processes

MDPI

Article

Day-Ahead Economic Dispatch Optimization for Cascade Hydro–Photovoltaic Complementary Power Generation Systems Based on Dual-Layer Decoupling (DP-MAPPO)

Dongyang He [1,2], Qingyan Yang [1], Haiku Zhang [2], Yujie Gao [1,3], Alfredo Iranzo [4] and Lei Han [1,*]

[1] School of Energy Science and Engineering, Harbin Institute of Technology, Harbin 150001, China
[2] Smart Energy Information Integration Research Center Water Power Research Institute, China Datang Corporation Science & Technology Research Institute, Beijing 100043, China
[3] Nanjing Nanzi HuaDun Digital Technology Co., Ltd., Nanjing 210003, China
[4] Energy Engineering Department, School of Engineering, Universidad de Sevilla, 41092 Sevilla, Spain; airanzo@us.es
* Correspondence: hithanlei@hit.edu.cn

Abstract

To address the deep coupling of unit commitment and load allocation, nonlinear unit characteristics, and non-convex feasible regions caused by vibration zones in intraday dispatching of cascade hydropower under high photovoltaic penetration, this paper proposes a dual-layer decoupled DP-MAPPO method for hydro-photovoltaic complementary systems, aiming to minimize total water consumption. The upper layer uses dynamic programming to optimize unit commitment, while the lower layer employs Multi-Agent Proximal Policy Optimization (MAPPO) for continuous load allocation. A case study is conducted on a hydro–photovoltaic power system comprising five cascaded hydropower plants and two photovoltaic plants, with a 24-h scheduling horizon and a 15-min time step. The results show that MAPPO exhibits stable convergence under fixed evaluation scenarios. For a typical day scenario, the objective function value obtained by DP-MAPPO is essentially consistent with the benchmark optimal value. Compared with the actual dispatch schedules, DP-MAPPO reduces total water consumption by 9.15%, 3.42%, and 0.65% across the three typical days analysed, respectively, with no violations of turbine vibration-zone constraints in any scenario. These results demonstrate that the proposed method achieves optimized scheduling that balances water conservation with safe and stable operation.

Keywords: DP-MAPPO; cascade hydropower; hydro-photovoltaic complementarity; intraday economic dispatch; vibration-zone constraints

Check for updates

Academic Editor: Michael C. Georgiadis

Received: 10 October 2026
Revised: 22 October 2026
Accepted: 23 October 2026
Published: 26 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Against the backdrop of the global deployment of next-generation power systems, the share of variable renewable energy sources (VRE)—such as wind and solar power—integrated into the grid has risen sharply; however, the highly stochastic nature and significant volatility of power output pose severe challenges to the safe and stable operation of power grids. Cascade hydropower plants, owing to their rapid response capability and flexible regulation characteristics, have become a core supporting power source within multi-energy complementary systems [1]. Nevertheless, with the increasing integration of high-proportion renewable energy sources and the ongoing reform of electricity markets,

the traditional dispatching paradigm is evolving toward a "high-frequency, near-real-time active response" approach. This shift imposes higher demands for agile response and lean operation on both the unit dispatching within power plants and the load allocation. The economic dispatching within cascade hydropower plants is, in essence, a complex optimization problem involving mixed integer and strongly non-convex coupling: the unit start-up and shutdown processes involve discrete variables, while the load allocation among units involves continuous variables, being these two aspects deeply coupled; moreover, the water consumption behavior of actual turbines is highly nonlinear, constrained by both dynamic head and power output, and their feasible region is often partitioned by multiple non-convex physical constraints [2].

In the field of hydroelectric power dispatch modeling, Parvez et al. [3] provided a comprehensive review of the modeling approaches and solution techniques for hydroelectric power dispatch problems; they noted that while dynamic programming-based methods are widely used, the number of random states in such models grows exponentially, making them impractical for multi-unit dispatch problems; Huang et al. [4] developed an integrated short-term multi-objective dispatch framework encompassing cascade stations and generating units, employing mixed-integer (non-)linear programming methods for modeling and solving the problem, thereby obtaining dispatch solutions with high accuracy; Wang et al. [5] established a two-objective day-ahead dispatch model aimed at minimizing both the residual system load variance and operational risk, utilizing multi-objective stochastic optimization to capture the uncertainty associated with both photovoltaic generation and load demand; Shi et al. [6] proposed a two-stage stochastic programming model in which system configuration constitutes the first-stage decision and operational dispatch constitutes the second-stage decision, employing the scenario-based approach to characterize the uncertainty arising from load growth; Shen et al. [7] introduced a multi-time-scale dispatch modeling method that accounts for temporal and spatial correlations, using nonparametric kernel density estimation combined with Frank–Copula correlation analysis to generate joint hydro-photovoltaic output scenarios; Hu et al. [8] formulated a high-dimensional, multivariable mixed-integer nonlinear programming model and employed modeling techniques—including piecewise linear approximation, introduction of integer variables, and head discretization—to transform the original model into a form suitable for efficient solution; Su et al. [9] developed a real-time dispatch framework based on model predictive control, converting the original mixed-integer nonlinear programming model into a mixed-integer linear programming model via linearization techniques for rolling horizon solution, thereby realizing a closed-loop dispatch process encompassing "prediction—decision—feedback adjustment".

Once a model is established, the efficient and accurate solution of the model is crucial to determining whether it can be effectively implemented in practice. Early research often relied on analytical methods—such as dynamic programming or the equal incremental rate method—to solve models directly; these methods are grounded in rigorous mathematical derivations and can yield theoretically optimal solutions for small-scale, low-dimensional problems. However, when dealing with real-world engineering problems involving multiple units, high precision, or cases where the feasible region is repeatedly partitioned by non-convex constraints (e.g., the "vibration region"), it becomes challenging to guarantee both computational efficiency and the feasibility of the solutions [10]. As a result, researchers have increasingly turned to heuristic algorithms that do not depend on gradient information or model convexity, leveraging swarm intelligence search mechanisms to explore satisfactory solutions within high-dimensional non-convex spaces. Yuan et al. [11] replaced the original local search strategy of the White Whale Algorithm with a spiral motion strategy based on the Whale Algorithm to enhance its local search capability, and

applied this modified algorithm to solve a hybrid hydro-photovoltaic power dispatching model; Xie et al. [12] introduced the Hyper-Dominance Evolutionary Algorithm into the field of multi-energy complementary dispatching for hydropower, wind, and solar power systems. By employing hyper-dominance dynamic control to manage selection pressure, combined with an angle diversity strategy and a population reselection mechanism, they balanced convergence and diversity in high-dimensional multi-objective optimization problems; Li et al. [13] proposed the Hyper-Quality Variation Artificial Bee Colony Algorithm for solving a coordinated dispatching model for hydropower, thermal, wind, solar, and pumped-storage power systems under carbon emission constraints; this algorithm demonstrated significant advantages over classical methods in both convergence speed and solution quality; He et al. [14] developed a joint hydropower–wind–solar–pumped-storage power dispatching model for the synergistic optimization of pumped-storage and battery energy storage systems (BESS), and proposed a Particle Swarm Optimization Algorithm based on reverse learning, providing a practical reference for the coordinated dispatching of multi-type energy storage systems; Li et al. [15] employed a Non-Dominance Sorting Genetic Algorithm incorporating an elite strategy to solve the problem under various scenarios, thereby validating the positive role of pumped-storage systems in enhancing system flexibility; Mingbo [16] proposed a two-level solution framework—where an Upper-Level Cuckoo Search Algorithm is used to optimize the unit start-stop combination, and a Lower-Level Dynamic Programming Approach is employed to optimize load distribution—thereby decomposing the complex integrated dispatching problem into two relatively simpler sub-problems for independent solution; Xie et al. [17] proposed a daily power generation dispatching method for cascade hydropower stations that takes peak-shaving constraints into account, and utilized a heuristic algorithm to determine the unit start-stop combination.

Although heuristic algorithms demonstrate strong adaptability when tackling high-dimensional non-convex optimization problems, they fundamentally operate under an "offline single-run" paradigm; furthermore, their handling of safety constraints—such as vibration zones—largely relies on penalty-based soft constraints, which cannot provide a mechanism-based guarantee for the deterministic avoidance of physical constraint violations. To balance real-time performance with decision-making quality, research interest in recent years has increasingly shifted toward deep reinforcement learning. By adopting an "offline centralized training, online rapid inference" approach, the iterative search process inherent in heuristic algorithms is moved forward to the training phase, enabling the model to generate decisions at nearly real-time speed during actual scheduling operations. Zeng et al. [18] proposed a real-time load dispatching method for hydropower stations based on D3QN-PER, integrating deep reinforcement learning into the unit load allocation problem; the results demonstrated that this method outperforms traditional approaches in both computational accuracy and solving efficiency. Wang et al. [19] developed a real-time dispatching method for hybrid hydroelectric and pumped-storage power generation systems based on deep reinforcement learning to address the uncertainty associated with photovoltaic power output. Ge et al. [20] introduced a knowledge-embedded deep reinforcement learning method and applied it to the real-time dispatching of hybrid wind-, solar-and hydro-power generation systems; the results showed that this method can circumvent the vibration zone constraint and accelerate the convergence speed. Zhou et al. [21] applied multi-agent deep reinforcement learning to optimal power flow calculations for power systems with high spatiotemporal resolution and high renewable energy penetration rates, validating the suitability of the multi-agent approach for solving large-scale, high-dimensional power system problems. Fang et al. [22] proposed a multi-layer energy management and policy learning method for micro-grids based on Proximal Policy

Optimization (PPO). Zuo et al. [23] proposed a day-ahead economic dispatching strategy for multiple types of distributed resources in distribution networks using an improved MAPPO algorithm.

