	55.1	91.8	134.0	171.2	189.1	179.0
h	13	14	15	16	17	18	19	20	21	22	23	
$\hat{\gamma}_{4h}^{(p)}(\cdot 10^{-3})$	177.8	160.5	154.0	176.3	234.2	223.8	214.3	183.5	120.3	57.8	-4.6	

spikes for each $\bar{t} \in \overline{\mathcal{T}}$: The price $p_{\bar{t}+1}$ contains a spike if $\hat{j}_{\bar{t}+1}$ is zero and $\left|\hat{p}'_{\bar{t}+1} - \mathbb{E}[p'_{\bar{t}+1}|\hat{\rho}_{\bar{t}}]\right|$ is no less than 50, where $p'_{\bar{t}+1}|\rho_{\bar{t}}$ is assumed to follow a Johnson SU distribution (Johnson 1949) with mean $-\ell \exp\left[0.5(\sigma^{(\bar{p})})^2\right] \sinh\left[-\rho_{\bar{t}}(1-\kappa^{(\bar{p})}) - s_{\bar{t}+1}\right]$. If the price $p_{\bar{t}+1}$ contains a spike, $\hat{j}_{\bar{t}+1}$ is updated to $\hat{p}'_{\bar{t}+1} - \mathbb{E}[p'_{\bar{t}+1}|\hat{\rho}_{\bar{t}}]$. Finally, $\hat{p}'_{\bar{t}+1}$ is replaced with $\mathbb{E}[p'_{\bar{t}+1}|\hat{\rho}_{\bar{t}}]$ and $\hat{\rho}_{\bar{t}+1}$ is replaced with $\sinh^{-1}(\mathbb{E}[p'_{\bar{t}+1}|\hat{\rho}_{\bar{t}}]/\ell) - s_{\bar{t}+1}$. With our model calibration, the mean absolute error (MAE) is only \$4.75/MWh for the despiked price.

As we assume hourly periods for our MDP, we convert the five-minute price model to an hourly price model that again consists of the components of seasonality (s_t), mean reversion (ρ_t), and spike (j_t): We estimate the hourly seasonality effect, s_t, from the right hand side of the equation in (4). We model the hourly mean-reverting component, ρ_t, as the AR(1) process:

$$\rho_t = \left(1-\kappa^{(p)}\right)\rho_{t-1} + \sigma^{(p)}\epsilon_t, \tag{6}$$

where $\kappa^{(p)}$ is the speed of mean reversion and $\sigma^{(p)}$ is the volatility of white noise. We obtain these parameters by recursively iterating the equation in (5):

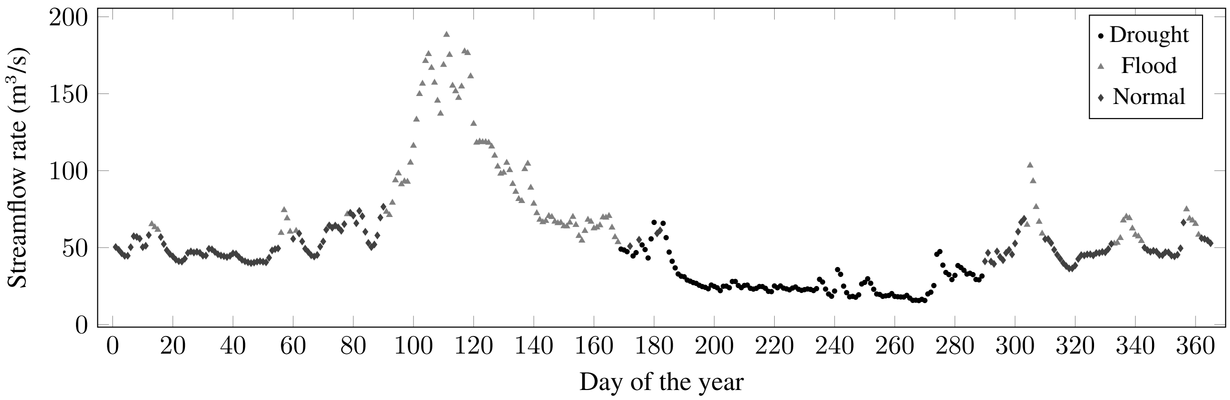
$$\rho_{\bar{t}+12} = \left(1-\kappa^{(\bar{p})}\right)^{12}\rho_{\bar{t}} + \sum_{\eta=1}^{12}\left(1-\kappa^{(\bar{p})}\right)^{12-\eta}\sigma^{(\bar{p})}\epsilon_{\bar{t}+\eta}. \tag{7}$$

Equating the mean and standard deviation of the right-hand side of (7) with those of (6) yields $\kappa^{(p)} = 1-\left(1-\kappa^{(\bar{p})}\right)^{12}$ and $\sigma^{(p)} = \sigma^{(\bar{p})}\sqrt{\left[1-\left(1-\kappa^{(\bar{p})}\right)^{2\times 12}\right]/\left[1-\left(1-\kappa^{(\bar{p})}\right)^{2}\right]}$, where the parameter estimates are $\hat{\kappa}^{(p)} = 0.827$ and $\hat{\sigma}^{(p)} = 0.330$. Lastly, the hourly spike component, j_t, is governed by a compound Bernoulli process. The spike occurrence probability is λ and the spike size follows an empirical distribution. The probability estimate $\hat{\lambda}$ is the ratio of the number of identified spikes to the number of five-minute length periods. The empirical distribution is based on the estimated spikes $\{\hat{j}_{\bar{t}}\}_{\bar{t}\in\overline{\mathcal{T}}}$ in the five-minute price model.

5.2. Time Series Model for the Streamflow Rate

We consider the streamflow data available for the Hudson River at North Creek in which the streamflow rate has been recorded every fifteen minutes between 2007 and 2019. For this data set from USGS (2020), the average, median, minimum, and maximum values are 55.46, 39.01, 2.18, and 988.25, in m^3/s, respectively. We redefine $\bar{t}$ as the index for fifteen-minute length periods, and define $f_{\bar{t}}$ as the streamflow rate in period $\bar{t}$.

Figure 4 Daily streamflow rates averaged over the years 2007-2019.



We construct our fifteen-minute streamflow model by adopting the periodic autoregressive (PAR) modeling approach of Toufani et al. (2022a): We fit the PAR model to our streamflow data by partitioning the 365 days of the year into three disjoint clusters with the fuzzy clustering method of Wang et al. (2006). Building on the log-transformation of daily average streamflow rates and the autocorrelation values at different lag times (from one day to ten days), the three clusters found correspond to the 'normal', 'drought', and 'flood' flow conditions that can be observed throughout the year. See Figure 4 for these clusters. We model the streamflow rate in each cluster $i \in \{$normal, drought, flood$\}$ as a different AR(1) process:

$$f_{\bar{t}} = \delta_i^{(\bar{r})} + \phi_i^{(\bar{r})} f_{\bar{t}-1} + \sigma_i^{(\bar{r})} \epsilon_{\bar{t}},\ \forall \bar{t} \in \overline{\mathcal{T}}_i, \tag{8}$$

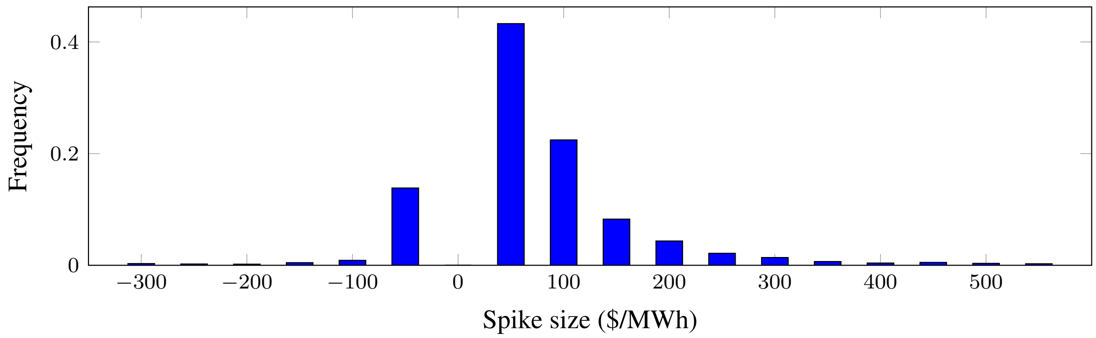
where $\delta_i^{(\bar{r})}$ is a constant, $\phi_i^{(\bar{r})}$ is the autoregressive coefficient, $\sigma_i^{(\bar{r})}$ is the volatility of white noise, and $\overline{\mathcal{T}}_i$ is the set of fifteen-minute length periods that belong to cluster i. With our model calibration, the MAE is only 0.35 m^3/s for the streamflow rate.

The frequency spectrum analysis indicates the significance of daily patterns, while there is no substantial fluctuation within each day. We thus convert the above fifteen-minute model to a daily model. We define $\underline{t}$ as the index for daily periods and $f_{\underline{t}}$ as the streamflow rate in period $\underline{t}$. For the daily streamflow model, we formulate the AR(1) process in each cluster $i \in \{$normal, drought, flood$\}$ as follows:

$$f_{\underline{t}} = \delta_i^{(\underline{r})} + \phi_i^{(\underline{r})} f_{\underline{t}-1} + \sigma_i^{(\underline{r})} \epsilon_{\underline{t}},\ \forall \underline{t} \in \underline{\mathcal{T}}_i, \tag{9}$$

where $\delta_i^{(\underline{r})}$ is a constant, $\phi_i^{(\underline{r})}$ is the autoregressive coefficient, $\sigma_i^{(\underline{r})}$ is the volatility of white noise, and $\underline{\mathcal{T}}_i$ is the set of daily periods in cluster i. Following similar steps to those in Section 5.1, we obtain $\delta_i^{(\underline{r})} = \delta_i^{(\bar{r})}$, $\phi_i^{(\underline{r})} = \left(\phi_i^{(\bar{r})}\right)^{96}$, and $\sigma_i^{(\underline{r})} = \sigma_i^{(\bar{r})} \sqrt{\left[1 - \left(\phi_i^{(\bar{r})}\right)^{2\times 96}\right] / \left[1 - \left(\phi_i^{(\bar{r})}\right)^2\right]}$, by iterating the expression in (8) recursively and equating the mean and standard deviation of the right hand side of the resulting expression with those of (9). The parameter estimates are $\hat{\delta}_n^{(\underline{r})} = 49.97$, $\hat{\phi}_n^{(\underline{r})} = 0.900$, $\hat{\sigma}_n^{(\underline{r})} = 14.17$, $\hat{\delta}_d^{(\underline{r})} = 28.25$, $\hat{\phi}_d^{(\underline{r})} = 0.885$, $\hat{\sigma}_d^{(\underline{r})} = 14.27$, $\hat{\delta}_f^{(\underline{r})} = 92.03$, $\hat{\phi}_f^{(\underline{r})} = 0.979$, and $\hat{\sigma}_f^{(\underline{r})} = 17.86$. As we assume hourly periods for our MDP, we incorporate the daily model by restricting the intraday hourly streamflow rates to stay constant.

