Electronic Supplementary Materials

For https://doi.org/10.1631/jzus.A2500598

An optimization scheduling strategy for electric-heat-hydrogen integrated energy systems based on memory-enhanced deep reinforcement learning

Zhongli BAI[1,2], Qiang GAO[1,2,3], Hongzhi ZHANG[1,2], Junjie LIU[1,2], Yuehui JI[1,2], Yu SONG[1,2], Xu CHENG[1,4]

[1]*Tianjin Key Laboratory of New Energy Power Conversion, Transmission, and Intelligent Control, Tianjin University of Technology, Tianjin 300384, China*

[2]*School of Electrical Engineering and Automation, Tianjin University of Technology, Tianjin 300384, China*

[3]*Maritime College, Tianjin University of Technology, Tianjin 300384, China*

[4]*School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China*

Section S1 Unit Mathematical Model

A. Hydrogen energy system:

The hydrogen production of EL is determined by its input electrical power, expressed as:

$$m_{el}^{H_2}(t) = \eta_{el} P_{el}(t) \tag{S1}$$

where $P_{el}(t)$ is the operating power of the EL, η_{el} is the hydrogen production efficiency.

$$Q_{fc}(t) = \eta_{fc}^{h} P_{fc}(t) \tag{S2}$$

$$P_{fc}(t) = m_{fc}^{H_2}(t) \eta_{fc}^{e} LHV_{h2} \tag{S3}$$

where LHV_{h2} is the calorific value of hydrogen gas, $m_{fc}^{H_2}(t)$ is the hydrogen consumption of FC, η_{fc}^{e}, η_{fc}^{h} are the electrical generation and heat generation efficiencies, respectively.

$$Q_{fc}(t) = \eta_{fc}^{h} P_{fc}(t) \tag{S4}$$

$$P_{fc}(t) = m_{fc}^{H_2}(t) \eta_{fc}^{e} LHV_{h2} \tag{S5}$$

where LHV_{h2} is the calorific value of hydrogen gas, $m_{fc}^{H_2}(t)$ is the hydrogen consumption of FC, η_{fc}^{e}, η_{fc}^{h} are the electrical generation and heat generation efficiencies, respectively.

B. Cogeneration system

$$P_{gt}(t) = G_{gt}(t) \eta_{gt}^{e} \cdot LHV_{gas} \tag{S6}$$

$$Q_{whb}(t) = \eta_{gt}^{h} \cdot P_{gt}(t) \tag{S7}$$

where LHV_{gas} is the calorific value of gas, $G_{gt}(t)$ is gas consumption, $\eta_{gt}^{e}, \eta_{gt}^{h}$ are the power generation and heat generation efficiencies, respectively.

C. Gas boiler

$$Q_{gb}(t) = G_{gb}(t)\eta_{gb} \cdot LHV_{gas} \tag{S8}$$

where $G_{gb}(t)$ is gas consumption, η_{gb} is heat production efficiency.

D. Energy storage devices

$$SOC_{ES}(t+1) = SOC_{ES}(t) + \frac{\eta_{ES}^{ch} P_{ES}^{ch}(t) - \frac{P_{ES}^{dis}(t)}{\eta_{ES}^{dis}}}{E_{ES}^{cap}} \tag{S9}$$

where $SOC_{ES}(t+1)$, $SOC_{ES}(t)$ represent the stored energy at the next moment and the current moment, respectively, η_{ES}^{ch}, η_{ES}^{dis} is the charging and discharging efficiency. $P_{ES}^{ch}(t), P_{ES}^{dis}(t)$ is the charging and discharging power, E_{ES}^{cap} is the energy storage capacity. ES indicates energy storage categories, including EES, HS and H_2S.

Define the total charge-discharge intensity as:

$$P_{EES}^{\Sigma}(t) = P_{EES}^{ch} + P_{EES}^{dis} \tag{S10}$$

The SOC change amount is:

$$\Delta SOC_{EES}(t) = SOC_{EES}(t+1) - SOC_{EES}(t) \tag{S11}$$

For units with high safety margins for heat storage and hydrogen storage, the process can be simplified as follows [28]:

$$C_{HS}(t) = \alpha_{HS}(P_{HS}^{ch} + P_{HS}^{dis}) \tag{S12}$$

$$C_{H_2S}(t) = \alpha_{H_2S}(P_{H_2S}^{ch} + P_{H_2S}^{dis}) \tag{S13}$$

Total energy storage cost:

$$C_{ES}(t) = C_{EES}(t) + C_{HS}(t) + C_{H_2S}(t) \tag{S14}$$

Section S2 Proof of the mathematical properties of the B-IDR nonlinear mapping.