To address these challenges, this paper proposes a dual-layer decoupled optimization framework combining upper-level dynamic programming (DP) and lower-level multi-agent proximal policy optimization (MAPPO) for cascade hydro–photovoltaic dispatch. The upper layer optimizes unit start-stop trajectories while eliminating infeasible combinations, whereas the lower layer allocates unit outputs within a non-convex feasible region. A constraint-handling mechanism combining feasible-segment parameterized encoding and Euclidean projection is developed to ensure power balance and vibration zone avoidance, without relying on reward-based soft constraints.

2. Methods

2.1. Economic Dispatch Model for In-Plant Units in a Cascade Hydropower–Solar Energy Complementary Power Plant

2.1.1. Objective Function

This model focuses on 96 intraday time intervals (i.e., with a scheduling step size of $\Delta t = 15$ min) and aims to minimize the total water consumption across the entire watershed. The objective function is defined as follows:

$$\min F_{total} = \sum_{t=1}^{T} \sum_{i=1}^{M} V_{i,t}^{plant} \tag{1}$$

Here, the subscript i denotes the i-th power plant, j denotes the j-th generating unit, and t denotes the t-th time interval. F_{total} represents the total water consumption of the water-photovoltaic complementary power generation system, $T = 96$ denotes the total number of scheduling periods; M represents the total number of cascade power stations; and $V_{i,t}^{plant}$ denotes the total water consumption volume of the power station during period t.

2.1.2. Constraints

(1) Power Balance Constraint

In a cascaded hydropower complementary system, the cascaded hydropower plants must dynamically accommodate the fluctuations in photovoltaic power generation in real time and meet the total load demand specified by the grid:

$$P_{target,t} = P_{hydro,t} + P_{pv,t} = \sum_{i=1}^{M} \sum_{j=1}^{Ni} P_{i,j,t} \cdot u_{i,j,t} + \sum P_{pvi} \tag{2}$$

In this formula, $P_{target,t}$ denotes the target total load of the hydro-photovoltaic complementary power generation system at a given time instant; $P_{pv,t}$ and P_{pvi} represent the total photovoltaic power output and the power output of the i-th photovoltaic power station, respectively; and $P_{hydro,t}$ represents the net load that the hydroelectric power cluster must bear, N_i denotes the number of generating units at the i-th power plant; $u_{i,j,t}$ denotes the operating status of the j-th generating unit at the i-th power plant during the t-th time interval (0 for shut down, 1 for operation).

(2) Unit Operational Characteristic Constraints

The relationship between the power generation flow rate, operating water head, and active power output of hydroelectric units exhibits a significant nonlinear, multidimensional coupling behavior. In the hydro-photovoltaic complementary system case study considered in this paper, the operating characteristic curves of the individual power plant units do not possess convexity. To avoid introducing a convex approximation that would alter the

actual operating characteristics of the units, this paper employs a bilinear interpolation method to discretize these characteristic curves, thereby deriving the mapping relationship between the power generation flow rate and the operating state as follows:

$$q = q(P, H) \tag{3}$$

$$V_{i,t}^{\text{plant}} = \sum_{j=1}^{N_i} \left(Q_{i,j,t} \cdot \Delta t \cdot u_{i,j,t} \right) \tag{4}$$

$$H_{i,t} = Z_{i,t}^{up} - Z_{i,t}^{down} - \Delta Z_{i,t}^{loss} \tag{5}$$

In this formula, q denotes unit power generation flow rate; P represents the actual output of the hydroelectric unit; H is the power plant water head; $Q_{i,j,t}$ denotes average outbound flow rate (power generation flow plus other flows, such as curtailed water flow and ecological flow); Δt denotes the duration of the scheduling interval; $Z_{i,t}^{up}$ is the upstream water level at the power station;$Z_{i,t}^{down}$ is the downstream water level at the power station; and $\Delta Z_{i,t}^{loss}$ is the head loss.

(3) Consideration of unit output constraints in the vibration zone

$$P_{i,j}^{\min} \le P_{i,j,t} \le P_{i,j}^{\max} \tag{6}$$

$$P_{i,j,t} \notin \bigcup_{k=1}^{K} \left[\underline{P}_{i,j,t}(H), \overline{P}_{i,j,t}(H) \right] \tag{7}$$

In this formula, $P_{i,j}^{\min}$ denotes the minimum output for the j-th unit at the i-th power station, and $P_{i,j}^{\max}$ denotes the maximum output for the j-th unit at the i-th power station. K represents the total number of vibration zones; $\underline{P}_{i,j,k}$ and $\overline{P}_{i,j,k}$ denote, respectively, the lower and upper bounds of the k-th vibration zone for the j-th unit at the i-th power station, depending on the head.

(4) Unit Output Variation Constraint

Due to the water hammer effect and the mechanical regulation rate of the turbine guide vanes, the continuous output adjustment range of the unit within an adjacent dispatch interval (Δt = 15 min) must comply with strict upper and lower limits:

$$-\Delta R_{i,j}^{down} \le P_{i,j,t} - P_{i,j,t-1} \le \Delta R_{i,j}^{up} \tag{8}$$

In this formula, $\Delta R_{i,j}^{up}$ and $\Delta R_{i,j}^{down}$ represent the absolute values of the maximum allowable ramp-up (increase in output) and ramp-down (decrease in output) variables for the j-th unit of the i-th power plant within a single dispatch step. Based on the regulation strategy employed in the reference example system, the load variation process is modeled on the principle of completing the output adjustment within a single scheduling period; that is, the unit's output changes in a single-step manner between adjacent scheduling periods, while the implicit regulation rate associated with this change must not exceed the unit's inherent physical regulation rate limit.

Additionally, the model is subject to conventional constraints, such as the unit discharge flow rate constraint and the power station reservoir capacity constraint.

2.2. A Two-Layer Collaborative Solution Method Based on DP + Multi-Agent PPO

This paper employs a dual-layer architecture organized around a "discrete–continuous" spatial decomposition approach (depicted in Figure 1). The upper-layer DP performs time-series optimization within a discrete state space comprising 96 scheduling time slots; it no longer relies on the simplified empirical water consumption formula but instead invokes

the lower-layer policy to obtain the actual physical water consumption cost, and then uses the Bellman recursion to derive the start/stop path that minimizes the total daily cost. The lower-layer layer encapsulates the trained MAPPO network as a "stateless physical policy" with frozen parameters, which performs load allocation among generating units given a specified combination of total load, head, and active units, strictly avoids the vibration zone, and calculates the physical water consumption. The two layers form a closed-loop through cross-layer bidirectional interaction: for each legitimate state transition explored by the DP—i.e., upon issuing an instruction (load, head, and active units combination) to the policy—the policy returns an allocation scheme and the corresponding water consumption cost within the safe operating region that satisfies the vibration zone and ramp-up constraints, which is then used to update the Bellman equation. As a result, the upper-layer retains the structural optimality inherent in the time-series combination search, while the lower-layer provides an efficient approximation in the continuous space, enabling decoupled coordination.

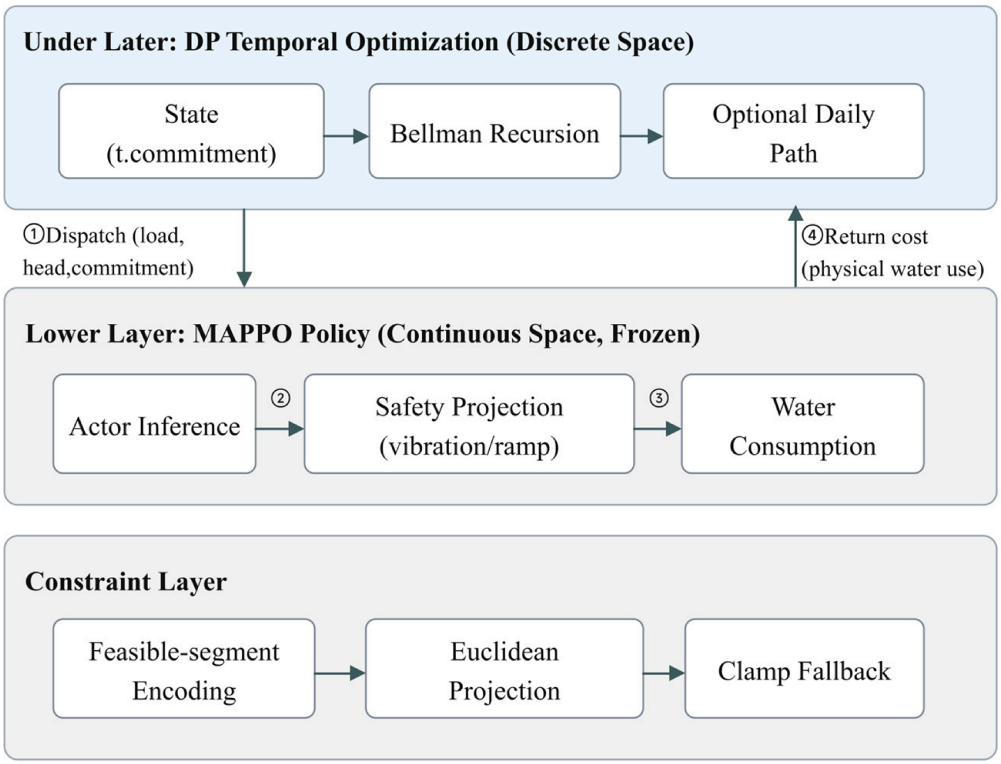


Figure 1. Two-layer decoupled optimization framework: DP-based temporal commitment search in the discrete space and MAPPO-based load allocation in the continuous space, coupled in a closed loop.