Figure 5 Empirical distribution of the spikes.



We note that the fifteen-minute model can also be converted to an hourly model. However, the volatility of white noise is too low for hourly streamflow rates so that our discrete-state approximations for the AR(1) processes (required for numerical calculations) would induce deterministic streamflow rates. Also, the accuracy of our PAR model can be improved via higher-order autoregressive processes. Such a relaxation can reduce the MAE by at most a few percent for our streamflow data. Since higher-order autoregressive processes lead to an exponential growth of the state space in our MDP, we choose to implement the AR(1) process into our PAR model. Finally, our further investigations have shown that fitting the PAR model to our data with a different number of clusters (other than three) provides no significant improvement in accuracy.

5.3. Discretization for the Numerical Study

For our numerical study, we now provide discrete-state approximations for the continuous-state AR(1) processes embedded in the tuple $\bar{y}_t = (f_t, \rho_t)$. For the electricity price, we characterize the hourly AR(1) process as a finite-state Markov chain via the trinomial lattice method of Hull and White (1994). We consider a three-hour trinomial lattice by following Hull and White (1994) and Jaillet et al. (2004). This leads to a Markov chain with the state space $\mathcal{P} := \{-0.57, 0, 0.57\}$. The transition matrix of this Markov chain is

$$\begin{array}{c} \\ {\scriptstyle -0.57} \\ {\scriptstyle 0.0} \\ {\scriptstyle 0.57} \end{array} \begin{array}{c} \begin{array}{ccc} {\scriptstyle -0.57} & {\scriptstyle 0.0} & {\scriptstyle 0.57} \end{array} \\ \begin{bmatrix} .268 & .637 & .095 \\ .167 & .667 & .167 \\ .095 & .637 & .268 \end{bmatrix} \end{array}.$$

On the other hand, the spikes can take values from the set $\mathcal{J} := \{-300, -250, \ldots, 550\}$ and the spike occurrence probability is 9.3% in each period. Figure 5 illustrates the empirical distribution of the spikes.

For the streamflow rate, after we remove the lowest and highest ten percent of the data in each cluster, the streamflow rates vary between 21 and 88 in the normal flow cluster, between 8 and 57 in the drought flow cluster, and between 24 and 196 in the flood flow cluster. We thus restrict the streamflow rate in our MDP to values from the sets $\mathcal{R}^n := \{20, 40, 60, 80\}$, $\mathcal{R}^d := \{10, 30, 50\}$, and $\mathcal{R}^f := \{30, 50, \ldots, 190\}$ in these three clusters, respectively. We characterize the daily AR(1) process in each cluster as a different Markov chain and follow Tauchen (1986) to calculate the transition probabilities. The transition matrix in normal, drought, and flood flow clusters, respectively, are

$$
\begin{array}{cc} & \begin{array}{cccc} 20 & 40 & 60 & 80 \end{array} \\ \begin{array}{c} 20 \\ 40 \\ 60 \\ 80 \end{array} & \begin{bmatrix} .801 & .187 & .012 & .0 \\ .336 & .502 & .153 & .008 \\ .045 & .344 & .482 & .129 \\ .002 & .059 & .384 & .556 \end{bmatrix} \end{array},
\quad
\begin{array}{cc} & \begin{array}{ccc} 10 & 30 & 50 \end{array} \\ \begin{array}{c} 10 \\ 30 \\ 50 \end{array} & \begin{bmatrix} .784 & .202 & .014 \\ .327 & .503 & .17 \\ .046 & .343 & .61 \end{bmatrix} \end{array},
\quad \text{and} \quad
\begin{array}{cc} & \begin{array}{ccccccccc} 30 & 50 & 70 & 90 & 110 & 130 & 150 & 170 & 190 \end{array} \\ \begin{array}{c} 30 \\ 50 \\ 70 \\ 90 \\ 110 \\ 130 \\ 150 \\ 170 \\ 190 \end{array} & \begin{bmatrix} .724 & .233 & .041 & .002 & .0 & 0 & 0 & 0 & 0 \\ .307 & .424 & .228 & .039 & .002 & .0 & 0 & 0 & 0 \\ .055 & .26 & .424 & .222 & .037 & .002 & .0 & 0 & 0 \\ .003 & .054 & .266 & .423 & .217 & .036 & .002 & .0 & 0 \\ .0 & .004 & .056 & .271 & .422 & .212 & .034 & .002 & .0 \\ 0 & .0 & .004 & .058 & .277 & .421 & .206 & .032 & .002 \\ 0 & 0 & .0 & .004 & .061 & .282 & .419 & .201 & .032 \\ 0 & 0 & 0 & .0 & .004 & .064 & .288 & .418 & .226 \\ 0 & 0 & 0 & 0 & .0 & .005 & .066 & .293 & .636 \end{bmatrix} \end{array}.
$$

Let $\mathcal{R}_t \in \{\mathcal{R}^n, \mathcal{R}^d, \mathcal{R}^f\}$ denote the state space of the AR(1) process in period t. If the system transitions to a different cluster in period t ($\mathcal{R}_{t-1} \neq \mathcal{R}_t$) and the AR(1) process defined on $\mathcal{R}_{t-1}$ in period $t-1$ transitions to a state in period t that falls outside of $\mathcal{R}_t$, we set the streamflow rate in period t to the closest state in $\mathcal{R}_t$.

With the above modifications, $y_t = (f_t, \rho_t, j_t) \in \mathcal{Y}_t := \mathcal{R}_t \times \mathcal{P} \times \mathcal{J}$ and $\bar{y}_t = (f_t, \rho_t) \in \overline{\mathcal{Y}}_t := \mathcal{R}_t \times \mathcal{P}$. We restrict the water levels in the upper and lower reservoirs to values from the sets $\mathcal{X}_u := \{n\zeta_a \in [0, C_U] : n \in \mathbb{Z}\}$ and $\mathcal{X}_l := \{n\zeta_a \in [0, C_L] : n \in \mathbb{Z}\}$, respectively, where ζ_a is a prespecified constant. For the discrete-state version of our MDP, we define $\mathbb{U}^D(x_{ut}, x_{lt}, y_t)$ as the admissible set of action pairs (a_t, b_t) in state $(x_{ut}, x_{lt}, y_t) \in \mathcal{X}_u \times \mathcal{X}_l \times \mathcal{Y}_t$. The set $\mathbb{U}^D(x_{ut}, x_{lt}, y_t)$ contains the set $\{(n\zeta_a, m\zeta_a) \in \mathbb{U}(x_{ut}, x_{lt}, y_t) : n \in \mathbb{Z}, m \in \mathbb{Z}_+\}$ and extreme points of $\mathbb{U}(x_{ut}, x_{lt}, y_t)$. Note that the water level in the upper reservoir (or lower reservoir) may take a value $\tilde{x}_{ut} \notin \mathcal{X}_u$ (or $\tilde{x}_{lt} \notin \mathcal{X}_l$) in any period t since the amount of water runoff need not be a multiple of ζ_a. In such cases, to solve the recursion in (1) for period $t-1$, we linearly interpolate the optimal values in period t for the two states in $\mathcal{X}_u$ (or $\mathcal{X}_l$) adjacent to $\tilde{x}_{ut}$ (or $\tilde{x}_{lt}$). Using these data-calibrated settings, we next evaluate the optimality gaps and computation times of the PA and FPA methods, and examine the resulting operations of the PHES facility.

6. Numerical Experiments

We consider instances in which the planning horizon spans the first week of January, April, or August ($T = 168$ hours); the upper and lower vertical distances are both 30 meters; the PHES facility is open-loop with upstream flow ($r_{ut} \geq 0$ and $r_{lt} = 0$, $\forall t$) or downstream flow ($r_{ut} = 0$ and $r_{lt} \geq 0$, $\forall t$); $C_U = C_L = 1000$ MWh; $C_{Ru} = C_P = 100$ MWh; $C_{Rl} = 200$ MWh; $\theta = 0.88$; and $\zeta_a = 20$ MWh. These parameter values are realistic for small to medium-sized PHES facilities (Hayes 2009, Catalão 2017, and Kocaman and Modi 2017). The negative price occurrence frequency (NPF) takes values from the set $\{0\%, 0.80\%, 1.57\%\}$ in January and from the set $\{0\%, 1.26\%, 2.48\%\}$ in April and August. (The electricity price can be negative in our time series model: the NPF is 0.80% in January, and 1.26% in April and August.) The initial water levels x_{u1} and x_{l1} are the closest states to $C_U/2$ and $C_L/2$, respectively, in all instances. The initial exogenous state $y_1 = (f_1, \rho_1, j_1)$ is $(60, 0, 0)$ in January, $(110, 0, 0)$ in April, and $(30, 0, 0)$ in August. All computations were executed on a dual 3.7 GHz Intel Xeon W-2255 CPU server with 96 GB of RAM. Table 2 exhibits the optimality gaps and computation times of our heuristic methods.

Table 2 Optimality gaps and computation times (in CPU minutes) for the solution methods.