This appendix provides proofs of the mathematical properties stated for the B-IDR nonlinear mapping in Eqs. (S14) – (S16). Consider $x \in \{e, h\}$. Let $u_x \in [-1,1], p_x \in \mathbb{R}_+, \tau_x \in [0,1], \mathcal{L}_x(t) \in \mathbb{R}_+, \overline{p}_x > 0, \beta_x > 1$ .The mapping $\mathcal{F}_x(u_x, p_x, \tau_x)$ in Eq. (S16) is

$$\mathcal{F}_x(u_x, p_x, \tau_x) = \begin{cases} \tau_x \cdot \mathcal{L}_x(t) \cdot \left(\frac{p_x + u_x - \overline{p}_x}{\overline{p}_x}\right)^{\beta_x}, & p_x + u_x > \overline{p}_x \\ 0, & otherwise \end{cases} \tag{S15}$$

Define:

$$\Delta x(u_x, p_x) \triangleq \frac{p_x + u_x - \overline{p}_x}{\overline{p}_x}, [a]_+ \triangleq \max\{a, 0\} \tag{S16}$$

Then Eq. (S1) can be equivalently written as:

$$\mathcal{F}_x(u_x, p_x, \tau_x) = \tau_x \cdot \mathcal{L}_x(t) \cdot [\Delta x(u_x, p_x)]_+^{\beta_x} \tag{S17}$$

Since $\mathcal{R}(\mathbf{u}, \mathbf{p}, \boldsymbol{\tau}) = [\mathcal{F}_e(\cdot), \mathcal{F}_h(\cdot)]^{\mathrm{T}}$ in Eq. (S15) is a component-wise stacking of $\mathcal{F}_x$, it suffices to prove the properties for $\mathcal{F}_x$.

Proposition 1 (non-negativity): $\mathcal{F}_x(u_x, p_x, \tau_x) \geq 0$ for all admissible u_x, p_x, τ_x.

Proof: By definition, $\tau_x \geq 0, \mathcal{L}_x(t) \geq 0$, and $[\Delta x]_+ \geq 0$. For $\beta_x > 1, [\Delta x]_+^{\beta_x} > 0$. Hence $\mathcal{F}_x(u_x, p_x, \tau_x) \geq 0$.

Proposition 2 (Continuity): For any $\beta_x > 1$, $\mathcal{F}_x$ is continuous in (u_x, p_x, τ_x).

Proof: $\Delta x(u_x, p_x)$ is affine in (u_x, p_x) and thus continuous. The rectifier $[\cdot]_+$ is continuous. The mapping $y \to y^{\beta_x}$ is continuous for $y \geq 0$. Multiplication by $\tau_x \mathcal{L}_x(t)$ preserves continuity.

Proposition 3 (Local Lipschitz on bounded sets): Assume $\beta_x > 1$. For any bounded domain:

$$\mathcal{D} = \{(u_x, p_x, \tau_x) : |u_x| \leq U, 0 \leq p_x \leq P, 0 \leq \tau_x \leq 1\} \tag{S18}$$

there exists a constant $L_{\mathcal{D}} > 0$ such that for all $(u_1, p_1, \tau_1), (u_2, p_2, \tau_2) \in \mathcal{D}$

$$|\mathcal{F}_x(u_1, p_1, \tau_1) - \mathcal{F}_x(u_2, p_2, \tau_2)| \leq L_{\mathcal{D}}(|u_1 - u_2| + |p_1 - p_2| + |\tau_1 - \tau_2|) \tag{S19}$$

Proof: On $\mathcal{D}$, Δx is bounded since $|u_x| \leq U$ and $0 \leq p_x \leq P$:

$$|\Delta x(u_x, p_x)| \leq \frac{P + U + \overline{p}_x}{\overline{p}_x} \triangleq M \tag{S20}$$