Additionally, alongside the primary method presented in this paper, five other methods are selected for comparison. Except for the actual engineering allocation method and the unit full-activation–output equalization method, the remaining four methods all employ the same upper-level dynamic programming framework to determine the unit start-up and shutdown combinations for each time period. The only difference lies in the lower-level output allocation rules. Detailed descriptions of the methods are provided in the Appendix A.

2.2.1. Time-Stepping Coupled Solution for the Upper-Level Unit Configuration

(1) Problem Description and Bellman State Transition

The scheduling period is divided into several time slots; for time slot t, let n_t denote the operational combination state of a power plant. The DP algorithm recursively computes

the optimal start/stop sequence over the entire day using the state transition equation. The optimal cost function for state transition is given by the Bellman Equation:

$$F_t(S_t) = \min\{F_{t-1}(S_{t-1}) + C_{water}(t, S_t)\} \tag{9}$$

$F_t(S_t)$ denotes the cumulative minimum cost for selecting the S_t unit activation time during period t (in this model, the cost is equivalent to the water consumption); $C_{water}(t, S_t)$ denotes the physical water consumption at the current time, obtained from the forward inference output of the lower-level MAPPO. By performing recursive and backtracking searches over all candidate combinations, the full-day unit start-up and shutdown trajectory is derived.

(2) Combined Pre-Screening Based on Dimensionality Reduction via Feasible Region Analysis

During the DP state generation phase, performing complex constraint validation on each candidate combination individually still results in high computational costs. To preclude invalid startup combinations before initiating the recursive search, this paper proposes a combinatorial pre-screening strategy based on analytical dimensionality reduction within the non-convex feasible region. Mathematically, by deriving the Minkowski sum of the discrete feasible set for a single machine and combining it with interval merging, highly complex mixed-integer constraints are pre-mapped into a continuous valid output range, thereby enabling dimensionality reduction for subsequent decision optimization.

For a given number of operating units n, the total output feasible region Ω_n (The feasible output region of a power plant refers to the existence of a unit combination and a load allocation scheme such that all unit output range constraints are satisfied) of the power station is mathematically equal to the Minkowski sum of the feasible regions Φ for each of the n independent single-unit units:

$$\Omega_n = \oplus_{i=1}^{n} \Phi_i = \left\{ \sum_{i=1}^{n} P_i | P_i \in \Phi_i, \forall i = 1, \ldots, n \right\} \tag{10}$$

For the common scenario in which a hydroelectric turbine unit possesses only one primary vibration zone (where this study also falls under this scenario). The continuous feasible output range for a single unit is divided by this vibration region $\left(\underline{P}_{j,k}, \overline{P}_{j,k}\right)$ into two operating intervals; the feasible region can thus be expressed as $\Phi = \left[P_j^{\min}, \underline{P}_{j,k}\right] \cup \left[\overline{P}_{j,k}, P_j^{\max}\right]$. For n identical units operating simultaneously, if k units (where $k \in \{0, 1, \ldots, n\}$) are operating within the upper interval, then the remaining $(n - k)$ units must be operating within the lower interval. Consequently, the feasible region for the total output Ω_n is actually composed of $n + 1$ subintervals $\omega_{n,k}$, whose upper and lower bounds can be expressed analytically as:

$$\omega_{n,k} = \left[(n-k)P_j^{\min} + k \cdot \overline{P}_{j,k}, k \cdot P_j^{\max} + (n-k) \cdot \underline{P}_{j,k}\right], \Omega_n = \cup_{k=0}^{n} \omega_{n,k} \tag{11}$$

Since the subintervals corresponding $\omega_{n,k}$ to different k-values may overlap, an interval merging calculation can be employed to merge and simplify these intervals, transforming them into several disjoint, continuous feasible output sets $\widetilde{\Omega}_n$.

Based on the aforementioned analytical derivation, for any given total plant target load P_{load}, by determining whether P_{load} falls within the merged continuous set $\widetilde{\Omega}_n$, the number of units to be activated that satisfy the physical constraints can be pre-emptively identified through a reverse screening process.

2.2.2. Lower-Layer Multi-Agent PPO Training and Inference

The lower-level problem involves allocating the total load among the various units within the power station, given a specified startup configuration and the overall plant target load. It is essentially a continuous decision-making problem subject to constraints imposed by a non-convex operational region. This problem is modeled as a multi-agent sequential decision-making problem. Its modeling and solving process is illustrated in Figure 2 and consists of two stages: training and inference.

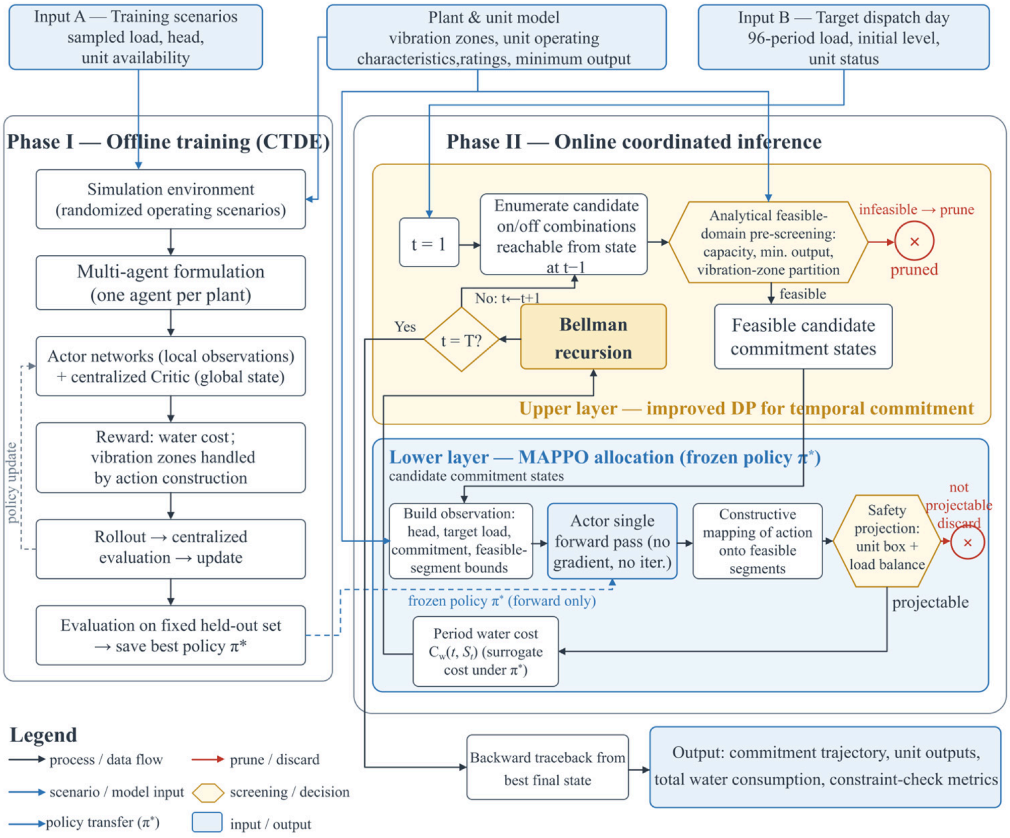


Figure 2. Framework of the proposed two-phase hydropower dispatch approach: offline CTDE-based multi-agent training and online coordinated inference.

– Action space construction: The vibration zone, as a physical quantity, exhibits a time-varying nonlinear coupling with the dimensionless actions output by the agent, where this coupling is determined by the total load. Treating this coupling as a reward or penalty term can easily lead to reward oscillations or convergence difficulties. To address this, this paper eschews the soft constraint approach and instead employs a two-stage mapping strategy based on "constructive parameterization–safe projection": First, the feasible output ranges for each unit under its current head are merged, and the vibration-prohibited regions are removed. The continuous action parameters are then mapped onto the sum of the lengths of these feasible segments, ensuring that once an action is generated, it inherently avoids the vibration zone. Next, all feasible segment combinations capable of supporting the target load for the given number of operating units are enumerated. Within the corresponding box constraints, the nearest point satisfying both the plant-wide load balance and the non-convex constraints is rapidly solved using a binary search method with Lagrange multipliers. As a result, both load balancing and vibration zone avoidance are strictly enforced at the action construction level, without relying on soft constraints imposed by a reward function.