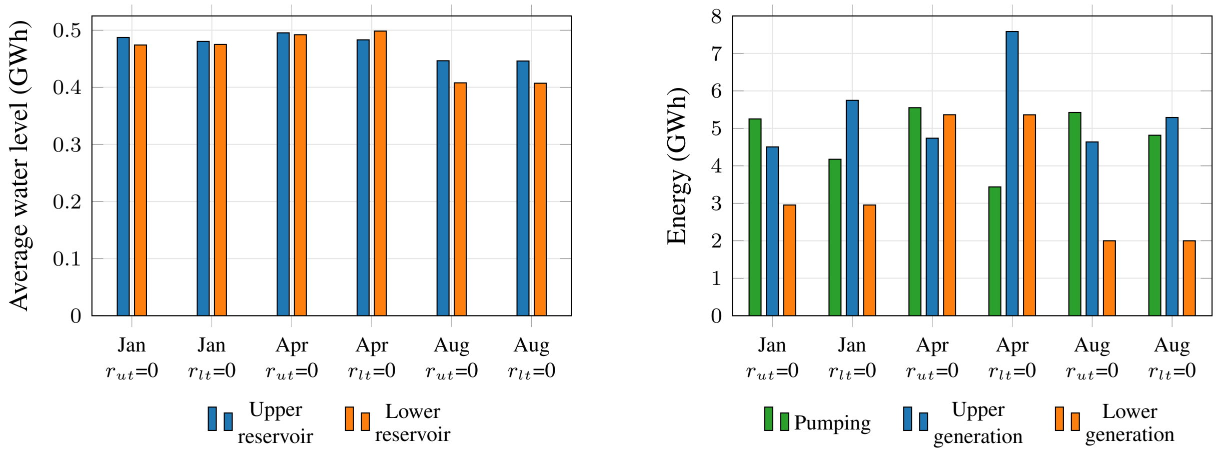
Runoff	Season	NPF	v_1^*	Optimality gaps		Computation times		
				PA	FPA	Optimal policy	PA	FPA
$r_{lt} = 0$	January	0%	0.804	0.00%	1.20%	140.9	16.8	8.0
		0.80%	0.812	0.00%	1.20%	140.5	16.8	8.0
		1.57%	0.821	0.01%	1.21%	140.9	16.8	8.0
	April	0%	0.912	0.00%	0.87%	309.5	31.1	11.4
		1.26%	0.920	0.01%	0.88%	309.9	31.1	11.4
		2.48%	0.928	0.02%	0.89%	310.0	31.0	11.4
	August	0%	0.559	0.00%	1.88%	106.3	13.8	7.2
		1.26%	0.569	0.00%	1.94%	105.8	13.7	7.2
		2.48%	0.579	0.01%	1.98%	106.0	13.8	7.2
$r_{ut} = 0$	January	0%	0.689	0.00%	1.23%	140.3	16.6	7.9
		0.80%	0.699	0.00%	1.24%	140.7	16.7	7.9
		1.57%	0.708	0.01%	1.27%	140.4	16.5	7.9
	April	0%	0.753	0.00%	0.65%	314.3	30.4	11.3
		1.26%	0.762	0.01%	0.70%	313.9	30.4	11.3
		2.48%	0.772	0.01%	0.73%	313.9	30.3	11.3
	August	0%	0.516	0.00%	1.91%	106.2	13.7	7.2
		1.26%	0.526	0.00%	1.97%	106.3	13.6	7.2
		2.48%	0.537	0.01%	2.02%	106.9	13.7	7.2

Note. v_1^* is the optimal expected total cash flow (in million dollars).

The results reveal the near-optimal performance of our PA method. It yields solutions with a maximum distance of only 0.02% from optimal and reduces the computation time of the optimal algorithm by 88.5% on average and by up to 90.3%. Our FPA method, on the other hand, causes only a very small drop in objective value. It yields solutions with an average distance of 1.32% and a maximum distance of 2.02% from optimal. Our FPA method offers further significant computational savings. It reduces the computation time of the optimal algorithm by 94.6% on average and by up to 96.4%. All these results underscore the practical value of our structural analysis. Finally, the optimality gaps of our FPA method are lowest in April and highest in August. Imposing a single pair of target levels that is independent of the total amount of water in the PHES facility is less restrictive in April with higher streamflow availability than in August with lower streamflow availability. See the feasible region of optimal target levels $\Phi(x_t^r)$ in equation (2).

Figure 6 compares instances with an NPF of 0.80% in January and 1.26% in April and August. This comparison examines the average water level in each reservoir (Figure 6a), alongside the total energy used for pumping and the total energy generation in the upper and lower reservoirs (Figure 6b). Since each reservoir has a storage capacity of 1 GWh, average water levels of approximately 0.4–0.5 GWh indicate that the reservoirs are not, on average, operated near either their lower or upper capacity limits. As expected, average water levels are higher in April (when the streamflow rate is large) than in August (when the streamflow rate is small).

Figure 6 Average water levels, and total pumping and generation volumes.



(a) Average water levels (b) Total pumping and generation

Note. The instances have an NPF of 0.80% in January and 1.26% in April and August.

Figure 6b clarifies the different operational roles of the two reservoirs. Generation from the lower reservoir is relatively insensitive to which reservoir receives natural inflows. This is because such energy generation represents a net release of water from the system and is therefore driven mainly by total streamflow availability. Its magnitude is highest in April and lowest in August. By contrast, pumping and generation from the upper reservoir are more sensitive to which reservoir receives natural inflows. Under downstream flow, pumping exceeds generation, and the difference between them remains relatively similar across seasons. This suggests that when streamflow feeds the lower reservoir, the system uses pumping mainly to move water to the upper reservoir and to maintain storage availability. Under upstream flow, generation becomes more dominant, especially in April. This is intuitive because inflow directly replenishes the upper reservoir, allowing more water to be used for generation, especially when the streamflow rate is large. Overall, the results suggest that generation from the lower reservoir is mainly governed by total water availability, while the operation of the upper reservoir is more strongly shaped by the location of inflow.

In the tested instances, spillage occurs only from the upper reservoir and only as a result of pumping. The maximum observed spill is very small (2.3 MWh, occurring in April when $r_{lt} = 0$). Consistent with the average water levels shown in Figure 6a, this indicates that the reservoirs are rarely operated near their maximum capacity and that pumping operations do not result in significant water losses.

To examine the effect of selected system parameters, we perform a one-at-a-time sensitivity analysis for August with NPF = 1.26%. While keeping all other parameters fixed at their baseline values, we vary $C_U \in \{900, 1000, 1100\}$ MWh, $C_P \in \{80, 100, 120\}$ MWh, $C_{Ru} \in \{80, 100, 120\}$ MWh, and $\theta \in \{0.80, 0.88, 0.95\}$. Since C_L and C_{Rl} remain fixed, varying C_U changes the reservoir-capacity ratio and varying C_{Ru} changes the release-capacity asymmetry. Considering both open-loop configurations with upstream flow and downstream flow, the sensitivity analysis contains 18 unique instances. Table 3 exhibits the performance of our heuristic methods across these instances.

Table 3 Optimality gaps and computation times (in CPU minutes) for the sensitivity analysis.

Runoff	Parameter	Value	v_1^*	Optimality gaps		Computation times		
				PA	FPA	Optimal policy	PA	FPA
$r_{lt} = 0$	Baseline		0.569	0.00%	1.94%	105.8	13.7	7.2
	C_U	900	0.563	0.01%	1.55%	95.8	12.7	7.0
	C_U	1100	0.577	0.00%	3.99%	115.6	14.4	7.4
	C_P	80	0.558	0.00%	1.83%	97.1	13.6	7.1
	C_P	120	0.579	0.01%	1.99%	114.8	13.6	7.2
	C_{Ru}	80	0.535	0.00%	1.95%	96.0	13.6	7.2
	C_{Ru}	120	0.601	0.00%	1.86%	115.0	13.5	7.2
	θ	0.80	0.506	0.00%	1.68%	105.8	13.6	7.2
	θ	0.95	0.628	0.00%	2.14%	105.7	13.5	7.2
$r_{ut} = 0$	Baseline		0.526	0.00%	1.97%	106.0	13.5	7.2
	C_U	900	0.521	0.00%	1.61%	95.7	12.6	6.9
	C_U	1100	0.534	0.00%	4.24%	116.4	14.3	7.4
	C_P	80	0.514	0.00%	1.79%	97.0	13.5	7.1
	C_P	120	0.537	0.00%	2.03%	114.6	13.4	7.2
	C_{Ru}	80	0.494	0.00%	2.01%	96.4	13.5	7.1
	C_{Ru}	120	0.557	0.00%	1.80%	115.1	13.4	7.1
	θ	0.80	0.463	0.00%	1.62%	106.1	13.5	7.1
	θ	0.95	0.585	0.00%	2.20%	106.0	13.5	7.1

Note. v_1^* is the optimal expected total cash flow (in million dollars).

The sensitivity results are consistent with the main computational findings. Across all sensitivity instances, the performance of our PA method is near-optimal and robust to changes in storage capacity, operational capacities, and system efficiency. Our FPA method also performs well, although its optimality gaps are more sensitive to parameter changes. The gaps of our FPA method are highest when C_U is increased to 1100 MWh, suggesting that a single pair of target levels becomes more restrictive when the upper reservoir has relatively greater storage capacity. The computation times also follow the same pattern as in the main results: both methods are substantially faster than the optimal policy. The changes in the optimal expected total cash flow are intuitive. Larger values of C_U, C_P, and C_{Ru} improve cash flow by increasing storage or operational flexibility, while larger values of θ improve cash flow through higher conversion efficiency.

Finally, we test the robustness of our heuristic solution methods in a setting where Assumption 2 is violated. In all data-calibrated instances considered earlier, this assumption is satisfied: the largest runoff into either reservoir is 55.92 MWh, which occurs in April. To create a violation of the assumption, we modify the April instance by increasing the reservoir head from 30 meters to 60 meters. This doubles the energy value of runoff, making the maximum values of r_{ut} and r_{lt} equal to 111.84 MWh and violating Assumption 2 in the corresponding upstream- and downstream-flow cases as $C_{Ru} = C_P = 100$ MWh and $C_{Rl} = 200$ MWh. Our PA and FPA methods continue to perform well in these modified instances, yielding solutions with a maximum distance of only 0.03% and 1.02% from optimal, respectively. These results indicate that Assumption 2, while useful for deriving the structural results, does not appear to be critical for the numerical performance of our solution methods.

7. Concluding Remarks

We have studied the energy generation and storage problem for an open-loop PHES facility within a wholesale-market framework. Modeling this problem as an MDP, we characterize the optimal policy structure when electricity prices are assumed to be positive throughout the finite horizon. Leveraging our structural results, we construct two policy-approximation algorithms as heuristic solution methods for the generalized setting when prices are allowed to be negative. Numerical experiments reveal that our policy-approximation algorithms enable substantial computational savings with no significant drain on cash flows.

Our results carry practical implications for PHES operators and project developers. The threshold policy provides a structured operating rule that can guide real-time dispatch decisions: the operator releases water when the water level in the upper reservoir exceeds a state-dependent target, pumps water when the level falls below a lower target, and keeps the level unchanged otherwise. Importantly, in the open-loop setting, the optimal operating decisions depend not only on the total amount of water in the system but also on its distribution across the two reservoirs. From a computational perspective, the proposed algorithms reduce solution times by an order of magnitude relative to standard DP, which is relevant for operational settings that require frequent re-optimization as updated price and streamflow forecasts become available.

Future research may extend our analysis to more complex PHES facilities that can still be modeled as MDPs with potentially larger state and/or action spaces. Examples include cascaded systems with multiple connected reservoirs subject to variable heads, or configurations where excess spillage from an upper reservoir feeds directly into a lower one. Future extensions of this study may also consider PHES facilities integrated with other renewable energy sources (e.g., wind, solar, and biomass). Lastly, future research could study the energy generation and storage problem for PHES facilities participating in forward electricity markets by incorporating energy commitment decisions across different time frames. Such an extension may entail developing multi-stage stochastic programming approaches in addition to MDPs.

Acknowledgments

This research was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) [Grant 118M419].