Consider $g(y) = [y]_+^{\beta_x}$. When $\beta_x > 1$, g is Lipschitz on $[-M, M]$. In particular, for $y > 0, g'(y) = \beta_x y^{\beta_x - 1} \leq \beta_x M^{\beta_x - 1}$, and for $y \leq 0, g(y) = 0$. Thus for all $y_1, y_2 \in [-M, M]$:

$$|g(y_1) - g(y_2)| \leq \beta_x M^{\beta_x - 1} |y_1 - y_2| \tag{S21}$$

Moreover,

$$|\Delta x(u_1, p_1) - \Delta x(u_2, p_2)| = \frac{|(u_1 - u_2) + (p_1 - p_2)|}{\overline{p}_x} \leq \frac{|(u_1 - u_2)| + |(p_1 - p_2)|}{\overline{p}_x} \tag{S22}$$

Using Eq. (S3) and the triangle inequality,

$$\begin{aligned}\left|\mathcal{F}_x(u_1,p_1,\tau_1)-\mathcal{F}_x(u_2,p_2,\tau_2)\right| &\le \left|\tau_1-\tau_2\right|\mathcal{L}_x(t)g(\Delta_1)+\tau_2\mathcal{L}_x(t)\left|g(\Delta_1)-g(\Delta_2)\right| \\ &\le \mathcal{L}_x(t)M^{\beta_x}\left|\tau_1-\tau_2\right|+\mathcal{L}_x(t)\beta_x M^{\beta_x-1}\left|\Delta_1-\Delta_2\right|\end{aligned} \tag{S23}$$

where $\Delta_i=\Delta_x(u_i,p_i)$ and we used $\tau_2\le 1, \left|g(\Delta_1)\right|\le M^{\beta_x}$, and Eq. (S7). Substituting Eq. (S8) yields Eq. (S5) with a finite constant $L_{\mathcal{D}}$.

Proposition 4 (First-order differentiability at the boundary): If $\beta_x>1$, then $\mathcal{F}_x$ is continuously differentiable (C^1) with respect to (u_x,p_x), including at the boundary $u_x, p_x=\overline{p}_x$.For $\Delta_x>0$:

$$\frac{\partial\mathcal{F}_x}{\partial u_x}=\tau_x\mathcal{L}_x(t)\beta_x\frac{1}{\overline{p}_x}\Delta_x^{\beta_x-1}, \frac{\partial\mathcal{F}_x}{\partial p_x}=\tau_x\mathcal{L}_x(t)\beta_x\frac{1}{\overline{p}_x}\Delta_x^{\beta_x-1} \tag{S24}$$

and for $\Delta_x\le 0, \dfrac{\partial\mathcal{F}_x}{\partial u_x}=\dfrac{\partial\mathcal{F}_x}{\partial p_x}=0$.

Proof: For $\Delta_x\le 0, \mathcal{F}_x=0$, so the derivatives are 0. For $\Delta_x>0, \mathcal{F}_x=\tau_x\mathcal{L}_x(t)\Delta_x^{\beta_x}$, and Eq. (S10) follows by the chain rule. At $\Delta_x\downarrow 0^+$, the right-hand side of Eq. (S10) tends to 0 if and only if $\beta_x>1$, which matches the derivative from the $\Delta_x\le 0$ side. Hence the first derivatives exist and are continuous at the boundary.

Proposition 5 (Second-order behavior): For $\Delta_x>0$, the second partial derivatives satisfy:

$$\frac{\partial^2\mathcal{F}_x}{\partial u_x^2}=\tau_x\mathcal{L}_x(t)\beta_x(\beta_x-1)\frac{1}{\overline{p}_x^2}\Delta_x^{\beta_x-2}, \frac{\partial^2\mathcal{F}_x}{\partial p_x^2}=\tau_x\mathcal{L}_x(t)\beta_x(\beta_x-1)\frac{1}{\overline{p}_x^2}\Delta_x^{\beta_x-2} \tag{S25}$$

Therefore, $\mathcal{F}_x$ is generally not twice continuously differentiable at the boundary when $1<\beta_x<2$: if $1<\beta_x<2$, Eq. (S11) diverges as $\Delta_x\downarrow 0^+$; if $\beta_x=2$, Eq. (S11) approaches a positive constant on the right side while it equals 0 on the left side, yielding a jump discontinuity. When $\beta_x>2$,the right-side second derivatives tend to 0, and the mapping becomes C^2 at the boundary.