- Training phase (Centralized Training—Decentralized Execution): Training is conducted in an offline environment. Each agent makes decisions solely based on local observations. However, a centralized Critic, which takes the global state as its input, is introduced to provide a unified evaluation of the joint action value, enabling the individual plants to "see" each other's decisions during training and thereby mitigating non-stationarity. Each plant shares the water consumption efficiency and load balancing rewards. By leveraging the global value function, it obtains gradients related to its own decisions, thereby facilitating cross-unit coordination. The training environment mirrors the actual dispatching problem, but the training process itself does not participate in the actual dispatching operations.
- Reasoning phase (forward allocation using the freezing strategy). Upon completion of training, the lower layer transitions from an "learner" to an "allocator": only the trained Actor is retained, while the global Critic is no longer used. Given an observation, each agent performs only a single forward computation and directly outputs an output schedule that satisfies the constraints via a safety projection—without exploration, sampling, or iteration. This feature enables the lower layer to serve as a rapid water consumption estimation module for the upper-layer DP: when the DP enumerates candidate combinations, it passes the current time-period load, headwater level, and operating status to this module, which then returns the corresponding water consumption value, providing efficient support for the upper-layer time-series optimization task.

3. Results

This section constructs a case study based on a cascade hydropower–photovoltaic complementary power generation system in a river basin in southwestern China, and details the physical topology of the system, the hydro-meteorological boundary conditions, and the equipment operating parameters.

3.1. Water-Photovoltaic Complementary Energy System Research Case Study

This case study involves a hydro-photovoltaic complementary power generation system comprising five cascade hydropower stations along Shuoku River (denoted as Plant1, Plant2, Plant3, Plant4, and Plant5) and two photovoltaic power stations (denoted as Pv1 and Pv2) (Figure 3), all located within the same watershed. The parameters for each station are provided in Appendix A.

On the physical topology, the cascade hydropower stations are closely hydraulically coupled. Additionally, Plant 1 and Pv1, as well as Plant 5 and Pv2, share a single transmission line, enabling bundled power transmission; the entire hydro–photovoltaic complementary system is under unified control via a centralized dispatch center.

This paper uses the typical daily dispatching process of this hydro-photovoltaic complementary energy system as a case study. The dispatching cycle is 24 h with a time step of 15 min, comprising a total of 96 dispatching intervals. Given the load curves of each power station and the water level fluctuation curves of each hydropower station in the typical-day system, the optimization objective is to minimize the total water consumption of the hydropower stations while satisfying the system power balance and unit operational safety constraints.

The typical daily power generation scheduling curve is shown in Figure 4. Based on the load variation characteristics, this day encompasses typical operational phases—including off-peak periods, ramp-up periods, peak periods, and off-peak periods—where the photovoltaic power station's output profile is relatively representative, and the load levels across the cascade of power stations exhibit spatially heterogeneous distribution

(higher at the beginning and end, and lower in the middle), thereby effectively capturing the dynamic process of unit grouping and output allocation during intra-day dispatching. Consequently, this typical day is well suited for comparing the differences in both economic efficiency and operational safety among various dispatching strategies.

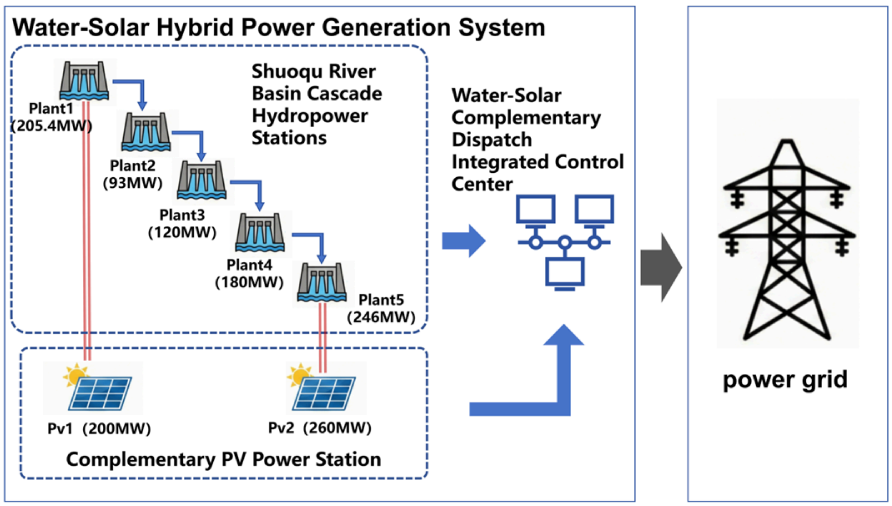


Figure 3. Schematic diagram of the water–photovoltaic complementary clean energy power generation system for the Shuoku River Basin.

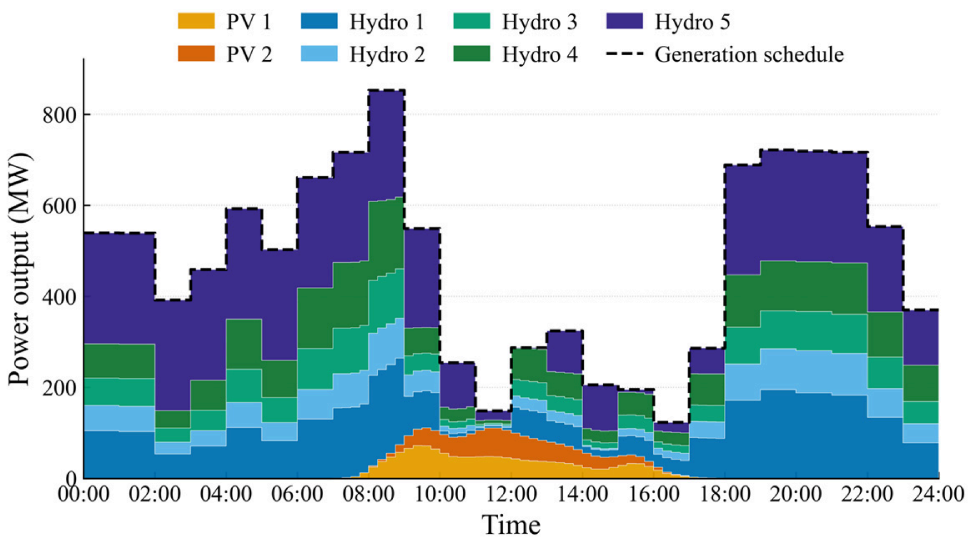


Figure 4. Typical-day generation schedule of a hydro-solar hybrid system.

To verify the adaptability of the proposed method under various operating scenarios of a cascaded hydro-photovoltaic complementary energy system, this study comprehensively considers indicators such as the hydroelectric load profile and the photovoltaic power output profile, and selects representative daily time series—which characterize the distinct operating characteristics of the hydro-photovoltaic complementary system under different load levels—as test scenarios. The operating characteristics of each representative day are presented in Table 1.

Table 1. Operational characteristics of typical days and hydro-photovoltaic complementary characteristics.

Typical Day	Hydrological Period	Photovoltaic Contribution Rate/%	Average Water and Electricity Net Load/MW	Maximum Net Water and Electricity Load/MW	Minimum Water and Electricity Net Load/MW	Net Load Peak-to-Valley Difference/MW	Pearson Correlation Coefficient	Primary Operational Characteristics
Day 1	low water period	39.85	119.5	242.4	12.9	229.5	−0.59	Photovoltaic power generation accounts for a high share of the total generation mix, while hydropower operates under low-load conditions.
Day 2	Pingshui Period	5.57	448.64	805.1	37.4	767.7	−0.58	Photovoltaic power generation accounts for a low share of the total energy mix, while hydropower operates under high load conditions and experiences relatively frequent fluctuations.
Day 3	wet season	18.48	474.4	679.5	329.5	350	−0.75	Hydropower plant operating at medium load with relatively stable performance.

3.2. Training Process Analysis and Strategy Selection

The model is trained offline using the PPO multi-agent reinforcement learning algorithm (the model parameters are provided in Appendix A). During the training phase, the agent interacts with the environment under randomly generated load scenarios to learn an unit output allocation strategy. In order to assess the effectiveness of the training and identify the optimal strategy, an offline evaluation is conducted using a deterministic strategy every 10 training rounds on a fixed operating condition dataset (which includes scenarios generated using fixed random numbers as well as single-period scenarios from a typical day), and the model's quality metrics for each evaluation round are recorded. Upon completion of training, the strategy with the best overall performance, as determined from the results of all previous evaluations, is selected as the final model for subsequent scheduling analysis.

Figure 5a and Figure 5b, respectively, show the evolution curves of the training sampling reward curve and the evaluation metrics over the number of training epochs during the model training process. It can be observed that the model's performance improves rapidly in the early stages of training. After approximately 2000 epochs, the model enters a stable phase and the curves become relatively flat, indicating that the model has fully converged and that the training was effective. Meanwhile, the model's performance on the aforementioned fixed evaluation scenario dataset gradually approaches the ideal scheduling benchmark (i.e., the DP-Equal scenario) as training progresses. At the final convergence point the deviation rate is approximately 0.21%, demonstrating that the trained policy exhibits excellent scheduling quality under the evaluation scenarios, and is capable of meeting the requirements of subsequent engineering analyses.