References

Abbaspour M, Satkin M, Mohammadi-Ivatloo B, Lotfi FH, Noorollahi Y (2013) Optimal operation scheduling of wind power integrated with compressed air energy storage (CAES). *Renew. Energy* 51:53–59.

Al-Swaiti MS, Al-Awami AT, Khalid MW (2017) Co-optimized trading of wind-thermal-pumped storage system in energy and regulation markets. *Energy* 138:991–1005.

Ardizzon G, Cavazzini G, Pavesi G (2014) A new generation of small hydro and pumped-hydro power plants: Advances and future challenges. *Renew. Sustain. Energy Rev.* 31:746–761.

Blakers A, Stocks M, Lu B, Cheng C (2021) A review of pumped hydro energy storage. *Progress in Energy* 3(2):022003.

Brown PD, Lopes JP, Matos MA (2008) Optimization of pumped storage capacity in an isolated power system with large renewable penetration. *IEEE Trans. Power Syst.* 23(2):523–531.

Castronuovo ED, Lopes JP (2004) On the optimization of the daily operation of a wind-hydro power plant. *IEEE Trans. Power Syst.* 19(3):1599–1606.

Catalão J (2017) *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling* (CRC Press).

Connolly D, Lund H, Finn P, Mathiesen BV, Leahy M (2011) Practical operation strategies for pumped hydroelectric energy storage (PHES) utilising electricity price arbitrage. *Energy Policy* 39(7):4189–4196.

Ding H, Hu Z, Song Y (2014) Rolling optimization of wind farm and energy storage system in electricity markets. *IEEE Trans. Power Syst.* 30(5):2676–2684.

DOE (2023a) Pumped storage hydropower. `https://www.energy.gov/eere/water/pumped-storage-hydropower`, accessed May 12, 2025.

DOE (2023b) Pumped storage hydropower technology strategy assessment. Technical report, The U.S. Department of Energy.

Duque ÁJ, Castronuovo ED, Sánchez I, Usaola J (2011) Optimal operation of a pumped-storage hydro plant that compensates the imbalances of a wind power producer. *Electr. Power Syst. Res.* 81(9):1767–1777.

Garcia-Gonzalez J, de la Muela RMR, Santos LM, Gonzalez AM (2008) Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. *IEEE Trans. Power Syst.* 23(2):460–468.

Harsha P, Dahleh M (2014) Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy. *IEEE Trans. Power Syst.* 30(3):1164–1181.

Hayes SJ (2009) Technical analysis of pumped storage and integration with wind power in the Pacific Northwest. Technical report, U.S. Army Corps of Engineers Hydroelectric Design Center.

Hull J, White A (1994) Numerical procedures for implementing term structure models I: Single-factor models. *J. Derivatives* 2(1):7–16.

IHA (2024a) 2024 World Hydropower Outlook: Opportunities to advance net zero. Technical report, International Hydropower Association.

IHA (2024b) Enabling new pumped storage hydropower: A guidance note for decision-makers to de-risk investments. Technical report, International Hydropower Association.

IHA (2025a) Pumped storage hydropower: Water batteries for the renewable energy sector. `https://www.hydropower.org/factsheets/pumped-storage`, accessed May 12, 2025.

IHA (2025b) Pumped storage tracking tool. `https://www.hydropower.org/hydropower-pumped-storage-tool`, accessed May 12, 2025.

Jaillet P, Ronn EI, Tompaidis S (2004) Valuation of commodity-based swing options. *Management Sci.* 50:909–921.

Johnson NL (1949) Systems of frequency curves generated by methods of translation. *Biometrika* 36(1/2):149–176.

Jurasz J, Mikulik J, Krzywda M, Ciapała B, Janowski M (2018) Integrating a wind-and solar-powered hybrid to the power system by coupling it with a hydroelectric power station with pumping installation. *Energy* 144:549–563.

Karakoyun EC, Avci H, Huh WT, Kocaman AS, Nadar E (2026) Optimal hour-ahead commitment and storage decisions of wind power producers. *Int. J. Prod. Econ.* 292:109696.

Kocaman AS, Modi V (2017) Value of pumped hydro storage in a hybrid energy generation and allocation system. *Appl. Energy* 205:1202–1215.

Kunzig R (2024) Water batteries. *Science* 383(6681):358–363.

Kusakana K (2016) Optimal scheduling for distributed hybrid system with pumped hydro storage. *Energy Convers. Manag.* 111:253–260.

Kusakana K (2018) Optimal operation scheduling of grid-connected pv with ground pumped hydro storage system for cost reduction in small farming activities. *J. Energy Storage* 16:133–138.

Liu J, Ou M, Sun X, Chen J, Mi C, Bo R (2022a) Implication of production tax credit on economic dispatch for electricity merchants with storage and wind farms. *Appl. Energy* 308:118318.

Liu J, Sun XY, Bo R, Wang S, Ou M (2022b) Economic dispatch for electricity merchant with energy storage and wind plant: State of charge based decision making considering market impact and uncertainties. *J. Energy Storage* 53:104816.

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

Nadar E, Akan M, Scheller-Wolf A (2016) Experimental results indicating lattice-dependent policies may be optimal for general assemble-to-order systems. *Production Oper. Management* 25(4):647–661.

NYISO (2020) Energy market & operational data. `https://www.nyiso.com/energy-market-operational-data`, accessed November 30, 2020.

Parker GG, Tan B, Kazan O (2019) Electric power industry: Operational and public policy challenges and opportunities. *Production Oper. Management* 28(11):2738–2777.

Porteus EL (2002) *Foundations of stochastic inventory theory* (Stanford University Press).

Saulsbury JW (2020) A comparison of the environmental effects of open-loop and closed-loop pumped storage hydropower. Technical report, Pacific Northwest National Lab., Richland, WA (United States).

Shu Z, Jirutitijaroen P (2013) Optimal operation strategy of energy storage system for grid-connected wind power plants. *IEEE Trans. Sustain. Energy* 5(1):190–199.

Steffen B, Weber C (2016) Optimal operation of pumped-hydro storage plants with continuous time-varying power prices. *Eur. J. Oper. Res.* 252(1):308–321.

Tauchen G (1986) Finite state markov-chain approximations to univariate and vector autoregressions. *Econ. Lett.* 20(2):177–181.

The Center for Land Use Interpretation (2025) Pumped storage facilities in the USA. https://clui.org/projects/offstream/pumped-storage-facilities-usa/blenheim-gilboa, accessed May 12, 2025.

Topkis DM (1998) *Supermodularity and Complementarity* (Princeton University Press).

Toufani P, Karakoyun EC, Nadar E, Fosso OB, Kocaman AS (2023) Optimization of pumped hydro energy storage systems under uncertainty: A review. *J. Energy Storage* 73:109306.

Toufani P, Nadar E, Kocaman AS (2022a) Operational benefit of transforming cascade hydropower stations into pumped hydro energy storage systems. *J. Energy Storage* 51:104444.

Toufani P, Nadar E, Kocaman AS (2022b) Short-term assessment of pumped hydro energy storage configurations: Up, down, or closed? *Renew. Energy* 201:1086–1095.

USGS (2020) USGS water data for the nation help. https://waterdata.usgs.gov/, accessed May 4, 2022.

Vattenfall (2025a) Bleiloch. https://powerplants.vattenfall.com/bleiloch, accessed May 12, 2025.

Vattenfall (2025b) Geesthacht. https://powerplants.vattenfall.com/geesthacht, accessed May 12, 2025.

Vattenfall (2025c) Goldisthal. https://powerplants.vattenfall.com/goldisthal, accessed May 12, 2025.

Vespucci MT, Maggioni F, Bertocchi MI, Innorta M (2012) A stochastic model for the daily coordination of pumped storage hydro plants and wind power plants. *Ann. Oper. Res.* 193(1):91–105.

Wang W, Van Gelder PH, Vrijling J, Ma J (2006) Forecasting daily streamflow using hybrid ANN models. *J. Hydrol.* 324(1-4):383–399.

Yurter G, Nadar E, Kocaman AS (2024) The impact of pumped hydro energy storage configurations on investment planning of hybrid systems with renewables. *Renew. Energy* 222:119906.

Zhao Z, Yuan Y, He M, Jurasz J, Wang J, Egusquiza M, Egusquiza E, Xu B, Chen D (2022) Stability and efficiency performance of pumped hydro energy storage system for higher flexibility. *Renewable Energy* 199:1482–1494.

Zhao Z, Zhang M, He M, Gao Y, Li J, Han S, Liu Z, Lei L, Díaz JIP, Cavazzini G, et al. (2024) Beyond fixed-speed pumped storage: A comprehensive evaluation of different flexible pumped storage technologies in energy systems. *Journal of Cleaner Production* 434:139994.

Zhou Y, Scheller-Wolf A, Secomandi N, Smith S (2016) Electricity trading and negative prices: storage vs. disposal. *Management Sci.* 62(3):880–898.

Zhou Y, Scheller-Wolf A, Secomandi N, Smith S (2019) Managing wind-based electricity generation in the presence of storage and transmission capacity. *Production Oper. Management* 28(4):970–989.

A. Proofs of the Analytical Results

We summarize the notation of our problem formulation in Table 4.

Table 4 Summary of notation.