Proof: Differentiate Eq. (S10) once more for $\Delta_x>0$ to obtain Eq. (S11), and examine the limit as $\Delta_x\downarrow 0^+$. The left-side second derivatives are identically 0 because $\mathcal{F}_x=0$ for $\Delta_x\le 0$.

Section S3 The cost of the sub-items

A. Energy purchase costs:

$$C_{ebe}(t)=p_e(t)\cdot P_{grid}(t)+p_{gas}(t)\cdot Q_{gas}(t)+p_h(t)\cdot Q_{hs}(t) \tag{S26}$$

where $P_{grid}(t)$, $Q_{gas}(t)$, and $Q_{hs}(t)$ represent purchased electricity, gas, and heat, respectively.

B. Carbon emission costs:

$$C_{CO_2}(t) = \pi_{CO_2} \cdot [\lambda_{gas}(G_{gt}(t) + G_{\mathrm{gb}}(t)) + \lambda_{grid} P_{grid}(t)] \quad \text{(S27)}$$

where π_{CO_2} is the unit carbon price, λ_{gas} and λ_{grid} are the carbon emission coefficients for gas and the power grid, and $G_{\bullet}$ is the gas consumption.

C. Equipment operating and maintenance (O&M) costs:

$$C_{dev}(t) = C_{gt}(t) + C_{gb}(t) + C_{whb}(t) + C_{fc}(t) + C_{el}(t) + C_{ES}(t) \quad \text{(S28)}$$

$$C_{gt}(t) = a_{gt} P_{gt}(t)^2 + b_{gt} P_{gt}(t) + c_{gt} \quad \text{(S29)}$$

$$C_{gb}(t) = b_{gb} Q_{gb}(t) + c_{gb} \quad \text{(S30)}$$

$$C_{whb}(t) = b_{whb} Q_{whb}(t) \quad \text{(S31)}$$

$$C_{fc}(t) = a_{fc} P_{fc}(t)^2 + b_{fc} P_{fc}(t) + c_{fc} \quad \text{(S32)}$$

$$C_{el}(t) = b_{el} P_{el}(t) \quad \text{(S33)}$$

where $a_{\bullet}, b_{\bullet}, c_{\bullet}$ is the cost parameters.

D. B-IDR cost items:

$$C_{DR}(t) = \sum_{x \in \{e,h\}} [C_x^{inc}(t) + C_x^{unsat}(t)] \quad \text{(S34)}$$

E. Penalties for energy abandonment:

$$C_{curt}(t) = \theta_{ren} \cdot P_{curt(t)} + \theta_{heat} \cdot Q_{curt(t)} + \theta_{H_2} \cdot H_{curt(t)} \quad \text{(S35)}$$

The curtailment penalty weight θ is set through a preliminary tuning study under the adopted reward scaling, with the purpose of discouraging avoidable renewable curtailment while keeping the learning signal well-conditioned and avoiding an overly conservative policy dominated by penalty terms. Sensitivity checks in a reasonable neighborhood around the tuned value suggest that the key economic conclusions are not materially affected: moderate variations primarily shift the absolute curtailment–cost trade-off (i.e., slightly changing the realized curtailment level and the associated purchase/carbon cost components), while leaving the main comparative outcomes and overall performance trends unchanged.

Section S4 The cost of the sub-items

A. Electrical power balance constraint: the system's electrical load is supplied by a combination of renewable energy, local generators, EES discharge, and purchased electricity, satisfying the following power conservation requirements:

$$P_{grid}(t) + P_{wt}(t) + P_{pv}(t) + P_{gt}(t) + P_{fc}(t) + P_{bat}^{dis}(t) = \mathcal{L}_e^{adj}(t) + P_{bat}^{ch}(t) + P_{el}(t) \quad \text{(S36)}$$

B. Heat power balance constraint: the heat demand side is coupled with WHB, FC waste heat, GB, HS, and external heat supply to meet:

$$Q_{grid}(t)+Q_{whb}(t)+Q_{fc}(t)+Q_{gb}(t)+P_{hs}^{dis}(t)=\mathcal{L}_{hs}^{adj}(t)+P_{hs}^{ch}(t) \tag{S37}$$