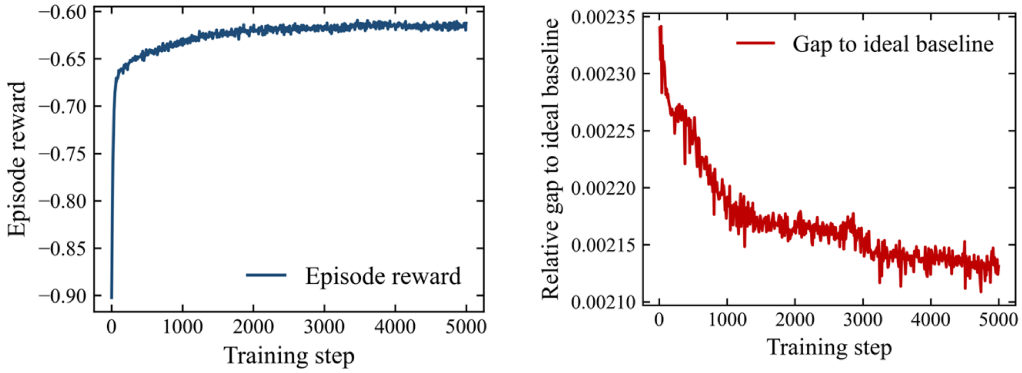


Figure 5. Training process and strategy selection. (**a**) Episode reward curves during training; (**b**) relative gap to the ideal baseline on the fixed evaluation scenario set. Smaller gap indicates better scheduling quality.

3.3. Single Typical Day Comparison Results

To verify the effectiveness of the proposed DP-MAPPO method in typical daily scheduling problems, this study conducts a comparative analysis of various scheduling methods under given daily load curves and water level boundary conditions. Figure 6 illustrates the power generation scheduling process for a typical day within the entire hydro-photovoltaic complementary energy system. Figure 7. illustrates the comparison of typical plant power output trajectories and vibration zones under different methods. The scheduling results for each method on a typical day are presented in Table 2.

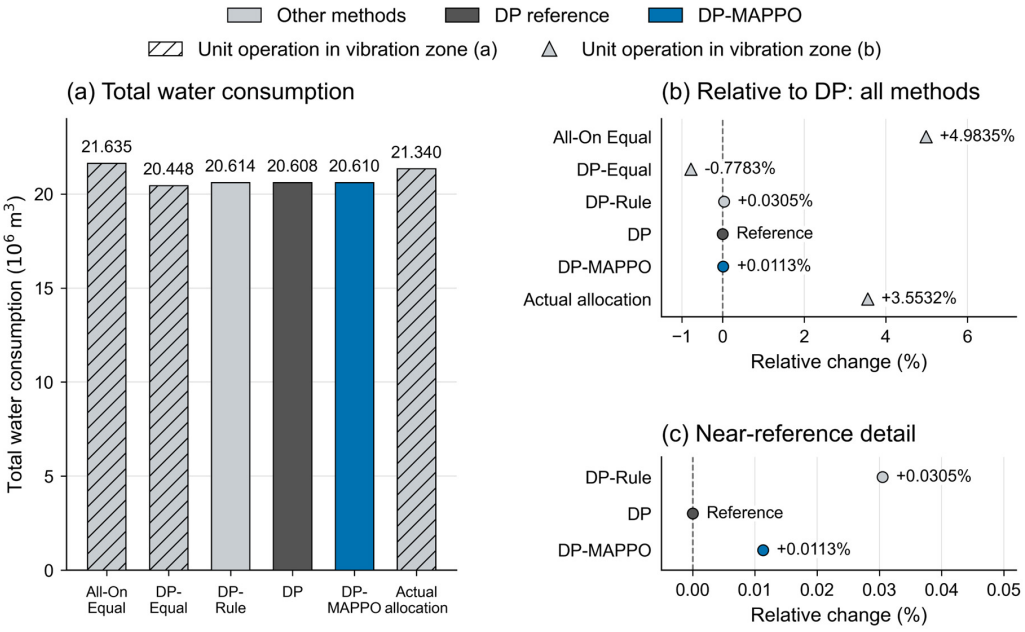


Figure 6. Comparison of water consumption and vibration-zone occurrence among allocation methods. (**a**) Total water consumption. (**b**) Reported relative changes with respect to the DP reference for all methods. (**c**) An expanded view of the near-reference results. Hatched bars in (**a**) and triangular markers in (**b**) indicate that at least one unit operates within its vibration zone during the evaluated period; these symbols do not quantify the duration or extent of such operation.

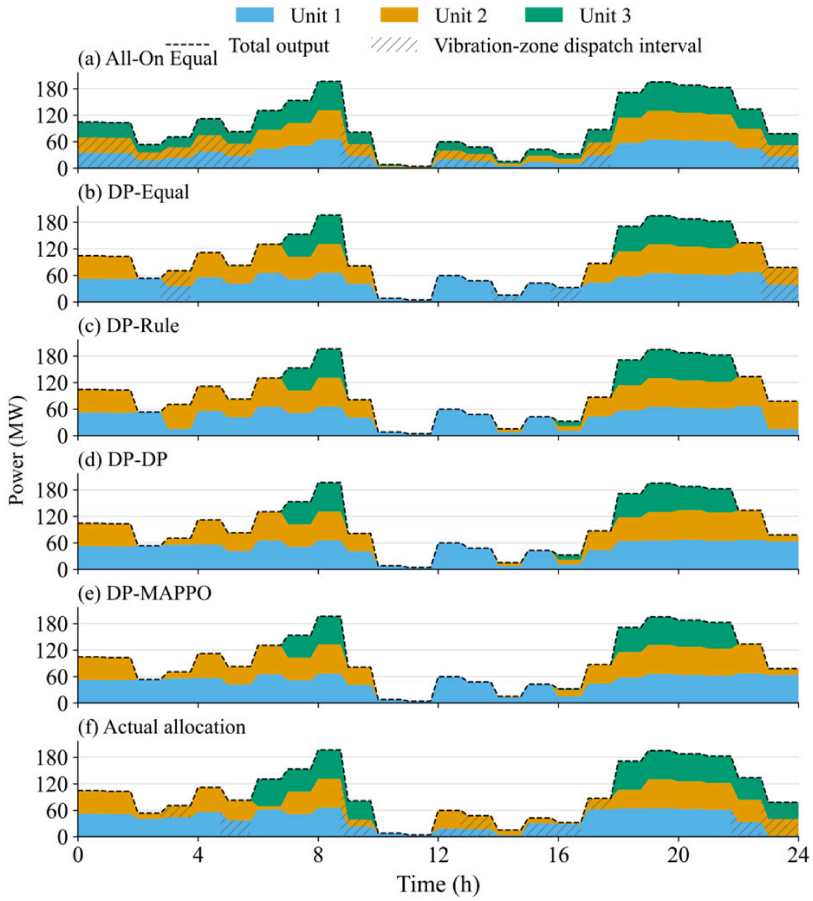


Figure 7. Comparison of typical plant power output trajectories and vibration zones under different methods (plant1).

Table 2. Scheduling Results under Different Methods.

Method	Total Water Consumption (m^3)	Water Consumption Relative Change Rate (Relative DP-DP Method)	Disrupts Stable Operating Conditions	Lower-Level Solution Type	Computation Time for a Single Scheduling Instance
All-On Equal	21,635,043.9	4.9835%	Units at Power Station 1–Power Station 5 operating within multi-period vibration range	Explicit computation	Subsecond
DP-Equal	20,447,660.10	−0.7783%	Units at Power Station 1–Power Station 5 operating within multi-period vibration range	Explicit computation	≈1 s
DP-Rule	20,614,347.00	0.0305%	None	Explicit computation	≈1 s
DP-DP (Feasible Optimal Solution)	20,608,051.9	/	None	Discrete optimization	Hundreds of seconds
DP-MAPPO	20,610,372.77	0.0113%	None	Deep Reinforcement Learning	Seconds
Actual allocation	21,340,303.00	3.5532%	Units at Power Station 1–Power Station 5 operating within multi-period vibration range	Recorded operation	/

Note: All algorithms were tested on the same computing platform. The host machine was equipped with an Intel Core i5 processor (12 cores and 16 threads) and 16 GB of RAM. The programs were executed in a Windows Subsystem for Linux 2 (WSL2) environment running Ubuntu 22.04, with 16 logical processors and approximately 7.6 GB of available memory. Python 3.11 was used. All scheduling optimization tasks were executed in a single-process mode, with the number of threads used by numerical computing libraries uniformly set to 1. No other high-load tasks were running during the tests. The offline training of the MAPPO policy took approximately 90 h. This training time can be substantially reduced by increasing training parallelism. It was recorded separately as a pre-deployment computational cost and was excluded from the solution time of individual scheduling cases. Provided that the power station and unit characteristics and constraint settings remain unchanged, the trained policy can be directly applied to subsequent scheduling tasks without retraining for each individual case. Subsequent scheduling is performed through upper-layer DP search and lower-layer policy forward inference, keeping the computational overhead of online solution procedures manageable and thereby balancing dispatching quality and computational efficiency.