C_U, C_L	:	Capacities of the upper and lower reservoirs, respectively.
C_{Ru}, C_{Rl}	:	Releasing capacities of the upper and lower reservoirs, respectively.
C_P	:	Pumping capacity of the lower reservoir.
θ	:	Efficiency of the PHES facility, $\theta \in (0,1)$.
$\mathcal{T}$	:	Set of periods; $\mathcal{T} = \{1, 2, \ldots, T\}$.
x_{ut}, x_{lt}	:	Water levels in the upper and lower reservoirs, respectively, at the beginning of period t; $x_{ut} \in [0, C_U]$ and $x_{lt} \in [0, C_L]$.
p_t	:	Electricity price in period t; $p_t > 0$ (Assumption 1).
r_{ut}, r_{lt}	:	Water runoff to the upper and lower reservoirs, respectively, in period t; $r_{ut} \leq C_{Ru}$ and $r_{lt} \leq C_{Rl} - C_{Ru}$ (Assumption 2).
y_t	:	Exogenous state tuple in period t; $y_t = (r_{u\eta}, r_{l\eta}, p_\eta)_{\eta \leq t}$.
x^r_{ut}, x^r_{lt}	:	Water levels in the upper and lower reservoirs, respectively, after observing the water runoff in period t; $x^r_{ut} = x_{ut} + r_{ut}$ and $x^r_{lt} = x_{lt} + r_{lt}$.
x^r_t	:	Total amount of water in the PHES facility after observing the water runoff in period t; $x^r_t := x^r_{ut} + x^r_{lt}$.
$x_{u(t+1)}, x_{l(t+1)}$	:	Water levels in the upper and lower reservoirs, respectively, at the end of period t; $x_{u(t+1)} = \min\{x^r_{ut} - a_t, C_U\}$ and $x_{l(t+1)} = \min\{x^r_{lt} + a_t - b_t, C_L\}$.
a_t	:	Amount of water released from or pumped to the upper reservoir in period t.
b_t	:	Amount of water released from the lower reservoir in period t.
$\mathbb{U}(x_{ut}, x_{lt}, y_t)$	:	Set of action pairs (a_t, b_t) that are admissible in state (x_{ut}, x_{lt}, y_t); if $(x_{ut}, x_{lt}, y_t) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$, then $-C_P \leq a_t \leq \min\{x^r_{ut}, C_{Ru}\}$, $0 \leq b_t \leq C_{Rl}$, and $b_t - a_t \leq x^r_{lt}$.
$(a^*_t(\cdot), b^*_t(\cdot))$	:	Optimal action pair in period t.
$E(\cdot)$	:	Amount of energy generated by releasing water or required to pump water in period t; $E(a_t) = \min\{\theta a_t, a_t/\theta\}$.
$R(\cdot, \cdot, y_t)$	:	Payoff in period t; $R(a_t, b_t, y_t) = (E(a_t) + \theta b_t) p_t$.
$v^*_t(x_{ut}, x_{lt}, y_t)$	:	Optimal value function in state (x_{ut}, x_{lt}, y_t) in period t.
z_t, z_{ut}	:	Amounts of water in the PHES facility and the upper reservoirs, respectively, at the end of period t if the action pair (a_t, b_t) is taken in period t; $z_t = x^r_t - b_t$ and $z_{ut} = x^r_{ut} - a_t$.
$S^{(\nu)}_t, S^{(\nu)}_{ut}$	:	Optimal state-dependent target levels that are associated with the amounts of water in the PHES facility and the upper reservoir, respectively, for $\nu \in \{\mathsf{P}, \mathsf{R}\}$.

A.1. Proof of Proposition 1

We note that $v^*_T(\cdot, \cdot, y_T)$ is jointly concave for all y_T. Pick an arbitrary $t < T$ and fix y_t. Assuming $v^*_{t+1}(\cdot, \cdot, y_{t+1})$ is jointly concave for all y_{t+1}, we will prove $v^*_t(\cdot, \cdot, y_t)$ is also jointly concave. We reformulate our problem into an equivalent form with linear constraints (see, e.g., Zhou et al. 2019). Accordingly, we define the decision variables: a^P_t denotes the amount of water pumped to the upper reservoir and a^R_t denotes the amount released from it. We define $\Gamma(x_{ut}, x_{lt}, y_t)$ as the set of decision variable tuples $\left(a^P_t, a^R_t, b_t, x_{u(t+1)}, x_{l(t+1)}\right) \in \mathbb{R}^5_+$ that satisfy the following constraints:

- Upper-reservoir decisions: $a^P_t \leq C_P$ and $a^R_t \leq \min\{x^r_{ut}, C_{Ru}\}$.

- Lower-reservoir decisions: $b_t \le C_{Rl}$ and $b_t + a_t^P - a_t^R \le x_{lt}^r$.
- State transition: $x_{u(t+1)} \le C_U$, $x_{u(t+1)} \le x_{ut}^r + a_t^P - a_t^R$, $x_{l(t+1)} \le C_L$, and $x_{l(t+1)} \le x_{lt}^r - a_t^P + a_t^R - b_t$.

We now consider the following problem:

$$\max_{\left(a_t^P, a_t^R, b_t, x_{u(t+1)}, x_{l(t+1)}\right) \in \Gamma(x_{ut}, x_{lt}, y_t)} \left\{ \left(\theta a_t^R - a_t^P/\theta + \theta b_t\right) p_t + \mathbb{E}\left[v_{t+1}^* \left(x_{u(t+1)}, x_{l(t+1)}, y_{t+1} \right) \right] \right\}. \tag{10}$$

We show that the above problem is equivalent to ours by constructing an optimal solution to (10) that satisfies $a_t^P = 0$ or $a_t^R = 0$. Let $\left(\hat{a}^P, \hat{a}^R, \hat{b}, \hat{x}_u, \hat{x}_l\right) \in \Gamma(x_{ut}, x_{lt}, y_t)$ denote a feasible solution to (10). Suppose $\hat{a}^P > 0$ and $\hat{a}^R > 0$. We define $\Delta = \min\{\hat{a}^P, \hat{a}^R\} > 0$ and note that $\left(\hat{a}^P - \Delta, \hat{a}^R - \Delta, \hat{b}, \hat{x}_u, \hat{x}_l\right) \in \Gamma(x_{ut}, x_{lt}, y_t)$. Since $0 < \theta < 1$ and $p_t > 0$, we obtain $\left(\theta\left(\hat{a}^R - \Delta\right) - \left(\hat{a}^P - \Delta\right)/\theta + \theta \hat{b}\right)p_t = \left(\theta \hat{a}^R - \hat{a}^P/\theta + \theta\hat{b}\right)p_t + \Delta\left(1/\theta - \theta\right)p_t > \left(\theta\hat{a}^R - \hat{a}^P/\theta + \theta\hat{b}\right)p_t$. Since $\left(\hat{a}^P, \hat{a}^R, \hat{b}, \hat{x}_u, \hat{x}_l\right)$ yields a smaller objective value than $\left(\hat{a}^P - \Delta, \hat{a}^R - \Delta, \hat{b}, \hat{x}_u, \hat{x}_l\right)$, $\left(\hat{a}^P, \hat{a}^R, \hat{b}, \hat{x}_u, \hat{x}_l\right)$ cannot be an optimal solution to (10) if $\hat{a}^P > 0$ and $\hat{a}^R > 0$. Hence $v_t^*(x_{ut}, x_{lt}, y_t)$ is equal to the optimal objective value of (10). We next show that $|v_t^*(x_{ut}, x_{lt}, y_t)| < \infty$. Note that $-C_P/\theta \le E(a_t) \le \theta C_{Ru}$ and $-C_P/\theta \le E(a_t) + b_t \le \theta(C_{Ru} + C_{Rl})$. This implies that $|R(a_t, b_t, y_t)| \le C|p_t|$ where $C = \max\{C_P/\theta, \theta(C_{Ru} + C_{Rl})\} < \infty$. Since $\mathbb{E}_{p_\kappa|y_t}[|p_\kappa|] < \infty$, $\forall \kappa \in \mathcal{T}$ with $\kappa \ge t$, $|v_t^*(x_{ut}, x_{lt}, y_t)| \le \sum_{\kappa=t}^{T-1} |\mathbb{E}_{p_\kappa|y_t}[p_\kappa]| C \le \sum_{\kappa=t}^{T-1} \mathbb{E}_{p_\kappa|y_t}[|p_\kappa|] C < \infty$. We define $\mathcal{C} := \left\{ \left(x_{ut}, x_{lt}, a_t^P, a_t^R, b_t, x_{u(t+1)}, x_{l(t+1)}\right) \mid (x_{ut}, x_{lt}) \in \Theta,\ \left(a_t^P, a_t^R, b_t, x_{u(t+1)}, x_{l(t+1)}\right) \in \Gamma(x_{ut}, x_{lt}, y_t) \right\}$ where $\Theta := \{(x_{ut}, x_{lt}) \mid 0 \le x_{ut} \le C_U, 0 \le x_{lt} \le C_L\}$. Note that $\mathcal{C}$ is a convex set since Θ and $\Gamma(x_{ut}, x_{lt}, y_t)$ are polyhedral and thus convex sets. The objective function of problem (10) is a concave function on $\mathcal{C}$ as we assume $v_{t+1}^*(\cdot, \cdot, y_{t+1})$ is jointly concave. Since $v_t^*(x_{ut}, x_{lt}, y_t) < \infty$, Theorem A.4 in Porteus (2002) implies that $v_t^*(x_{ut}, x_{lt}, y_t)$ is a concave function on Θ.

A.2. Proofs of Lemmas 1 and 2

Let $a = a_t^*(x_{ut}, x_{lt}, y_t)$ and $b = b_t^*(x_{ut}, x_{lt}, y_t)$. Assume to the contrary that $a < x_{ut}^r - C_U$. We show that $(x_{ut}^r - C_U, b) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$: Since $(a, b) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$ and $r_{ut} \le C_{Ru}$ under Assumption 2, we obtain $-C_P \le a < x_{ut}^r - C_U = x_{ut} + r_{ut} - C_U \le \min\{x_{ut}^r, C_{Ru}\}$ and $b - (x_{ut}^r - C_U) \le b - a \le x_{lt}^r$. Hence $(x_{ut}^r - C_U, b) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$. Since $R(\cdot, b, y_t)$ is an increasing function under Assumption 1, $R(x_{ut}^r - C_U, b, y_t) > R(a, b, y_t)$. Since the optimal value function increases with the system water level, $v_t^*(x_{ut}, x_{lt}, y_t) = R(a, b, y_t) + \mathbb{E}\left[v_{t+1}^*\left(C_U, \min\{x_{lt}^r + a - b, C_L\}, y_{t+1}\right)\right] < R(x_{ut}^r - C_U, b, y_t) + \mathbb{E}\left[v_{t+1}^*\left(C_U, \min\{x_{lt}^r + x_{ut}^r - C_U - b, C_L\}, y_{t+1}\right)\right]$. Since the action pair $(x_{ut}^r - C_U, b) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$ yields a larger value function, we have a contradiction. Thus $a \ge x_{ut}^r - C_U$.

Now assume to the contrary that $b - a < x_{lt}^r - C_L$. We show that $(a, x_{lt}^r + a - C_L) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$: Since $(a, b) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$ and $r_{lt} \le C_{Rl} - C_{Ru}$ under Assumption 2, we obtain $0 \le b < x_{lt}^r + a - C_L \le r_{lt} + C_{Ru} \le C_{Rl}$ and $x_{lt}^r + a - C_L - a = x_{lt}^r - C_L < x_{lt}^r$. Hence $(a, x_{lt}^r + a - C_L) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$. We note that $R(a, x_{lt}^r + a - C_L, y_t) > R(a, b, y_t)$ and $x_{lt}^r + a - (x_{lt}^r + a - C_L) = C_L$. Since the optimal value function increases with the system water level,

$v_t^*(x_{ut}, x_{lt}, y_t) = R(a,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(\min\{x_{ut}^r - a, C_U\}, \min\{x_{lt}^r + a - b, C_L\}, y_{t+1}\right)\right] < R(a, x_{lt}^r + a - C_L, y_t) + \mathbb{E}\left[v_{t+1}^*\left(\min\{x_{ut}^r - a, C_U\}, C_L, y_{t+1}\right)\right]$. Since the action pair $(a, x_{lt}^r + a - C_L) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$ yields a larger value function, we have a contradiction. Thus $a - b \leq C_L - x_{lt}^r$.