C. Hydrogen energy production and consumption constraints: hydrogen production by EL, hydrogen consumption by FC, and H_2S are dynamically coupled to form the following flow balance relationship:

$$m_{el}^{H_2}(t)+P_{H_2S}^{dis}(t)=m_{fc}^{H_2}(t)+P_{H_2S}^{ch}(t) \tag{S38}$$

When the hydrogen consumption of the FC is less than the hydrogen production of the EL, the excess hydrogen is stored in the H_2S. When the hydrogen supply is insufficient, the system first draws on the H_2S to supplement the supply. If the total hydrogen source still cannot meet the demand, the FC will automatically reduce its output.

D. Operating limits and ramp rate limits: to ensure that critical units of the system operate within safe limits, upper and lower limits are set for output power, and ramp rate restrictions are introduced for gas-fired equipment with inertia. The output power of all types of units should meet the following boundary conditions:

$$\begin{cases} P_{gt}^{\min}(t)\le P_{gt}(t)\le P_{gt}^{\max}(t) \\ P_{fc}^{\min}(t)\le P_{fc}(t)\le P_{fc}^{\max}(t) \\ P_{el}^{\min}(t)\le P_{el}(t)\le P_{el}^{\max}(t) \\ Q_{gb}^{\min}(t)\le Q_{gb}(t)\le Q_{gb}^{\max}(t) \end{cases} \tag{S39}$$

$$\begin{cases} \triangle P_{gt}^{\min}(t)<P_{gt}(t)-P_{gt}(t-1)<\triangle P_{gt}^{\max}(t) \\ \triangle Q_{gb}^{\min}(t)<Q_{gb}(t)-Q_{gb}(t-1)<\triangle Q_{gb}^{\max}(t) \end{cases} \tag{S40}$$

For energy storage units, their charging and discharging power must simultaneously meet the rated capacity of the equipment and mutual exclusion logic constraints. All energy storage units meet the upper and lower limit requirements:

$$SOC_{ES}^{\min}(t)<SOC_{ES}(t)<SOC_{ES}^{\max}(t) \tag{S41}$$

$$\begin{cases} P_{EES}^{ch,\min}(t)<P_{EES}^{ch}(t)<P_{EES}^{ch,\max}(t) \\ P_{EES}^{dis,\min}(t)<P_{EES}^{dis}(t)<P_{EES}^{dis,\max}(t) \\ P_{EES}^{ch}(t)\cdot P_{EES}^{dis}(t)=0 \\ P_{HS}^{ch,\min}(t)<P_{HS}^{ch}(t)<P_{HS}^{ch,\max}(t) \\ P_{HS}^{dis,\min}(t)<P_{HS}^{dis}(t)<P_{HS}^{dis,\max}(t) \\ P_{HS}^{ch}(t)\cdot P_{HS}^{dis}(t)=0 \\ P_{H_2S}^{ch,\min}(t)<P_{H_2S}^{ch}(t)<P_{H_2S}^{ch,\max}(t) \\ P_{H_2S}^{dis,\min}(t)<P_{H_2S}^{dis}(t)<P_{H_2S}^{dis,\max}(t) \\ P_{H_2S}^{ch}(t)\cdot P_{H_2S}^{dis}(t)=0 \end{cases} \tag{S42}$$

Although specifies mutual exclusion at the model level, a continuous-action RL policy may output small simultaneous charge/discharge values due to approximation or numerical noise. In the environment step, we therefore apply an explicit

feasibility layer: the storage action is mapped to a mutually exclusive charging/discharging command, and then clipped to satisfy power limits and SOC bounds before state transition and cost settlement. Concretely, given the raw SAC output $a_{EES}^{ch}, a_{EES}^{dis}$, we enforce mutual exclusion by keeping only a single direction (e.g. $P_{EES}^{ch} = \max(a_{EES}^{ch} - a_{EES}^{dis}, 0)$). And then clipped to satisfy power limits and SOC bounds before state transition and cost settlement. This guarantees that the executed dispatch trajectory respects mutual exclusion and operational limits during the offline evaluation.