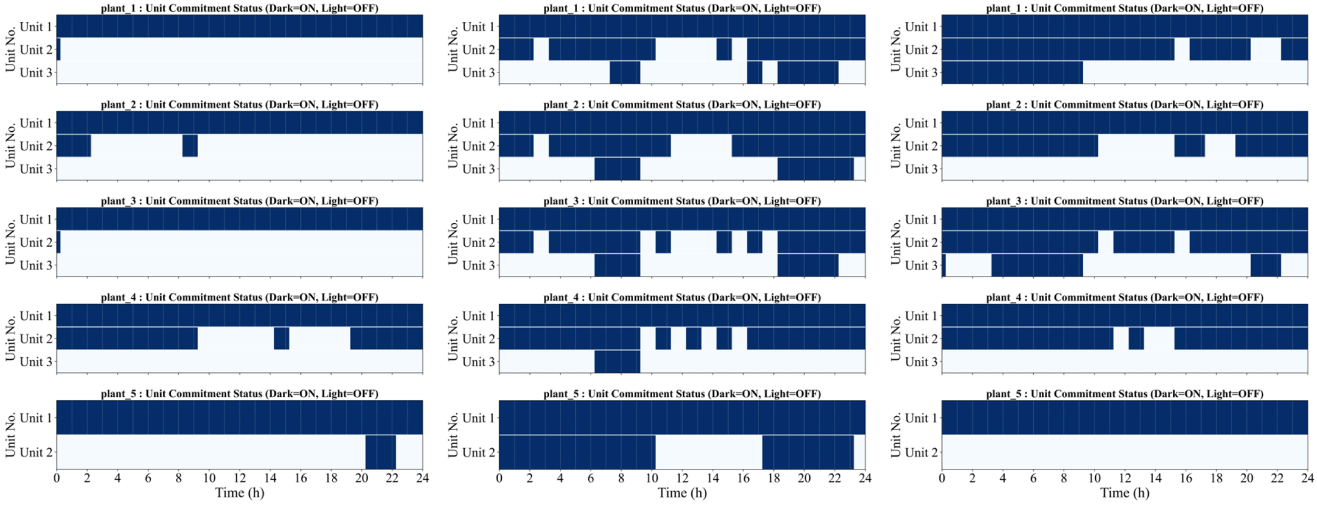
3.4. Calculation Results for Multiple Typical Days

Table 3 compares the actual scheduling results from the original engineering approach with those obtained using the method proposed in this study across various typical days. It can be observed that the proposed method yields scheduling solutions that satisfy the unit safety constraints for all these typical days. Compared to the original scheduling approach, the proposed method overall reduces the total daily water consumption. This indicates that the methodology—where the unit configuration for the power station is determined via upper-level dynamic programming (Figure 8) and the multi-unit output allocation is performed by the lower-level MAPPO (Figure 9)—can further unlock the synergistic operation potential among the generating units.

Table 3. Comparison of historical dispatching results under different typical-day scenarios with the dispatching results obtained using the proposed method.

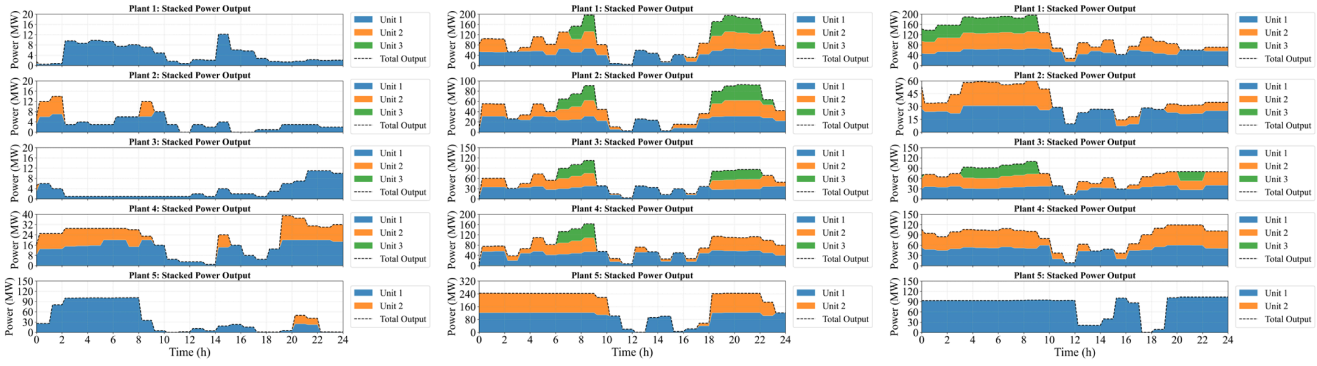
Typical Day	Method	Total Water Consumption (m^3)	Relative Water-Saving Rate/%	Total Number of Operational Cycles in the Vibration Zone of the Water–Energy Complementary System Units
Day 1	Actual project allocation	5,123,629.50	/	8
	Proposed method	4,654,565.43	9.15	0
Day 2	Actual project allocation	21,340,303.00	/	276
	Proposed method	20,610,372.77	3.42	0
Day 3	Actual project allocation	17,715,920.80	/	140
	Proposed method	17,600,898.02	0.65	0

Note: The relative water-saving rate here represents the relative value by which the proposed method reduces water consumption compared to the actual water allocation in the engineering project.



(**a**) (**b**) (**c**)

Figure 8. Optimized unit commitment schedules for five hydropower plants in a hydro–photovoltaic complementary system on three representative days: (**a**) Day 1; (**b**) Day 2; (**c**) Day 3. Each row corresponds to a hydropower plant. Dark and light-colored cells indicate the ON and OFF states of individual units, respectively., respectively.



(**a**) (**b**) (**c**)

Figure 9. Optimized unit power outputs for five hydropower plants in a hydro–photovoltaic complementary system on three representative days: (**a**) Day 1; (**b**) Day 2; (**c**) Day 3. Each row corresponds to a hydropower plant, and different colors distinguish individual units.

4. Discussion

4.1. Single Typical Day Results Analysis

This section uses the typical Day 2 scenario as an example to conduct a comparative analysis of the upper-level combinatorial optimization, the vibration zone constraints, and the lower-level allocation method.

(1) Role of upper-level unit configuration optimization: A comparison between the All-On Equal and DP-Equal approaches reveals that, in both cases, the lower-level units operate with equal output power without considering the vibration zone constraints. The sole difference lies in whether the upper-level unit configuration undergoes DP-based optimization. The full-activation scenario results in a water consumption of 21,635,043.9 m^3, which is 4.98% higher than the feasible optimal solution; after incorporating upper-level DP configuration optimization, the water consumption decreases to 20,447,660.10 m^3, representing a reduction of 0.778% relative to the feasible optimal solution. The difference between these two values is approximately 1,187,400 m^3, which corresponds to a potential water savings of about 5.76% compared to the feasible optimal solution. This demonstrates that by implementing appropriate start-stop scheduling at each time interval to ensure that the units operate as extensively as possible within their high-efficiency output range, significant water-saving benefits can be achieved.

(2) Impact of vibration zone constraints: The DP-Equal model neglects vibration zone constraints. Its "water-saving" result represents merely a theoretical upper bound and lacks practical engineering feasibility. When vibration zone constraints are incorporated (DP-Rule), the water consumption amounts to 20,614,347.00 m^3—a reduction of only 0.03% compared to the feasible optimal solution. The difference between the two values is approximately 166,700 m^3, which corresponds to about 0.81% of the feasible optimal solution. This indicates that incorporating non-convex vibration zone constraints to ensure the engineering feasibility of a proposed scheme entails an additional water consumption cost.

(3) Comparison of lower-level allocation methods: Under the condition that the upper-layer dynamic programming (DP) unit commitment optimization remains unchanged, three feasible lower-layer allocation strategies are compared. The rule-based mapping method, DP-Rule, prioritizes equal load allocation and then maps the resulting outputs to the nearest feasible boundaries. It yields a water consumption of 20,614,347.00 m^3. Although this method is straightforward to implement, highly interpretable, and computationally inexpensive, it does not perform a global search over combinations of feasible operating segments, leaving approximately 0.03% of the optimization potential unexplored. Its water consumption performance depends jointly on unit characteristics and the allocation rules. Therefore, it generally cannot guarantee optimality under arbitrary operating conditions, and its generalization capability is limited. The discrete exact method, DP-DP, enumerates all feasible-segment combinations across individual units to obtain the global optimum, achieving a water consumption of 20,608,051.9 m^3. However, among the compared methods, it requires the longest computation time. Although it provides a benchmark for the achievable allocation accuracy, its computational complexity and runtime increase exponentially with the number of units, making it difficult to apply to online analysis. In contrast, the proposed DP-MAPPO method retains the same upper-layer DP framework and learns the load allocation policy through offline training. During online operation, policy forward inference replaces rule-based allocation and optimization search in the lower layer. DP-MAPPO achieves a water consumption of 20,610,372.77 m^3, only 0.011% higher than the feasible optimum and approximately 0.019% lower than that of DP-Rule. These results indicate that the allocation policy learned by MAPPO through offline training closely approaches the optimum obtained by exact optimization in terms of solution quality, while substantially outperforming the heuristic rule-based method. In

terms of online computational efficiency, DP-MAPPO also offers a significant advantage over DP-DP, thereby balancing dispatching economy and computational efficiency.