A.3. Proof of Proposition 2

We note that $v_T^*(\cdot)$ satisfies properties (a)–(d). Pick an arbitrary $t < T$. Assuming $v_{t+1}^*(\cdot)$ satisfies properties (a)–(d), we will show $v_t^*(\cdot)$ also satisfies properties (a)–(d). Our key proof technique is to construct feasible actions that, when taken in period t, provide lower bounds for the optimal value function on the right-hand side of property (a) and for those in the second and third terms of properties (b)–(d). By selecting these actions appropriately, we will establish the desired properties.

(a) Let $a = a_t^*(x_{ut}, x_{lt} + \alpha, y_t)$ and $b = b_t^*(x_{ut}, x_{lt} + \alpha, y_t)$. Also, let $\hat{a} = \min\{a + \alpha, C_{Ru}\}$. We show that $(\hat{a}, b) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$: Since $(a, b) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$, we obtain $b - \hat{a} = b - a - \alpha \leq x_{lt}^r$ if $\hat{a} = a + \alpha$ and $b - \hat{a} = b - C_{Ru} = b - C_{Rl} \leq 0 \leq x_{lt}^r$ if $\hat{a} \neq a + \alpha$. Also, note that $-C_P \leq a \leq \hat{a} = \min\{a + \alpha, C_{Ru}\} \leq \{x_{ut}^r + \alpha, C_{Ru}\}$. Hence $(\hat{a}, b) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$. Since $R(\cdot, b, y_t)$ is an increasing function under Assumption 1, $R(\hat{a}, b, y_t) \geq R(a, b, y_t)$. Lemmas 1 and 2 imply that $C_U \geq x_{ut}^r - a$ and $C_L \geq x_{lt}^r + \alpha + a - b \geq x_{lt}^r + \hat{a} - b$, respectively. Thus:

$$\begin{aligned} v_t^*(x_{ut}, x_{lt} + \alpha, y_t) &= R(a,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r - a, x_{lt}^r + \alpha + a - b, y_{t+1}\right)\right] \\ &\leq R(\hat{a}, b, y_t) + \mathbb{E}\left[v_{t+1}^*\left(\min\{x_{ut}^r + \alpha - \hat{a}, C_U\}, x_{lt}^r + \hat{a} - b, y_{t+1}\right)\right] \leq v_t^*(x_{ut} + \alpha, x_{lt}, y_t). \end{aligned}$$

If $\hat{a} = a + \alpha$, the first inequality above holds trivially. If $\hat{a} = C_{Ru} \leq a + \alpha$, Assumption 2 implies $x_{ut}^r - a \leq x_{ut} + r_{ut} + \alpha - \hat{a} \leq x_{ut} + \alpha \leq C_U$ and the same inequality holds by the induction hypothesis.

(b) We now prove that $a_{lhs} := v_t^*(x_{ut} + \alpha, x_{lt}, y_t) - v_t^*(x_{ut}, x_{lt} + \alpha, y_t) \leq v_t^*(x_{ut} + \alpha, x_{lt} + \beta, y_t) - v_t^*(x_{ut}, x_{lt} + \alpha + \beta, y_t) =: a_{rhs}$. Let $a = a_t^*(x_{ut} + \alpha, x_{lt}, y_t)$ and $c = a_t^*(x_{ut}, x_{lt} + \alpha + \beta, y_t)$. Also, let $b = b_t^*(x_{ut} + \alpha, x_{lt}, y_t)$ and $d = b_t^*(x_{ut}, x_{lt} + \alpha + \beta, y_t)$. Lemma 1 implies that $C_U \geq x_{ut}^r + \alpha - a$ and $C_U \geq x_{ut}^r - c$. Lemma 2 implies that $C_L \geq x_{lt}^r + a - b$ and $C_L \geq x_{lt}^r + \alpha + \beta + c - d$. We consider the following three scenarios to prove the statement:

(b1) Suppose $\alpha + c \leq a$: We show that $(a - \alpha, b) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$: Since $(a, b) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$, we obtain $a - \alpha \leq x_{ut}^r$, $a - \alpha < a \leq C_{Ru}$, and $b - a + \alpha \leq x_{lt}^r + \alpha$. Since $(c, d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha + \beta, y_t)$, we obtain $-C_P \leq c \leq a - \alpha$. Hence $(a - \alpha, b) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$. We also show that $(c + \alpha, d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt} + \beta, y_t)$: Since $(c, d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha + \beta, y_t)$, we obtain $c + \alpha \leq x_{ut}^r + \alpha$, $-C_P \leq c \leq c + \alpha$, and $d - c - \alpha \leq x_{lt}^r + \beta$. Since $(a, b) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$, we obtain $c + \alpha \leq a \leq C_{Ru}$. Hence $(c + \alpha, d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt} + \beta, y_t)$. Thus:

$$\begin{aligned} a_{lhs} &\leq R(a,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r + \alpha - a, x_{lt}^r + a - b, y_{t+1}\right)\right] \\ &\quad - R(a - \alpha, b, y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r + \alpha - a, x_{lt}^r + a - b, y_{t+1}\right)\right] \\ &= (E(a) - E(a - \alpha))p_t \leq (E(c + \alpha) - E(c))p_t \end{aligned}$$

$$
\begin{aligned}
&= R(c+\alpha,d,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+\beta+c-d,y_{t+1}\right)\right]\\
&\quad -R(c,d,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+\beta+c-d,y_{t+1}\right)\right]\le a_{rhs}.
\end{aligned}
$$

The second inequality above holds since $E(\cdot)$ is concave under Assumption 1.

(b2) Suppose $\beta+b\le d$: We show that $(c,d-\beta)\in\mathbb{U}(x_{ut},x_{lt}+\alpha,y_t)$: Since $(c,d)\in\mathbb{U}(x_{ut},x_{lt}+\alpha+\beta,y_t)$, we obtain $d-\beta\le d\le C_{Rl}$ and $d-\beta-c\le x^r_{lt}+\alpha$. Since $(a,b)\in\mathbb{U}(x_{ut}+\alpha,x_{lt},y_t)$, we obtain $0\le b\le d-\beta$. Hence $(c,d-\beta)\in\mathbb{U}(x_{ut},x_{lt}+\alpha,y_t)$. We also show that $(a,b+\beta)\in\mathbb{U}(x_{ut}+\alpha,x_{lt}+\beta,y_t)$: Since $(a,b)\in\mathbb{U}(x_{ut}+\alpha,x_{lt},y_t)$, we obtain $0\le b\le b+\beta$ and $b+\beta-a\le x^r_{lt}+\beta$. Since $(c,d)\in\mathbb{U}(x_{ut},x_{lt}+\alpha+\beta,y_t)$, we obtain $b+\beta\le d\le C_{Rl}$. Hence $(a,b+\beta)\in\mathbb{U}(x_{ut}+\alpha,x_{lt}+\beta,y_t)$. Thus:

$$
\begin{aligned}
a_{lhs}&\le R(a,b,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-a,x^r_{lt}+a-b,y_{t+1}\right)\right]\\
&\quad -R(c,d-\beta,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+\beta+c-d,y_{t+1}\right)\right]\\
&= R(a,b+\beta,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-a,x^r_{lt}+a-b,y_{t+1}\right)\right]\\
&\quad -R(c,d,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+\beta+c-d,y_{t+1}\right)\right]\le a_{rhs}.
\end{aligned}
$$

(b3) Suppose $\alpha+c>a$ and $\beta+b>d$: Since $(a,b)\in\mathbb{U}(x_{ut}+\alpha,x_{lt},y_t)$, we obtain $b-c\le b-a+\alpha\le x^r_{lt}+\alpha$ and $d-a<b+\beta-a\le x^r_{lt}+\beta$. Hence $(c,b)\in\mathbb{U}(x_{ut},x_{lt}+\alpha,y_t)$ and $(a,d)\in\mathbb{U}(x_{ut}+\alpha,x_{lt}+\beta,y_t)$. Thus:

$$
\begin{aligned}
a_{lhs}&\le R(a,b,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-a,x^r_{lt}+a-b,y_{t+1}\right)\right]\\
&\quad -R(c,b,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+c-b,y_{t+1}\right)\right]\\
&\le R(a,d,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-a,x^r_{lt}+\beta+a-d,y_{t+1}\right)\right]\\
&\quad -R(c,d,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c,x^r_{lt}+\alpha+\beta+c-d,y_{t+1}\right)\right]\le a_{rhs}.
\end{aligned}
$$

Note that $R(a,b,y_t)-R(c,b,y_t)=(E(a)-E(c))p_t=R(a,d,y_t)-R(c,d,y_t)$. The second inequality above holds as we assume $v^*_{t+1}(\cdot)$ satisfies property (b).

(c) We next prove that $b_{lhs}:=v^*_t(x_{ut}+\alpha+\beta,x_{lt},y_t)-v^*_t(x_{ut}+\beta,x_{lt}+\alpha,y_t)\le v^*_t(x_{ut}+\alpha,x_{lt},y_t)-v^*_t(x_{ut},x_{lt}+\alpha,y_t)=:b_{rhs}$. Let $a=a^*_t(x_{ut}+\alpha+\beta,x_{lt},y_t)$ and $c=a^*_t(x_{ut},x_{lt}+\alpha,y_t)$. Also, let $b=b^*_t(x_{ut}+\alpha+\beta,x_{lt},y_t)$ and $d=b^*_t(x_{ut},x_{lt}+\alpha,y_t)$. Lemma 1 implies that $C_U\ge x^r_{ut}+\alpha+\beta-a$ and $C_U\ge x^r_{ut}-c$. Lemma 2 implies that $C_L\ge x^r_{lt}+a-b$ and $C_L\ge x^r_{lt}+\alpha+c-d$. We consider the following three scenarios to prove the statement:

(c1) Suppose $\alpha+c>a$ and $d>b$: Since $(c,d)\in\mathbb{U}(x_{ut},x_{lt}+\alpha,y_t)$, we obtain $c\le x^r_{ut}<x^r_{ut}+\beta$ and $(c,d)\in\mathbb{U}(x_{ut}+\beta,x_{lt}+\alpha,y_t)$. Since $(c,d)\in\mathbb{U}(x_{ut},x_{lt}+\alpha,y_t)$, we obtain $a<c+\alpha\le x^r_{ut}+\alpha$. Hence, and since $(a,b)\in\mathbb{U}(x_{ut}+\alpha+\beta,x_{lt},y_t)$, we obtain $(a,b)\in\mathbb{U}(x_{ut}+\alpha,x_{lt},y_t)$. Thus:

$$
v^*_t(x_{ut}+\alpha+\beta,x_{lt},y_t)-v^*_t(x_{ut}+\alpha,x_{lt},y_t)
$$

$$\le R(a,b,y_t) + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + a - b, y_{t+1}\right)\right]$$
$$- R(a,b,y_t) - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + a - b, y_{t+1}\right)\right]$$
$$\le \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \beta - \max\{c+b-d, a-\alpha\}, x^r_{lt} + \alpha + c - d, y_{t+1}\right)\right]$$
$$- \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} - \max\{c+b-d, a-\alpha\}, x^r_{lt} + \alpha + c - d, y_{t+1}\right)\right]$$
$$\le R(c,d,y_t) + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \beta - c, x^r_{lt} + \alpha + c - d, y_{t+1}\right)\right]$$
$$- R(c,d,y_t) - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} - c, x^r_{lt} + \alpha + c - d, y_{t+1}\right)\right]$$
$$\le v^*_t(x_{ut} + \beta, x_{lt} + \alpha, y_t) - v^*_t(x_{ut}, x_{lt} + \alpha, y_t).$$

The third inequality above holds since $-\max\{c+b-d, a-\alpha\} > -c$ and $v^*_{t+1}\left(\cdot, x^r_{lt} + \alpha + c - d, y_{t+1}\right)$ is concave. The second inequality can also be shown to hold:

(1) If $c + b - d \ge a - \alpha$, as we assume $v^*_{t+1}(\cdot)$ satisfies property (c), $v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + a - b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + a - b, y_{t+1}\right) \le v^*_{t+1}\left(x^r_{ut} + \beta - c - b + d, x^r_{lt} + \alpha + c - d, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut} - c - b + d, x^r_{lt} + \alpha + c - d, y_{t+1}\right)$.

(2) If $c + b - d < a - \alpha$, as we assume $v^*_{t+1}(\cdot)$ satisfies property (d), $v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + a - b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + a - b, y_{t+1}\right) \le v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + \alpha + c - d, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + \alpha + c - d, y_{t+1}\right)$.

(c2) Suppose $\alpha + c > a$ and $d \le b$: We show that $(c,b) \in \mathbb{U}(x_{ut} + \beta, x_{lt} + \alpha, y_t)$: Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$, we obtain $c \le x^r_{ut} < x^r_{ut} + \beta$. Since $(a,b) \in \mathbb{U}(x_{ut} + \alpha + \beta, x_{lt}, y_t)$, we obtain $b - c \le b - a + \alpha \le x^r_{lt} + \alpha$. Hence $(c,b) \in \mathbb{U}(x_{ut} + \beta, x_{lt} + \alpha, y_t)$. We also show that $(a,d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$: Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$, we obtain $a < c + \alpha \le x^r_{ut} + \alpha$. Since $(a,b) \in \mathbb{U}(x_{ut} + \alpha + \beta, x_{lt}, y_t)$, we obtain $d - a \le b - a \le x^r_{lt}$. Hence $(a,d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$. Thus:

$$b_{lhs} \le R(a,b,y_t) + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + a - b, y_{t+1}\right)\right]$$
$$- R(c,b,y_t) - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \beta - c, x^r_{lt} + \alpha + c - b, y_{t+1}\right)\right]$$
$$\le (E(a) - E(c))p_t + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + a - b, y_{t+1}\right)\right] - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} - c, x^r_{lt} + \alpha + c - b, y_{t+1}\right)\right]$$
$$\le R(a,d,y_t) + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha - a, x^r_{lt} + a - d, y_{t+1}\right)\right]$$
$$- R(c,d,y_t) - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} - c, x^r_{lt} + \alpha + c - d, y_{t+1}\right)\right] \le b_{rhs}.$$

The second and third inequalities hold as $v^*_{t+1}(\cdot)$ satisfies properties (c) and (b), respectively.

(c3) Suppose $\alpha + c \le a$: We show that $(a - \alpha, b) \in \mathbb{U}(x_{ut} + \beta, x_{lt} + \alpha, y_t)$: Since $(a,b) \in \mathbb{U}(x_{ut} + \alpha + \beta, x_{lt}, y_t)$, we obtain $a - \alpha \le x^r_{ut} + \beta$, $a - \alpha < a \le C_{Ru}$, and $b - (a - \alpha) \le x^r_{lt} + \alpha$. Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$, we obtain $-C_P \le c \le a - \alpha$. Hence $(a - \alpha, b) \in \mathbb{U}(x_{ut} + \beta, x_{lt} + \alpha, y_t)$. We also show that $(c + \alpha, d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$: Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt} + \alpha, y_t)$, we obtain $c + \alpha \le x^r_{ut} + \alpha$, $-C_P \le c \le c + \alpha$, and $d - (c + \alpha) \le x^r_{lt}$. Since $(a,b) \in \mathbb{U}(x_{ut} + \alpha + \beta, x_{lt}, y_t)$, we obtain $c + \alpha \le a \le C_{Ru}$. Hence $(c + \alpha, d) \in \mathbb{U}(x_{ut} + \alpha, x_{lt}, y_t)$. Thus:

$$b_{lhs} \le R(a,b,y_t) + \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut} + \alpha + \beta - a, x^r_{lt} + a - b, y_{t+1}\right)\right]$$

$$
\begin{aligned}
&- R(a-\alpha,b,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r+\alpha+\beta-a, x_{lt}^r+a-b, y_{t+1}\right)\right] \\
&= (E(a)-E(a-\alpha))p_t \le (E(c+\alpha)-E(c))p_t \\
&= R(c+\alpha,d,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\alpha+c-d, y_{t+1}\right)\right] \\
&\quad - R(c,d,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\alpha+c-d, y_{t+1}\right)\right] \le b_{rhs}.
\end{aligned}
$$

The second inequality above holds since $E(\cdot)$ is concave under Assumption 1.

(d) Last we prove that $c_{lhs} := v_t^*(x_{ut}+\alpha, x_{lt}+\beta, y_t) - v_t^*(x_{ut}, x_{lt}+\beta, y_t) \le v_t^*(x_{ut}+\alpha, x_{lt}, y_t) - v_t^*(x_{ut}, x_{lt}, y_t) =: c_{rhs}$. Let $a = a_t^*(x_{ut}+\alpha, x_{lt}+\beta, y_t)$ and $c = a_t^*(x_{ut}, x_{lt}, y_t)$. Also, let $b = b_t^*(x_{ut}+\alpha, x_{lt}+\beta, y_t)$ and $d = b_t^*(x_{ut}, x_{lt}, y_t)$. Lemma 1 implies that $C_U \ge x_{ut}^r+\alpha-a$ and $C_U \ge x_{ut}^r - c$. Lemma 2 implies that $C_L \ge x_{lt}^r+\beta+a-b$ and $C_L \ge x_{lt}^r+c-d$. We consider the following scenarios:

(d1) Suppose $\beta+d \le b$: We show that $(c, d+\beta) \in \mathbb{U}(x_{ut}, x_{lt}+\beta, y_t)$: Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$, we obtain $0 \le d \le d+\beta$ and $d+\beta-c \le x_{lt}^r+\beta$. Since $(a,b) \in \mathbb{U}(x_{ut}+\alpha, x_{lt}+\beta, y_t)$, we obtain $d+\beta \le b \le C_{Rl}$. Hence $(c, d+\beta) \in \mathbb{U}(x_{ut}, x_{lt}+\beta, y_t)$. We also show that $(a, b-\beta) \in \mathbb{U}(x_{ut}+\alpha, x_{lt}, y_t)$: Since $(a,b) \in \mathbb{U}(x_{ut}+\alpha, x_{lt}+\beta, y_t)$, we obtain $b-\beta \le b \le C_{Rl}$ and $b-\beta-a \le x_{lt}^r$. Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$, we obtain $0 \le d \le b-\beta$. Hence $(a, b-\beta) \in \mathbb{U}(x_{ut}+\alpha, x_{lt}, y_t)$. Thus:

$$
\begin{aligned}
c_{lhs} &\le R(a,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+\beta+a-b, y_{t+1}\right)\right] \\
&\quad - R(c,d+\beta,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+c-d, y_{t+1}\right)\right] \\
&= R(a,b-\beta,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+\beta+a-b, y_{t+1}\right)\right] \\
&\quad - R(c,d,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+c-d, y_{t+1}\right)\right] \le c_{rhs}.
\end{aligned}
$$

(d2) Suppose $\beta+d > b$ and $a > c$: Since $(c,d) \in \mathbb{U}(x_{ut}, x_{lt}, y_t)$, $b-c < d+\beta-c \le x_{lt}^r+\beta$ and $d-a < d-c \le x_{lt}^r$. Hence $(a,d) \in \mathbb{U}(x_{ut}+\alpha, x_{lt}, y_t)$ and $(c,b) \in \mathbb{U}(x_{ut}, x_{lt}+\beta, y_t)$. Thus:

$$
\begin{aligned}
&v_t^*(x_{ut}+\alpha, x_{lt}+\beta, y_t) - v_t^*(x_{ut}+\alpha, x_{lt}, y_t) \\
&\le R(a,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+\beta+a-b, y_{t+1}\right)\right] - R(a,d,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+a-d, y_{t+1}\right)\right] \\
&\le \theta(b-d)p_t + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\beta+\min\{\alpha+c,a\}-b, y_{t+1}\right)\right] - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\min\{\alpha+c,a\}-d, y_{t+1}\right)\right] \\
&\le R(c,b,y_t) + \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\beta+c-b, y_{t+1}\right)\right] - R(c,d,y_t) - \mathbb{E}\left[v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+c-d, y_{t+1}\right)\right] \\
&\le v_t^*(x_{ut}, x_{lt}+\beta, y_t) - v_t^*(x_{ut}, x_{lt}, y_t).
\end{aligned}
$$

Note that $R(a,b,y_t) - R(a,d,y_t) = \theta(b-d)p_t = R(c,b,y_t) - R(c,d,y_t)$. The third inequality above holds since $\min\{\alpha+c,a\} > c$ and $v_{t+1}^*\left(x_{ut}^r-c, \cdot, y_{t+1}\right)$ is concave. The second inequality can also be shown to hold:

(1) If $\alpha+c \ge a$, as we assume $v_{t+1}^*(\cdot)$ satisfies property (d), $v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+\beta+a-b, y_{t+1}\right) - v_{t+1}^*\left(x_{ut}^r+\alpha-a, x_{lt}^r+a-d, y_{t+1}\right) \le v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+\beta+a-b, y_{t+1}\right) - v_{t+1}^*\left(x_{ut}^r-c, x_{lt}^r+a-d, y_{t+1}\right)$.