Section S5 Experimental Data and Parameters

The PV/WT generation profiles (Fig. S1) and the demand-side load profiles (Fig. S2) are derived from measured operational data collected in a representative operating day in October 2024 (local time, UTC+8). To protect data confidentiality, the measured profiles are uniformly rescaled, and we have verified that this proportional rescaling preserves the internal consistency of the time-series patterns and their relative relationships. The underlying measured dataset can be made available upon reasonable request by contacting the corresponding author, subject to an approval process. The time-of-use energy prices used for settlement are shown in Fig. S3.

To simulate uncertainty, a 10% relative random disturbance is applied to the input profiles (PV, WT, electrical load, and heat load) in a pointwise multiplicative manner. For each perturbed series x_t and each hour t, we sample an independent disturbance factor $\delta_{x,t} \sim \mathcal{U}(-0.1, 0.1), \text{seed} = 2025$, construct the perturbed profile as $\tilde{x}_t = x_t(1+\delta_{x,t})$, and the perturbed values are then projected to a physically feasible range by $\tilde{x}_t \leftarrow \min\{\max(\tilde{x}_t, 0), x^{\max}\}$. where $x^{\max}$ denotes the corresponding upper bound. All results reported under the "10% disturbance" setting use the above perturbation rule and the fixed seed to ensure reproducibility. Equipment capacity, efficiency, and other constraints are shown in Tables S1 and S2; cost coefficients are shown in Table S3; other parameter settings are shown in Table S4.

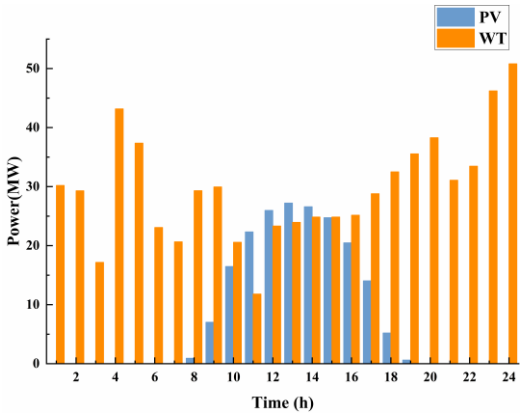


Fig. S1 Renewable energy inputs. (Representative day in October 2024; 1-h resolution.)

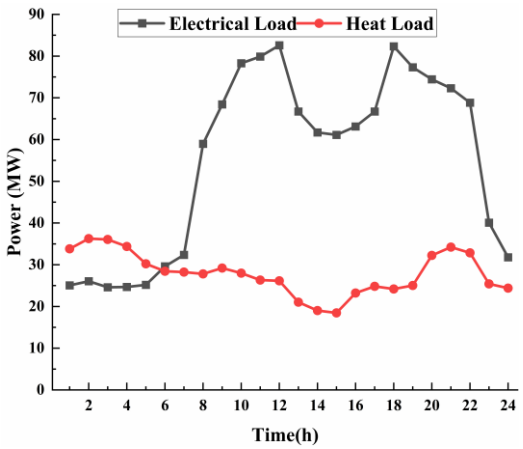


Fig. S2 The demand-side hourly load. (Representative day in October 2024; 1-h resolution.)

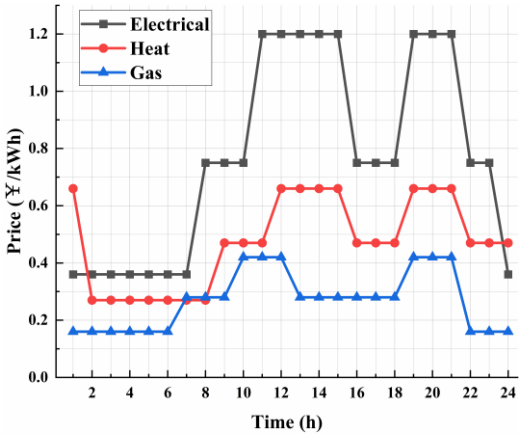


Fig. S3 The time-of-use energy prices. (Representative day in October 2024; 1-h resolution.)