(4) Comparison with actual operational performance. The actual water allocation scheme for this representative day consumed 21,340,303.00 m^3 of water, representing a 3.55% degradation compared to the feasible optimal solution. However, there was a potential risk that multiple units operating over multiple time periods would fall within the vibration zone. In contrast, the DP-MAPPO scheme in this study consumed 20,610,372.77 m^3 of water, achieving water savings of approximately 729,900 m^3 (a 3.42% reduction compared to the actual operational scenario) and eliminated all instances of operating outside the vibration zone. This demonstrates that the proposed framework offers significant room for optimization relative to the actual field operating conditions and possesses substantial practical value for engineering applications.

4.2. Results Analysis for Multiple Typical Days

To evaluate the adaptability and generalization capability of the proposed framework under different hydrological conditions and various solar–hydropower output combinations, three representative days were selected—namely, the dry-season day with a high PV penetration rate (Day 1), the average-water-season day with high load (Day 2), and the wet-season day with steady operation (Day 3)—and their performance was compared against the actual project allocation scheme; the results are presented in Table 3.

(1) The water-saving benefit exhibits significant dependence on operating conditions and is positively correlated with the required regulation margin. The relative water-saving rates for the three typical days are 9.15% (Day 1), 3.42% (Day 2), and 0.65% (Day 3), respectively. These rates decrease as the hydropower's net load level increases and the available regulation margin shrinks. Day 1 corresponds to the dry season, during which the photovoltaic generation contribution rate reaches as high as 39.85%, the average hydropower net load is only 119.5 MW, and the minimum net load is as low as 12.9 MW. The generating unit operates in a low-load regime over an extended period, deviating the most from its high-efficiency operating range, resulting in the greatest deviation between the actual dispatch schedule and the optimal schedule. Consequently, the potential for optimization under this method is greatest (9.15%). Day 3 falls within the wet season, during which the hydropower unit operates at a stable mid-range load (the peak-to-valley difference in net load is only 350 MW, the smallest among the three days). The ample inflow allows the unit to operate under relatively optimal conditions, and the actual operating scenario is already close to the economically optimal operating point. The water-saving rate is thus the lowest (0.65%). On Day 2, the peak-to-valley difference in net load is the largest (767.7 MW), involving frequent peak-shaving operations and abrupt changes in the unit's operating mode. The actual operation deviates from the high-efficiency optimal region, yielding a water-saving rate of 3.42%. Due to the highest absolute water consumption baseline (21.34 million m^3), the corresponding absolute water savings are also the most substantial. These patterns indicate that the water-saving benefit achieved by this method is not constant but increases in magnitude depending on the complexity of the dispatching conditions and the degree of deviation from the optimal operating scenario. This method demonstrates its greatest advantage under conditions characterized by significant hydropower peak-shaving pressure and the forced deviation of the generating unit from its high-efficiency operating range.

(2) The vibration zone safety constraints are strictly satisfied under all operating conditions. During the actual engineering scheduling scenarios for the three typical days, the units experienced operation within the vibration zone 8, 276, and 140 times, respectively. Notably, on Day 2, the number of exceedances reached as high as 276 times, indicating

that there is a widespread safety risk of units falling into the prohibited vibration zone during field operations under high-load and frequent peak-shaving conditions. In contrast, the method proposed in this paper yields zero exceedances within the vibration zone across all three types of typical days, and it does not require any empirical calibration tailored to specific operating conditions. This demonstrates that the constructive constraint mechanism—based on "feasible segment parameterization + segment combination projection" —fundamentally ensures the physical feasibility of the control actions. Its safety performance does not degrade even when the distribution of operating conditions changes, and it exhibits greater robustness and generalization capability compared to rule-based allocation approaches that rely on empirical data.

(3) Impact of solar–hydropower complementarity characteristics on dispatch economic efficiency. The Pearson correlation coefficients between net load and photovoltaic (PV) output across three typical days are -0.59, -0.58, and -0.75, respectively—all indicating a clear negative correlation: the higher the PV output, the lower the required net hydropower load for the system, highlighting the pronounced solar–hydropower complementarity feature. Among these, Day 3 exhibits the strongest negative correlation (-0.75), indicating that on this day the PV generation exerts the most significant "peak shaving" effect on hydropower; the hydropower ramp-up demand is thereby mitigated by the PV output, resulting in the most stable system operation—which explains the phenomenon of the lowest water conservation potential observed on this day.

5. Conclusions

This study proposes a dual-layer DP-MAPPO framework for the day-ahead economic dispatch of cascade hydro–photovoltaic complementary systems, combining dynamic programming (DP) for unit commitment optimization with multi-agent proximal policy optimization (MAPPO) for unit output allocation. The main findings are summarized as follows.

(1) The proposed method exhibits stable convergence and high solution accuracy. On the evaluation scenario set, the average deviation rate from DP-Equal after convergence is approximately 0.21%. Under the typical Day 2 operating condition, the deviation rate between the feasible optimal solution and DP-Equal is 0.78%, whereas that between DP-MAPPO and the feasible optimum is only 0.011%. Although DP-Equal provides a relaxed benchmark, it ignores vibration zone constraints and therefore does not always yield feasible operating solutions.

(2) DP-MAPPO balances solution accuracy and computational efficiency while satisfying hard constraints. Across all evaluated operating conditions, the number of vibration zone violations is zero. In the lower-level allocation comparison, the accuracy difference between MAPPO and the discrete exact method (DP-DP) is only 0.011%, while DP-MAPPO reduces water consumption by approximately 0.019 percentage points relative to DP-Rule. Unlike DP-DP, which enumerates feasible-segment combinations with exponentially increasing computational complexity, MAPPO uses offline training and online policy inference to obtain high-quality allocation decisions efficiently.

(3) Compared with historical actual dispatch, DP-MAPPO reduces water consumption by 9.15%, 3.42%, and 0.65% across the three typical operating periods representing dry, average, and wet conditions, respectively. Historical dispatch includes unit operation within vibration zones, whereas the proposed method produces schedules with zero vibration zone violations in all three cases. These results demonstrate improved water-use efficiency and operational feasibility for the evaluated scenarios.

(4) The current method is trained offline using operational data. Future research will incorporate intraday rolling prediction and online adaptive policy updating, while

considering additional factors such as unit start-stop frequency, ramping stability, and photovoltaic curtailment risk to improve adaptability under highly stochastic inflow and market conditions.

Author Contributions: Conceptualization, D.H., H.Z. and L.H.; Methodology, D.H., Q.Y. and Y.G.; Software, D.H. and H.Z.; Validation, D.H., Q.Y. and Y.G.; Formal analysis, H.Z. and Y.G.; Investigation, D.H. and H.Z.; Resources, Q.Y., H.Z. and L.H.; Data curation, Q.Y.; Writing—original draft, Q.Y.; Writing—review & editing, Q.Y. and A.I.; Visualization, D.H.; Supervision, H.Z. and L.H.; Project administration, H.Z.; Funding acquisition, H.Z. and L.H. All authors have read and agreed to the published version of the manuscript.

Funding: Study on Dynamic Instability and Abrasive Mechanisms of High-Head Impulse Hydraulic Machinery No. 52476026; Key Technologies and Applications for Collaborative Operation and Disaster Prevention and Control in Renewable Energy Facilities Based on Precision Micro-meteorological Monitoring and Prediction No. DTKY-2023-20532.

Data Availability Statement: For publicly available data, certain operational parameters of the hydropower stations, settings for dispatching model parameters, and selected research methodologies employed in this study are presented in the main text and the appendices; Regarding non-public data, some of the hydropower plant operating data, inflow data, and related operating parameters used in this study constitute actual operational records of specific power stations within the research region, rather than publicly available benchmark datasets generated by the authors. Consequently, access to such data is restricted by the data owners and the operational management authorities of the power stations, as well as by confidentiality requirements; therefore, the authors are unable to directly upload this data to a public data repository.

Acknowledgments: The research results presented in this paper are supported by Harbin Institute of Technology and China Datang Corporation Science & Technology Research Institute.

Conflicts of Interest: Authors Dongyang He and Haiku Zhang were employed by China Datang Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that this study received funding from China Datang Corporation. The funder had the following involvement with the study: Conceptualization, Validation, Formal Analysis, Resources, Visualization, Supervision, Project Administration, Funding Acquisition.

Appendix A

Table A1. Basic Parameters for Each Power Station.

Power Station Name	Number of Units	Total Installed Capacity (MW)	Vibration Zone Distribution (MW)	Regulatory Performance
Plant1	3	201	(15, 40)	year
Plant2	3	93	(10, 20)	daily
Plant3	3	120	(12, 25)	daily
Plant4	3	180	(20, 40)	daily
Plant5	2	246	(40, 80)	quarter
PV1	—	200	—	—
PV2	—	210	—	—

Table A2. Primary Parameter Settings for the Multi-Agent PPO Scheduling Model.