(2) If $\alpha+c<a$, as we assume $v^*_{t+1}(\cdot)$ satisfies property (b), $v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+\beta+a-b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+a-d, y_{t+1}\right) \leq v^*_{t+1}\left(x^r_{ut}-c, x^r_{lt}+\alpha+\beta+c-b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}-c, x^r_{lt}+\alpha+c-d, y_{t+1}\right)$

(d3) Suppose $\beta+d>b$ and $c\geq a$: Since $(c,d)\in\mathbb{U}(x_{ut},x_{lt},y_t)$, we obtain $a\leq c\leq x^r_{ut}$. Hence, and since $(a,b)\in\mathbb{U}(x_{ut}+\alpha,x_{lt}+\beta,y_t)$, we obtain $(a,b)\in\mathbb{U}(x_{ut},x_{lt}+\beta,y_t)$. Also, since $(c,d)\in\mathbb{U}(x_{ut},x_{lt},y_t)$, we obtain $(c,d)\in\mathbb{U}(x_{ut}+\alpha,x_{lt},y_t)$. Thus:

$$
\begin{aligned}
c_{lhs} &\leq R(a,b,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+\beta+a-b, y_{t+1}\right)\right] \\
&\quad - R(a,b,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-a, x^r_{lt}+\beta+a-b, y_{t+1}\right)\right] \\
&\leq \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-\max\{c+b-d-\beta,a\}, x^r_{lt}+c-d, y_{t+1}\right)\right] \\
&\quad - \mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-\max\{c+b-d-\beta,a\}, x^r_{lt}+c-d, y_{t+1}\right)\right] \\
&\leq R(c,d,y_t)+\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}+\alpha-c, x^r_{lt}+c-d, y_{t+1}\right)\right] \\
&\quad - R(c,d,y_t)-\mathbb{E}\left[v^*_{t+1}\left(x^r_{ut}-c, x^r_{lt}+c-d, y_{t+1}\right)\right] \leq c_{rhs}.
\end{aligned}
$$

The third inequality above holds since $-\max\{c+b-d-\beta,a\}>-c$ and $v^*_{t+1}\left(\cdot, x^r_{lt}+c-d, y_{t+1}\right)$ is concave. The second inequality can also be shown to hold:

(1) If $c+b-d-\beta\geq a$, as we assume $v^*_{t+1}(\cdot)$ satisfies property (c), $v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+\beta+a-b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}-a, x^r_{lt}+\beta+a-b, y_{t+1}\right) \leq v^*_{t+1}\left(x^r_{ut}+\alpha-c-b+d+\beta, x^r_{lt}+c-d, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}-c-b+d+\beta, x^r_{lt}+c-d, y_{t+1}\right)$.

(2) If $c+b-d-\beta<a$, as we assume $v^*_{t+1}(\cdot)$ satisfies property (d), $v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+\beta+a-b, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}-a, x^r_{lt}+\beta+a-b, y_{t+1}\right) \leq v^*_{t+1}\left(x^r_{ut}+\alpha-a, x^r_{lt}+c-d, y_{t+1}\right) - v^*_{t+1}\left(x^r_{ut}-a, x^r_{lt}+c-d, y_{t+1}\right)$.

Hence $v^*_t(\cdot)$ satisfies properties (a), (b), (c), and (d).

A.4. Proof of Theorem 1

Let $a=a^*_t(x_{ut},x_{lt},y_t)$ and $b=b^*_t(x_{ut},x_{lt},y_t)$. To characterize the optimal energy generation and storage decision in the upper reservoir, first suppose that $a\geq 0$ and consider the following problem:

$$
\max_{(z_t,z_{ut})\in\Phi(x^r_t):z_{ut}\leq x^r_{ut}} \left\{\left(E^{(\mathsf{R})}(-z_{ut})-\theta z_t\right)p_t+\mathbb{E}_{y_{t+1}|y_t}\left[v^*_{t+1}\left(z_{ut}, z_t-z_{ut}, y_{t+1}\right)\right]\right\}.
$$

Proposition 1 implies that (z^*_t, z^*_{ut}) yields the maximum value in this problem where $z^*_{ut}=\min\left\{S^{(\mathsf{R})}_{ut}, x^r_{ut}\right\}$. Incorporating the capacity constraints, we obtain $a=\min\left\{x^r_{ut}-S^{(\mathsf{R})}_{ut}, C_{Ru}\right\}$ if $x^r_{ut}>S^{(\mathsf{R})}_{ut}$ and $a=0$ if $x^r_{ut}\leq S^{(\mathsf{R})}_{ut}$. Now suppose that $a\leq 0$ and consider the following problem:

$$
\max_{(z_t,z_{ut})\in\Phi(x^r_t):z_{ut}\geq x^r_{ut}} \left\{\left(E^{(\mathsf{P})}(-z_{ut})-\theta z_t\right)p_t+\mathbb{E}_{y_{t+1}|y_t}\left[v^*_{t+1}\left(z_{ut}, z_t-z_{ut}, y_{t+1}\right)\right]\right\}.
$$

Proposition 1 implies that (z_t^*, z_{ut}^*) yields the maximum value in this problem where $z_{ut}^* = \max\left\{S_{ut}^{(\mathsf{P})}, x_{ut}^r\right\}$. Incorporating the capacity constraints, we obtain $a = -\min\left\{S_{ut}^{(\mathsf{P})} - x_{ut}^r, C_P\right\}$ if $x_{ut}^r \le S_{ut}^{(\mathsf{P})}$. These results together provide the optimal energy generation and storage decision in the upper reservoir, as in Theorem 1. We specify the optimal energy generation decision in the lower reservoir with the following problem:

$$\max_{z_t \in \left[\max\left\{z_{ut}, x_{ut}^r + x_{lt}^r - C_{Rl}\right\}, \min\left\{z_{ut}+C_L, x_{ut}^r + x_{lt}^r\right\}\right]} \left\{-\theta z_t p_t + \mathbb{E}_{y_{t+1}|y_t}\left[v_{t+1}^*\left(z_{ut}, z_t - z_{ut}, y_{t+1}\right)\right]\right\}$$

where $z_{ut} = x_{ut}^r - a$. Since $S_t(x_{ut}^r - a)$ yields the maximum value here, $b = x_{ut}^r + x_{lt}^r - S_t(x_{ut}^r - a)$.

We now show that $S_t(\overline{z}_{ut}) \ge S_t(\underline{z}_{ut})$ for $x_t^r - C_L \le \underline{z}_{ut} < \overline{z}_{ut} \le x_t^r - C_{Rl}$: Assume to the contrary that $S_t(\overline{z}_{ut}) < S_t(\underline{z}_{ut})$. By definition of $S_t(\cdot)$, the following inequalities hold.

$$\mathbb{E}\left[v_{t+1}^*\left(\overline{z}_{ut}, S_t(\overline{z}_{ut}) - \overline{z}_{ut}, y_{t+1}\right)\right] - p_t\theta S_t(\overline{z}_{ut}) \ge \mathbb{E}\left[v_{t+1}^*\left(\overline{z}_{ut}, S_t(\underline{z}_{ut}) - \overline{z}_{ut}, y_{t+1}\right)\right] - p_t\theta S_t(\underline{z}_{ut}),$$
$$\mathbb{E}\left[v_{t+1}^*\left(\underline{z}_{ut}, S_t(\underline{z}_{ut}) - \underline{z}_{ut}, y_{t+1}\right)\right] - p_t\theta S_t(\underline{z}_{ut}) \ge \mathbb{E}\left[v_{t+1}^*\left(\underline{z}_{ut}, S_t(\overline{z}_{ut}) - \underline{z}_{ut}, y_{t+1}\right)\right] - p_t\theta S_t(\overline{z}_{ut}).$$

The above inequalities yield $\mathbb{E}\left[v_{t+1}^*\left(\overline{z}_{ut}, S_t(\overline{z}_{ut}) - \overline{z}_{ut}, y_{t+1}\right)\right] - \mathbb{E}\left[v_{t+1}^*\left(\underline{z}_{ut}, S_t(\overline{z}_{ut}) - \underline{z}_{ut}, y_{t+1}\right)\right] \ge \mathbb{E}\left[v_{t+1}^*\left(\overline{z}_{ut}, S_t(\underline{z}_{ut}) - \overline{z}_{ut}, y_{t+1}\right)\right] - \mathbb{E}\left[v_{t+1}^*\left(\underline{z}_{ut}, S_t(\underline{z}_{ut}) - \underline{z}_{ut}, y_{t+1}\right)\right]$, contradicting property (b) of Proposition 2. Thus $S_t(\overline{z}_{ut}) \ge S_t(\underline{z}_{ut})$. We also show that $S_{ut}^{(\mathsf{P})} \le S_{ut}^{(\mathsf{R})}$ and $S_t^{(\mathsf{P})} \le S_t^{(\mathsf{R})}$. By definitions of $S_t^{(\nu)}$ and $S_{ut}^{(\nu)}$, and letting $u_t^*\left(z_t, z_{ut}\right) = \mathbb{E}_{y_{t+1}|y_t}\left[v_{t+1}^*\left(z_{ut}, z_t - z_{ut}, y_{t+1}\right)\right]$, the following hold.

$$u_t^*\left(S_t^{(\mathsf{P})}, S_{ut}^{(\mathsf{P})}\right) - p_t S_{ut}^{(\mathsf{P})}/\theta - p_t\theta S_t^{(\mathsf{P})} \ge u_t^*\left(S_t^{(\mathsf{R})}, S_{ut}^{(\mathsf{R})}\right) - p_t S_{ut}^{(\mathsf{R})}/\theta - p_t\theta S_t^{(\mathsf{R})},$$
$$u_t^*\left(S_t^{(\mathsf{R})}, S_{ut}^{(\mathsf{R})}\right) - p_t\theta S_{ut}^{(\mathsf{R})} - p_t\theta S_t^{(\mathsf{R})} \ge u_t^*\left(S_t^{(\mathsf{P})}, S_{ut}^{(\mathsf{P})}\right) - p_t\theta S_{ut}^{(\mathsf{P})} - p_t\theta S_t^{(\mathsf{P})}.$$

The sum of the above inequalities implies that $p_t(\theta - 1/\theta)S_{ut}^{(\mathsf{P})} \ge p_t(\theta - 1/\theta)S_{ut}^{(\mathsf{R})}$. Since $0 < \theta < 1$, $S_{ut}^{(\mathsf{P})} \le S_{ut}^{(\mathsf{R})}$. Since $S_{ut}^{(\mathsf{P})} \le S_{ut}^{(\mathsf{R})}$ and $S_t(\cdot)$ is nondecreasing, $S_t^{(\mathsf{P})} = S_t\big(S_{ut}^{(\mathsf{P})}\big) \le S_t\big(S_{ut}^{(\mathsf{R})}\big) = S_t^{(\mathsf{R})}$.