Table S1 Energy coupling device parameters

Device	Capacity (MW)	Efficiency (Electric/Heat)	Climbing rate (MW)
GT	20	0.3/0.85	4
GB	30	-/0.92	6
EL	15	0.87/-	-
FC	10	0.6/0.4	-
WT	130	-	-
PV	50	-	-

Table S2 Energy storage equipment parameters

Device	Capacity (MW)	SOC constraints	Maximum power constraints	Charge-discharge efficiency
EES	10	0.2-0.8	3	0.95/0.90
HS	20	0.2-0.8	10.5	0.95/0.90
H_2S	30	0.2-0.8	9	0.95/0.90

Table S3 Energy coupling equipment cost parameters

Device	Second coefficient (CNY/kWh)	Single coefficients (CNY /kWh)	Constant
GT	0.0009	0.063	0.01
GB	-	0.002	0.0125
FC	0.0001	0.004	0.001
WHB	-	0.029	-
EL	-	0.026	-

Table S4 The other parameters (kWh)

Parameters	Value (CNY /kWh)	Parameters	Value (CNY /kWh)
λ_{gas}	2.16	α_{HS}	0.05
λ_{grid}	0.53	α_{H_2}	0.01
π_{CO_2}	0.3	β_x	1.5
LHV_{gas}	9.72	κ_x	1e-3
LHV_{h2}	3	θ_{ren}	3.08e-3
α_{ESS}	0.1	θ_{heat}	9.4e-4
β_{ESS}	1.2	θ_{H_2}	5.4e-4
ϖ_{ESS}	-0.8		

Table S5 Key hyperparameters for algorithmic comparison

Item	SAC	DDPG	TD3	PPO
learning rate	1e-3	1e-3	1e-3	3e-4
Batch size	256	256	256	256
Discount factor	0.98	0.98	0.98	0.98
Soft update	0.006	0.006	0.006	/
Replay buffer size	1e6	1e6	1e6	/
Exploration noise	policy stochasticity	Gaussian (0.15)	Gaussian (0.15)	policy stochasticity
TD3 policy delay	/	/	2	/
TD3 target policy noise / clip	/	/	0.25/0.6	/
PPO clip	/	/	/	0.25
PPO GAE	/	/	/	0.95

Section S6 Sensitivity to instantaneous load reduction cap

The backlog state variable $\mathrm{B}_x(t)$ characterizes the cumulative deferred flexible demand (i.e., the remaining unrecovered reduction) induced by demand-side load reduction and its subsequent replenishment. To examine whether the reported economic outcomes depend on the chosen instantaneous reduction bound, we vary the per-time-step reduction cap to 10%/20%/30% of the flexible load under identical data, equipment constraints, and learning budget, while keeping the cumulative rebound (backlog) capacity fixed at 60% of the flexible load. The resulting cost breakdown and backlog utilization are summarized in Table S6.

Table S6 Sensitivity to instantaneous load reduction cap ($\times 10^4$ CNY)

Instantaneous cap	Total cost	Carbon cost	Energy procurement costs	Incentive cost	O&M	Penalty	Max $\mathrm{B}_x(t)$
10%	22.76	3.422	13.10	0.989	4.623	0.622	0.22
20%	21.82	3.168	12.36	1.237	4.577	0.478	0.28
30%	21.17	2.978	11.87	1.422	4.544	0.359	0.35

As shown in Table S6, tightening the instantaneous load reduction cap (10%) restricts short-term flexibility, which increases reliance on grid purchase and weakens peak-shaving capability; consequently, Purchase and Carbon costs rise and the Penalty term increases, even though Incentive payouts decrease because the admissible load reduction is more limited.

Conversely, relaxing the cap (30%) enlarges the feasible demand-side shifting space, reducing Purchase/Carbon costs and lowering Penalty at the expense of higher Incentive payments, reflecting a clear economics–flexibility trade-off. Meanwhile, max $B_x(t)$ stays well below the bound in (42), indicating that the 60%-of-flexible-load limit mainly serves as an intertemporal recoverability constraint (to avoid excessive rebound pressure and demand-side latency), whereas the per-step reduction cap in (41) more directly governs the realized trade-off. Although the 30% cap achieves the lowest total cost, we select 20% as the baseline because it offers a more conservative and implementable flexibility envelope—delivering most of the economic benefit while avoiding a disproportionately higher incentive budget and more aggressive demand-side intervention implied by the 30% setting.