Project	Set Up
Reinforcement Learning Framework	Ray RLlib Multi-Agent PPO Framework
Agent Definition	Each power station acts as an agent.
state variable	Power plant load, unit start/stop status, and power plant head; constraint information is provided by the environmental model.
Action Variable	Initial output allocation for each unit; the action space is a continuous space.
Action Correction	Crop → Normalize → Map to feasible segment length
Safety Constraints	Unit output upper and lower limit constraints; vibration zone constraints
Safe handling procedure	Constructive Parametric Mapping + Euclidean Projection Correction
reward function	Water conservation rate (relative to baseline operating conditions)
network topology	A two-layer fully connected Actor-Critic network with hidden layer sizes of [256, 256].
learning rate	5×10^{-4}
discount factor	0.99
Cropping coefficient	0.2
Number of optimization iterations per round	10
Training batch size/Small batch size	24,000/4000
Strategy Implementation and Evaluation	During training, regular evaluations are performed on a fixed evaluation set; the optimal policy π^* (best checkpoint) and multiple feasible candidate policies are saved.

Table A3. Comparison Table of Economic Operation Scheduling Methods.

Method	Upper Unit Assembly	Lower-Level Force Distribution	Should Vibration Zone Constraints Be Considered?	Method Description
All-On Equal	All-On	Micro-increase Rate/Average Distribution	deny	Assume that all units in the power plant are operational throughout all dispatch periods, and that the total plant load is evenly distributed among the units according to their respective rated capacities. The dispatch optimization model is often used as a simplified representation of the power plant's operational characteristics.
DP-Equal	dynamic programming	Micro-increase Rate/Average Distribution	deny	The upper layer employs dynamic programming to solve the time-step-by-time-step unit start-up and shutdown scheduling problem for power plants; the lower layer, based on the equal incremental cost allocation principle, distributes the total plant target load proportionally among all units. In the example presented in this paper, all units within each power plant have equal capacity.

Table A3. *Cont.*

Method	Upper Unit Assembly	Lower-Level Force Distribution	Should Vibration Zone Constraints Be Considered?	Method Description
DP-Rule	dynamic programming	Micro-increase Rate/Average + Vibrational Zone Mapping	yes	This method builds upon the DP-Equal approach by further incorporating constraints from the vibration zone; specifically, if an allocated value falls within the prohibited vibration zone, it is corrected to the endpoint of the feasible segment closest to it in Euclidean distance—a commonly used engineering practice.
DP-DP		dynamic programming (accuracy 0.01 MW)	yes	The upper layer employs dynamic programming to determine the unit commitment; the lower layer further utilizes dynamic programming for the precise allocation of generation output among the generating units.
DP-MAPPO		MAPPO	yes	The primary method employed in this paper
Actual project allocation	Actual Procedures	Actual Procedures	yes	The actual historical operational dispatch records of the power station shall serve as the allocation plan.

References

1. Jin, X.; Liu, B.; Liao, S.; Cheng, C.; Zhang, Y.; Jia, Z. Assessing hydropower capability for accommodating variable renewable energy considering peak shaving of multiple power grids. *Energy* **2024**, *305*, 132283. [CrossRef]
2. Cheng, C.; Shen, J.; Wu, X. Short-term scheduling for large-scale cascaded hydropower systems with multivibration zones of high head. *J. Water Resour. Plan. Manag.* **2012**, *138*, 257–267. [CrossRef]
3. Parvez, I.; Shen, J.; Khan, M.; Cheng, C. Modeling and Solution Techniques Used for Hydro Generation Scheduling. *Water* **2019**, *11*, 1392. [CrossRef]
4. Huang, J.; Qin, H.; Yang, X.; Shen, K.; Yao, H.; Chang, X.; Li, G.; Gao, Y. Framework for short-term hydropower cascade–station–unit integrated multi-objective scheduling: Considering unit safety and economic efficiency. *J. Hydrol.* **2025**, *653*, 132756. [CrossRef]
5. Wang, Z.; Li, Y.; Wu, F.; Wu, J.; Shi, L.; Lin, K. Multi-objective day-ahead scheduling of cascade hydropower-photovoltaic complementary system with pumping installation. *Energy* **2024**, *290*, 130258. [CrossRef]
6. Shi, Y.; Wang, H.; Li, C.; Negnevitsky, M.; Wang, X. Stochastic optimization of system configurations and operation of hybrid cascade hydro-wind-photovoltaic with battery for uncertain medium-and long-term load growth. *Appl. Energy* **2024**, *364*, 123127. [CrossRef]
7. Shen, L.; Wang, Q.; Wan, Y.; Xu, X.; Liu, Y. Multi-timescale scheduling optimization of cascade hydro-solar complementary power stations considering spatio-temporal correlation. *Sci. Technol. Energy Transit.* **2025**, *80*, 17. [CrossRef]
8. Hu, S.; Miao, J.; Wu, J.; Zhao, L.; Wang, Y.; Meng, F.; Wei, C.; Zhang, X.; Zhu, B. Short-Term Optimal Scheduling of a Cascade Hydro-Photovoltaic System for Maximizing the Expectation of Consumable Electricity. *Processes* **2025**, *13*, 328. [CrossRef]
9. Su, C.; Li, L.; Zhang, T.; Sui, Q.; Yang, Y. A real-time scheduling framework of cascade hydropower-photovoltaic power complementary systems based on model predictive control. *Appl. Energy* **2025**, *392*, 126023. [CrossRef]
10. Shang, Y.; Lu, S.; Gong, J.; Liu, R.; Li, X.; Fan, Q. Improved genetic algorithm for economic load dispatch in hydropower plants and comprehensive performance comparison with dynamic programming method. *J. Hydrol.* **2017**, *554*, 306–316. [CrossRef]
11. Yuan, X.; Qin, H.; Cao, W.; Zhang, T.; Niu, X. Optimal scheduling study of hydro–solar complementary system based on improved beluga whale algorithm. *Water* **2025**, *17*, 878. [CrossRef]
12. Xie, M.; Liu, B.; Peng, Y.; Wu, D.; Qian, R.; Yang, F. Multi-Objective Optimal Dispatch of Hydro-Wind-Solar Systems Using Hyper-Dominance Evolutionary Algorithm. *Water* **2025**, *17*, 2127. [CrossRef]
13. Li, Y.; Hua, X.; Wang, L.; Lv, R.; Ouyang, C.; Zhang, F.; Yuan, F. Optimal Scheduling of Hydro–Thermal–Wind–Solar–Pumped Storage Multi-Energy Complementary Systems Under Carbon-Emission Constraints: A Coordinated Model and SVBABC Algorithm. *Electronics* **2025**, *14*, 4896. [CrossRef]

14. He, Y.; Xian, N. Short-term optimal scheduling of hydro–wind–PV and multi-storage complementary systems based on opposition-based learning PSO algorithm. *Appl. Energy* **2025**, *394*, 126125. [CrossRef]
15. Li, J.; Shi, L.; Fu, H. Multi-objective short-term optimal dispatching of cascade hydro–wind–solar–thermal hybrid generation system with pumped storage hydropower. *Energies* **2023**, *17*, 98. [CrossRef]
16. Bo, M. *Research on Whole-Lifecycle Coordinated Operation of Large-Scale Hydropower-Solar Hybrid Systems*; Wuhan University: Wuhan, China, 2019.
17. Xie, M.; Zhou, J.; Li, C.; Lu, P. Daily generation scheduling of cascade hydro plants considering peak shaving constraints. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04015072. [CrossRef]
18. Zeng, Y.; Wen, X.; Tan, Q.; Liu, Y.; Chen, X. Real-time load dispatch in hydropower plant based on D3QN-PER. *J. Hydrol.* **2023**, *625*, 130019. [CrossRef]
19. Wang, Z.; Li, Y.; Wu, F.; Shi, L.; Ding, R.; He, S. Deep reinforcement learning real-time dispatch approach for cascade hydropower with hybrid pumped-storage mitigating photovoltaic uncertainties. *Appl. Energy* **2026**, *408*, 127403. [CrossRef]
20. Ge, Y.; Xie, J.; Feng, S.; Chang, J.; Fu, D. Knowledge-embedded deep reinforcement learning for real-time dispatch of wind-solar-hydro hybrid generation systems. *J. Clean. Prod.* **2026**, *558*, 148255. [CrossRef]
21. Zhou, L.; Huo, L.; Liu, L.; Xu, H.; Chen, R.; Chen, X. Optimal power flow for high spatial and temporal resolution power systems with high renewable energy penetration using multi-agent deep reinforcement learning. *Energies* **2025**, *18*, 1809. [CrossRef]
22. Fang, X.; Hong, P.; He, S.; Zhang, Y.; Tan, D. Multi-layer energy management and strategy learning for microgrids: A proximal policy optimization approach. *Energies* **2024**, *17*, 3990. [CrossRef]
23. Zuo, J.; Ai, Q.; Wang, W.; Tao, W. Day-ahead economic dispatch strategy for distribution networks with multi-class distributed resources based on improved MAPPO algorithm. *Mathematics* **2024**, *12*, 3993. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.